Technische Universität München
TUM School of Computation, Information and Technology

Finite Blocklength Codes for Secure Communication

Anna Frank

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades einer:

Doktorin der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr.-Ing. Georg Sigl
Prüfer der Dissertation:

1. Prof. Dr.-Ing. Dr.rer.nat. Holger Boche
2. Prof. Dr.-Ing. Aydin Sezgin

Die Dissertation wurde am 02.06.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 12.12.2022 angenommen.

Abstract

We are interested in secure communication without sharing a secret key. Currently, the construction of an explicit and practical secure encoder and decoder with an optimal performance is still an unsolved problem in the general case, except for some special cases.

In the first part of the thesis, we constructed codes for secure transmission for a so-called wiretap channel model. In the model there is a third party listening to the conversation between sender and receiver. The channels between the transmitter and receivers are noisy in the general case, assuming that for physical reasons the eavesdropper's channel is noisier than that of the legitimate receiver. We have constructed codes with a small Galois field size (which are preferably binary) for simplified wiretap models. In the simplified wiretap model, bursts of erasure occur in the channel to the legitimate transmitter, and the eavesdropper is able to observe an interval of a prescribed length noiselessly. Purposefully, codes were constructed that provide perfect security (strongest security metric) and error-free decoding, and can be transmitted at the maximum possible secrecy rate.

In the second part of the thesis, we considered the so-called modular wiretap coding scheme for secure transmission. The modular wiretap coding scheme consists of three layers. The first layer is for secure transmission, realized by a randomized function. The second and third layers are for reliable transmission, realized by a conventional error-correcting code and a modulation scheme, respectively. The advantage of the modular scheme is that no new error-correcting codes need to be constructed and it can be integrated into existing systems without the need for costly system modifications. We analyzed the modular wiretap code for the AWGN channel and implemented it in Matlab. We used the 3GPP standard for reliable transmission and a "Universal Hash Function" (UHF) in the first layer. The eavesdropper uses the maximum likelihood (ML) test as an attack strategy. We considered the distinguishing security which is equivalent to the semantic security in the asymptotic case. The distinguishing security can be assessed by the probability of error. We have seen that for a given signal-to-noise ratio (SNR) with increasing randomness at the encoder, the error probability converges towards the maximum possible error probability.

Zusammenfassung

Wir sind an einer sicheren Kommunikation ohne gemeinsame Nutzung eines geheimen Schlüssels interessiert. Derzeit ist die Konstruktion expliziter und praktischer sicherer Einkodierer und Dekodierer mit optimaler Leistung im allgemeinen Fall noch ein ungelöstes Problem, abgesehen von einigen Spezialfällen.

Im ersten Teil der Thesis konstruieren wir Kodes zur sicheren Übertragung für ein sogenanntes Wiretap Kanal Model. In dem Model gibt es eine dritte Partei die der Unterhaltung zwischen Sender und Empfänger lauscht. Die Kanäle zwischen Sender und Empfängern sind im allgemeinen Fall verrauscht, wobei angenommen wird, dass aus physikalischen Gründen der Kanal des Lauschers verrauschter ist als der des legitimen Empfängers. Wir haben für vereinfachte Wiretap Modelle Kodes mit kleiner Galois-Feldgröße (die vorzugsweise binär sind) konstruiert. In dem vereinfachten Wiretap Model treten im Kanal zum legitimen Sender gebündelte Löschungen auf, und der Lauscher ist in der Lage ein Intervall einer vorgeschriebenen Länge rauschfrei zu beobachten. Gezielt wurden Kodes konstruiert die perfekte Sicherheit (stärkste Sicherheitsmetrik) und fehlerfreie Dekodierung gewährleisten und mit maximal möglicher sicheren Rate übertragen werden können.

Im zweiten Teil der Thesis haben wir das sogenannte modulare Wiretap Kodierungs Schema zur sicheren Übertragung betrachtet. Das modulare Wiretap Kodierungs Schema besteht aus drei Schichten. Die erste Schicht dient der sicheren Übertragung, realisiert durch eine randomisierte Funktion. Die zweite und dritte Schicht dient der zuverlässigen Übertragung, realisiert durch jeweils einem konventionellen Fehlerkorrigierenden Kode und einem Modulationsschema. Der Vorteil des modularen Schemas ist, dass keine neuen fehlerkorrigierenden Kodes konstruiert werden müssen und es integriert werden kann in bestehende Systeme ohne das System aufwendig anpassen zu müssen. Wir haben den modularen Wiretap Kode für den AWGN Kanal analysiert und in Matlab implementiert. Zur zuverlässigen Übertragung haben wir den 3GPP Standard verwendet und in der ersten Schicht eine Universal Hash Function (UHF). Der Lauscher verwendet den Maximum-Likelihood (ML) Test als Angriffsstrategie. Wir haben die differenzierende Sicherheit betrachtet die äquivalent zur semantischen Sicherheit im asymptotischen Fall ist. Die differenzierende Sicherheit kann über die Fehlerwahrscheinlichkeit bewertet werden. Wir haben gesehen, dass für gegebenes Signal Rausch Verhältnis (SNR) mit wachsendem Zufall am Einkodierer die Fehlerwahrscheinlichkeit gegen die maximal mögliche Fehlerwahrscheinlichkeit konvergiert.

Acknowledgements

First and foremost I thank my advisor Prof. Holger Boche, for giving me the opportunity to do a phd at his chair and for all his advice and motivation. I would like to express my profound and sincere gratitude to Dr.rer.nat. Moritz Wiese and Dr., Dipl.-Math. Harutyun Aydinyan for their useful comments, remarks, engagement, and patience through the learning process of this thesis.
I want to thank all my colleagues for the pleasant social atmosphere and helpful discussions. Furthermore, I would also like to thank all my friends and in particular Anton Heronimus, Elena Zibart and Christin Abt for mental support.
Last but not least, I take this opportunity to express the profound gratitude from my deep heart to my beloved parents, Vera Schmidt and Andrej Frank and my sister Dr.rer.nat. Svetlana Frank for giving me courage.

Basic Notations

Unless otherwise specified, we define the following notation.

$\mathbb{N}$	natural numbers
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$\mathcal{X}$	alphabet or set
x^+	$\max\{x, 0\}$, $x \in \mathbb{R}$
X	random variable implicitly defined on alphabet $\mathcal{X}$
$\|\mathcal{X}\|$	cardinality of $\mathcal{X}$
x^n	sequence $(x_1, \ldots, x_n)$
P_X	probability distribution of the random variable X
p_X	probability density of the random variable X
$P_{X\|Y}$	probability distribution of X conditioned on Y
$Pr(\cdot)$	probability of the event $(\cdot)$
P_e	per bit error probability
$P_e^{(n)}$	the average probability of error
$h(\cdot)$	binary entropy function
$H(X)$	entropy of the discrete random variable X
$H(X\|Y)$	entropy of X conditioned on Y
$I(X;Y)$	mutual information between X and Y
$\mathbb{F}_q$	finite field with q elements
$\mathbb{F}_q^n$	n-dimensional space over $\mathbb{F}_q$
$\mathbb{F}_q^{n\times k}$	the set of $n \times k$ matrices over $\mathbb{F}_q$
I_k	$k \times k$ identity matrix
$d_H(x^n, y^n)$	Hamming distance between x^n and y^n
$d(C)$	minimum distance of a code C
G_C	generator matrix of a linear code C
H_C	parity check matrix of a linear code C
E_X	expectation with respect to X

Contents

1. Introduction

The need for reliable and secure data communication over wireless networks is greater than ever before. It is increasingly possible to listen into the communication between a computer and a wireless router for example. Information-theoretic security is becoming more significant as computation hardware is getting drastically cheaper every day; this means that computational security schemes that are currently considered secure will no longer be secure in the future. Information-theoretic security assumes that the eavesdropper has limited access to the transmission, but an unlimited power to process it. On the other hand, cryptographic security assumes that the eavesdropper of a secure transmission has unlimited access to the transmission, but a limited processing power. The idea of using information theory to analyze cryptosystems was first introduced by Shannon in his 1949 paper [1], in which a secret key is considered to protect confidential messages. Wyner proposed an alternative approach to secure communication schemes in his seminal paper [2], where he introduced the so-called wiretap channel model. He demonstrated that secure communication is possible without sharing a secret key and determined the secrecy capacity for a wiretap channel. Wyner's model was later generalized by Csiszar and Körner [3] and was further developed in [4]. Authors in [4] introduced the wiretap channel II model, in which the legitimate transmitter communicates over a noiseless main channel, while the eavesdropper has access to μ noiseless bits (of his own choice) of the length-n binary codeword. Authors in [4] showed that perfect security is attainable provided that μ is not too large, and proposed a randomized coset coding scheme, where the partition of the binary code $C = \{0,1\}^n$ corresponds to a group code and its cosets, and showed that it achieves the capacity-equivocation region.

Part II

Several recent papers have studied various wiretap channels and provided results on secrecy capacity, e.g. by Thangaraj et al. [5] and Liu et al. [6]. Also among them are wiretap channel models with delay constraints, e.g. [7],[8], [9], [10]. Most of the literature is concerned with wiretap channel models in which the eavesdropper only overhears the transmission but does not try to modify the transmission. The wiretap channel with an active eavesdropper was first considered by Lai et al. [11], where the goal of the receiver is to detect whether the transmitted packet has been modified or not. Aggarwal et al. [12] were the first who studied the model, where the eavesdropper not only noiselessly

overhears a subset of the transmitted bits, but also modifies the bits, so that the legitimate receiver receives a corrupted version of the sender's codeword. In this model, they designed a scheme that achieves a secrecy rate of $(1 - \epsilon - h(\epsilon))^+$, where $\epsilon = \mu/n$ is the portion of the bits observed and erased by the eavesdropper and $h(\epsilon)$ is the binary entropy. However, existence of better achievable rates for the described channels remains an open problem for arbitrary fields and code lengths.

Moreover, design of efficient coding schemes for both the wiretap channel and the model of wiretap channel II with an active eavesdropper is also an open problem. This motivates us first to introduce and study a model of wiretap channel II, where the abilities of the eavesdropper are more restricted compared to the one in [12]. In our models, the eavesdropper can observe an interval of μ symbols from n transmitted symbols. In addition, the active eavesdropper can erase the symbols in any interval of length B of the transmitted codeword. It is worth noting that code constructions for this model also work for the wiretap channel with an eavesdropper, where the main channel causes erasures in any interval of length B. In both cases, the designer of the encoder has to proceed on the assumption that the worst case can occur. In addition, neither the transmitter nor the intended destination knows in advance which interval of length B has been erased. However, we assume that the legitimate receiver of the message has a physical advantage over the eavesdropper. In addition, we have constructed burst-erasure wiretap codes for the streaming case where the legitimate receiver has to meet a decoding delay deadline. In many emerging communication systems such as interactive voice and video communication, internet of things, etc., low-delay is an important task along with reconstruction of corrupted or lost data. The goal of our work is to design practical coding schemes which achieve the maximum secrecy rate, perfect security, i.e., the adversary's observations are completely decoupled from the message, and perfect reliability, i.e., zero error decoding.

Part III

We consider the wiretap channel, where Alice (the sender) wants to convey messages from a finite message set to Bob (the legitimate receiver) over a noisy channel. Eve (an eavesdropper) observes a different noisy version of the channel input. Alice has to encode the message so that Bob is able to decode the channel output, and so that Eve learns as little as possible about the message from her observation. There are different ways to measure security under a given security paradigm (weak, strong, perfect, semantic), e.g. the total variation distance, the mutual information, the equivocation rate or the advantage. We use the advantage at Eve as the security measure. The target value for the advantage is also zero. We consider three communication scenarios, each reflecting the operational meaning of different security measures and different assumptions about Eve's strengths. In the first two scenarios, the message distribution may be arbitrary,

so these setups would be variants of "semantic security" in common terminology. In the third scenario, the advantage is measured under the assumption of a uniformly distributed message. This is usually referred to as "strong security".

We consider a seeded modular code for the additive white Gaussian noise (AWGN) wiretap channel consisting of a security layer, an error-correction layer and a modulation layer for the reliable transmission from Alice to Bob. In the security layer, a function f_s of certain properties is used, which depends on a randomly chosen seed s. We can assume that before the transmission begins, the seed s is known to all participants. Practically, the seed could have been sent by Alice before the communication started. Using seed recycling, [13] showed that the rate loss can be asymptotically neglected. Note that a seed is different from a key, because the seed is public and does not have to be kept secret from Eve. When encoding the message, in the security layer the randomized inverse function f_s^{-1} maps the message M together with a randomly chosen seed s and a randomly generated vector r to the input vector v of the forward error-correction (FEC) code. Any FEC code adapted for the channel can be used. Then the codeword is modulated. At the receiver, the channel output is demodulated, decoded and then the message is reconstructed using the seed and f_s. The error probability of the seeded modular code is at most as high as that of the FEC code and of the modulation scheme. Eve knows the coding procedure, the channel, the seed and the distribution of the message P_M. The artificial randomness used in the randomized inverse serves to confuse Eve.

The seeded modular coding scheme has the advantage that already-existing and long-researched FEC codes can be used. Additionally, embedding in existing wireless systems is associated with low refitting costs. The security aspect of wireless communications in 6G is of paramount importance to combat cybercriminal activities. This is especially true because more and more people are using wireless networks (e.g. mobile networks and WLAN) for online banking and personal e-mails, due to the widespread use of smartphones. But also in machine-to-machine communication in industry 4.0, the security of wireless communication is enormously important for personal protection and to enable a smooth production workstation. For further applications and more details, we refer to [14], [15].

The functions we use for the security layer are *universal hash functions* (UHF). We call a modular scheme that uses the UHF as the function in the security layer a *modular UHF scheme*.

1.1. Outline and Contribution

Part I: Chapter 2 contains a review of fundamental results in information theory and coding theory needed for the rest of the thesis. In Chapter 3, we first introduce the notion of the wiretap channel and Wyner's random encoding strategy which achieves the secrecy capacity. Then we study the binary erasure wiretap channel II and describe an

information-theoretic analysis of the coset coding scheme for this model.

Part II: In chapter 4, we consider a wiretap channel II with an active eavesdropper. The eavesdropper is able to observe any interval of μ symbol positions and erase the symbols in any interval of B positions of a transmitted codeword. We present an explicit construction of nested linear codes that achieve maximum secrecy rate for the finite length coding regime with perfect security and zero-error decoding for any admissible code parameters.

In Chapter 5, we consider transmission of secure messages over a burst-erasure wiretap channel under decoding delay constraint. For block codes we introduce and study delay-optimal secure burst-erasure correcting (DO-SBE) codes that provide perfect security and recover a burst of erasures of a limited length with minimum possible delay. Our explicit constructions of DO-SBE block codes achieve maximum secrecy rate. We also consider a model of a burst erasure wiretap channel for the streaming setup, where in any sliding window of a given size, in a stream of encoded source packets, the eavesdropper is able to observe packets in an interval of a given size. For that model we obtain an information-theoretic upper bound on the secrecy rate for delay-optimal streaming codes. We show that our block codes can be used for construction of delay-optimal burst-erasure correcting streaming codes which provide perfect security and meet the upper bound for a certain class of code parameters.

In Chapter 6, for streaming applications, we consider parallel burst erasure channels in the presence of an eavesdropper. The legitimate receiver must perfectly recover each source symbol subject to a decoding delay constraint without the eavesdropper gaining any information from his observation. For a certain class of code parameters, we propose delay-optimal M-link codes that recover a certain number of bursts of erasures of a limited length each occurring on a separate link, and where the codes provide perfect security even if the eavesdropper can observe a link of his choice. Our codes achieve the maximum secrecy rate for the channel model.

Part III: In Chapter 7, we consider a seeded modular code for the additive white Gaussian noise (AWGN) wiretap channel consisting of a security layer, an error-correction layer and a modulation layer. For reliable transmission, we use any forward error-correction (FEC) code and modulation method. In the security layer, a universal hash function (UHF) is used, which depends on a randomly chosen seed s. We consider three communication scenarios in which the advantage (the security measure) at the eavesdropper is measured in different ways. To assess the security performance, we derive the operational meaning of the advantages in terms of the error probability.

In Chapter 8, we experimentally verify the information-theoretic security of a seeded modular code for the AWGN wiretap channel consisting of a security layer, an error-correction layer and a modulation layer. In the security layer, a universal hash function (UHF) is used, which depends on a randomly chosen seed s. In the error-correction layer and the modulation layer we use polar codes and quadrature amplitude modulation QAM, respectively. The eavesdropper uses the maximum likelihood (ML) test as an attack

strategy. We analyze the security in different communication scenarios using simulations. Since the ML test as proposed in Chapter 7 goes with a high level of complexity, we compare the performance of different attack strategies.

In Chapter 9, we use a seeded modular code as proposed in Chapter 7 for implementing physical layer security in a wiretap scenario. We evaluate the performance of the seeded modular code in an experimental setup with software defined radios and compare these results to simulation results. In order to assess the security level of the scheme, we employ the distinguishing security metric. In our experiments, we compare the distinguishing error rate for different seeds and block lengths.

1.2. List of Publications

Part II: Codes for the Burst-Erasure Wiretap Channel in the Finite Blocklength Regime
Chapter 4 is based on the following publication:

- A. Frank, H. Aydinian, and H. Boche, "Type II wiretap channel with an active eavesdropper in finite blocklength regime," IEEE Wireless Communications and Networking Conference, pp. 1-6, Apr 2016.

Chapter 5 is based on the following publication:

- A. Frank, H. Aydinyan, H. Boche, "Delay Optimal Coding for Secure Transmission over a Burst Erasure Wiretap Channel," IEEE Wireless Communications and Networking Conference, pp. 1-7, Apr 2019.

Chapter 6 is based on the following publication:

- A. Frank, "Delay-Optimal Coding for Secure Transmission over Parallel Burst Erasure Channels with an Eavesdropper," 2020 IEEE International Symposium on Information Theory, pp. 960-965, 2020

Part III: Modular Codes for the Wiretap Channel in the Finite Blocklength Regime
Chapters 7 and 8 is based on the following publications:

- A. Frank, J. Voichtleitner, M. Wiese, and H. Boche, "Implementation of a Modular Code for Secure Communication," 2022 IEEE International Conference on Communications (ICC).

Chapter 9 is based on the following publication:

- L. Torres-Figueroa, U. J. Mönich, J. Voichtleitner, A. Frank, V. C. Andrei, M. Wiese, and H. Boche, "Experimental evaluation of a modular coding scheme for physical layer security," 2021 IEEE Global Communications Conference (GLOBECOM), pp. 1-6, 2021.

Part I.

Preliminaries

2. Basics of Information Theory and Coding Theory

In this chapter, the first section contains a review of the basic notions of information theory. In the second section we introduce basic notions of coding needed for the rest of the thesis.

2.1. Elements of Information Theory

We need some definitions from information theory that will be used in the subsequent chapters. Most of them can be found in the textbook [16]. In this section we use logarithms to the base 2 and set $0 \log 0$ to 0.

Definition 2.1 (Shannon Entropy). *Let X be a discrete random variable taking values in a finite alphabet $\mathcal{X}$ and probability distribution $P_X(\cdot)$. The Shannon entropy or uncertainty of X is defined as*

$$H(X) = \sum_{x \in \mathcal{X}} -P_X(x) \log P_X(x).$$

The units of the entropy in this case are bits.

The entropy is a measure of the average uncertainty in the random variable.

Definition 2.2 (Binary Entropy Function $h(p)$). *Consider the entropy $H(X)$ of a Bernoulli random variable X where $X = 1$ with probability p and $X = 0$ with probability $1 - p$. The entropy of X is*

$$h(p) = H(X) = -p \log p - (1 - p) \log(1 - p).$$

Definition 2.3 (Joint Entropy). *The joint entropy of X and Y is defined by considering the concatenation XY as a new discrete random variable, i.e., we have*

$$H(X, Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} -P_{XY}(x, y) \log P_{XY}(x, y).$$

Definition 2.4 (Conditional Entropy). *Given a joint distribution $P_{XY}(\cdot)$ and two random variables X and Y take on values in the finite alphabets $\mathcal{X}$ and $\mathcal{Y}$, respectively, the conditional entropy of Y given the event $X = x$ with probability $Pr(X = x) > 0$ is defined as*

$$H(Y|X = x) = \sum_{y \in \mathcal{Y}} -P_{Y|X}(y|x) \log P_{Y|X}(y|x),$$

where

$$H(Y|X) = \sum_{x \in \mathcal{X}} P_X(x) H(Y|X = x) = H(X, Y) - H(X).$$

Definition 2.5 (Mutual Information). *The mutual information between two discrete random variables X and Y is defined as*

$$I(X;Y) = \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{XY}(x, y) \log \frac{P_{XY}(x, y)}{P_X(x) P_Y(y)} = H(X) - H(X|Y).$$

The mutual information $I(X;Y)$ is a symmetric function, that is

$$I(X;Y) = I(Y;X),$$

which shows the dependence between the two random variables or, in other words, the amount of information obtained about X by observing Y.

The mutual information $I(X;Y) = 0$ iff the random variables X and Y are statistically independent.

Chain rule for entropy: The chain rule for entropy is equivalent to

$$H(X_1, X_2, \ldots, X_n) = H(X_1) + H(X_2|X_1) + \cdots + H(X_n|X_1, X_2, \ldots, X_{n-1}).$$

Definition 2.6 (Markov Chain). *A discrete stochastic process $X_1, X_2, \ldots$ is said to be a Markov chain or a Markov process if for $n = 1, 2, \ldots,$*

$$Pr(X_{n+1} = x_{n+1}|X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_1 = x_1) = Pr(X_{n+1} = x_{n+1}|X_n = x_n)$$

for all $x_1, x_2, \ldots, x_n, x_{n+1} \in \mathcal{X}$.

Data processing inequality: If $X \to Y \to Z$ forms a Markov chain, then we have

$$I(X;Y) \geqslant I(X;Z) \text{ and } I(Y;Z) \geqslant I(X;Z).$$

Equality iff $I(X;Y|Z) = 0$.

Fano's inequality: Suppose both X and $\hat{X}$ take on values in the alphabet $\mathcal{X}$, and let $P_e = Pr(\hat{X} \neq X)$. We have

$$H(X|\hat{X}) \leqslant h(P_e) + P_e \log(|\mathcal{X}| - 1).$$

Definition 2.7 (Discrete Channel). *Let $\mathcal{X}$ and $\mathcal{Y}$ be discrete alphabets, and $P(y|x)$ (or $W(y|x)$) be a transition probability matrix from $\mathcal{X}$ to $\mathcal{Y}$. A discrete channel $P(y|x)$ is a single-input single-output system with input random variable X taking values in $\mathcal{X}$ and output random variable Y taking values in $\mathcal{Y}$ such that*

$$Pr(X = x, Y = y) = Pr(X = x)P(y|x)$$

for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

Definition 2.8 (Continuous Channel). *A continuous channel $p(y|x)$ is a system with input random variable X and output random variable Y taking values in $\mathbb{R}$ such that Y is related to X through $p(y|x)$.*

Definition 2.9 (Discrete Memoryless Channel (DMC)).
A sequence of channels $\{W_n : \mathcal{X}^n \to \mathcal{Y}^n\}_{n=1}^{\infty}$ is called a discrete memoryless channel (DMC) with transition probability matrix W if

$$W_n(y^n|x^n) = \prod_{i=1}^{n} W(y_i|x_i).$$

Definition 2.10 (Channel Code for a DMC). *An $(n, |\mathcal{M}|)$ code for a DMC consists of an encoding function*

$$f : \mathcal{M} \to \mathcal{X}^n$$

and a decoding function

$$g : \mathcal{Y}^n \to \mathcal{M}.$$

The sequence $f(i) \in \mathcal{X}^n$ with $i \in \{1, 2, \ldots, |\mathcal{M}|\}$ is called a codeword. The set of codewords is called the codebook.

Definition 2.11 (Rate of a Channel Code). *The rate of an $(n, |\mathcal{M}|)$ code for the $(\mathcal{X}, P(y|x), \mathcal{Y})$ channel is*

$$R = \frac{\log |\mathcal{M}|}{n}$$

and is measured in terms of bits/transmission (i.e. channel use).

Definition 2.12 (Capacity). *The capacity of a DMC with input X and output Y is defined by*

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

The capacity of a DMC is the supremum of all achievable rates.

Definition 2.13 (Conditional Probability of Error). *Let*

$$\lambda_i = Pr(g(Y^n) \neq i | f(i)) = \sum_{y^n} W_n(y^n|f(i)) I(g(y^n) \neq i)$$

be the conditional probability of error given that index i was sent, where $I(\cdot)$ is the indicator function.

Definition 2.14. *The maximal probability of error of an $(n, |\mathcal{M}|)$ code is*

$$\hat{P}_e = \max_i \lambda_i.$$

Definition 2.15. *The average probability of error $P_e^{(n)}$ for an $(n, |\mathcal{M}|)$ code is defined as*

$$P_e^{(n)} = \frac{1}{|\mathcal{M}|} \sum_{i=1}^{|\mathcal{M}|} \lambda_i.$$

Definition 2.16 (Achievability of a Rate). *A rate R is said to be achievable if there exists a sequence of $(n, 2^{nR})$ codes such that the maximal probability of error λ_{max} tends to 0 as $n \to \infty$.*

2.2. Elements of Error Correcting Codes

For our purposes we need to introduce only linear block codes. All definitions and statements presented below can be found in a standard textbook on coding theory, e.g., [17] or [16]. Throughout the thesis we use the following notation. $\mathbb{F}_q$ denotes a finite field with q elements. $\mathbb{F}_q^n$ is an n-dimensional vector space over $\mathbb{F}_q$ and $\mathbb{F}_q^{n\times k}$ is the set of $n \times k$ matrices over $\mathbb{F}_q$.

Definition 2.17. *A linear code with length n over $\mathbb{F}_q$ is a subspace of $\mathbb{F}_q^n$.*

Definition 2.18. *The weight of a codeword c, denoted by $wt(c)$, is defined as the number of non-zero entries of c.*

Definition 2.19. *The Hamming distance $d_H(u^n, v^n)$ between two vectors $u^n, v^n \in \mathbb{F}^n$ is the number of coordinates in which they differ.*

Definition 2.20. *Let C be a code with at least two codewords. The minimum distance $d(C)$ of C is the smallest distance between distinct codewords, that is*

$$d(C) = \min \{d_H(u^n, v^n) | u^n, v^n \in C; u^n \neq v^n\}.$$

If C is a linear code, then $d(C) = \min_{c \in C, c \neq 0} wt(c)$.

Definition 2.21. *If C is a linear code over $\mathbb{F}_q$ with length n, dimension k, and minimum distance $d(C) = d$, then we say that C is an $[n, k, d]_q$ code, or $[n, k]_q$ code if $d(C)$ is not specified. The numbers n, k, and d are called the parameters of the linear code.*

Definition 2.22. *The dual code of an $[n, k]_q$ code C denoted by $C^\perp$ is a null space of C.*

Definition 2.23 (Generator matrix). *A generator martix for a linear code C is a matrix G whose rows form a basis for C.*

Definition 2.24 (Parity-check matrix). *A parity check matrix H for C is a generator matrix for the dual code $C^\perp$.*

Let C be an $[n, k]_q$ code. Then C can be given by its generator matrix G_C, or the parity check matrix H_C as follows

$$C = \left\{v^n \in \mathbb{F}_q^n : u^k G = v^n;\ u^k\ \in \mathbb{F}_q^k\right\}$$
$$C = \left\{v^n \in \mathbb{F}_q^n : H(v^n)^T = 0\right\}.$$

Theorem 2.1. *Let H be a parity check matrix for a linear code C of length n. Then C has distance d if and only if every subset of $d-1$ columns of H are linearly independent, and at least one set of d columns of H are linearly dependent.*

Theorem 2.2 (Singleton bound). *Let C be an $[n, k, d]_q$ code. Then $|C| \leqslant q^{n-d+1}$, or equivalently $d \leqslant n - k - 1$.*

Definition 2.25 (MDS code). *An $[n, k, d]_q$ code achieving the singleton bound is called a maximum distance separable (MDS) code.*

Theorem 2.3 (Properties of MDS codes). *Let C be a linear $[n, k, d]_q$ code. Let G and H be respectively the generator and parity check matrices for C. The following claims are equivalent:*

- *C is an MDS code.*
- *Every subset of $n - k$ columns in H is linearly independent.*
- *Every subset of k columns in G is linearly independent.*
- *$C^\perp$ is an MDS code.*

An important class of MDS codes is the Reed-Solomon (RS) code. RS codes are $[q-1, k, q-k]_q$ MDS codes. A generator matrix of an RS code can be given with the help of Vandermonde matrices. A Vandermonde matrix of order $q-1$ over $\mathbb{F}_q$ is defined as

$$V = \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_{q-1} \\ a_1^2 & a_2^2 & \dots & a_{q-1}^2 \\ \vdots & \vdots & \dots & \vdots \\ a_1^{q-2} & a_2^{q-2} & \dots & a_{q-1}^{q-2} \end{pmatrix},$$

where $a_1, \dots, a_{q-1}$ are the nonzero elements of $\mathbb{F}_q$. Every first k ($k \leqslant q-1$) rows of V results in a generator matrix of the $[q-1, k, q-k]_q$ RS code. Thus RS codes have a so called nested structure, that is every $[q-1, k, q-k]_q$ RS code, where $2 \leqslant k \leqslant q-1$, contains the $[q-1, k-1, q-k+1]_q$ RS code as a subcode.

3. The Wiretap Channel

In the first section, we consider the wiretap channel model and summarize the notions of information-theoretic security on this channel. In addition, we explain the secrecy coding method for the wiretap channel. In Section 3.2, we introduce the binary erasure wiretap channel II (BEWC-II) model and describe an information-theoretic analysis of the coset coding scheme for this model.

3.1. Wiretap Channel and Information-theoretic Security

Wyner [2] introduced the notion of a wiretap channel in 1975. It is the most basic channel model that takes security into account. In Wyner's model of secure communication and its generalization to a broadcast scenario [3], a transmitter (Alice) wants to convey a secret message to a legitimate receiver (Bob) through a discrete memoryless channel (DMC). The message must be kept secret from an eavesdropper (Eve) who has a degraded version of the legitimate receiver's observation. Wyner's original work showed that communication with (asymptotic) perfect security and reliability is possible if the eavesdropper's channel is noisier than the main channel. Importantly, security is information-theoretic and does not require a pre-shared secret key.

3.1.1. The Wiretap Channel Model

Consider the communication system, in Fig. 3.1. This system consists of three parties,

- Alice - the transmitter
- Bob - the legitimate receiver
- Eve - the eavesdropper.

The eavesdropper cannot influence Alice or the channel in any way.

Information-theoretic security usually considers the case where the wiretap channel is memoryless, and has a discrete input alphabet and a discrete output alphabet. The input alphabet is $\mathcal{X}$, and the output alphabets are $\mathcal{Y}$ and $\mathcal{Z}$ for Bob and Eve, respectively. The alphabets $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ are finite. For a memoryless channel, successive transmissions are independent of each other and the channel is defined by its joint transition probability $P_{YZ|X}(y, z|x)$. In a wiretap channel, Alice communicates a message S^k to Bob through

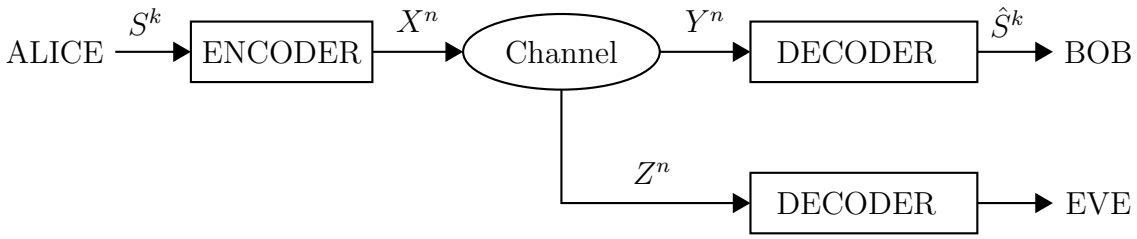


Figure 3.1.: The wiretap channel model.

the main channel, which is chosen uniformly at random from the message set $\mathcal{S}^k$. Alice performs this task by encoding S^k as a vector X^n of length n and transmitting X^n. Bob and Eve receive noisy versions of S^k, which we denote by Y^n and Z^n, via their respective channels. The encoding of a message S^k by Alice should be such that Bob is able to decode S^k reliably and Z^n provides as little information as possible to Eve about S^k.

3.1.2. Historical Background

Wyner considered a physically degraded wiretap channel where the eavesdropper (Eve) observes a degraded version of the signal obtained by the legitimate receiver. Thus, $X^n \to Y^n \to Z^n$ forms a Markov chain.

Degraded Wiretap Channel

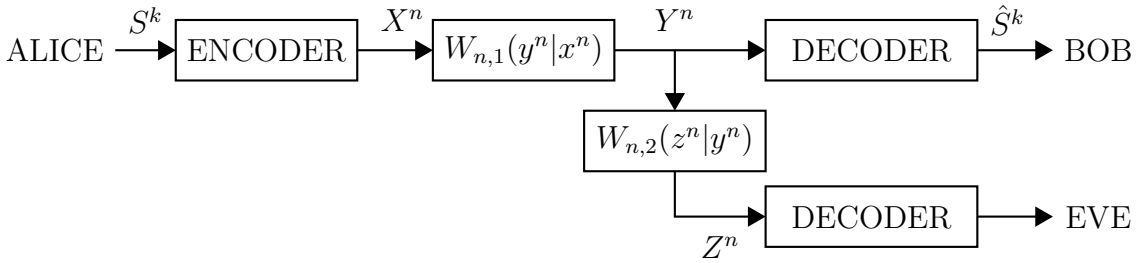


Figure 3.2.: The general wiretap channel model.

A degraded wiretap channel W_n is one in which for every $n \in \mathbb{N}$, and for every $(x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$,

$$W_n(y^n, z^n|x^n) = W_{n,1}(y^n|x^n)W_{n,2}(z^n|y^n),$$

where $W_{n,1} : \mathcal{X}^n \to \mathcal{Y}^n$ and $W_{n,2} : \mathcal{Y}^n \to \mathcal{Z}^n$.

A discrete memoryless wiretap channel is stationary and memoryless in the sense that

$$W_{n,1}(y^n|x^n) = \prod_{i=1}^{n} W_1(y_i|x_i), \text{ and } W_{n,2}(z^n|y^n) = \prod_{i=1}^{n} W_2(z_i|y_i)$$

for every $(x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$.

Non-degraded Wiretap Channel

Csiszár and Körner [3] generalized Wyner's model, where Eve's observation Z^n need not be a degraded version of Bob's observation Y^n. The channel is denoted by $X^n \to (Y^n, Z^n)$ and is depicted in Fig. 3.3. This channel and the channels $\{W_{n,1} : \mathcal{X}^n \to \mathcal{Y}^n\}$ and $\{W_{n,2} : \mathcal{X}^n \to \mathcal{Z}^n\}$ from Alice to Bob and Alice to Eve, respectively, are discrete and memoryless.

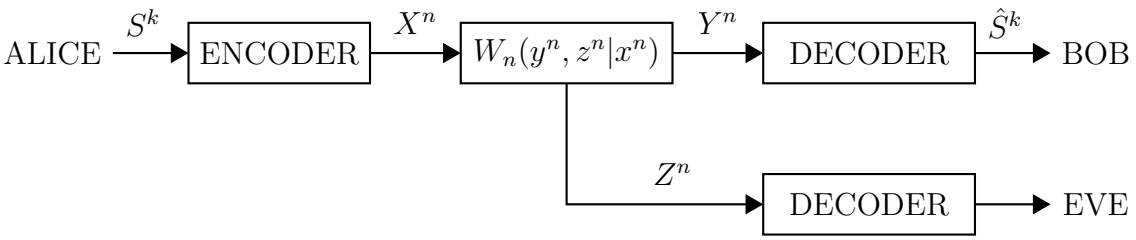


Figure 3.3.: The generalized wiretap channel model.

3.1.3. Secrecy Capacity of the Wiretap Channel

In the information-theoretic approach, the secrecy performance of a code C of length n is measured in terms of the mutual information between the secret and Eve's observation $\frac{1}{n} I(S^k; Z^n)$ or by the equivocation rate at Eve

$$R_e^{(n)} = \frac{1}{n} H(S^k|Z^n).$$

The equivocation rate is a measure of how much uncertainty Eve has about the message S^k after observing Z^n. Because the encoder is assumed to be one-to-many mapping, the equivocation $H(S^k|Z^n)$ is a positive number. A code of rate $R^{(n)}$ with block length n for the wiretap channel is given by a message set S^k of cardinality $|\mathcal{S}^k| = 2^{nR^{(n)}}$, and a collection of disjoint subcodes $\{C_{s^k} \subset \mathcal{X}^n\}_{s^k \in \mathcal{S}^k}$. To encode a message S^k, Alice chooses one of the codewords in C_{s^k} uniformly at random and transmits it. Bob uses a decoder $g : Y^n \to S^k$ to determine which message was sent. We assume that the message S^k is

uniformly distributed over $\mathcal{S}^k$. The average probability of error for the secrecy code is defined as $P_e^{(n)} = Pr(\hat{S}^k \neq S^k)$.

Remark 3.1. *The $(n, 2^{nR^{(n)}})$ code C is assumed to be known by Alice, Bob and Eve, although the source is only available to Alice and thus, the realizations of the discrete memoryless source (DMS) used for encoding.*

Definition 3.1. *A rate-equivocation pair (R, R_e) is said to be achievable for the wiretap channel, if for every $\epsilon > 0$ there exists a sequence of codes of rate $R^{(n)}$ with the average probability of error $P_e^{(n)} < \epsilon$ as the code length n goes to infinity, and with the equivocation rate $R_e^{(n)}$ satisfying*

$$\begin{aligned} \lim_{n\to\infty} R^{(n)} &> R - \epsilon, \\ \lim_{n\to\infty} R_e^{(n)} &> R_e - \epsilon. \end{aligned}$$

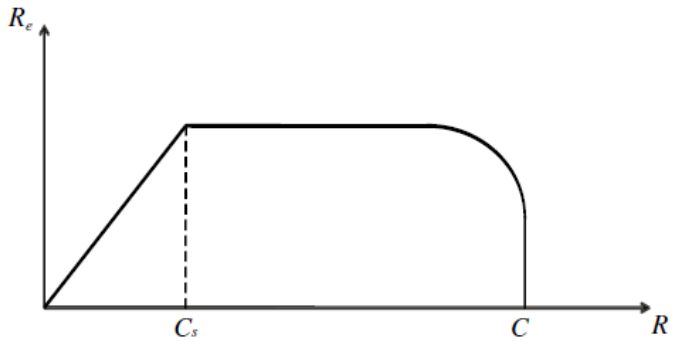


Figure 3.4.: A typical (R, R_e) region.

We want R_e to be as high as possible, and ideally it should equal the rate R.

Definition 3.2 (Perfect security). *An encoder for the wiretap model achieves perfect security in Shannon's sense if the probability of error in Bob's estimate $\hat{S}^k$ is zero and the mutual information between Eve's observation Z^n and the secret S^k is zero; that is,*

$$\begin{aligned} P_e^{(n)} &= 0, && (Reliability) \\ I(S^k; Z^n) &= 0. && (Security) \end{aligned}$$

Hence, S^k and Z^n have to be independent random variables and we can obtain this requirement if all messages are equally likely, that is $S^k \sim unif(\mathcal{S}^k)$, so that Eve can not indicate the message.

Thus, perfect security can be obtained if

$$R_e = \lim_{n\to\infty} \frac{1}{n} H(S^k|Z^n) = \lim_{n\to\infty} \frac{H(S^k)}{n} = R.$$

Definition 3.3 (Secrecy capacity). *The maximum rate at which both objectives are attainable is called the secrecy capacity C_s of the wiretap channel.*

Theorem 3.2 (Csiszár and Körner[3]). *The maximum perfect secrecy rate, i.e., the secrecy capacity C_s for a discret memoryless wiretap channel can be calculated as follows:*

$$C_s = \max_{U \to X \to (YZ)} [I(U;Y) - I(U;Z)].$$

The notation $U \to X \to (YZ)$ forms a Markov chain in this order with the random variables U, X, Y and Z. The auxiliary random variable U is used for calculation purposes with $|\mathcal{U}| \leqslant |\mathcal{X}|$.

For a degraded wiretap channel, i.e., $P(y,z|x) = P(y|x)P(z|y)$, follows

$$I(U;Y) - I(U;Z) = I(U;Y|Z) \leqslant I(X;Y|Z) = I(X;Y) - I(X;Z).$$

Hence, the secrecy capacity simplifies to

$$C_s = \max_{X \to (YZ)} [I(X;Y) - I(X;Z)],$$

and also holds for a general wiretap channel if Y is more capable than Z.

The secrecy capacity C_s is always positive unless, channel $X \to Y$ is less noisy than channel $X \to Z$.

Theorem 3.3 ([3]). *If channel $X \to Y$ is less noisy than channel $X \to Z$, the rate-equivocation region of the wiretap channel $X \to (YZ)$ contains all rate-equivocation pairs (R, R_e) that satisfy*

$$\begin{aligned} 0 &\leqslant R_e \leqslant I(X;Y) - I(X;Z), \\ R_e &\leqslant R \leqslant I(X;Y). \end{aligned}$$

In Wyner's work the secrecy capacity was determined under weak security conditions. Later, Csiszár [18], and independently Maurer and Wolf [19], defined the notion of strong security, and argued that this is a much better security condition compared to weak security.

Definition 3.4 (Weak security).

$$\lim_{n \to \infty} \frac{1}{n} I(S^k; Z^n) = 0, \qquad (S^k \sim unif(\mathcal{S}^k)).$$

The rate of information leaked about S^k through observing Z^n goes to zero as n goes to infinity.

Definition 3.5 (Strong security).

$$\lim_{n\to\infty} I(S^k; Z^n) = 0, \qquad (S^k \sim unif(\mathcal{S}^k)).$$

The total amount of information leaked about S^k through observing Z^n goes to zero as n goes to infinity.

The more stringent information-theoretic security metric was formalized by Bellare et al. [20] by adapting the notion of semantic security used in computational cryptography [21].

Definition 3.6 (Semantic security).

$$\lim_{n\to\infty} \max_{P_{S^k}} I(S^k; Z^n) = 0.$$

Definition 3.7 (Perfect security).

$$I(S^k; Z^n) = 0.$$

The total amount of information leaked about S^k through observing Z^n is zero.

3.1.4. Coding for the Wiretap Channel according to Wyner

Wyner [2] introduced the stochastic encoding scheme to achieve the secrecy capacity C_s of the wiretap channel. The stochastic encoding scheme serves to confuse the eavesdropper by allocating a message to many codewords at random. In the secrecy coding scenario, deterministic encoders, in general, have a poorer secrecy performance compared to stochastic encoders. Due to this, almost all secrecy coding makes use of stochastic encoders.

Suppose, Alice wants to transmit one out of $|\mathcal{S}|^k$ equally likely messages, i.e., a message denoted S^k is such that $S^k \in \left\{s_1^k, s_2^k, \ldots, s_{|\mathcal{S}|^k}^k\right\}$ and $Pr(S^k = s_i^k) = 1/|\mathcal{S}|^k$, where $1 \leqslant i \leqslant |\mathcal{S}|^k$. Consider a codebook C' of length n which is randomly partitioned into $|\mathcal{S}|^k$ subcodes C_i, i.e. $C' = \bigcup_i C_i$. Each message s^k is associated with one subcode C_{s^k}. For the case where $|\mathcal{S}| = 2$, Fig. 3.5 shows the encoding process for a wiretap channel.

A message s^k is encoded into x^n which is chosen uniformly at random from the subcode C_{s^k}. The receiver on the main channel (Bob) decodes a word y^n of length n with respect to the overall code C' into $\hat{s}^k$. One such decoding method is the maximum likelihood (ML) decoding.

Alice's objective is to design a secure and reliable encoder. To guarantee *reliability*, the legitimate receiver should be able to decode the message with error probability which approaches zero for $n \to \infty$.

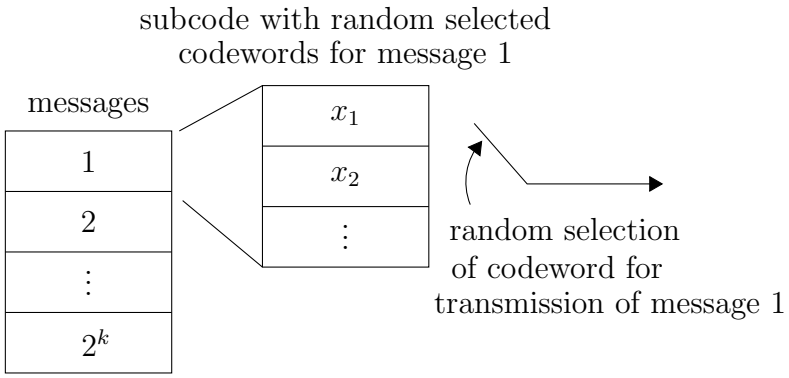


Figure 3.5.: Encoding process for a wiretap channel.

$$\lim_{n\to\infty} P_e^{(n)} = 0.$$

To guarantee *security*, Eve should not gather any information from her observation.

$$\lim_{n\to\infty} \frac{1}{n} I(S^k; Z^n) = 0. \qquad (weak\ security)$$

If an encoder with $R_s = \log |\mathcal{S}|^k / n$ satisfies the security and reliability constraints for a given wiretap channel, then such an encoder is said to achieve a secrecy rate R_s.

A detailed information-theoretic overview of general wiretap channels can be found in [22].

3.2. The Binary Erasure Wiretap Channel II

Before constructing efficient coding schemes for our channel models in Part II where the main and wiretapper's channel are both erasure channels, we first study the binary erasure wiretap channel II (BEWC-II) model, because it is a fundamental model and its analysis is extendable to a lot of different wiretap models. The wiretapper's channel is a binary erasure channel (BEC) and the main channel is noiseless, as shown in Fig. 3.6.

The BEWC-II model is a special case of the wiretap channel model. Thangaraj, et al. [5] were the first who constructed explicit codes for the BEWC-II model. The two legitimate nodes, Alice and Bob, want to communicate in the presence of an eavesdropper, Eve.

We denote the channel between Alice and Eve by BEC$(1-\epsilon)$, i.e. the probability of erasure in the wiretapper's channel is $1-\epsilon$. The BEC is a memoryless channel, which means that bits sent successively are erased independently. Alice's objective is again to convey a secret message S^k to Bob without revealing it to Eve. Therefore, Alice encodes

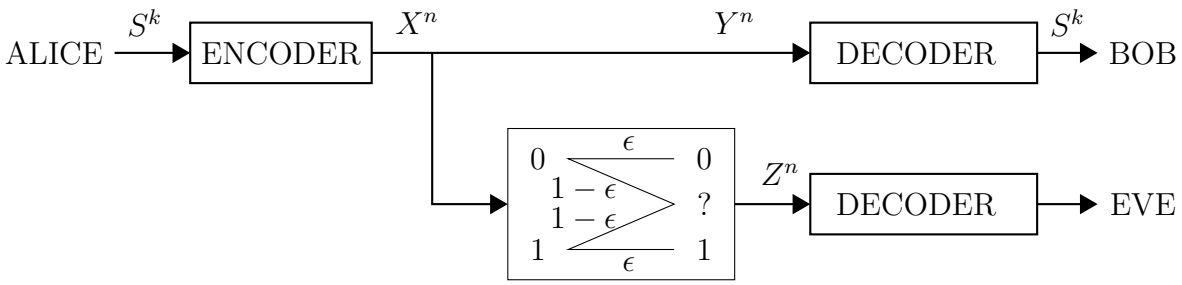


Figure 3.6.: The binary erasure wiretap channel II (BEWC-II) model.

S^k to a random variable X^n and then sends X^n over the BEWC-II. The secrecy capacity C_s of the BEWC-II is [23]

$$C_s = Capacity(X \to Y) - Capacity(X \to Z) = 1 - \epsilon.$$

3.2.1. Coset Coding for the Binary Erasure Wiretap Channel II

A coset coding scheme, which was introduced by Wyner [2] and further studied by both Ozarow and Wyner [4], is based on a linear code and its cosets. Given the blocklength n and the rate R of the coset coding scheme, a binary linear block code C of length n is used as a starting point.

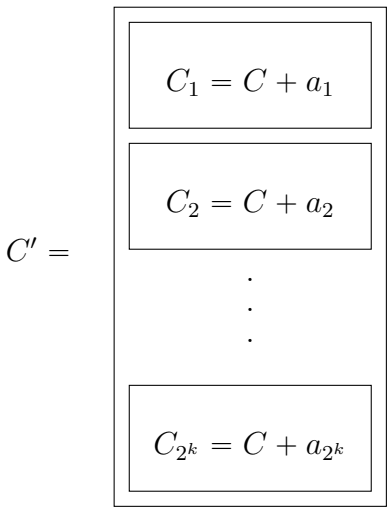


Figure 3.7.: The partitioning of the code C' of all possible output vectors according to the input message S^k, where $a_i \in C^*$ and C^* is an $[n, k]$ code generated by G^* and the input message. Each coset C_i, where $i = 1, \ldots, 2^k$, represents output codewords corresponding to a certain message.

The stochastic encoding scheme called the coset coding scheme is illustrated in Fig. 3.7. To transmit k-bit messages, consider an $[n, n-k]$ linear code C as the base code. Let G be the generator matrix of C with rows $\mathbf{g}_1, \ldots, \mathbf{g}_{n-k}$ and let G^* be the generator matrix for the code C^* with rows $\mathbf{g}_1^*, \ldots, \mathbf{g}_k^*$. The rows of G and G^* form a basis for $\{0,1\}^n$ so that $C \oplus C^* = \{0,1\}^n$. The coset corresponding to a k-bit message $s^k = (s_1, s_2, \cdots, s_k)$

is determined as follows:

$$s^k \rightarrow s_1\mathbf{g}_1^* + s_2\mathbf{g}_2^* + \cdots + s_k\mathbf{g}_k^* + C.$$

A secret k-bit message s^k is mapped to a codeword x^n using the transformation

$$x^n = s_1\mathbf{g}_1^* + s_2\mathbf{g}_2^* + \cdots + s_k\mathbf{g}_k^* + e_1\mathbf{g}_1 + e_2\mathbf{g}_2 + \cdots + e_{n-k}\mathbf{g}_{n-k},$$

where $e^{n-k} = (e_1, e_2, \cdots, e_{n-k})$ is a uniformly random $(n-k)$-bit vector. The correspondence is deterministic but the encoding procedure has a random component in the selection of the transmitted codeword. A k-bit message s^k is encoded into an n-bit codeword randomly selected from the coset of C corresponding to s^k. The encoding operation can be described as a matrix multiplication:

$$x^n = \begin{bmatrix} s^k & e^{n-k} \end{bmatrix} \begin{bmatrix} G^* \\ G \end{bmatrix},$$

where x^n belongs to the code C' generated by G'. The goal of both the legitimate receiver and the eavesdropper is to determine s^k from their respective received vectors.
Restating the desired twofold objectives, the design of the codes C and C' should be such that (1) s^k can be determined without error across the main channel, and (2) every s^k is equally likely across the wiretapper's channel. Since the channel between Alice and Bob is error-free, i.e., $Pr(\hat{S^k} \neq S^k) = 0$, Bob is able to find the syndrome s^k of C by $s^k = H\,(x^n)^T$, where C is an $[n, n-k]$ code and H is a carefully constructed $k \times n$ parity-check matrix. How to provide security will be discussed in the next subsection.

3.2.2. Security Criterion for the Binary Erasure Wiretap Channel II

Consider an eavesdropper's observation Z^n with μ unerased bits in positions $(i_1, \ldots, i_\mu)$. The number and the position of these erasures may be random. To develop a security criterion for the choice of C, we calculate the eavesdropper's uncertainty $H(S^k|Z^n)$ by first evaluating $H(S^k|Z^n = z^n)$. We assume that the eavesdropper has infinite computational power and complete knowledge of the code C. But the knowledge of the allocation of the codeword to the message is secret. As mentioned before, the code C is an $[n, n-k]$ code and the code C' is chosen to be the entire vector space $\{0,1\}^n$. If a coset of code C contains at least one vector that agrees with $z^n \in \{0, 1, ?\}^n$ in the unerased positions, we say that the coset is consistent with z^n. Each coset corresponds to a possible message for the eavesdropper.

Lemma 3.4 ([5]). *All cosets of C that are consistent with z^n contain the same number of sequences consistent with z^n.*

Proof. Let v^n be a vector consistent with z^n in the coset $v^n + C$. Let S be the set of all vectors in $v^n + C$ consistent with z^n. Then, $v^n + S$ is the set of all vectors in C with zeros in the positions revealed in z^n. That is,

$$v^n + S = \{u^n \in C : u_i = 0 \text{ whenever } z_i^n \neq ?\}.$$

Note that $|S| = |v^n + S|$, and this holds for any v^n which is consistent with z^n. ∎

Proposition 3.5 ([5]). *The total number of cosets of C consistent with z^n is denoted by $N(C, z^n)$. Since each message is equally likely a priori, we get*

$$H(S^k|Z^n = z^n) = \log N(C, z^n).$$

Proof.

$$\begin{aligned} H(S^k|Z^n = z^n) &= H(S^k X^n|z^n) - H(X^n|S^k z^n) \\ &= H(X^n|z^n) - H(X^n|S^k z^n). \end{aligned}$$

The first term $H(X^n|z^n)$ is the uncertainty in the codeword that was sent given the observation z^n. Suppose N is the number of sequences that are consistent with z^n, then $H(X^n|z^n) = \log N = \log 2^{n-\mu}$, since all codewords are used with equal probability. For the second term, holds

$$H(X^n|S^k z^n) = \sum_{s^k} H(X^n|S^k = s^k, z^n) P_{S^k|Z^n}(s^k|z^n).$$

Here, $H(X^n|S^k z^n)$ is the uncertainty in the codeword that was sent given the observation z^n and the coset corresponding to s^k that was used. Since all codewords are used with equal probability, and by Lemma 3.4 all cosets consistent with z^n contain the same number of sequences consistent with z^n, the term is reduced to $H(X^n|S^k = s^k, z^n) = \log N_c$, where N_c is the number of sequences consistent with z^n in a coset consistent with z^n. Hence,

$$H(S^k|z^n) = \log N - log N_c = \log \tfrac{N}{N_c} = \log N(C, z^n).$$ ∎

The total number of cosets we have is 2^k, which implies that the total number of cosets of C consistent with z^n, $N(C, z^n) \leqslant 2^k$. If $N(C, z^n) = 2^k$, we say that z^n is secured by C since the eavesdropper's $Pr(S^k = s^k|Z^n = z^n) = 1/2^k$ for every possible message s^k. In other words, if all cosets of C are consistent with z^n and all cosets have the same number

of vectors that match with z^n, we obtain:

$$\begin{aligned} I(S^k; Z^n) &= H(S^k) - H(S^k|Z^n) \\ &= \log 2^k - \sum_{z^n} P_{z^n}(z^n) H(S^k|Z^n = z^n) \\ &= k - H(S^k|Z^n = z^n) \\ &= k - \log N(C, z^n) \\ &= k - k = 0. \end{aligned}$$

The following theorem states a condition for a vector z^n to be secured by a code C.

Theorem 3.6 ([4], Lemma 4.1). *Let G be the generator matrix of an $[n, n-k]$ code C, and let $\mathbf{g}_i$ denote the $i-th$ column of G, where $i \in \{1, \ldots, n\}$. The eavesdropper can observe μ of n bits of the transmitted codeword and the unerased positions are given by $\{i_1, i_2, \cdots, i_\mu\}$. Then z^n is secured by C if and only if the matrix $G_\mu = (\mathbf{g}_{i_1}, \mathbf{g}_{i_2}, \cdots, \mathbf{g}_{i_\mu})$ has rank μ.*

Sketch of Proof. Suppose μ are the unerased positions of any n bit vector x^n. If G_μ has rank μ then the code C has codewords with all 2^μ possible sequences in the μ unerased positions. Since cosets are obtained by translating C, all cosets also have codewords with all possible binary sequences in the μ unerased positions. Therefore, $N(C, z^n) = 2^k$ and $I(S^k; Z^n) = 0$. If G_μ has rank less than μ, the code C does not have all μ-tuples in the μ unerased positions. So there exists at least one coset that does not contain a given μ-tuple in the μ unerased positions, and $N(C, z) < 2^k$. We obtain a necessary and sufficient condition for communication in perfect security with respect to an eavesdropper who observes any set of μ unerased bits. ∎

Corollary 3.7. *Let C be an $[n, n-k]$ binary linear code with generator matrix G. Coset coding with C guarantees perfect security against an eavesdropper who observes any set of μ unerased bits, if and only if all submatrices of G with μ columns have rank μ.*

In the next chapter we will discuss a wiretap channel model in which the eavesdropper is known to access no more than μ of n transmitted bits. This model differs from the BEWC-II of Fig. 3.6 in that the eavesdropper can, in principle, choose which μ bits are observed.

Part II.

Codes for Secure Transmission over the Burst-Erasure Wiretap Channel in the Finite Blocklength Regime

4. Block Codes for a Burst-Erasure Wiretap Channel

4.1. Introduction

In this chapter we present the wiretap channel II model with an active eavesdropper. Ozarow and Wyner [4] introduced the wiretap channel II model, in which the transmitter communicates over a noiseless main channel, while the eavesdropper can observe μ bits of the n-bit binary codeword transmitted to the legitimate receiver. They showed that information-theoretic security can be achieved over this channel, introducing a stochastic encoding scheme, called coset coding. Since then, researchers have studied various types of wiretap channels and have provided fundamental results on secrecy capacity (see [24], [22]). Most of the studies in this direction consider a passive eavesdropper model in which the eavesdropper only overhears the transmission.

Contribution

In this chapter we present the wiretap channel II model with an active eavesdropper, where the eavesdropper is not only able to overhear, but can also modify the transmission sent to the legitimate receiver. In general, Bob observes a sequence Y^n, which is a function of a codeword X^n and of the eavesdropper's transformation T^n.

In the following, the eavesdropper is able to observe any interval of μ symbol positions and erase the symbols in any interval of B positions of a transmitted codeword. We present explicit constructions of binary and non binary nested linear codes that achieve the maximum secrecy rate for the finite length coding regime, with perfect security and zero-error decoding for any admissible code parameters. It is worth to mention that our construction works for both the burst-erasure wiretap channel model with an eavesdropper and for the wiretap channel II model with an active eavesdropper that can cause a burst of erasures.

Related Work

The wiretap channel with an active eavesdropper was first considered by Lai et al. [11], where the goal of the receiver is to detect whether the transmitted packet has been modified or not. Aggarwal et al. [12] were the first who studied the model where the receiver

not only needs to detect the changes made by the eavesdropper, but also to correct the errors introduced by the eavesdropper. Boche and Schaefer [25] introduced and studied arbitrarily varying wiretap channels with active eavesdroppers. Recently, a model of the wiretap channel called the adversarial wiretap channel has been studied by Wang and Safavi-Naini [26]. Rouayheb et al. [27] showed that the secure network coding problem can be viewed as a network generalization of the wiretap channel II. A Wiretap network of type II has been further studied by other authors (see a survey in [28] and its references). Aggarwal et al. [12] considered two models for the wiretap channel II over a binary alphabet, where the eavesdropper can observe up to μ bits noiselessly from n transmitted bits and erase/replace the bits he observes. For the first model, they designed a coding scheme that achieves a secrecy rate of $(1 - \epsilon - h(\epsilon))^+$, where $\epsilon = \mu/n$ is the portion of the bits observed and erased by the eavesdropper and $h(\epsilon)$ is the binary entropy. For the second modification they showed that a secrecy rate $R_s = (1 - \epsilon - h(2\epsilon))^+$ is achievable. Deriving better achievable secrecy rates, as well as developing practical channel codes for these models, is an open and seemingly difficult problem. In fact, the problem of designing codes with the best perfect secrecy rates for both modification models is related to the classical open problem of the best trade-off between rate and distance (see e.g. [29]).

Some Notes

Although the results in [12] show the existence of channel codes that achieve a positive secrecy rate, developing practical channel codes for the models considered in [12] remains an open problem. First, their approach for error correction in the main channel is based on a random (Varshamov's construction) coding argument. Second, to achieve the equivocation rate of the eavesdropper, the latter code is partitioned into subcodes, where the existence of a "good partition" is shown again by a probabilistic argument (used in Ozarow-Wyner [4]). We also note that deriving better bounds for the secrecy capacity of the binary erasure wiretap channel with an active eavesdropper is an open problem. The reason is that in the considered model, the channel is no i.i.d. and we need to consider a worst case scenario.

It is worth mentioning that the problem discussed above becomes much easier in the case where we allow the alphabet size q to grow with the code length $n \leqslant q + 1$. In this case one can use MDS codes to achieve the maximum secrecy rate with perfect security and zero error. However, the same can not be achieved for a fixed alphabet size and growing n. This will be discussed in Section 4.4 in more detail.

All this motivates us to introduce and study another model of the wiretap channel II with an active eavesdropper, where the abilities of the eavesdropper are more restricted.

Outline

In Section 4.2, we introduce the notions burst of erasures and burst-erasure correcting codes, needed for the study of our model, where the eavesdropper is able to cause only bursts of erasures in the main channel. We also characterize the limitations for linear burst-erasure correcting codes over finite fields. In Section 4.3, we introduce a model of wiretap channel II with an active eavesdropper and discuss our main objectives. Furthermore, we specify the properties of the secure nested code pairs (C', C), which are necessary to fulfill the desirable objectives and determine an upper bound of the maximum equivocation. In Section 4.4, we state our main results. Section 4.6 gives constructions of binary and non binary linear nested codes achieving maximum secrecy rate for all admissible parameters n, B, μ. In Section 4.7 and 4.8, we present encoding and decoding procedures for secure nested codes. Section 4.9 concludes with a discussion and open problems.

4.2. Burst-Erasure Correcting Codes

Burst-error correction is an important part of error control coding, as in many communication and storage systems errors tend to occur in clusters rather than independently of each other. Two main types of bursts are typical in most communication systems: bursts of erasures and bursts of errors (see [30]). Erasure bursts often occur in recording, jammed, and some fading channels. For instance, in applications such as recording, an important requirement is that the code used should be capable of correcting bursts of erasures (in addition to random errors) caused by media defects such as scratches. The correction of burst erasures also has application in wireless communication systems limited by interference.

In the following, we will concentrate only on burst-erasure correcting codes. The notion of a burst of erasure is defined in a natural way.

Definition 4.1. *If the interval in a received sequence, formed by the first and the last erased positions, is of length B, we say that a burst of erasure of length B or B-burst erasure for short has been occurred. The pattern corresponding to this interval is called a burst erasure pattern. If all cyclic shifts of bursts of length B are also considered as burst patterns, we speak about wrap-around bursts of length B. In other words, all cyclic shifts of bursts of length B are also considered as B-bursts.*

A code capable of correcting all bursts of length B or less, is called a *B-burst-erasure correcting code*. Correspondingly, we speak about a code capable of correcting B-burst erasures including wrap-around bursts. The burst-erasure correction capabilities of linear codes follow from a more general statement, for erasure correcting codes. The following proposition follows from the proofs provided in [30].

Proposition 4.1. *Let C be a linear $[n,k]_q$ code and let E_L be an erasure pattern with coordinate positions $L \subset I = \{1,\dots,n\}$. Then C can correct E_L (with zero error) iff the columns of a parity check matrix H_C corresponding to indices in L are linearly independent, or equivalently, iff the columns of a generator matrix G_C corresponding to indices in $I\backslash L$ have rank k.*

Proof. Let y^n be the received sequence when the codeword x^n has been sent. Denote by x_L^n the subsequence of x^n with indices in L. Thus, in our case we have $x_{I\backslash L}^n = y_{I\backslash L}^n$ (the unerased subsequence of x^n). Clearly x_L^n can be uniquely recovered from $x_{I\backslash L}^n$ iff there exists a unique codeword x^n with $x_{I\backslash L}^n = y_{I\backslash L}^n$, satisfying $H_C(x^n)^T = \mathbf{0}$. It is easy to see that the latter is possible iff the columns of H_C with indices in L are linearly independent. Also note that y^n can be uniquely decoded to x^n, iff all patterns $\check{x}_{I\backslash L}^n$ with $\check{x}^n \in C$ are distinct, that is $|\left\{\check{x}_{I\backslash L}^n : \check{x}^n \in C\right\}| = q^k$. This clearly means that the columns of G_C with indices in $I\backslash L$ have rank k. ∎

Corollary 4.2. *A linear $[n,k]_q$ code C is capable of correcting up to $|L| = B$ erasures iff any B columns of an H_C are linearly independent, or equivalently, iff any $n-B$ columns of a G_C have rank k.*

Remark 4.3. *Note that the corollary implies that $B \leqslant n-k$, and in the case where $B = n-k$, we have an MDS code.*

Corollary 4.4. *A linear $[n,k]_q$ code C is B-burst-erasure correcting iff every B consecutive columns of H_C are linearly independent. Correspondingly, C can correct B-burst-erasures, including wrap-around bursts, iff every B cyclically consecutive columns of H_C are linearly independent.*

Corollary 4.5. *If the $[n,k]_q$ code C is capable of correcting B-burst erasures, then we have $B \leqslant n-k$ and hence $|C| \leqslant q^{n-B}$.*

Definition 4.2. *An $[n,k]_q$ code C capable of correcting B-burst erasures is called an optimal burst-erasure correcting code if $B = n-k$. If C can correct all burst erasures of length $n-k$, including cyclic (wrap-around) bursts, then C is called cyclically-optimal burst-erasure correcting, or c-optimal for short.*

Remark 4.6. *Later we will see that for our purposes we need to design coding schemes with optimal (respectively c- optimal) burst-erasure correcting codes.*

Note that Proposition 4.1 implies that the following holds.

Proposition 4.7. *If an $[n,k]_q$ code C is a c-optimal burst-erasure correcting code, then the dual code $C^\perp$ is a c-optimal burst-erasure correcting $[n,n-k]_q$ code.*

Proof. C is capable of correcting any $(n-k)$-burst erasures including wrap-around bursts. This with Proposition 4.1 implies that any k cyclically consecutive columns of a G_C must be linearly independent. Since G_C is a parity check matrix for the dual code $C^{\perp}$, the statement follows. ∎

Remark 4.8. *We note that the statement does not extend to optimal codes, namely the optimality of C does not imply the optimality of $C^{\perp}$. The reason is that the linear independence of all $n-k$ consecutive columns in H_C does not imply that every k consecutive columns of G_C are also linearly independent.*

4.3. The Channel Model

In our model the abilities of the eavesdropper are restricted, compared to the model of Aggarwal et al. [12], as follows. The eavesdropper can observe an interval of μ symbols from n transmitted symbols. In addition, it can erase the symbols in any interval of length B of the transmitted codeword over the main channel. The channel under consideration is depicted in Fig. 4.1.

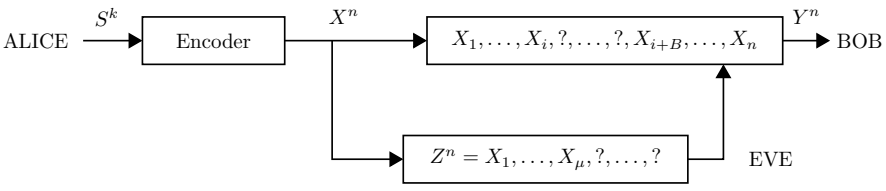


Figure 4.1.: A wiretap channel II model with an active eavesdropper.

Alice has a uniformly distributed k-symbol random message $S^k \in \mathbb{F}_q^k$ that must be conveyed to Bob by transmitting an n-symbol vector $X^n \in \mathbb{F}_q^n$ over the main channel. Eve has noiseless access to the Alice - Bob communication channel with the ability to observe any interval of μ symbols and to erase the symbols in any interval of B positions of her choice. In other words, Eve can cause any burst of erasures of length B in the channel. Thus, the output is $Y^n \in (\mathbb{F}_q \cup \{?\})^n$. Alice does not know anything about the erasures on the main channel or the symbols being tapped by Eve. The only thing she knows is that at most B-burst erasure can occur in the channel. Her task is to choose an encoding scheme which ensures that Bob can decode the message with zero error, while Eve must have complete equivocation over the message in spite of knowing the encoding procedure and the μ symbols observed by her own choice. We note that in our model Eve is able to erase an arbitrary interval of positions up to length B, unlike the model considered in Aggarwal et al. [12], where Eve can erase only the symbols she observes.

Our objective is to design a coding scheme which fulfills the tasks of Alice stated above. These are:

$$H(S^k|Y^n) = 0, \qquad \text{(perfect reliability)} \tag{4.1}$$

$$H(S^k|Z^n) = H(S^k), \qquad \text{(perfect security)} \tag{4.2}$$

where the entropy is computed using base-q logarithms. At the same time, we wish to achieve the maximum secrecy rate $R_s = k/n$ by explicit construction of a code (i.e. encoder and decoder), given parameters n, B, μ. In the following, we give a construction of so called nested linear codes which carry out all the tasks.

4.3.1. Secure Linear Nested Codes

Informally, a linear code pair (C', C) is called a nested code if $C \subset C'$. The main code C', also called the mother code, is partitioned into K cosets of C which is called a coarse code, thus $K = |C'|/|C|$. Each coset corresponds to a secret message and the transmission scheme is the same as for the Ozarow-Wyner coset coding described in Section 3.2.1. Of course for the purposes of reliability and security both codes C and C' must satisfy certain properties. Note also that the mother code serves for the reliability and the coarse code is used for stochastic encoding to provide security. Thus, the nested code approach is just a generalization of the Ozarow-Wyner coset coding method, where C' is the whole space. A nested code (C', C) is called secure if it satisfies the conditions (4.1), (4.2) and its secrecy rate $R_S = R' - R$, where R' is the rate of C' and R is the rate of C. The nested code approach has been used by many authors (see, for example, [6] and [31]), for the design of secure coding schemes in different models of the wiretap channel. For our model of wiretap channel II with an active eavesdropper, we also use the nested code approach. In the following we will specify the properties of C' and C that must be satisfied in order to achieve perfect security and reliability in our model of the wiretap channel.

4.3.2. Alice-Bob Communication

Let (C', C) be a nested code which we need to fulfill our tasks (4.1) and (4.2). Let C' also be an $[n, m]_q$ code. Suppose that a codeword $x^n \in C'$ has been transmitted over the channel and denote by $C(x^n)$ the coset to which x^n belongs. Then C' must be chosen in such a way that for every received vector y^n, Bob can determine the coset $C(x^n)$, and hence the message sent. Clearly, for this goal it is sufficient to recover x^n. In this case, regardless of a B-burst erasure introduced by Eve (possibly based on her observation), Bob should be able to decode y^n to x^n, that is C' must be an B-burst-erasure correcting code. Moreover, it is desirable that C' has the maximum rate, that is $m = n - B$. Thus, we suppose that we can take an optimal B-burst-erasure correcting code as a mother

code C'. We will see later that this is really the case, and moreover, this is a necessary condition to achieve our task.

4.3.3. Alice-Eve Communication

We are now interested in how large the equivocation $H(S^k|Z^n)$ can be and what the tradeoffs between parameters μ, B, and n are. Let $I = \{1, \ldots, n\}$ be the coordinate set and let $E \subset I$ be an interval of positions that Eve observes. Let $I \backslash M$ be an interval chosen by Eve for erasures. Thus, Eve observes an interval denoted by X_E^n and Bob observes the subsequence X_M^n (with the index set M) of the transmitted sequence X^n. The number of symbols μ observed by Eve must be smaller than $n - B$, which is the number of positions that Bob observes, for otherwise conditions (4.1) and (4.2) do not hold. Indeed, if $\mu \geqslant n - B$, then Eve can choose B positions to erase, such that $M \subset E$ which in view of 4.1 implies that $H(S^k|Z^n) = 0$.

For $\mu \leqslant n - B$, suppose now that Eve chooses first the pattern $X_{I \backslash M}$ to be erased and then observes an interval X_E such that $E \subset M$. Then we have $X^n \to X_M^n \to X_E^n$ and hence

$$\begin{aligned} H(S^k|Z^n) &= H(S^k|X_E^n) - H(S^k|X_M^n) \\ &= H(S^k|X_E^n) - H(S^k|X_E^n, X_{M\backslash E}^n) \\ &= I(S^k; X_{M\backslash E}^n|X_E) \\ &\leqslant H(X_{M\backslash E}^n|X_E^n) \\ &\leqslant H(X_{M\backslash E}^n) \\ &\leqslant (n - B) - \mu. \end{aligned}$$

This together with (4.2) implies that the number of symbols k that can be securely transmitted is upper bounded by $k \leqslant n - B - \mu$. Thus, we have the following.

Theorem 4.9. *For the wiretap channel II with an active eavesdropper that can observe a fraction $\epsilon = \mu/n$ of consecutive positions and erase a fraction $\vartheta = B/n$ of consecutive positions from transmitted symbols, the secrecy rate $R_s = \frac{k}{n}$ is upper bounded by $R_s \leqslant (1 - \vartheta - \epsilon)^+$.*

Remark 4.10. *It is obvious that the same upper bound holds for the case where the eavesdropper can respectively observe and erase arbitrary μ and B positions.*

4.4. Performance Criteria and Main Result

We are going now to analyze security constraints for the codes with a nested structure. We note that although Ozarow and Wyner [4] consider only the binary case, their results

on coset coding directly extend to codes over any finite field $\mathbb{F}_q$. In other words we want to resolve how to choose C and C' to provide perfect security and maximum equivocation with a nested code (C', C). Let us turn for a moment to the Ozarow-Wyner coset coding scheme, that is consider the case of noiseless main channels, thus $C' = \mathbb{F}_q^n$.

The algebraic secrecy criterion in [4] applied to our model says that perfect security is achieved iff a generator matrix G_C for C satisfies the following property:

- *Every μ consecutive columns of G_C are linearly independent.*

The fulfillment of this condition implies that for each Z^n with μ consecutive unerased positions, the following holds:

- *Every coset of C has the same number of vectors which are consistent with Z^n, that is vectors from which Z^n can be obtained by $n - \mu$ erasures.*

This means that we have perfect security, since every message is equally probable. Suppose now we choose any q^k (out of $q^n/|C|$) cosets of C for a secure transmission over a noiseless channel. Then every Z^n is again *secure* (that is the condition above holds again for every Z^n) and we can transmit q^k messages with perfect security. Thus, the security depends only on C. On the other hand, it is clear that maximum equivocation can be achieved with the noiseless main channel if there exists an $[n, \mu]$ code C satisfying the property above. Let C' be an $[n, m]_q$ B-burst-erasure correcting code with $C \subset C'$. Then regardless of the choice of a B-burst pattern and μ (consecutive) positions, to be observed by Eve, Bob is able to correctly reconstruct the codeword sent by Alice. This situation is actually equivalent to a scenario when the main channel is noiseless and only $q^m/|C|$ cosets are chosen for encoding. Thus to achieve maximum equivocation, we have to choose a nested code (C', C) where $|C'|$ is as large as possible and $|C|$ is as small as possible. In other words, if we can choose as C' an optimal B-burst-erasure correcting code, i.e. an $[n, n - B]_q$ code and an $[n, \mu]_q$ code C satisfying the property stated above, then we achieve the upper bound for the equivocation $k \leqslant n - B - \mu$, fulfilling both tasks (4.1) and (4.2).

Clearly, these conditions for C and C' are necessary and sufficient. Let k denote the maximum equivocation, given parameters μ, B, and n. Our observation is summarized in the following theorem.

Theorem 4.11. *In a wiretap channel II, with an active eavesdropper that can observe any interval of μ symbols, out of n transmitted symbols from $\mathbb{F}_q$, and erase any interval of B symbols, one can convey securely and with zero error, at most $k = (n - B - \mu)^+$ symbols.*
To achieve the positive secrecy rate $R_s = k/n$ with a nested linear code pair (C', C), where

C' is a mother code and C is a coarse code, the following three conditions are necessary and sufficient:

- $n - B > \mu$.
- *The mother code C' is an $[n, n-B]_q$ optimal burst-erasure (i.e. B-burst erasure) correcting code.*
- *The coarse code $C \subset C'$ is an $[n, \mu]_q$ code such that its dual code $C^{\perp}$ is an $[n, n-\mu]_q$ optimal burst-erasure (i.e. μ-burst erasure) correcting code. The equivalent condition is that every μ consecutive columns of a generator matrix of C are linearly independent.*

The next theorem shows the existence of secure nested codes (C', C) satisfying the conditions of Theorem 4.11. For ease of description, we denote $m = n - B$.

Theorem 4.12. *(i) For arbitrary admissible parameters n, m, μ, that is for $1 \leqslant \mu < m \leqslant n$, and a finite field F_q with the non-binary alphabet, there exist explicit constructions of secure nested codes (C', C) that achieve the maximum secrecy rate R_s, (i.e. codes satisfying the conditions of Theorem 4.11).*

(ii) Such binary codes (C', C) exist for the following cases:
1) $1 \leqslant \mu < m \leqslant n/2$,
2) $n/2 \leqslant \mu < m < n$,
3) $1 \leqslant \mu < n/2 < m < n$,
where $n = 2Bt$ if $\mu \leqslant B$, and $n = 2\mu t$ if $\mu > B$, with $t \in \mathbb{N}$.

This theorem will be proved in Section 4.6.

We note here that similar arguments, as for Theorem 4.11 (together with Theorem 4.9), give us the following necessary and sufficient conditions for achieving $R_s = k/n$, in the case when the active eavesdropper is able to observe μ symbols and erase B symbols by their own choice:

- C' and C are optimal respectively $[n, n-B]_q$ and $[n, \mu]_q$ erasure correcting codes.

This means that both C' and C are MDS codes (see Remark 4.3). The condition above can be achieved if $n \leqslant q + 1$. In particular, for $n \leqslant q - 1$ we can use Reed-Solomon codes [17] which are known to have a nested structure (see Section 2.2). However, there are no known nontrivial MDS codes with $n > q + 2$ (see [17]). Therefore, it is impractical to use MDS codes for the purpose mentioned above, since in this case q must grow with n.

4.5. Preparations for Code Construction

In this subsection we study matrices over finite fields which have specified properties required for construction of secure nested codes. We start with some new definitions.

Definition 4.3.

- *An $m \times n$ $(m \leqslant n)$ matrix G over a given finite field is called good if every m consecutive columns in it are linearly independent.*
- *We call an $m \times n$ matrix G cyclically good (or c-good for short) if any m cyclically consecutive columns of G are linearly independent.*

The following observation is obvious.

Proposition 4.13. *Let G be an $m \times n$ c-good matrix. Then:*

- *$(G\ G)$ is also a c-good matrix.*
- *If $G = (I_m\ A)$, then $(I_m\ I_m\ A)$ is a c-good matrix.*

Let (C', C) be a secure nested code with given parameters n; $m = n - B$ and μ satisfying the properties in Theorem 4.11. Then we can rephrase these properties in terms of the matrices defined above as follows:

- A parity check matrix $H_{C'}$ of the mother code C' is an $(n-m) \times n$ good matrix.
- A generator matrix G_C of the coarse code C is a $\mu \times n$ good matrix.

Clearly (by Proposition 6.11), these properties are fulfilled if both $G_{C'}$ and G_C are c-good matrices. In this case C' is also capable of correcting all wrap-around B-bursts. Moreover, (C', C) remains secure if Eve also observes all cyclically consecutive intervals of length μ. In our construction of nested (C', C) codes, we essentially use c-good matrices. Therefore we are now interested in how to construct c-good matrices over finite fields. This problem has been solved by Hollmann and Tolhuizen in [32]. They gave explicit constructions of c-good $k \times n$ matrices over $\mathbb{F}_q$ for all parameters k, n and q. The following result shows that c-good matrices can be constructed recursively.

Theorem 4.14 ([32]). *For every c-good $m \times n$ matrix, one can add a column such that the resulting $m \times (n+1)$ matrix is c-good.*

Another construction in [32] is given by means of the $2^r \times 2^r$ binary matrix M_r defined as follows. Let M_1 be the matrix

$$M_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

and for $r \geqslant 1$, M_{r+1} is defined as $\begin{pmatrix} M_r & 0 \\ M_r & M_r \end{pmatrix}$. In other words, M_r is the r-th Kronecker power $M_1^{\otimes r}$ of the matrix M_1. Thus, for example

$$M_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Note that M_r is a lower triangular matrix and it is symmetric with respect to the second diagonal. The following property of the matrix M_r allows to construct c-good $k \times n$ matrices for all k and n.

Theorem 4.15 ([32]). *Let k and $n-k$ be positive integers, and let r be the smallest integer such that $k, n-k \leqslant 2^r$. Let Q be a $k \times (n-k)$ matrix residing in the lower left corner of M_r. Then $(I_k \; Q)$ is a $k \times n$ c-good matrix.*

Note that the matrix Q is not necessarily c-good, although it is a good matrix (in view of Theorem 4.15). The following property of M_r is not mentioned in [32].

Proposition 4.16. *Every $k \times 2^r$ submatrix formed by the last k rows of M_r is a c-good matrix.*

The proposition follows from a more general statement given below.

Definition 4.4. *An $m \times n$ $(m \leqslant n)$ matrix M is called a nested c-good (resp. good) matrix if every $k \times n$ $(k \leqslant m)$ submatrix formed by its last k rows is a c-good (resp. good) matrix.*

Proposition 4.17. *Let M be an $n \times n$ nested c-good matrix. Then so is the matrix*

$$\begin{pmatrix} M & 0 \\ M & M \end{pmatrix}.$$

This proposition is a special case of the following theorem.

Lemma 4.18. *Let A and D be respectively $k \times k$ and $n \times n$ nested c-good matrices over $\mathbb{F}_q$. Then $A \otimes D$ is a $kn \times kn$ nested c-good matrix, where "$\otimes$" is the Kronecker product.*

Proof. Let $A = (a_{ij})_{i,j=1,\ldots,k}$ and $D = (d_{ij})_{i,j=1,\ldots,n}$, thus

$$A \otimes D = \begin{pmatrix} a_{11}D & a_{12}D & \ldots & a_{1k}D \\ a_{21}D & a_{22}D & \ldots & a_{2k}D \\ \vdots & \vdots & \ldots & \vdots \\ a_{k1}D & a_{k2}D & \ldots & a_{kk}D \end{pmatrix}.$$

Let $\mathbf{d}_1, \ldots, \mathbf{d}_n$ be the columns of D and denote $T_r = (a_{r1}D \; a_{r2}D \; \ldots \; a_{rk}D)$, where $r \in \{1, \ldots, k\}$. Now suppose there exists a nonzero vector $\alpha^{kn} = (\alpha_1, \ldots, \alpha_{kn})$ such that $T_r(\alpha^{kn})^T = 0$. Since the columns of D are linearly independent, it follows that there exists a column $\mathbf{d}_j$ of D and a nonzero subsequence $\alpha_j, \alpha_{j+n}, \alpha_{j+2n}, \ldots, \alpha_{j+(k-1)n}$ of α^{kn}, such that

$$\mathbf{d}_j \sum_{i=0}^{k-1} a_{r,i+1}\alpha_{j+in} = 0$$

and hence

$$\sum_{i=0}^{k-1} a_{r,i+1}\alpha_{j+in} = 0.$$

Let T be the submatrix of $A \otimes D$ formed by its last m rows. We have to show that T is a c-good matrix. Note first that this is the case if $m \leqslant n$. This clearly follows from the fact that D and hence T_k is a nested c-good matrix. Now let $m = nr + t$, where $1 \leqslant r \leqslant k-1$ and $0 \leqslant t < n$. Thus,

$$T = \begin{pmatrix} T'_{k-r} \\ T_{k-r+1} \\ \vdots \\ T_k \end{pmatrix},$$

where T'_{k-r} consists of the last t rows of T_{k-r}.

Let Q be an $m \times m$ submatrix of T formed by m cyclically consecutive columns of T. Note first that the set $\mathcal{D}_n := \{d_{n1}, d_{n2}, \ldots, d_{nn}\}$ (the elements of the last row in D) consists of nonzero elements, since D is a nested c-good matrix. Now suppose there exists a nonzero vector $\beta^m = (\beta_1, \ldots, \beta_m)$ such that $Q(\beta^m)^T = 0$. Then it is not hard to see that our observation above implies the following. There exists a nonzero subsequence $\beta_s, \beta_{s+n}, \ldots, \beta_{s+nr}$ of β^m, where $1 \leqslant s \leqslant n$ and an element $d_{ns} \in \mathcal{D}_n$, such that

$$d_{ns} \sum_{i=0}^{r} a_{j,s+i} \cdot \beta_{s+in} = 0, \;\; j = k-r, \ldots, k,$$

where the indices of β are taken modulo kn and the indices of a are taken modulo k. But

this means that there are $r+1$ (if $t \geqslant 1$) cyclically consecutive columns in the last $r+1$ rows of A which are linearly dependent, which is a contradiction. Similarly, for t $= 0$ we have r linearly dependent columns in A. This completes the proof. ∎

Lemma 4.19. *Let M be an $m \times m$ nested c-good matrix which is also symmetric with respect to the second diagonal. Then for any $k \times (n-k)$ submatrix Q of M, residing in the lower left corner, $(I_k\ Q)$ is a $k \times n$ c-good matrix.*

Proof. To prove the statement we have to show that both $(I_k\ Q)$ and $(Q\ I_k)$ are good matrices. This can be easily demonstrated with the help of the figure below.

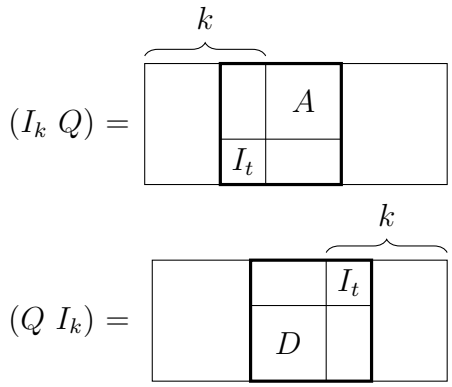


Figure 4.2.: Good matrices $(I_k\ Q)$ and $(Q\ I_k)$.

Consider the $k \times k$ matrices M_1 and M_2 indicated in the Fig. 4.2 with bold line shapes. It follows from the properties of M that both $(k-t) \times (k-t)$ matrices A and D, where $\max\{2k-n, 1\} \leqslant t \leqslant k-1$, are invertible. This clearly implies that both $k \times k$ matrices M_1 and M_2 are invertible as well, which completes the proof. ∎

We note that Lemma 4.18 (together with Lemma 4.19) gives a proof for Theorem 4.15, which differs from the one in [32]. Moreover, it gives a possibility to construct a wider class of nested $n \times n$ c-good matrices than that of M_r matrices (with $n = 2^r$). For example consider the matrices $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ and $M_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ over $\mathbb{F}_3$. Clearly, both are nested c-good. Then by Lemma 4.18, every matrix $A^{\otimes m} \otimes M_1^{\otimes k}$ is nested c-good, where $m, k \in \mathbb{N} \cup \{0\}$. In general, it can be shown that for a given prime power q, there exist c-good $n \times n$ matrices over $\mathbb{F}_q$ for all $2 \leqslant n \leqslant q$. This together with Theorem 4.18 gives us a new class of c-good matrices over $\mathbb{F}_q$. We will go into this in more detail in Section 5.3.

4.6. Proof of Theorem 4.12

We give now explicit constructions of secure nested codes (C', C) that achieve the maximum secrecy rate. The nested codes are given by means of generator matrices for C' and

C which have properties clarified in the previous subsection. We denote by $m = n - B$ the dimension of the mother code C'.

i) First we give a recursive construction of required codes, over every non binary finite field aphabet, for all admissible parameters μ, B, n. We use Theorem 4.14 and adapt it to our nested codes. Let $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$ be an $m \times n$ c-good matrix, where G is a $\mu \times n$ good matrix. Such a matrix G' clearly exists for $n = m$ and $1 \leqslant \mu < m$. Our goal is to show that we can add a column $\mathbf{x}$ to G' such that the resulting matrix $(G'\ \mathbf{x})$ gives us a secure nested code with parameters $n+1, m, \mu$.

Let $\mathbf{g_1}, \ldots, \mathbf{g_n}$ be the columns of G'. Consider then the submatrices $S_1, \ldots, S_m$ of G', with $S_i = (\mathbf{g_{n-m+i+1}}, \mathbf{g_{n-m+i+2}}, \ldots, \mathbf{g_{n+i-1}})$; $i = 1, \ldots, m$, where the column indices greater than n are taken modulo n. In other words, the index sets of the columns in S_i consist of the interval $\{n-m+2, \ldots, n\}$ and its $m-1$ right cyclic shifts. Note that $(G'\ \mathbf{x})$ is a c-good matrix if all matrices $(S_i\ \mathbf{x})$ are invertible. Let now D be an $m \times m$ matrix D with rows $\{\mathbf{d_1}, \ldots, \mathbf{d_m}\}$, such that $\mathbf{d_i}^T S_i = \mathbf{0}$ with $i = 1, \ldots, m$. It can be shown that the matrix D is invertible (see Lemma 4.24) and hence the matrix DG' is a c-good matrix. By definition of D, in each submatrix DS_i of DG', the ith row consists of zeros. The latter clearly implies that by adding any column vector $\mathbf{v} = (v_1, \ldots, v_m)^T$ with nonzero entries to DG' we get an $m \times (n+1)$ c-good matrix. Thus, for each $\mathbf{v}$ with nonzero entries there exists a unique column vector $\mathbf{x}$ such that $D\mathbf{x} = \mathbf{v}$. This in turn implies that for every such a vector $\mathbf{x}$, the matrix $(G'\ \mathbf{x})$ is c-good as well.

Let X denote the set of all such column vectors $\mathbf{x}$. We now show that there exists an $\mathbf{x} \in X$ such that by adding it to G' we get an $m \times (n+1)$ matrix where the last μ rows form a good matrix. Without loss of generality, we may assume that $G = (A\ I_\mu)$. Observe then that we are done, if there exists an $\mathbf{x} = (x_1, \ldots, x_m)^T \in X$ with $x_{m-\mu+1} \neq 0$. Suppose for a contradiction that $x_{m-\mu+1} = 0$ for every $\mathbf{x} \in X$. Next, we note that the set of all vectors $\mathbf{v} \in \mathbb{F}_q^m$ with nonzero coordinates spans $\mathbb{F}_q^m$ if $q \neq 2$. This clearly implies that X spans $\mathbb{F}_q^m$ as well, which is a contradiction, in view of our assumption on X. Thus, there exists an $\mathbf{x} \in X$ with $x_{m-\mu+1} \neq 0$. To find such an $\mathbf{x}$ we proceed as follows. Let $\mathbf{d} = (d_{1,m-\mu+1}, \ldots, d_{m,m-\mu+1})^T$ be the $(m-\mu+1)$th column of D. Let $\mathbf{u} = (u_1, \ldots, u_m)^T$ be a column vector with the nonzero coordinates, such that $u_i \neq -d_{i,m-\mu+1}$; $i = 1, \ldots, m$. Furthermore, let $\mathbf{x}' = (x'_1, \ldots, x'_m)^T$ be such that $D\mathbf{x}' = \mathbf{u}$. Now if $x'_{m-\mu+1} \neq 0$, then $\mathbf{x} = \mathbf{x}'$. Otherwise, we take $\mathbf{x} = (x_1, \ldots, x_m)^T$, where $x_{m-\mu+1} = 1$ and $x_i = x'_i$ elsewhere. Then we have $D\mathbf{x} = \mathbf{v}$, where $\mathbf{v} = \mathbf{u} + \mathbf{d}$ has nonzero coordinates. This completes the proof.

Remark 4.20. *We note that the code C' generated by G' also tolerates wrap-around B-burst erasures, that is reliable and secure transmission is provided when the eavesdropper is able to cause any B-burst erasure including wrap-around bursts.*

ii) We turn now to the binary case. We start with a special case $n = 2^r$ where we can

directly apply Proposition 4.16. In this case, for any given parameters m and μ, we just take the last m rows of M_r for the generator matrix G' of C'. This matrix is c-good, that is C' is a c-optimal burst-erasure correcting $[n, m]$ code. The matrix G' in turn contains the c-good $\mu \times n$ submatrix G formed by the last μ rows. This clearly gives us a secure nested code (C', C), where C is the $[n, \mu]$ (c-optimal) code, satisfying the properties of Theorem 4.11. Note that the same construction works with any $n \times n$ nested c-good matrix. In fact, this is the simplest way to construct a secure nested code (C', C). However, nested c-good $n \times n$ matrices do not exist for every $n \in \mathbb{N}$. For example, in the binary case, it can be easily shown that such matrices do not exist when n is odd.

1) Case $\mathbf{1 \leqslant \mu < m \leqslant n/2}$: For the construction of a secure nested code (C', C) we need two auxiliary results.

Lemma 4.21. *Let C be an $[n, k]_q$ code with a generator matrix G_C and a parity check matrix H_C. Let $J \subset I$ be a subset of the index set $I = \{1, \ldots, n\}$ with $|J| = k$. If the columns of G_C with indices in J are linearly independent, then the columns of H_C with indices in $I \backslash J$ are linearly independent.*

Proof. This is a direct consequence of Proposition 4.1. ∎

Lemma 4.22. *Let $G' = (A\ D)$ be a $k \times n$ good matrix over $\mathbb{F}_q$ with $1 \leqslant k \leqslant n/2$, where A is a $k \times k$ matrix and D is a $k \times (n-k)$ matrix. Then the code C with generator matrix $G = (D\ A)$ is an optimal burst-erasure correcting code.*

Proof. Let $I' = \{1, \ldots, n\}$ be the (ordered) set of column indices in G' and let C' be the code generated by G'. Let also $\mathcal{L}$ be the set of all intervals of length k in I'. Since for every $L \in \mathcal{L}$ the columns of G' with indices in L are linearly independent, Lemma 4.21 implies that the columns of $H_{C'}$ with indices in $I' \backslash L$ are linearly independent. Then, by Proposition 4.1, for every $L \in \mathcal{L}$ the code C' generated by G' can correct the burst of erasures in positions $I' \backslash L$. Note now that in the new ordering of the columns $I = \{k+1, \ldots, n, 1, \ldots, k\}$ the set of subsets $\{I' \backslash L : L \in \mathcal{L}\}$ contains all intervals of length $n-k$ in I. This means that C can correct all bursts of erasures of length $n-k$, and hence C is optimal. ∎

Given positive integers r, n, m with $n - \mu \leqslant 2^r$, let M_r be the matrix defined in the previous subsection. Let $\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ be the $m \times (n-\mu)$ submatrix of M_r, residing in the lower left corner, where A_1 is an $(m-\mu) \times (n-\mu)$ matrix and A_2 is an $\mu \times (n-\mu)$ matrix. Define the $m \times n$ matrix $\bar{G}$ as

$$\bar{G} = \begin{pmatrix} \mathbf{0} & A_1 \\ I_\mu & A_2 \end{pmatrix}.$$

In view of Theorem 4.15, $(I_\mu\ A_2)$ is an $\mu \times n$ c-good matrix and $\bar{G}$ is an $m \times n$ good matrix. Let now $\mathbf{g}_1, \ldots, \mathbf{g}_n$ be the columns of $\bar{G}$ and denote by G' the matrix defined as

$$G' = (\mathbf{g}_{m+1} \dots \mathbf{g}_n \ \mathbf{g}_1 \dots \mathbf{g}_m) = \begin{pmatrix} G^* \\ G \end{pmatrix},$$

where G^* and G are the resulting submatrices obtained from $(\mathbf{0} \ A_1)$ and $(I_\mu \ A_2)$ after the corresponding permutation of columns. In fact, G' is obtained by m cyclic shifts of the columns of $\bar{G}$. By Lemma 4.22, the code C' generated by G' is an (optimal) B-burst-erasure correcting code. Denote by C the code generated by G. Since $(I_\mu \ A_2)$ is a c-good matrix, any cyclic shift of its columns also gives a c-good matrix. Thus, (C', C) is a secure nested code with maximum secrecy rate.

2) Case $\boldsymbol{n/2 \leqslant \mu < m < n}$: This case is the "dual" to the previous case and follows from the proposition below.

Proposition 4.23. *If (C', C) is a secure nested code satisfying the properties of Theorem 4.11, then so is the nested code $(C^\perp, C'^\perp)$.*

Proof. Let (C', C) be a secure nested code where C' is an $[n, m]$ code and C is an $[n, \mu]$ code. Thus, C' is an optimal burst-erasure correcting code, and $C^\perp$ is an optimal $[n, n-\mu]$ burst-erasure correcting code. Furthermore, $C'^\perp \subset C'$ is an $[n, n-m]$ code, such that its generator matrix is an $(n-m) \times n$ good matrix. Thus, we have a secure nested code $(C^\perp, C'^\perp)$ with new parameters $m_1 = n - \mu$, $\mu_1 = n - m$, $B_1 = \mu$, where $n/2 \leqslant \mu_1 < m_1 < n$. ∎

3) Case $\boldsymbol{1 \leqslant \mu < n/2 < m < n}$: In this case our construction extends only to specified parameters. We distinguish between two subcases.

(i) $\mu \leqslant B$.

Let $n = 2Bt$, with $t \in \mathbb{N}$. Consider the following c-good $B \times n$ matrix $H = (\underbrace{I_l \ I_l \dots I_l}_{2t})$. Let C' be the $[n, n-B]$ code with the parity check matrix H. By Proposition 4.7, a generator matrix of C' is also c-good and C' is a c-optimal B-burst-erasure correcting code. Note now that $HH^T = 0$, that is the dual code $C'^\perp$ is self orthogonal, i.e. $C'^\perp \subset C'$. Since for every $1 \leqslant \mu \leqslant B$ there exists a c-good $\mu \times B$ matrix A, the row space of H contains a $\mu \times n$ submatrix $G = (A \ A \dots A)$, which is also c-good in view of Proposition 4.13. This implies that there exists a $m \times n$ generator matrix G' of C', such that it contains a c-good $\mu \times n$ submatrix. Thus we have a secure nested code (C', C), where C is the $[n, \mu]$ code generated by G.

(ii) $\mu > B$.

In this case we take $n = 2\mu t$, with $t \in \mathbb{N}$ and proceed similarly. We consider a c-good matrix $G = (\underbrace{I_\mu \ I_\mu \dots I_\mu}_{2t})$ and note that $GG^T = 0$. Clearly, G can be transformed to a matrix $\hat{G}$ (by linear operations on rows) such that $\hat{G}$ contains a $B \times n$ c-good submatrix H. We now consider the $[n, n-B]$ code C' with parity check matrix H. Note that C' is a c-optimal burst-erasure correcting code since H is a c-good matrix. Thus

$$C' = \left\{ x^n \in \mathbb{F}_q^n : H(x^n)^T = 0 \right\}.$$

Since $GG^T = 0$ implies in particular that $HG^T = 0$, we conclude that the row space of G is a subspace of C'. Therefore, the code C' contains the rows of a c-good $\mu \times n$ submatrix G. Hence, there exists a generator matrix G' of C', which contains G, taken as a generator matrix for the coarse code C. Thus, we get a secure nested code (C', C) satisfying conditions of Theorem 4.11.

4.7. Encoding and Decoding Schemes

In this section we present an encoding and decoding procedure for the secure nested codes described in the previous section. By encoding we mean here the channel encoding and each message as before is identified with a k-vector over a fixed finite field. The decoding consists of two steps: (1) channel decoding, i.e. codeword recovering, and (2) message decoding. Let (C', C) be a linear nested code pair achieving maximum secrecy rate with zero-error probability. Let C' (the mother code) be an $[n, m = n - B]_q$ code and C (the coarse code) be an $[n, \mu]_q$ code. Recall that C' is a B-burst-erasure correcting code and $C \subset C'$ has the property that any μ consecutive columns of its generator matrix are linearly independent. The maximum number of symbols that can be securely transmitted equals $k = m - \mu = n - B - \mu$. Let us represent C' as $C' = C^* + C$, where C^* is an $[n, m - \mu]_q$ subcode of C' such that $C^* \cap C = 0$.

Furthermore, let $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$ be a generator matrix of C', where G^* and G are generator matrices of C^* and C respectively. Observe now that we can choose the generator matrices G^* and G having the form shown in Fig. 4.3, where $\mathbf{0}$ is a $k \times (m - k)$ all-zero matrix and A is an $(m - k) \times k$ matrix.

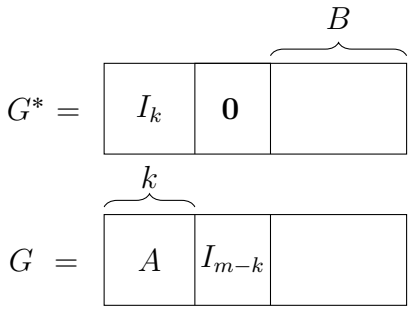


Figure 4.3.: Generator matrix G'.

This is clear because every $m - k = \mu$ and m consecutive columns of generator matrices of C and C' are linearly independent, respectively, and hence any generator matrix of C' can be transformed to G and G^* by elementary row operations.

Let us denote $G_1 = (A \; I_{m-k})$ and $H_1 = (I_k \; -A^T)$. Thus, H_1 is a parity check matrix of the code generated by G_1. We are prepared now to describe the encoding and decoding of a message tent through the main channel.

Encoding

A message $(s_1, \ldots, s_k)$, is encoded to the codeword

$$(x_1, \ldots, x_n) = (s_1, \ldots, s_k, e_1, \ldots, e_{m-k}) \begin{pmatrix} G^* \\ G \end{pmatrix},$$

where $(e_1, \ldots, e_{m-k}) \in \mathbb{F}_q^{m-k}$ is chosen uniformly at random.

Suppose $y^n = (y_1, \ldots, y_n)$ is a vector received by Bob when the codeword $x^n = (x_1, \ldots, x_n)$ has been sent. Let $\{i, i+1, \ldots, i+t-1\}$ be the coordinate positions where a burst of erasures of length $t \leqslant B$ have occurred.

Channel Decoding

Let H be a $B \times n$ parity check matrix of the code C' with the columns denoted by $\mathbf{h}_1, \ldots, \mathbf{h}_n$. Recall that any B consecutive columns of H are linearly independent. Considering the erased symbols $(y_i, \ldots, y_{i+t-1})$ as unknowns and taking into account that $H(x^n)^T = \mathbf{0}$, we have

$$\sum_{r=i}^{i+t-1} y_r \mathbf{h}_r = -\sum_{j \in T} x_j \mathbf{h}_j.$$

This system of linear equations with at most B unknowns $(y_i, \ldots, y_{i+t-1})$ has a unique solution, since the columns $\mathbf{h}_i, \ldots, \mathbf{h}_{i+t-1}$ are linearly independent.

Message Decoding

If the submitted codeword $(x_1, \ldots, x_m)$ is successfully recovered, then we claim that

$$(s_1, \ldots, s_k) = (x_1, \ldots, x_m) H_1^T.$$

To show that the equality holds, we note that

$$(x_1, \ldots, x_m) = (s_1, \ldots, s_k, \underbrace{0, \ldots, 0}_{m-k}) + (e_1, \ldots, e_{m-k}) G_1.$$

Then

$$(x_1, \ldots, x_m) H_1^T = (s_1, \ldots, s_k, 0, \ldots, 0) H_1^T + (e_1, \ldots, e_{m-k}) G_1 H_1^T.$$

Since $G_1 H_1^T = 0$, we have

$$(x_1, \ldots, x_m) H_1^T = (s_1, \ldots, s_k, 0, \ldots, 0) H_1^T = (s_1, \ldots, s_k).$$

4.8. Low Complexity Channel Decoding

We note that the channel decoding approach described above is a standard decoding technique, which is a kind of syndrome decoding and can be applied to any erasure correcting linear code capable of correcting a given number of erasures. This technique however, is not in general efficient. A suitable approach for erasure correction is the iterative decoding, which is a powerful technique, especially, when applied to low density parity check (LDPC) codes [33]. The basic idea of iterative decoding is to correct erasures one-by-one. In each step a parity check equation is used, which involves precisely one erasure position, thus allowing this erasure to be corrected. More specifically, let C be a binary $[n, k]$ code capable of correcting B erasures. Let H be a matrix whose rows span the dual code $C^{\perp}$. Thus H is a parity check matrix of C, possibly with some redundant vectors from $C^{\perp}$. Let also $\mathbf{h}_j = (h_{j,1}, \ldots, h_{j,n})$, with $j = 1, \ldots, r$ $(r \geqslant n - k)$ being the rows of H. Thus, for any submitted codeword $x^n = (x_1, \ldots, x_n)$, we have

$$\sum_{i=1}^{n} x_i h_{j,i} = 0, \;\; j = 1, \ldots, r.$$

Now let $y^n = (y_1, \ldots, y_n)$ be the received vector with B erased positions, when x^n has been sent. Without loss of generality we may assume that $y^n = (?, \ldots, ?, x_{B+1}, \ldots, x_n)$. Suppose now there exists an $\mathbf{h}_j \in H$ such that $\mathbf{h}_j$ contains precisely one 1 in the erased positions, for example $\mathbf{h}_j = (1, 0, \ldots, 0, h_{j,B+1}, \ldots, h_{j,n})$. Then clearly we can correct the first erasure in y^n since we have

$$1 \cdot y_1 + \sum_{i=B+1}^{n} x_i h_{j,i} = 0,$$

and hence $y_1 = a$, where $a = \sum_{i=B+1}^{n} x_i h_{j,i}$ is known. This procedure is repeated until all erasures in y^n are corrected, or the procedure stops if no parity check $\mathbf{h}_j$, with the above property, can be found for the set of current erasures. Therefore, we can correct all erasure patterns if for each such pattern there exists a parity check $\mathbf{h}_j$ which contains a single 1 in the corresponding positions. Thus, for this decoding method, the choice of a parity check matrix H (defined in a more general way) plays a crucial role. Recall that for the standard decoding mentioned above, the choice of H does not play any role. Notice now that for burst-erasure correcting codes, a weaker condition is required for the successful use of an iterative decoding approach. Namely, given a parity check matrix H, one can correct all B-burst erasures if for every burst erasure pattern of length B or less, there exists a parity check $\mathbf{h}_j \in H$ which contains a single 1 in an erased position. In [34], Fossorier showed that using iterative decoding approach to any binary B-burst-erasure correcting $[n, k]$ code, the decoder complexity is $O(n^2)$. In other words it is possible to choose a parity check matrix H such that at most $O(n^2)$ binary operations are needed for

successful decoding of any burst of length B or less.

We now show that in a special cases of our constructions of (C', C) nested codes where C' is an $[n, m]$ optimal code, we can achieve decoding complexity not exceeding n.
Let $n = 2Bt$, $t \in \mathbb{N}$ (i.e. $n = \frac{mt}{2t-1}$), $\mu \leqslant B$.
The construction of a secure nested code (C', C) for this case is described in subsection 4.12. C' is a c-optimal $[n, m]$ code given by the $B \times 2Bt$ parity check matrix $H = (I_B \ I_B \dots I_B)$. Obviously H satisfies the required property for correction of all burst erasures of length B or less. Moreover we need at most $2tB$ binary operations for the correction of any B-burst erasure.

4.9. Conclusion

A model of a wiretap channel II with an active eavesdropper has been introduced and studied. We have shown that with a coset coding approach, one can convey securely and with zero-error decoding at most $k = (n - B - \mu)^+$ symbols and consider necessary and sufficient conditions for achieving the maximum secrecy rate $R_s = k/n$. Linear nested codes achieving maximum secrecy rate have been constructed for all admissible parameters. The constructed nested codes provide also perfect security and zero error at the receiver. The nested code consists of a so-called mother code C' and a coarse code $C \subset C'$. Zero-error decoding and perfect security can be achieved for given parameters n, B and μ, if and only if C' is an optimal $[n, n - B]_q$ burst-erasure correcting code and the dual code $C^\perp$ of C is an optimal $[n, n - \mu]_q$ burst-erasure correcting code, respectively.

Further, we showed that an iterative decoding approach can be effectively applied for our constructions of nested codes.
We find it interesting to study other models of the wiretap channel II with an active eavesdropper. An initial problem in this direction could be the study of a model where the eavesdropper can observe any μ consecutive symbols and is able to cause any burst of errors of length B in the main channel. Furthermore, it would be interesting to consider the models in the streaming setup where the legitimate receiver is subject to a delay constraint.

4.10. Appendix

Lemma 4.24. *Let* $\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{m-1}, \mathbf{s}_m, \mathbf{s}_{m+1}, \dots, \mathbf{s}_{2m-2})$ *be an* $m \times m$ *matrix with values in* $\mathbb{F}_q$ *and with its columns* $\mathbf{s}_i$, $i = 1, \dots, 2m-2$ *of length* m. *Let*

$$\mathbf{S}_i = (\mathbf{s}_i, \mathbf{s}_{i+1}, \mathbf{s}_{i+2}, \dots, \mathbf{s}_{i+m-1}) \ , \qquad i = 1, 2, \dots, m \ ,$$

be the $m \times m$ submatrices obtained by m consecutive columns of $\mathbf{S}$ and let

$$\mathbf{Q}_i = (\mathbf{s}_{i+1}, \mathbf{s}_{i+2}, \ldots, \mathbf{s}_{i+m-1}) \ , \qquad i = 1, 2, \ldots, m \ ,$$

be the $m \times m-1$ submatrices obtained by deleting the first column in each $\mathbf{S}_i$, i.e. $\mathbf{S}_i = (\mathbf{s}_i, \mathbf{Q}_i)$. Assume $\mathbf{S}$ has the property that

$$\operatorname{rank}(\mathbf{S}_i) = m \qquad \textit{and} \qquad \operatorname{rank}(\mathbf{Q}_i) = m - 1 \ , \quad \textit{for all } i = 1, 2, \ldots, m \ .$$

Then there exists an invertible $m \times m$ matrix $\mathbf{D}$ with columns $\mathbf{d}_i^T$, $i = 1, \ldots, m$ so that

$$\mathbf{d}_i^T \mathbf{Q}_i = \mathbf{0} \, , \qquad \textit{for all } i = 1, 2, \ldots, m \ . \tag{4.3}$$

Proof. Each $\mathbf{Q}_i$ has rank $m-1$. Therefore it possess a non-trivial one-dimensional (left) null space $\mathcal{N}(\mathbf{Q}_i)$ and so we choose $\mathbf{d}_i \in \mathcal{N}(\mathbf{Q}_i^T)$ for each $i = 1, \ldots, m$. By this choice, it is clear that (4.3) it satisfied and so we only need to show that $\mathbf{D}$ is invertible.

To this end, let $\mathbf{A} = \mathbf{D}\mathbf{S}_1$ be the $m \times m$ matrix whose entry $a_{k,i}$ in row k and column i is given by the scalar product of $\mathbf{d}_k$ and $\mathbf{s}_i$, i.e.

$$a_{k,i} = \mathbf{d}_k^T \mathbf{s}_i \ , \qquad k, i \in \{1, 2, \ldots, m\} \ .$$

It follows from (4.3) that $\mathbf{A}$ is a lower triangular matrix, i.e. $a_{k,i} = 0$ whenever $i > k$. Moreover, all diagonal entries of $\mathbf{A}$ are nonzero, i.e. $a_{k,k} \neq 0$ for all $k = 1, 2, \ldots, m$. Indeed, $a_{k,k} = \mathbf{d}_k^T \mathbf{s}_k = 0$ would imply, in connection with (4.3), that $\mathbf{d}_k^T \mathbf{S}_k = 0$ contradicting the assumption that $\mathbf{S}_k$ has rank m. So since all diagonal entries of the lower triangular matrix $\mathbf{A}$ are nonzero, it follows that $\det(\mathbf{A}) = \det(\mathbf{D}) \det(\mathbf{S}_1) \neq 0$. Since $\operatorname{rank}(\mathbf{S}_1) = m$, we have $\det(\mathbf{S}_1) \neq 0$ and so it follows that $\det(\mathbf{D}) \neq 0$, i.e. $\mathbf{D}$ is invertible. ∎

5. Delay-Optimal Codes for a Burst-Erasure Wiretap Channel

5.1. Introduction

We consider transmission of secure messages over a burst-erasure wiretap channel under decoding delay constraint. In many emerging communication systems such as interactive voice and video communication, internet of things, etc., low-delay is an important task along with reconstruction of corrupted or lost data. Such systems are highly susceptible to sporadic burst packet losses. The transmitter must encode a source stream of packets sequentially, and the receiver must recover each source packet within a fixed playback deadline. This naturally motivates the study of codes that achieve fast recovery from burst losses. Moreover, communication systems that convey secret data, e.g. electronic payment systems, must be protected against eavesdropping. Classical encryption methods only offer security against eavesdropping if the encryption algorithms are sufficiently complex and the eavesdropper's computing power is limited. Since these security mechanisms can only be implemented at higher protocol layers, this leads to noticeable delays. To avoid these problems, security must be embedded in the physical layer.

Related Work

Martinian et al. [35],[7],[8] were the first to study low-delay burst-erasure correcting codes. Their bounds and constructions provided the basis for several follow-up works, which considered different scenarios of low-delay communication such as low-delay multiple bursts [36], multicasting [37], [38], average delay scenario [39], etc. Additional works devoted to low-delay coding can be found in [9], [10], [40], [41], [42],[43],[44],[45]. Martinian and Trott [8] presented a construction of delay optimal streaming codes for a burst-erasure channel. A stream of source packets $\{s[i]\}_{i\geqslant 0}$ arrives sequentially at the encoder and is mapped to a stream of channel packets $\{x[i]\}_{i\geqslant 0}$. Each source packet $s[i]$, respectively each encoded packet $x[i]$, is a vector of k symbols, resp. n symbols, from the same finite field. The rate of the code is defined as $R = k/n$. The channel can introduce a burst of erasures of length B, starting at any time i. The decoder is required to reconstruct each source packet with delay of at most T, i.e. after receiving T subsequent packets. The construction in [8] consists of two steps: first constructing a *delay optimal* $[T + B, T]$

binary systematic block code followed by diagonal interleaving applied to that code. The resulting code is a rate-$T/(T+B)$ convolutional code that achieves the delay-burst bound in [8]: $T/B \geqslant \max\left[1, \frac{R}{1-R}\right]$. A code that meets the bound is called a *delay-optimal* code.

Contribution

We propose delay-optimal block codes as well as streaming codes for secure transmission over a burst-erasure wiretap channel. The block codes are intended for a model of a B-burst-erasure channel where the eavesdropper can noiselessly observe any interval of at most μ symbols from n symbols transmitted to the legitimate receiver. This model can be viewed as a special case of the wiretap channel II introduced by Ozarow and Wyner [4], with an additional requirement of low delay. We give explicit constructions of block codes that achieve maximum secrecy rate, provide perfect security (i.e. the eavesdropper can obtain no information about the secret message) and provide zero-error decoding with minimum decoding delay. For the streaming setup, our model of a burst-erasure wiretap channel is as follows. In any sliding window of size W the eavesdropper is able to observe an interval of at most μ packets by his choice. We present constructions of delay optimal streaming codes that provide perfect security.

Outline

In Section 5.2, we introduce a model of a burst-erasure wiretap channel for a stream of encoded packets. Section 5.3 includes definitions and the construction of special matrices required for the construction of delay-optimal secure burst-erasure correcting (DO-SBE) block codes. In Section 5.4 we present explicit constructions of two classes of DO-SBE block codes over any finite field of order of at least three. The first is for systematic and non-systematic block codes where $B|T$ and $\mu \leqslant T - B$, and the second is for systematic block codes for arbitrary $T \geqslant 2B$ with $\mu = T - B$. In Section 5.5, we use our DO-SBE block codes to obtain delay-optimal burst-erasure convolutional codes by applying proper diagonal interleaving. The resulting codes are shown to have perfect security. We derive an upper bound for the secrecy rate of a delay-optimal streaming code and show that this bound is achieved for a certain class of code parameters. Section VI concludes with a discussion and problems for future research.

5.2. The Channel Model in the streaming setup

We consider the burst-erasure wiretap channel illustrated in Fig. 5.1, where each time $i \geqslant 0$ the randomized encoder observes a source packet $s[i]$ and transmits a channel packet $x[i]$. The source packet consists of k symbols, while the channel packet consists of n symbols over a common finite field $\mathbb{F}_q$. For each $i \in \mathbb{Z}^+$, the randomized encoding

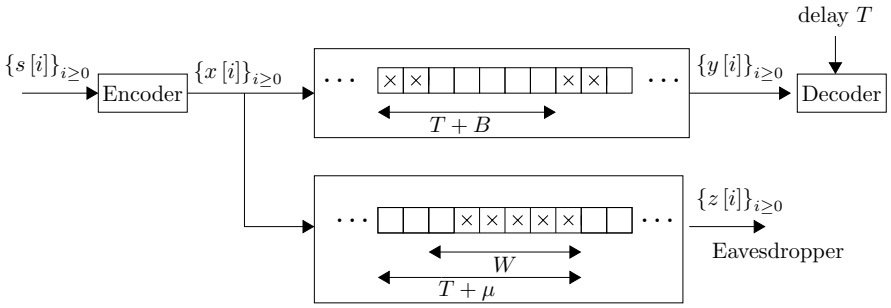


Figure 5.1.: The channel model with erased packets indicated by crossed squares and noiseless received packets by white squares.

function $\xi_i : \mathbb{F}_q^{k\cdot(i+1)} \times \mathbb{F}_q^{\mu\cdot(i+1)} \to \mathbb{F}_q^n$ of the $(n, k, \mu, T)_q$ streaming code maps causally[1] a sequence of source packets $\{s[i]\}_{i\geqslant 0}$ and a sequence of encoder packets $\{e[i]\}_{i\geqslant 0}$ into a channel packet $x[i]$. ξ_i is used by the source at time i to encode $s[i]$ according to

$$x[i] = \xi_i((s\,[0]\,, s\,[1]\,\ldots, s\,[i]), (e\,[0]\,, e\,[1]\,\ldots, e\,[i])). \tag{5.1}$$

The main channel causes erasures, i.e., the received channel packet $y\,[i]$ is either erased (denoted by ?) or passed to the legitimate receiver noiselessly, thus $y\,[i] \in \mathbb{F}_q^n \cup \{?\}$. The erasures occur in bursts of length B. Moreover, the eavesdropper is able to observe an interval of at most μ packets in any sliding window of size W, which implies that any two intervals of μ packets observed by the eavesdropper are separated by at least $W-1$ (undisclosed) packets. We assume that the packets $s[0], s[1], \ldots$ and $e[0], e[1], \ldots$ are realizations of i.i.d. sequences $S_0, S_1, \ldots$ and $E_0, E_1, \ldots$ of random variables which are uniformly distributed over $\mathbb{F}_q^k$ and $\mathbb{F}_q^\mu$, respectively. The eavesdropper's channel output induced by $S_0, \ldots, S_i$ and $E_0, \ldots, E_i$ is $Z_0, \ldots, Z_i$, with realization $z[i] \in \mathbb{F}_q^n \cup \{?\}$ where $i = 0, 1, \ldots$.

If a B-burst-erasure occurs, we require that an $(n, k, \mu, T)_q$ streaming code can reconstruct any source packet $s\,[i]$ with delay T, that is there exists a set of decoding functions φ_i such that $s\,[i] = \varphi_i(y\,[0]\,, \ldots, y\,[i+T])$. In other words, using notation $A_0^i = A_0, \ldots, A_i$, we have $H(S_0^i|Y_0^{i+T}) = 0$, with $i = 0, 1, \ldots$, where Y_i is the random variable that describes the receiver's input. Furthermore, we require perfect security, that is $H(S_0^i|Z_0^{i+T}) = H(S_0^i)$. Informally, the eavesdropper must have complete equivocation over the source packets (messages) in spite of knowing the encoding procedure and the observed packets. In this case we say that a streaming code has secrecy rate $R_s = \frac{k}{n}$. Note that in advance (in the initialization phase), i.e. for $i < 0$, a constant number of random packets $e[i]$ must be securely transmitted to ensure perfect reliability and perfect security in the first T transmitted packets. Since in the initialization phase the number of pre-transmitted packets is constant, the resulting rate loss quickly converges to 0 as

[1] The code is causal if in the encoding function the current channel packet is a function of the current and previous source/encoder symbols of the source/encoder packets.

the number of packet transmissions grows.

Definition 5.1. *We denote an $(n, k, \mu, T)_q$ streaming code as a $(T, B, \mu; W)_q$ streaming code if the code can reconstruct any source packet within delay T if any erasure burst of length B occurs, and if the code provides perfect security even if the eavesdropper is able to observe an interval of at most μ packets in any sliding window of size W.*

5.3. Preparation for the Code Constructions

In this section we provide matrices over finite fields, which have specific properties required for the construction of block codes that we convert into streaming codes for the channel model introduced in Section 5.2.

First we refer to the Definitions 4.2, 4.3 and 4.4.

Lemma 5.1. *For a prime power q and any integer $1 \leqslant n \leqslant q$ there exists an $n \times n$ nested c-good matrix over $\mathbb{F}_q$.*

Proof. Let V_n denote an $n \times n$ Vandermonde matrix over $\mathbb{F}_q$, where $1 \leqslant n \leqslant q-1$, with the rows written in reverse order, i.e. $v_{ij} = a_j^{n-i}$, and $a_1, \ldots, a_n$ are nonzero elements in $\mathbb{F}_q$. Note that V_n is a nested c-good matrix. For $n = q$ we take the $q \times q$ nested c-good matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ & & & \vdots \\ & V_{q-1} & & 0 \\ & & & 1 \end{pmatrix},$$

where V_{q-1} is a $(q-1) \times (q-1)$ Vandermonde matrix. ∎

Lemma 4.18 and 5.1 imply the following theorem.

Theorem 5.2. *For a prime power q, let $p_1, p_2, \ldots, p_{\pi(q)}$ be all primes less than or equal to q. Let $N(q) := \left\{ p_1^{k_1} p_2^{k_2} \cdots p_{\pi(q)}^{k_{\pi(q)}} : k_i \in \mathbb{N} \cup \{0\} \right\}$. Then for any $n \in N(q)$ there exists an $n \times n$ nested c-good matrix over $\mathbb{F}_q$.*

Remark 5.3. *Note that any $n \times n$ nested c-good matrix can be brought to a nested c-good lower triangular matrix. Recall that the Kronecker product of any two lower triangular matrices again gives a lower triangular matrix.*

Lemma 5.4. *Let $(I_{T-B}\ A)$ be a $(T-B) \times T$ c-good matrix. Then the matrix G is also c-good.*

$$G = \begin{pmatrix} I_B & \mathbf{0}_{B \times (T-B)} & I_B \\ \mathbf{0}_{(T-B) \times B} & I_{T-B} & A \end{pmatrix} \tag{5.2}$$

The proof is provided in the Appendix.

Matrix G has been used in [8] for the construction of delay-optimal burst-erasure correcting codes.

Theorem 5.5 ([8]). *A $[T+B,T]_q$ block code with a systematic generator matrix of the form (5.2) recovers a burst erasure of length B with delay T.*

5.4. DO-SBE Block Codes

We define block codes over $\mathbb{F}_q$, for small q, of dimension T and blocklength $T+B$ that we can apply in Section 5.5 to construct convolutional codes for the model introduced in Section 5.2.

DO-SBE block codes can be used for delay-optimal transmissions of secret messages over a burst-erasure wiretap channel, where the channel or the eavesdropper (in the case of an active eavesdropper) causes an erasure burst of length B and the eavesdropper observes an interval of at most μ symbols noiselessly from $n = T+B$ symbols. For the construction of DO-SBE block codes we use linear secure nested codes (see Section 4.3.1).

Definition 5.2. *We say that a $[T+B,T]_q$ code $C' = (\text{encoder } \xi, \text{decoder } \zeta)$ is a delay-optimal secure burst-erasure correcting (DO-SBE) block code and call it a $\mu-[T+B,T]_q$ DO-SBE code if*

(i) C' is an optimal burst-erasure correcting code.

(ii) The transmitter can convey $k = T-\mu$ symbols with perfect security and the secrecy rate of the code is $R_s = \frac{T-\mu}{T+B}$, which in fact is maximum possible (see Chapter 4).

(iii) Every source symbol can be reconstructed with delay of at most T (i.e. for a set of decoding functions ζ_i we have $s_i = \zeta_i(y_1,\dots,y_{i+T})$ with $i = 1,\dots,k$), that is the code is delay-optimal. In other words, for any B-burst erasure, all source symbols s_i must be recovered up to receiving y_{i+T}.

Codes satisfying conditions (i) and (ii) (without a delay constraint) are studied in Chapter 4 where the construction of such codes for all admissible parameters T, B, μ, q has been presented. For the construction, a nested code approach has been used. Recall that a nested linear code is a pair (C', C) of linear codes, $C \subset C'$, in $\mathbb{F}_q^n$, where C' is an outer code and C is called a coarse code. The outer code C' is partitioned into cosets of C and each of $|C'|/|C|$ cosets is put into correspondence with a secret message to be sent, by a fixed bijective map. The inner code serves for reliability and the coarse code is used for stochastic encoding, to provide security. The code is given by a matrix $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$, where G', respectively, G, is a generator matrix for the outer code, respectively, for the coarse code.

Remark 5.6. *We emphasize that the codes constructed in Chapter 4 are not appropriate for the given purposes. The codes in Chapter 4 do not guarantee low delay.*

Definition 5.3. *We call a generator matrix* $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$ *for a DO-SBE code* C' *systematic, if both* G^* *and* G *have a systematic form. Equivalently, we say that* C' *is a systematic DO-SBE code, or* C' *has a systematic encoder.*

Lemma 5.7. *For a prime power* $q > 2$*, let* $T \in N(q)$ *and* $\alpha \in \mathbb{F}_q \setminus \{0, 1\}$*. Then for any integer* $1 \leqslant B \leqslant T$ *there exists a* $T \times (T + B)$ *c-good matrix*

$$G' = \begin{pmatrix} I_B & \mathbf{0}_{B\times(T-B)} & I_B \\ A & C & \alpha A \end{pmatrix}, \tag{5.3}$$

where $G := (A \; C \; \alpha A)$ *is a* $(T - B) \times (T + B)$ *nested good matrix and* $(C \; \alpha A)$ *is a* $(T - B) \times T$ *c-good matrix.*

Proof. Let $Q = \begin{pmatrix} M & \mathbf{0}_{B\times(T-B)} \\ A & C \end{pmatrix}$ be a $T \times T$ lower triangular nested c-good matrix over $\mathbb{F}_q$ with $q > 2$. Such a matrix exists for any $T \in N(q)$ in view of Theorem 5.2 (and Remark 5.3). Note then that G is a nested good matrix and $(C \; \alpha A)$ is a $(T - B) \times T$ c-good matrix. Furthermore, $G'' := \begin{pmatrix} M & \mathbf{0}_{B\times(T-B)} & M \\ A & C & \alpha A \end{pmatrix}$, and hence G' in (5.3) are c-good matrices, since G'' can be brought to G' and to a matrix of the form (5.2) by elementary row operations. ∎

In the sequel we will show that there exist delay-optimal block codes that have the same maximum secrecy rate as the codes without delay constraint.

Lemma 5.8 ([46]). *Let* $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$ *be a* $T \times (T + B)$ *generator matrix for a linear block code* C'*, where* G *is a* $\mu \times (T + B)$ *submatrix of* G'*. Then* C' *satisfies requirements (i) and (ii) in Definition 5.2, if the following three conditions are fulfilled:*

(a) $T > \mu$

(b) G' *is a c-good matrix. (Recall that in this case* C' *is also capable of correcting all wrap-around* B*-burst erasures.)*

(c) G *is a good matrix.*

Remark 5.9. *1. In fact, we can replace condition (b) by the following weaker condition:*

(b′) a generator matrix for the dual code $C'^{\perp}$ *is good.*

In this case $(a), (b')$ *and* (c) *are also necessary conditions.*

2. It is also worth mentioning that if G *is a c-good matrix, then requirement (ii) (in Definition 5.2) is satisfied even if the eavesdropper is able to observe any cyclic interval (of codeword positions) of length* μ*.*

Remark 5.10. *When we consider the delay constraints we define the first B source symbols as the urgent symbols and the remaining $T-B$ source symbols as non-urgent.*

Next, we give constructions of $\mu-[T+B,T]_q$ DO-SBE block codes satisfying the conditions in Lemma 5.8 and the delay constraint.

We first analyze the tradeoff between parameters T,B,μ for systematic and non-systematic $\mu-[T+B,T]_q$ DO-SBE block codes, conditioned by construction.

Proposition 5.11. *(i) For a $\mu-[T+B,T]_q$ DO-SBE code C' we have $\mu\leqslant T-B$, or equivalently $B\leqslant k$.*

(ii) For a systematic $\mu-[T+B,T]_q$ DO-SBE code with $\mu>0$ we have $B\leqslant\mu\leqslant T-B$, or equivalently $B\leqslant k\leqslant T-B$. In particular, we have $T\geqslant 2B$.

Proof. Let $G'=\begin{pmatrix}G^*\\G\end{pmatrix}$ be a generator matrix for a $\mu-[T+B,T]_q$ DO-SBE code. Let M be the submatrix of G' formed by its first $T+1$ columns and M_1 be the submatrix of M formed by the deletion of the last B columns and the first row denoted by $\mathbf{m}$. Furthermore, for an output y^{T+B} let the erased set of positions $E(y^{T+B})=\{T-B+2,\ldots,T+1\}$. Observe then that the first source symbol s_1 can be recovered with delay T, only if the columns of M_1 are linearly dependent.

(i) Suppose that $\mu\geqslant T-B+1$. First note that $\mathbf{m}$ contains a nonzero entry in position $i\in\{1,\ldots,T\}$, since the first T columns of M are linearly independent. Clearly, if the nonzero positions of $\mathbf{m}$ are in $E(y^{T+B})$, then s_1 cannot be uniquely recovered. Thus, there exists a nonzero position of $\mathbf{m}$ in $\{1,\ldots,T-B+1\}$. Note then that M_1 has full rank in view of the property of submatrix G, which is a contradiction.

(ii) Now suppose there exists a systematic $\mu-[T+B,T]_q$ DO-SBE code with $\mu<B$. Observe then that again matrix M_1 has full rank, which completes the proof. ∎

Next we present a construction of systematic and non-systematic DO-SBE block codes for any $\mu\leqslant T-B$, in the case when $B|T$.

Theorem 5.12. *Let $T\in N(q)$ and $T=tB$, where $q>2$, $t\in\mathbb{N}$.*

(i) For $0\leqslant\mu\leqslant T-B$ there exists an explicit construction of a $\mu-[T+B,T]_q$ DO-SBE code.

(ii) For $t\geqslant 2$ and $\mu=iB$; $i\in\{1,\ldots,t-1\}$, we have a $\mu-[T+B,T]_q$ DO-SBE code with a systematic encoder.

Proof. (i) Consider the following $T\times(T+B)$ generator matrix

$$G'=\begin{pmatrix}a_{11}I_B & 0 & 0 & \cdots & a_{11}I_B\\ a_{21}M & a_{22}M & 0 & \cdots & \alpha a_{21}M\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ a_{t1}M & a_{t2}M & \cdots & a_{tt}M & \alpha a_{t1}M\end{pmatrix}$$

for a B-burst-erasure correcting code C', where

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \dots & a_{11} \\ a_{21} & a_{22} & 0 & \dots & \alpha a_{21} \\ \vdots & \vdots & \dots & \vdots & \\ a_{t1} & a_{t2} & \dots & a_{tt} & \alpha a_{t1} \end{pmatrix}$$

is a $t \times (t+1)$ c-good matrix over $\mathbb{F}_q$, such that its first t columns form a lower triangular nested c-good matrix. Note that such a matrix exists for any $T \in N(q)$, in view of Theorem 4.18. Furthermore, let M be a $B \times B$ nested c-good lower triangular matrix over $\mathbb{F}_q$ and $\alpha \in \mathbb{F}_q \backslash \{0, 1\}$.

Clearly, without loss of generality, we may assume that $a_{11} = 1$. Note that G' is a matrix of the same form as (5.3). Hence, by Lemma 5.7, G' is a $tB \times (t+1)B$ c-good matrix and its $\mu \times (t+1)B$ submatrix, formed by the last μ rows with $0 \leqslant \mu \leqslant (t-1)B$, is a good matrix. Thus, in view of Lemma 5.8, requirements (i) and (ii) in Definition 5.2 are satisfied.

Now we prove the delay constraint of the code. Suppose $y^{T+B} = (y_1, \dots, y_{(t+1)B})$ is an output of the channel. Then in each subvector $y_i^{t+1} = (y_i, y_{i+B}, \dots, y_{i+tB})$, where $i \in \{1, \dots, B\}$, there is at most one erased symbol. Also note also that y_i^{t+1} is a codeword of the code C_A with generator matrix A, and $y_i^{t+1} = (s_i, \beta_{i_1}, \dots, \beta_{i_{t-1}})A$, where s_i with $i = 1, \dots, B$ is the i-th source symbol in a codeword of C'. Since C_A is a single erasure correcting code with delay t, source symbol s_i can be reconstructed with delay of at most $tB = T$. In case $\mu < T - B$, the symbols $s_{B+1}, \dots, s_{T-\mu}$ are non urgent, that is they can be reconstructed with delay smaller than T. Notice that generator matrix G' is universal in the sense that it can be used for a DO-SBE block code for any $\mu \leqslant T - B$, however it is not systematic and thus can not be mapped to convolutional codes for the channel model given in Section 5.2 with the required properties.

(ii) Let $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ be a $t \times (t+1)$ c-good matrix of the form as (5.3), where $A_1 = (I_r \ \mathbf{0}_{r\times(t-r)} \ a^r)$ with $a^r = (a_1, \dots, a_r)^\top$, $1 \leqslant r \leqslant t-1$, and A_2 is a $(t-r) \times (t+1)$ systematic good matrix. Clearly we can assume that $a_1 = 1$. Then we take $G' = A \otimes I_B$ and obtain a $tB \times (t+1)B$ c-good systematic matrix $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$, where G^* is a $rB \times (t+1)B$ matrix and G is a $(t-r)B \times (t+1)B$ good matrix. We now have $T = tB$ and $\mu = (t-r)B$. Clearly, the delay for recovering each s_i is at most T (as in case (i)).

Example 5.13. *(i) We construct a* $3 - [8, 6]_3$ *DO-SBE code, where* $B = 2$ *and* $T = tB = 6$. *We choose the* 3×4 *matrix* $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 2 & 2 & 0 & 1 \\ 2 & 1 & 1 & 1 \end{pmatrix}$

and matrix $M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. We take $A \otimes M$ and convert by elementary row operations the first B rows in order to obtain

$$G' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 2 & 0 & 2 & 0 & 0 & 0 & 1 & 0 \\ 2 & 2 & 2 & 2 & 0 & 0 & 1 & 1 \\ 2 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 2 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Note that the submatrix of G' consisting of the last $T - B$ rows is nested good. We can convey securely $k = T - \mu = 3$ source symbols, that is, the channel input vector is $(s_1, s_2, s_3, e_1, e_2, e_3)$.

Suppose the burst erasure affects either of the positions $B + 1, \ldots, T$. For example, x_5 and x_6 are erased. For simplicity of description consider the channel output $y^8 = (x_1, \ldots x_4, ?, ?, x_7, x_8)$. Clearly the unerased received symbols correspond to the codeword symbols. The decoder has to reconstruct the urgent source symbols s_1 and s_2 with delay $T = 6$, that is, upon receiving y_7 and y_8, respectively. For $i = 1$, the codeword $y_1^4 = (x_1, x_3, x_5, x_7)$ of C_A has an erasure on position 3 that can be reconstructed upon receiving x_7, since C_A is a single erasure correcting code and thus s_1 can be recovered by solving the following linear system of equations

$x_1 = s_1 + \gamma$

$x_7 = s_1 + \alpha\gamma$,

where $\gamma = a_{21}\beta_{1_1} + a_{31}\beta_{1_2} = a_{21}(s_3 m_{11} + e_1 m_{21}) + a_{31}(e_2 m_{11} + e_3 m_{21})$.

Similarly, we can recover s_2 upon receiving y_8.

Next, suppose an erasure burst occurs in the first or the last B positions. Observe that we can reconstruct s_3, e_1, e_2, e_3 by $x_3, \ldots, x_6$ and the 4×4 nonsingular submatrix of G' residing at the bottom in the middle. Obviously, we can reconstruct s_1 and s_2 with delay less than T, respectively.

(ii) For $T = 6$ and $B = 2$ we choose $\mu = 2$. We take the same $t \times t+1$ matrix $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ $= \begin{pmatrix} 1 & 0 & 0 & 1 \\ 2 & 2 & 0 & 1 \\ 2 & 1 & 1 & 1 \end{pmatrix}$ as above, and bring the 2×4 matrix A_1 and the 1×4 matrix A_2 to a systematic form, such that

$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 2 & 1 & 2 \end{pmatrix}$. Note that A_2 is still a good matrix. Now we take $G' = A \otimes I_B$ to

obtain

$$G' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & 0 & 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 & 0 & 1 & 0 & 2 \end{pmatrix},$$

with $T = tB = 3 \cdot 2$ and $k = T - \mu = 4$. Note that for any burst erasure of length 2 s_i with $i = 1, 2$ can be reconstruct with delay of at most $T = 6$ as in case (i) of the example.

Our next construction of DO-SBE codes with a systematic encoder is suitable for every admissible T and B.

First we define a matrix $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$ and show that G' satisfies the conditions in Lemma 5.8.

Lemma 5.14. *For $q > 2$, integers $B \geqslant 1$ and $T \geqslant 2B$, let $\begin{pmatrix} I_{T-2B} & A \end{pmatrix}$ be a $(T-2B) \times (T-B)$ c-good matrix. Then the matrix*

$$G' = \begin{pmatrix} I_B & \mathbf{0}_{B\times(T-2B)} & \mathbf{0}_B & I_B \\ I_B & \mathbf{0}_{B\times(T-2B)} & I_B & \alpha I_B \\ \mathbf{0}_{(T-2B)\times B} & I_{T-2B} & A & \mathbf{0}_{(T-2B)\times B} \end{pmatrix} \tag{5.4}$$

is a c-good matrix and its submatrix G consisting of the last $T-B$ rows is a good matrix, where $\alpha \in \mathbb{F}_q \backslash \{0, 1\}$.

Proof. Recall that a $k \times n$ c-good matrix exists for every $k, n \in \mathbb{N}$ over any finite field (see 4.12). Thus, we have a $(I_{T-2B}\ A)$ c-good matrix for any B and $T \geqslant 2B$. Note then that (by Lemma 5.4) submatrix G_1 of matrix G

$$G_1 = \begin{pmatrix} I_B & \mathbf{0}_{B\times(T-2B)} & I_B \\ \mathbf{0}_{(T-2B)\times B} & I_{T-2B} & A \end{pmatrix} \tag{5.5}$$

is c-good, since $\begin{pmatrix} I_{T-2B} & A \end{pmatrix}$ is c-good. The latter implies (by Lemma 5.4) that G is a good matrix.

Now observe that by elementary row operations, matrix G' can be brought to the following systematic matrix

$$\begin{pmatrix} I_B & \mathbf{0}_{B\times(T-B)} & I_B \\ \mathbf{0}_{(T-B)\times B} & I_{(T-B)} & A' \end{pmatrix}, \tag{5.6}$$

where $A' = \begin{pmatrix} -A \\ I_B \end{pmatrix}$. Note that $\begin{pmatrix} I_{(T-B)} & A' \end{pmatrix}$ is c-good. Hence, by Lemma 5.4, G' is

c-good. ∎

Theorem 5.15. *For positive integers T, B, μ, where $T \geqslant 2B$ and $\mu = T - B$, we have an explicit construction of a $\mu - [T + B, T]_q$ systematic DO-SBE code for any $q \geqslant 3$.*

Proof. Lemma 5.14 together with Lemma 5.8 implies that the $[T+B, T]_q$ $(q > 2)$ code C' with generator matrix G' in (5.4) corrects any B-burst erasures, including wrap around bursts. Moreover, code C' provides perfect security and achieves maximum secrecy rate $R_s = \frac{T-\mu}{T+B}$. Thus, it remains to be shown that the code can reconstruct arbitrary source symbol s_i, $i = 1, \ldots, B$, with delay of at most T. We refer to Example 5.16.

Let $E(y^{T+B}) \subset \{1, \ldots, T + B\}$ be the interval of bursty positions in y^{T+B}. We have to show that for each $E(y^{T+B})$ with $\left|E(y^{T+B})\right| = B$ we can reconstruct every source symbol s_i, $i \in \{1, \ldots, B\}$, with delay at most T. We refer only to the urgent symbols. Let $i \in \{1, \ldots, B\}$.

Case 1: Suppose that $E(y^{T+B}) \subset \{B + 1, \ldots, T\}$. Then s_i can be reconstructed using i-th and $(i + T)$-th unerased positions in y^{T+B} and corresponding columns in G'.

Case 2: $E(y^{T+B}) \subset \{1, \ldots, 2B - 1\}$ and $E(y^{T+B}) = \{i, \ldots, i + B - 1\}$. If $i = 1$, we can determine $s_{B+1}, \ldots, s_T$, using the corresponding known symbols $y_{B+1}, \ldots, y_T$ in y^{T+B}. Observe that $s_i = y_{i+T} - \alpha s_{B+i}$. If $i > 1$, we first determine $s_1, \ldots, s_{i-1}$ as in Case 1. Consequently, we get $s_{B+1}, \ldots, s_{B+i-1}$. Now the source symbols in the erased positions $E(y^{T+B})$ can be determined thanks to G_1 in (5.5).

Case 3: $E(y^{T+B}) \subset \{T - B + 2, \ldots, T + B\}$ and $E(y^{T+B}) = \{T - B + 1 + i, \ldots, T + i\}$. Using $y_{B+1}, \ldots, y_{T-B+i}$ we can determine the erased symbols $y_{T-B+1+i}, \ldots, y_T$. Hence, we can reconstruct $s_1, \ldots, s_B$ from $y_1, \ldots, y_T$ with delay at most T. ∎

Here we note that unlike low-delay codes (e.g. codes in [8]), in a secure code, optimal decoding of source symbol s_i (i.e. ith secret symbol) is not guaranteed even if the corresponding code symbol x_i is not erased.

Example 5.16. *As an example of a code construction in Theorem 5.15, consider the following matrix*

$$G' = \begin{pmatrix} G^* \\ G \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}. \tag{5.7}$$

In view of the theorem (and Lemma 5.14), the code with generator matrix G' is a $3 - [7, 5]_3$ systematic DO-SBE code.

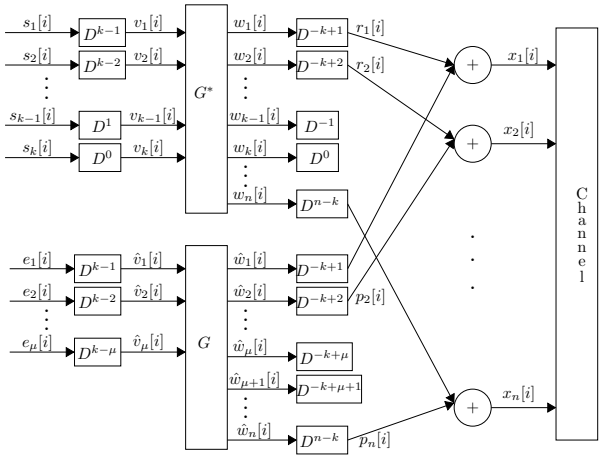


Figure 5.2.: A convolutional code structure based on diagonal interleaving. D^i denotes a delay of i packets.

$x[i] =$	$\underline{s_1[i] + e_1[i]}$	$s_2[i] + e_2[i]$	$e_3[i]$	$e_1[i-3] + e_3[i-1]$	$e_2[i-3] + e_3[i-2]$	$s_1[i-5] + 2e_1[i-5]$	$s_2[i-5] + 2e_2[i-5]$
$x[i+1] =$	$s_1[i+1] + e_1[i+1]$	$\underline{s_2[i+1] + e_2[i+1]}$	$e_3[i+1]$	$e_1[i-2] + e_3[i]$	$e_2[i-2] + e_3[i-1]$	$s_1[i-4] + 2e_1[i-4]$	$s_2[i-4] + 2e_2[i-4]$
$x[i+2] =$	$s_1[i+2] + e_1[i+2]$	$s_2[i+2] + e_2[i+2]$	$\underline{e_3[i+2]}$	$e_1[i-1] + e_3[i+1]$	$e_2[i-1] + e_3[i]$	$s_1[i-3] + 2e_1[i-3]$	$s_2[i-3] + 2e_2[i-3]$
$x[i+3] =$	$s_1[i+3] + e_1[i+3]$	$s_2[i+3] + e_2[i+3]$	$e_3[i+3]$	$\underline{e_1[i] + e_3[i+2]}$	$e_2[i] + e_3[i+1]$	$s_1[i-2] + 2e_1[i-2]$	$s_2[i-2] + 2e_2[i-2]$
$x[i+4] =$	$s_1[i+4] + e_1[i+4]$	$s_2[i+4] + e_2[i+4]$	$e_3[i+4]$	$e_1[i+1] + e_3[i+3]$	$\underline{e_2[i+1] + e_3[i+2]}$	$s_1[i-1] + 2e_1[i-1]$	$s_2[i-1] + 2e_2[i-1]$
$x[i+5] =$	$s_1[i+5] + e_1[i+5]$	$s_2[i+5] + e_2[i+5]$	$e_3[i+5]$	$e_1[i+2] + e_3[i+4]$	$e_2[i+2] + e_3[i+3]$	$\underline{s_1[i] + 2e_1[i]}$	$s_2[i] + 2e_2[i]$
$x[i+6] =$	$s_1[i+6] + e_1[i+6]$	$s_2[i+6] + e_2[i+6]$	$e_3[i+6]$	$e_1[i+3] + e_3[i+5]$	$e_2[i+3] + e_3[i+4]$	$s_1[i+1] + 2e_1[i+1]$	$\underline{s_2[i+1] + 2e_2[i+1]}$

Figure 5.3.: A secrecy rate-2/7 code constructed by diagonally interleaving the $3-[7,5]_3$ DO-SBE block code.

5.5. The Secure Streaming Codes

In the following, we consider codes for the model introduced in Section 5.2. We analyze the correction capability under delay constraint and the security condition of the convolutional code obtained by a proper diagonal interleaving applied to a systematic DO-SBE block code. The mapping from a nested block code (C, C') to a convolutional code is shown in Fig. 5.2. We note that the channel input packet at time i is $x[i] = r[i] + p[i]$, where $r[i] = f_i(s[0], s[1], \ldots, s[i])$ and $p[i] = h_i(e[0], e[1], \ldots, e[i])$. Furthermore, if $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$ is a generator matrix for C', then $r[i]$ is the packet obtained by diagonal interleaving applied to the block code generated by G^*. Correspondingly, $p[i]$ is the resulting packet obtained by diagonal interleaving applied to the coarse code C.

Fig. 5.3 shows the convolutional code obtained by diagonal interleaving applied to the $3-[7,5]_3$ DO-SBE code in Example 5.16. The ith line in the semi-infinite array represents channel packet $x[i]$. The codewords of the block code appear along the diagonals, as illustrated by the underlined symbols.

5.5.1. The Achievability

Our systematic DO-SBE block codes differ from the systematic block codes in [35], [10]. In contrast to the design of codes that serve only for reliable transmission, the source symbols of $s[i]$ are not immediately obtained from the channel symbols of $x[i]$. In our construction of a convolutional code, the channel packet is produced causally from the source stream and the randomly chosen stream.

In the following, we analyze the correction capability under the delay constraint and the security conditions of the convolutional code obtained by diagonally interleaving the systematic DO-SBE block code.

First, we give two definitions of convolutional codes.

Definition 5.4. *An $(n, k, \mu, \varpi, T)_q$ convolutional code with encoder memory ϖ and decoding delay T is an $(n, k, \mu, T)_q$ streaming code, constructed as follows: Let $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$. For any $i \geqslant 0$, we obtain the packet*

$$x[i] = \sum_{l=0}^{\varpi} (s\,[i-l]\,G_l^{*conv} + e\,[i-l]\,G_l^{conv}), \tag{5.8}$$

where G_l^{*conv} is a $k \times n$ matrix so that

$$G^* = \sum_{l=0}^{\varpi} G_l^{*conv} \tag{5.9}$$

and G_l^{conv} is a $\mu \times n$ matrix so that

$$G = \sum_{l=0}^{\varpi} G_l^{conv}. \tag{5.10}$$

By convention we choose $s\,[-1], \ldots, s\,[-\varpi] = \mathbf{0}_{1\times k}$ and $e\,[-1], \ldots, e\,[-\varpi]$, which correspond to i.i.d. sequences of random variables which are uniformly distributed over $\mathbb{F}_q^{\mu}$.

Definition 5.5. *The mapping from source sequence to code sequence can be defined by a multiplication with the generator matrix $G'^{conv} = \begin{pmatrix} G^{*conv} \\ G^{conv} \end{pmatrix}$ of the $(n, k, \mu, \varpi, T)_q$ convolutional code:*

$$G^{*conv} = \begin{pmatrix} G_0^{*conv} & G_1^{*conv} & \cdots & G_{\varpi}^{*conv} \\ \mathbf{0}_{k\times n} & G_0^{*conv} & \cdots & G_{\varpi-1}^{*conv} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{k\times n} & \mathbf{0}_{k\times n} & \cdots & G_0^{*conv} \end{pmatrix} \tag{5.11}$$

$$G^{conv} = \begin{pmatrix} G_0^{conv} & G_1^{conv} & \cdots & G_{\varpi}^{conv} \\ \mathbf{0}_{k\times n} & G_0^{conv} & \cdots & G_{\varpi-1}^{conv} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{k\times n} & \mathbf{0}_{k\times n} & \cdots & G_0^{conv}, \end{pmatrix} \tag{5.12}$$

where G_l^{*conv} and G_l^{conv} are respectively $k \times n$ and $\mu \times n$ matrices, $l \in [0, \varpi]$. Note that (5.11) and (5.12) are truncated matrices.

We review the standard argument of interleaving a blockcode into a convolutional code [47],[7], where we first address only the correctability of the convolutional code with delay constraint. A similar lemma was shown by Fong et al. [10].

Lemma 5.17. *Given an $\mu - [T+B, T]_q$ DO-SBE block code, we can construct an $(n, k, \mu, T, T)_q$ convolutional code that is able to reconstruct any B-burst erasure and recover source symbols with delay T. More specifically, let $G' = \begin{pmatrix} G^* \\ G \end{pmatrix}$ be the generator matrix of the $\mu - [T+B, T]_q$ DO-SBE block code. Let $g_{*i,j}$ with $0 \leqslant i \leqslant k-1$, $0 \leqslant j \leqslant n-1$ and $g_{i,j}$ with $0 \leqslant i \leqslant \mu - 1$, $0 \leqslant j \leqslant n-1$ be the entries situated in row i and column j of generator matrices G^* and G, respectively. Then we can construct the $n-1$ generator matrices of the $(n, k, \mu, n-1, T)_q$ convolutional code as follows*

$$G_l^{*conv} = \begin{cases} \left(\mathbf{0}_{k\times l} \;\middle|\; diag(g_{0,l}, \cdots, g_{k-1,l+k-1}) \;\middle|\; \mathbf{0}_{k\times(n-k-l)} \right) & \text{if } 0 \leqslant l \leqslant n-k \\ \left(\mathbf{0}_{k\times l} \;\middle|\; \begin{matrix} diag(g_{0,l}, g_{1,l+1} \cdots, g_{n-1-l,n-1}) \\ \mathbf{0}_{(k-n+l)\times(n-l)} \end{matrix} \right) & \text{if } n-k < l \leqslant n-1 \end{cases} \tag{5.13}$$

$$G_l^{conv} = \begin{cases} \left(\mathbf{0}_{\mu\times l} \;\middle|\; diag(g_{0,l}, \cdots, g_{\mu-1,l+\mu-1}) \;\middle|\; \mathbf{0}_{\mu\times(n-\mu-l)} \right) & \text{if } 0 \leqslant l \leqslant n-\mu \\ \left(\mathbf{0}_{\mu\times l} \;\middle|\; \begin{matrix} diag(g_{0,l}, g_{1,l+1} \cdots, g_{n-1-l,n-1}) \\ \mathbf{0}_{(\mu-n+l)\times(n-l)} \end{matrix} \right) & \text{if } n-\mu < l \leqslant n-1. \end{cases} \tag{5.14}$$

Proof. See Appendix 5.7. ∎

Intuitively, the codewords of the block code appear along the diagonals of the convolutional code. In other words, the parity check symbols are computed along diagonals of the convolutional code. Thus if a B-burst erasure occurs, each affected codeword has an erasure burst of length B or less in the diagonals. Since the codewords are elements of a delay-optimal B-burst-erasure correcting code, the source symbols can be reconstructed with delay of at most T, which is optimal.

Next we analyze the security constraint. In Theorem 5.18 we state that for $W \geqslant T+1$ in any time interval of length $i \geqslant 0, 1, \ldots$, the eavesdropper has full equivocation. Note

that the size W of the sliding window is chosen such that the intervals of at most μ packets, observed by the eavesdropper, are separated by T or more erased packets.

Theorem 5.18. *Consider a burst erasure wiretap channel where the eavesdropper can observe an interval of at most μ packets in a sliding window of size $T+1$. The $(n, k, \mu, T, T)_q$ convolutional code, obtained by diagonally interleaving a $\mu - [T+B, T]_q$ systematic DO-SBE block code, attains perfect security.*

Proof. Recall that each diagonal in the resulting streaming code is a codeword of the $\mu-[T+B, T]_q$ DO-SBE block code. Thus, the diagonals are independent random variables taking values from $\mathbb{F}_q^{T+B}$. Suppose the eavesdropper observes packets $x[i], \ldots, x[i+\mu-1]$, which in fact is a $\mu \times (T+B)$ matrix (as illustrated in Fig. 5.3), and consider all diagonals containing entries of this matrix. Let S^k and Z^{T+B} (with some abuse of notation) respectively, be the random variables corresponding to the source symbols and the eavesdropper's channel output of the block code which appears along one of the diagonals. By construction of the block code, we have that $H(S^k|Z^{T+B}) = k$, and this holds for every diagonal. It follows (from the structure of corresponding block codes) that the μ packets observed by the eavesdropper do not reveal any information about the source symbols in those packets, as well as in any other interval of μ previously observed packets, since they are separated by an interval of length at least T. Since $k = T - \mu$, we get the result. ∎

We summarize our observations in the following theorem.

Theorem 5.19. *We obtain by diagonally interleaving a $\mu - [T+B, T]_q$ systematic DO-SBE block code a $(T, B, \mu; T+1)_q$ streaming code with $k = T - \mu$, $n = T + B$, which achieves the secrecy rate $R_s = k/n = \frac{T-\mu}{T+B}$.*

5.5.2. An Upper Bound for the Secrecy Rate

We consider a periodic erasure channel in the presence of an eavesdropper who can noiselessly observe an interval of μ packets in any sliding window of size $T+1$ (see Fig. 5.4(a)). Every two successive B-bursts (respectively μ-intervals observed by an eavesdropper), starting from the first interval of length $T+B$, are separated by T packets. In fact, the packet length of the period is $\ell = lcm(T+B, T+\mu)$.

Let L_j be the index set for the packets in the jth period, $j = 1, 2, \ldots$. Let $M_j \subset L_j$ and $E_j \subset L_j$ be the index sets of the packets revealed respectively to the legitimate receiver and to the eavesdropper in the jth period. Correspondingly, $\mathcal{Y}_{M_j}$ and $\mathcal{Y}_{E_j}$ are the observations at the receiver and the eavesdropper (see Fig. 5.4(a)). Furthermore, $\mathcal{S}_j$ is a random variable representing the sequence of messages produced by the source in the jth period, thus $\mathcal{S}_j \in \mathbb{F}_q^{k\ell}$. For an integer $h \geqslant 1$, we use $\mathcal{Y}_{M_1}^h$ to denote $\mathcal{Y}_{M_1}, \ldots, \mathcal{Y}_{M_h}$ and $\mathcal{S}_1^h$ for $\mathcal{S}_1, \ldots, \mathcal{S}_h$. Let $\mathcal{U}_{M_{h+1}^*}$ and $\mathcal{U}_{E_{h+1}^*}$ respectively be observations of the receiver and

the eavesdropper in the interval of T successive outgoing packets after h periods in the stream.

Clearly, to provide optimal delay and perfect security, the following must hold $\forall h \geqslant 1$:

$$H(\mathcal{S}_1^h|\mathcal{Y}_{M_1}^h, U_{M_{h+1}^*}) = 0, \tag{5.15}$$

$$H(\mathcal{S}_1^h|\mathcal{Y}_{E_1}^h, \mathcal{U}_{E_{h+1}^*}) = H(\mathcal{S}_1^h). \tag{5.16}$$

Note that $H(\mathcal{S}_1^h) = k\ell h$, using the fact that all source packets have the same entropy. Denote $\mathcal{W}_{M(h)} = \mathcal{Y}_{M_1}^h, \mathcal{U}_{M_{h+1}^*}$ and $\mathcal{W}_{E(h)} = \mathcal{Y}_{E_1}^h, \mathcal{U}_{E_{h+1}^*}$, where $M(h) := \cup_{i=1}^h M_i \cup M_{h+1}^*$, $E(h) := \cup_{i=1}^h E_i \cup E_{h+1}^*$. Then (5.15) and (5.16) imply that

$$\begin{aligned} h\ell k &= H(\mathcal{S}_1^h|\mathcal{W}_{E(h)}) \leqslant H(\mathcal{S}_1^h, \mathcal{W}_{M(h)}|\mathcal{W}_{E(h)}) \\ &= H(\mathcal{W}_{M(h)\backslash E(h)}|\mathcal{W}_{E(h)}) + H(S_1^h|\mathcal{W}_{M(h)}, \mathcal{W}_{E(h)}) \\ &\leqslant H(\mathcal{W}_{M(h)\backslash E(h)}) = H(\mathcal{Y}_{M_1\backslash E_1}^h, \mathcal{U}_{M_{h+1}^*\backslash E_{h+1}^*}) \\ &\leqslant hH(\mathcal{Y}_{M_1\backslash E_1}) + n(T-\mu). \end{aligned}$$

Denote $a = |M_1\backslash E_1|$ (recall that $|M_1\backslash E_1| = |M_2\backslash E_2| = \dots$). Then, $H(\mathcal{Y}_{M_i\backslash E_i}) \leqslant na$. The latter implies that

$$R_s = \frac{k}{n} \leqslant \frac{ha + T - \mu}{\ell h} \overset{h\to\infty}{\to} \frac{a}{\ell}. \tag{5.17}$$

In the case $B = \mu$ and thus $\ell = T + B$, we get $H(\mathcal{Y}_{M_1\backslash E_1}) \leqslant n(T-\mu)$, and hence $R_s \leqslant \frac{T-\mu}{T+B}$, which matches the secrecy rate in Theorem 5.19.

Next we present a general upper bound. As an example see Fig. 5.4(b).

For positive integers T, B, μ with $B + \mu \leqslant T$, let $d = gcd(T + B, T + \mu)$. Thus, we have $B \equiv \mu \mod d$, that is $B = t_1 d + r, \mu = t_2 d + r$, for suitable nonnegative integers t_1, t_2, r, with $r < d$.

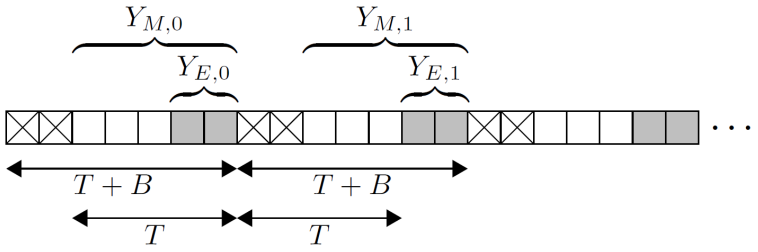
Theorem 5.20. *The secrecy rate R_s of a delay-optimal burst-erasure correcting $(T, B, \mu; T + 1)$ streaming code is upper bounded by*

$$R_s \leqslant \frac{T^2 - \rho(d)^2}{(T + B)(T + \mu)}, \tag{5.18}$$

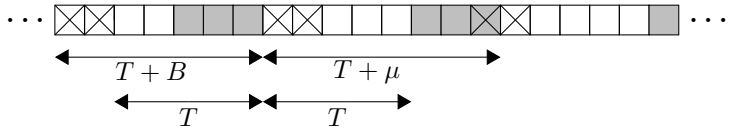
where $\rho(d) = r$, if $d \geqslant 2r$ and $\rho(d) = 2r - d$ otherwise.

Proof. Denote $J(s) = \{s, s + 1, \dots, s + B - 1\}, E(m) = \{m, m + 1, \dots, m + \mu - 1\}$, where $0 \leqslant s \leqslant T$ and $0 \leqslant m \leqslant T + B - \mu$. Let $L := L_1 = \{0, 1, \dots, \ell - 1\}$, $\mathcal{B}(s) = \{i \in L : i \mod (T + B) \in J(s)\}$, and $\mathcal{E}(m) = \{j \in L : j \mod (T + \mu) \in E(m)\}$.

Thus, $\mathcal{B}(s)$, respectively, $\mathcal{E}(m)$, is the set of erased packets, respectively, the set of packets observed by the eavesdropper, in a time slot corresponding to one period, that is



(a) The periodic erasure channel with periodic length $P = T + B$ and $B = \mu$.



(b) The periodic erasure channel with periodic length $P = lcm(T + \mu, T + B)$.

Figure 5.4.: The periodic erasure channel used in proving the upper bound, with indication of which symbols are observed by the eavesdropper $Y_{E,j}$ (gray squares) and by the legitimate receiver $Y_{M,j}$ (white squares). Crossed squares are erasures of length B.

within transmission of ℓ packets over the channel. Clearly
$|L| = \ell = \dfrac{(T+B)(T+\mu)}{d}$, $|\mathcal{B}(s)| = \dfrac{(T+\mu)B}{d}$, and $|\mathcal{E}(m)| = \dfrac{(T+B)\mu}{d}$.

Let $\widehat{a} := \min_{s,m} |L \backslash (\mathcal{B}(s) \cup \mathcal{E}(m))|$. Then, in view of (8), the secrecy rate $R_s \leqslant \dfrac{\widehat{a}}{\ell}$. Furthermore, observe that

$$\widehat{a} = |L| - |\mathcal{B}(s)| - |\mathcal{E}(m)| + \min_{s,m} |\mathcal{B}(s) \cap \mathcal{E}(m)|.$$

Hence we get $\widehat{a} = \dfrac{T^2 - B\mu}{d} + \min_{s,m} |\mathcal{B}(s) \cap \mathcal{E}(m)|$.

Our goal now is to determine $f(T, B, \mu, s, m) := \min_{s,m} |\mathcal{B}(s) \cap \mathcal{E}(m)|$ (later we use short notation f instead), given admissible parameters T, B, μ.

Note that $\mathcal{B}(s) \cap \mathcal{E}(m) = \{x \in L : x \equiv i \mod (T+B), x \equiv j \mod (T+\mu),\ (i,j) \in J(s) \times E(m)\}$.

Also it is not hard to see that

$$f = \min_{m} |\mathcal{B}(0) \cap \mathcal{E}(m)| = \min_{s} |\mathcal{B}(s) \cap \mathcal{E}(0)|. \tag{5.19}$$

The well-known Chinese Remainder Theorem, extended for non coprime moduli, tells us that a system of congruences $x \equiv i \mod (T+B)$, $x \equiv j \mod (T+\mu)$ has a solution iff $i \equiv j \mod d$, and such a solution is unique modulo ℓ. Hence,

$$\mathcal{B}(s) \cap \mathcal{E}(m) = \{(i,j) \in J(s) \times E(m) : i \equiv j \bmod d\}. \tag{5.20}$$

Obviously we have the following partitions into d classes:

$$J(0) = \bigcup_{s=0}^{d-1} J_s(0), \quad E(m) = \bigcup_{k=m}^{m+d-1} E_k(m),$$

where $J_s(0) = \{i \in J(0) : i \equiv s \mod d\}$ and $E_k(m) = \{j \in E(m) : j \equiv k \mod d\}$.

Note then that in turn (5.20) implies that

$$|\mathcal{B}(0) \cap \mathcal{E}(m)| = \sum_{s=0}^{d-1} |J_s(0)||E_{\sigma(s)}(m)|, \tag{5.21}$$

where $\sigma(s) \in \{m, \ldots, m+d-1\}$ with $\sigma(s) \equiv s \mod d$.

Next we claim that

$$f = |\mathcal{B}(0) \cap \mathcal{E}(r)|. \tag{5.22}$$

To show this, recall that
$B = r(t_1+1) + (d-r)t_1$ and $\mu = r(t_2+1) + (d-r)t_2$.

Observe then that the latter implies that for $1 \leqslant r \leqslant d-1$ we have
$|J_0(0)| = |J_1(0)| = \ldots = |J_{r-1}(0)| = t_1 + 1$,
$|J_r(0)| = |J_{r+1}(0)| = \ldots = |J_{d-1}(0)| = t_1$, and for any m
$|E_m(m)| = |E_{m+1}(m)| = \ldots = |E_{m+r-1}(m)| = t_2 + 1$,
$|E_{m+r}(m)| = |E_{m+r+1}(m)| = \ldots = |E_{m+d-1}(m) = t_2$; in particular this holds for $m = r$ (that is for the case we need). In the case where $r = 0$ we have $|J_i(0)| = t_1$ and $|E_i(m)| = t_2$ for $i = 0, 1, \ldots, d-1$ (and for any m).

Now (5.22) follows in view of (5.21) and the fact that for real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ the minimum of quantity $\sum_{i=1}^n a_{r_i} b_{s_i}$, taken over all rearrangments of these sequences, is attained when $a_{r_1} \geqslant \ldots \geqslant a_{r_n}$ and $b_{s_1} \leqslant \ldots \leqslant b_{s_n}$.

Furthermore, (5.22) and (5.21) imply that in the case $d \geqslant 2r$ we have the following:
$f = r(t_1+1)t_2 + (d-2r)t_1t_2 + r(t_2+1)t_1 = r(t_1+t_2) + dt_1t_2$
$=$ (via simple calculations) $\dfrac{B\mu - r^2}{d}$.

Similarly, in the case $d < 2r$ we have
$f = (d-r)(t_1+1)t_2 + (2r-d)(t_1+1)(t_2+1) + (d-r)(t_2+1)t_1 = r(t_1+t_2+2) + d(t_1t_2-1) = \dfrac{B\mu - (d-r)^2}{d}$, which implies that for integers $T, B, \mu \geqslant 1$, with $B + \mu \leqslant T$, we have

$$f = \frac{B\mu - \rho(d)^2}{d}. \tag{5.23}$$

Note that for $B = \mu$, we have $f = 0$, and for $d = 1$ the equality simplifies to $f = B\mu$.

Thus, we have

$$\widehat{a} = \frac{T^2 - B\mu}{d} + \frac{B\mu - \rho(d)^2}{d} = \frac{T^2 - \rho(d)^2}{d}$$

and hence

$$R_s \leqslant \frac{\widehat{a}}{\ell} = \frac{T^2 - \rho(d)^2}{(T+B)(T+\mu)}.$$

Remark 4. It is easy to see that in the case $\mu \leqslant d/2$ we have $\rho(d) = r = \mu$, which implies that $R_s \leqslant \dfrac{T-\mu}{T+B}$. Note that the latter holds in a special case when $\mu = B$. ∎

5.6. Conclusion and Discussion

We have provided two constructions of delay-optimal B-burst-erasure correcting streaming codes (or DO-SBE code for short) for a burst-erasure wiretap channel, where the eavesdropper observes an interval of at most μ packets in a sliding window of size $T+1$. The first construction is suitable for parameters $T = tB$ and $\mu = iB$, where $t \in \mathbb{N}\setminus\{0,1\}$ and $i \in \{1, \dots t-1\}$, and the second is suitable for $\mu = T - B$ and any $T \geqslant 2B$. While our DO-SBE block codes that we require for the construction of DO-SBE convolutional codes achieve the maximum secrecy rate R_s, our DO-SBE convolutional codes achieve the maximum secrecy rate for a special case, that is if $B = \mu$.

Clearly, if we vary the size W of the sliding window of the eavesdropper the maximum secrecy rate also changes. For example, in the case $\mu \leqslant B$, if we put $W = T + B - \mu + 1$, then R_s achieves the upper bound $\frac{T-\mu}{T+B}$ since in this case in each period of size $T+B$ the set of erased packets and observations of the eavesdropper do not overlap. Similarly, we require in case $\mu > B$ to choose $W = 2T + 2B - \mu + 1$ to obtain no intersections of the set of erased packets and observations of the eavesdropper in each period of size $2(T+B)$. The question arises whether better secrecy rates can be achieved by converting DO-SBE block codes into DO-SBE streaming codes for fix W. For future work it would be interesting to construct DO-SBE streaming codes with a secrecy rate that matches the upper bound for any $B, \mu \leqslant T$. Moreover it would be interesting to consider the same problem for a multi-link scenario, or for a model where the channel injects isolated erasures.

5.7. Appendix

5.7.1. Proof of Lemma 5.4

Since $\begin{pmatrix} I_{T-B} & A \end{pmatrix}$ is c-good, $\begin{pmatrix} -A^\top & I_B \end{pmatrix}$ is also c-good. We can observe that the parity check matrix $H = \begin{pmatrix} -A^\top & I_B & I_B \end{pmatrix}$ of the code with generator matrix G is c-good and thus G as well.

5.7.2. Proof of Lemma 5.17

Suppose we are given a $\mu-[T+B,T]_q$ systematic DO-SBE block code with the generator matrix $G'=\begin{pmatrix}G^*\\G\end{pmatrix}$. We can construct an $(n,k,\mu,n-1,T)_q$ convolutional code as follows:

For each $i\in\mathbb{Z}^+$ we can construct

$$\begin{aligned}&x_1[i],x_2[i+1],\ldots,x_n[i+n-1]\\&=(s_1[i],s_2[i+1],\ldots,s_k[i+k-1])G^*+(e_1[i],e_2[i+1],\ldots,e_\mu[i+\mu-1])G\end{aligned}\tag{5.24}$$

for each i. Here we are coding the source symbols diagonally, as illustrated in Fig. 5.3.

From (5.24) to the encoded packet $x[i]=x_1[i],x_2[i],\ldots,x_n[i]$ at time i, we get as follows:

$$x[i]=\sum_{l=0}^{n-1}(s_1[i-l],s_2[i+1-l],\ldots,s_k[i+k-1-l])\left(\begin{array}{c|c|c}\mathbf{0}_{k\times l}&\begin{matrix}g^*_{1,l+1}\\\vdots\\g^*_{k,l+1}\end{matrix}&\mathbf{0}_{k\times n-l-1}\end{array}\right)\tag{5.25}$$

$$+\sum_{l=0}^{n-1}(e_1[i-l],e_2[i+1-l],\ldots,e_\mu[i+\mu-1-l])\left(\begin{array}{c|c|c}\mathbf{0}_{\mu\times l}&\begin{matrix}g_{1,l+1}\\\vdots\\g_{\mu,l+1}\end{matrix}&\mathbf{0}_{\mu\times n-l-1}\end{array}\right)$$

$$=\sum_{l=0}^{n-1}s[i-l]\,G_l^{*conv}+\sum_{l=0}^{n-1}e[i-l]\,G_l^{conv},\tag{5.26}$$

where

$$G^*=(I_k\ P)\text{ with }G^*=\sum_{l=0}^{n-1}G_l^{*conv}\tag{5.27}$$

and

$$G=(I_\mu\ P)\text{ with }G=\sum_{l=0}^{n-1}G_l^{conv}.\tag{5.28}$$

Since the block code is causal, symbols from future packages can be considered as zero symbols.

Now, we want to show that the $(n,k,\mu,n-1,T)_q$ convolutional code whose encoding function at time i is specified by (5.26), is able to decode packet $s[i]$ with delay of at most T. For any burst erasure of length B, the $\mu-[T+B,T]_q$ systematic DO-SBE block code is able to reconstruct symbol s_t, $t=1,\ldots,k$ with delay T. According to (5.24), the source symbols are coded along diagonals as illustrated with underlined symbols in Fig. 6.2. This implies that the destination can reconstruct packet $s[i]$ up to time $i+T$ based on $(y[0],y[1],\ldots,y[i+T])$, since each source symbol in B-erased consecutive packets can be reconstructed separately, using the corresponding diagonals. We can see

that due to the form of the generator matrix G' it follows from (5.13) and (5.14) that $G_l^{*conv} = 0^{k\times n}, G_l^{conv} = 0^{\mu\times n}$ for any $l \geqslant T+1$, thus $\varpi = T$. We obtain an $(n, k, \mu, T, T)_q$ convolutional code.

6. Delay-Optimal Codes for Parallel Burst-Erasure Channels with an Eavesdropper

6.1. Introduction

For streaming applications, we consider parallel burst-erasure channels in the presence of an eavesdropper. Peer-to-peer networks are subjected to different performance constraints such as high throughput, low latency and high reliability. However, during transmission different types of errors can occur, such as clustered and bursty packet losses, which lead to low-quality video and high delay [48], [49]. The requirements on time-critical communication systems are challenging, particularly when private or sensitive data must be transmitted, which needs to be protected against eavesdropping attacks and active attacks (e.g. payment transmission in a smart shop or machine-to-machine communication in a smart factory).

Contribution

We consider block codes and streaming codes. We introduce a new channel model for the transmission of block codes, which we refer to as the block channel model. The block channel model consists of a sender, a legitimate receiver, M parallel channels and an eavesdropper. Z links can experience a burst of erasures of length B, while the remaining links are noiseless. The eavesdropper is able to observe a noiseless copy of any link of his choice. He is able to switch between the links at any time. His restriction may be due to the fact that he has access only to certain frequencies in a wireless system or to individual nodes in a distributed storage system, e.g. because of his location, or on purpose to remain undetected. For $T \geqslant B$ and $Z = M - 1$, we give explicit constructions of M-link codes over a small finite field $\mathbb{F}_q$ that provide perfect security (i.e. the eavesdropper can obtain no information about the message) and provide zero-error decoding with minimum possible delay. More precisely, we distinguish two cases in the construction: 1) M is odd and the code is binary and 2) M is even, where $q > 2$.

Our block codes can be mapped to M-link convolutional codes for the streaming channel model, where we assume that in each of the Z links a burst erasure of at most B packets

can occur in any sliding window of a fixed size, and where the eavesdropper can observe one link of his choice noiselessly. The M-link convolutional codes provide perfect security and provide zero-error decoding with minimum possible delay. Our codes achieve the maximum secrecy rate for the channel models.

Moreover, in Section 6.5 we consider two similar models as in Section 6.2, with the difference that the erasures in the main channel are caused by an active eavesdropper and not by the noise. Furthermore, in the first model the eavesdropper is able to cause a burst erasure of a fixed length in any single link. In the second model he can erase the complete link. In addition, in the first model the eavesdropper is able to observe any interval of length at most μ in any link and in the second model he is able to observe any complete link. For some admissible code parameters, again we first construct delay-optimal block codes that provide perfect security, and then covert them to delay-optimal convolutional codes that again provide perfect security in the respective setting. For both models we construct codes that achieve the maximum secrecy rate for the channel models.

The mapping of a delay-optimal block code for a single link scenario (i.e. for a burst-erasure wiretap channel) is shown in Section 5.5 and can be extended in a straightforward manner to a multi-link scenario. For a single link without an eavesdropper, the authors in [10] describe the mapping of delay-optimal erasure block codes to convolutional delay-optimal erasure codes in detail. As in the previous works, we use causal codes. This enables us to recover source symbols with a minimum possible delay (see [36]).

Related Work

In [35], Martinian considered an adversarial multi-link model where bursts of erasures are injected in a single link. In [44], delay-optimal burst-erasure codes for parallel links were designed for two types of errors - erasure burst and link outage. Additional works devoted to low-delay coding can be found in [41], [42], [43]. We refer to Section 5.1 where more works related to low-delay communication systems can be found.

Outline

In Section 6.2, we describe the channel model in the streaming setup. In Section 6.3, we describe the block code channel model and construct codes for that channel. The section is divided into the achievability part and the converse part. In the achievability part, we present explicit constructions of M-link codes for $T \geqslant B$ and $Z = M - 1$ over small finite fields for the channel model. The converse part is proved by using the entropy argument. In Section 6.4, we discuss the mapping from M-link block codes to M-link streaming codes for the channel model given in Section 6.2. Section 6.4 is also divided into the achievability and converse part. In Section 6.5, we consider two similar but simplified models as in Section 6.2.

6.2. The Channel Model

We consider an M-link channel consisting of a source and a sink to convey secure messages to the legitimate receiver in parallel. At time $i \geqslant 0$, the randomized encoder observes a source packet $s[i]$, and transmits each channel packet $x[i,j]$ on the corresponding link, where $j = 1, \ldots, M$. The source packet consists of k symbols, while each channel packet $x[i,j]$ assigned to link j consists of n symbols over a common finite field $\mathbb{F}_q$. In the streaming setup, we assume that in each of Z links a burst of erasures of length no longer than B packets can occur in any sliding window of size W[1], and where the eavesdropper can observe one link of his choice noiselessly. The sliding window model is considered in many previous works (e.g. [42], [10], etc.).

An $(n, k, \mu, T)_q$ M-link streaming code for an M-link channel model with an eavesdropper as given above consists of a set of M encoding functions $\{\xi_{i,j}\}_{j=1}^{M}$ and a decoding function ϕ_i, where i is a time unit.

Encoding: The random encoding function $\xi_{i,j} : \mathbb{F}_q^{k\cdot(i+1)} \times \mathbb{F}_q^{\mu\cdot(i+1)} \rightarrow \mathbb{F}_q^n$ takes in a source packet sequence $\{s[i]\}_{i\geqslant 0}$ together with an encoder packet sequence $\{e[i]\}_{i\geqslant 0}$ and maps them causally[2] into a packet $x[i,j]$, consisting of n symbols over the same finite field $\mathbb{F}_q$. In other words, $x[i,j] = \xi_{i,j}(s[0], s[1] \ldots, s[i], e[0], e[1] \ldots, e[i]) = f^*_{i,j}(s[0], s[1] \ldots, s[i]) + f'_{i,j}(e[0], e[1] \ldots, e[i])$. Each source packet and encoder packet consists of k symbols $s[i] = (s_1[i], s_2[i], \ldots, s_k[i]) \in \mathbb{F}_q^k$ and μ symbols $e[i] = (e_1[i], e_2[i], \ldots, e_\mu[i]) \in \mathbb{F}_q^\mu$, respectively.

Decoding: At the legitimate receiver, the decoding function $\phi_{i+T} : (\mathbb{F}_q^n \cup \{?\})^{M\cdot(i+1+T)} \rightarrow \mathbb{F}_q^k$ is defined as a packet decoder operating with delay T, that is, $s[i] = \phi_{i+T}(\{y[0,j], ..., y[i+T-1,j], y[i+T,j]\}_{j=1}^{M})$. The secrecy rate of the code is defined as: $R_s = k/n$ symbols per time unit. Note that in the initialization phase, i.e. for $i < 0$, where Eve is not able to observe the whole link noiselessly (i.e. $\mu < T + B$) a constant number of random packets $e[i]$ must be securely transmitted to ensure perfect reliability in the first T transmitted packets. Since in the initialization phase the number of pre-transmitted packets is constant, the resulting rate loss quickly converges to 0 as the number of packet transmissions grows.

Definition 6.1. *We denote an $(n, k, \mu, T)_q$ M-link streaming code as a $(T, B, \mu, Z; W)_q$ M-link streaming code if the code can recover source packets with delay T, even if in each of the Z links any burst of at most B erasures has occurred in a sliding window of size $W = T + 1$, and if the code provides perfect security, even if the eavesdropper is able to observe a complete link, that is, $\mu = n$.*

[1] In an M-link channel the windows can be seen as an $M \times W$ matrix with packets as its entries. If any of the Z bursts affect source symbols that are coded at time i in the corresponding links, the decoding deadline for all source symbols that are coded at time i is $i + T$, i.e., $W_i = \{i, i+1, \ldots, i+T\}$.

[2] The code is causal if in the encoding function the current channel packet is a function of the current and previous source/encoder symbols.

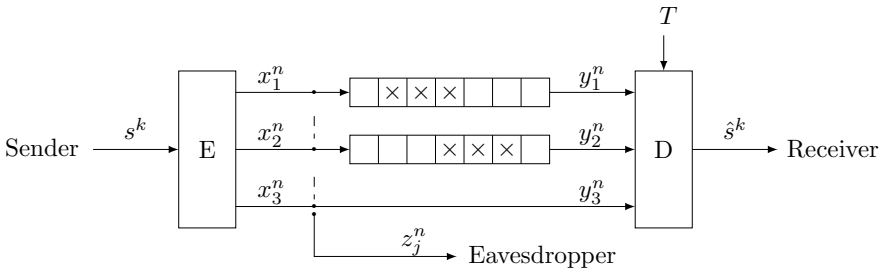


Figure 6.1.: The channel model described in Section 6.3 with $M = 3$, $Z = 2$, $T = 4$, $B = 3$ and $n = \mu = 7$. Crossed squares illustrate erasures. The eavesdropper observes any link $j \in \{1, \ldots, M\}$.

6.3. The Secure B-Burst-Erasure correcting M-Link Block Codes

We define M-link block codes over $\mathbb{F}_q$ for small q that we can apply to construct convolutional codes for the model introduced in Section 6.2.

We consider channel symbol blocks of length n for each link j, $j = 1, \ldots, M$. $Z = M - r$ links experience erasures while r links remain noiseless. We assume that the erasures occur as bursts of length B. Let $\mathscr{Z} \subset \{1, \ldots, M\}$ be the set of links where the erasure bursts have occurred, and $\mathscr{B}_j \subseteq \{1, \ldots, n\}$ the set of B consecutive positions where the erasures have occurred on link $j \in \mathscr{Z}$. Then the erasure pattern $\tilde{\mathscr{B}} = \{(j, \mathscr{B}_j)\}_{j \in \mathscr{Z}}$, where $|\mathscr{Z}| = Z$ and $|\tilde{\mathscr{B}}| = ZB$. The transmitted symbols $x[j, i] \in \mathbb{F}_q$, generated at time i, can either be erased or passed to the receiver noiselessly, so that the receiver observes either $y[j, i] =?$ or $y[j, i] = x[j, i]$, where $i = 1, \ldots, n$. Moreover, the eavesdropper is able to observe μ out of Mn symbols noiselessly. We consider an eavesdropper who observes a noiseless copy of any link of his choice, i.e. $\mu = n$. He can switch from $x[j, i]$ to $x[j', i + 1]$, where $j, j' \in \{1, \ldots, M\}$. Fig. 6.1 illustrates the channel model with three links.

It is assumed that a uniform source produces k symbols over the finite field $\mathbb{F}_q$. The code operates as follows.

Encoding: Consider a set of M random encoding functions $\{E_j\}_{j=1}^M$. The random encoding function $E_j : \mathbb{F}_q^{k+\mu} \to \mathbb{F}_q^n$ takes in a source vector $s^k \in \mathbb{F}_q^k$ and maps it together with a random encoder vector $e^\mu \in \mathbb{F}_q^\mu$ into $x_j^n \in \mathbb{F}_q^n$, where $k + \mu =: m$ and $j = 1, \ldots, M$. The secrecy rate of the code is defined as: $R_s := k/n$.

Let $s[i] = (s[1, i], \ldots, s[F_i, i])$ and $e[i] = (e[1, i], \ldots, e[L_i, i])$ be the sub-vectors of s^k and e^μ respectively, where F_i and L_i are the numbers of source symbols and encoder symbols respectively, injected at time i, $i = 1, \ldots, n$. At time i, the random encoding function $E_{j,i}$ generates the output $x[j, i] = E_{j,i}(s[i], e[i])$, where $E_j = \{E_{j,i}\}_{i=1}^n$, $k = \sum_{i=1}^n F_i$ and $\mu = \sum_{i=1}^n L_i$.

In the case where the code is causal, $x[j, i] = E_{j,i}(s[1], \ldots, s[i], e[1], \ldots, e[i])$, with $j = 1, \ldots, M$.

Decoding: At the legitimate receiver, the decoding function $D : (\mathbb{F}_q \cup \{?\})^{n\times M} \to \mathbb{F}_q^k$ maps the M channel outputs $y_j^n = (y[j,1], y[j,2], \ldots, y[j,n])$, $j = 1, \ldots, M$ into the reconstructed source symbol vector $\hat{s}^k \in \mathbb{F}_q^k$.

We define a decoder operating with delay T as $D_{i+T} : (\mathbb{F}_q \cup \{?\})^{(i+T)\times M} \to \mathbb{F}_q^{F_i}$, that is, $\hat{s}\,[i] = D_{i+T}(\{y[j,1], y[j,2], ..., y[j,i+T]\}_{j=1}^{M})$, $i = 1, \ldots, k$.

We require codes that recover source symbols with delay T and perform zero-error decoding, i.e. $\hat{s}^k = s^k$. In formal terms, $H(S^k|Y_1^n, \ldots, Y_M^n) = 0$. Moreover we require perfect security, that is $H(S^k|Z^n) = H(S^k)$, where $Z^n \in \mathbb{F}_q^n$ is the observation at the eavesdropper.

Definition 6.2. *An $[Mn, k]_q$ M-link code C (that is we spread the vector of length Mn into M equal sized sub-vectors to transmit them over M links), capable of correcting any set of Z erasure bursts of length B such that $Mn-k = ZB$, which in fact is the maximum possible, is called an optimal ZB-burst-erasure correcting M-link code.*

For our purposes, we reformulate Proposition 4.1 to specify the necessary condition for a burst-erasure correcting M-link block code.

Proposition 6.1. *Let C be a linear $[Mn, k]_q$ code and let $\mathscr{B} \subset \mathscr{I} = \{1, \ldots, Mn\}$ be an erasure pattern that is the set of positions where erasures occur. Let $E_{\mathscr{B}}$ be the erased symbols. Then C can recover $E_{\mathscr{B}}$ iff the columns of a parity check matrix H_C corresponding to indices in $\mathscr{B}$ are linearly independent, or equivalently, iff the columns of a generator matrix G_C corresponding to indices in $\mathscr{I}\backslash\mathscr{B}$ have rank k.*

Next we provide causal M-link block codes for the channel model given in Section III, where $Z = M - 1$.

Theorem 6.2. *For the admissible parameters $T \geqslant B, \mu, M, Z = M - 1$, and a suitable finite field $\mathbb{F}_q$, there exist causal delay-optimal M-link codes that are able to recover source symbols with optimal delay even if ZB-burst erasures occur, where each B-burst erasure occurs on a separate link, and perfect security is provided even if the eavesdropper observes any link noiselessly, with maximum secrecy rate*

$$R_s = \frac{k}{n} = \frac{m-\mu}{n} = M - \frac{ZB}{T+B} - 1 = \frac{(M-1)T}{T+B}. \tag{6.1}$$

Proof. We divide the proof into the achievability part and the converse part. ∎

Remark 6.3. *In (6.1), for the two last equations we use $\mu = n$. $\frac{m-\mu}{n}$ is the more general expression, where $m = Mn - ZB$.*

Remark 6.4. *We can also use the block codes for a delay-optimal M-link channel model, where the eavesdropper causes the erasure bursts of length B in any Z links and observes a copy of any link.*

For $\mu = 0$, R_s degrades to the rate in [44, Theorem 1, for $L = 0$], which is the maximum possible for the multi-link scenario (without security constraint).

6.3.1. The Achievability

We deal with secure nested block codes described in 4.3.1. The block code for the M-link channel has M samples for a given "time unit", thus we write the encoder output for link j with $j = 1, \ldots, M$ as $x_j^n = (x[j,1], x[j,2], \ldots, x[j,n])$. The pair (j,i) indexes the position of a symbol within the M-link coding block. Let $G^{M-link} = (G_1 \,||\cdots||\, G_M)$ with $G_j = \begin{pmatrix} G_j^* \\ G_j' \end{pmatrix}$ for $j = 1, \ldots, M$, or $G^{M-link} = \begin{pmatrix} G^* \\ G' \end{pmatrix}$ for short. We represent the encoding operation as $(x_1^n \,||\cdots||\, x_M^n) = (s^k, e^\mu) \cdot (G_1 \,||\cdots||\, G_M)$, where G_j is an $m \times n$ generator matrix for link $j = 1, \ldots, M$. G^{M-link} is an $m \times Mn$ generator matrix for the outer code C and $G' = (G_1' \,||\ldots||\, G_M')$ is the $\mu \times Mn$ generator matrix for the coarse code C', thus $C' \subset C$. G^{M-link} consists of equal sized submatrices $G_j = \begin{pmatrix} G_j^* \\ G_j' \end{pmatrix}$.

In the following, a slight modification of Theorem 4.11 yields the conditions for an optimal burst-erasure correcting nested M-link block code C that provides perfect security (without delay constraint).

Lemma 6.5. *Let $G^{M-link} = \begin{pmatrix} G^* \\ G' \end{pmatrix}$ be an $m \times Mn$ generator matrix for a linear M-link block code C, where G' is a $\mu \times Mn$ submatrix of G^{M-link}. Then C is an optimal $|\tilde{\mathscr{B}}|$-burst-erasure correcting M-link code that can convey $k = m - \mu$ symbols with perfect security and achieves the maximum secrecy rate $R_s = \frac{m-\mu}{n}$ if the following three conditions are fulfilled:*

(a) $m > \mu$,

(b) the columns of G^{M-link} for code C are linearly independent on positions $\mathscr{I} \backslash \tilde{\mathscr{B}}$, where $\mathscr{I} = \{1, \ldots, Mn\}$ and $|\tilde{\mathscr{B}}| = Mn - m$ (see Proposition 6.1),

(c) the columns of G' on position $\{i_1, \ldots, i_\mu\}$, the non-erased positions at the eavesdropper, are linearly independent.

In the sequel, we show that constructions for delay-optimal M-link codes satisfying (a), (b), (c) have the same maximum secrecy rate as the corresponding constructions for codes satisfying (a), (b), (c) without delay constraint.

We provide codes for $T \geqslant B$ and $n = \mu = T + B$.

Next we show the tradeoff between r, T, B, M, k.

Proposition 6.6. *Let $k = tT$, where $t \in \mathbb{N}$, and let r be a positive integer such that $|\tilde{\mathscr{B}}| = (M - r)B$. Then M-link block codes for our channel model, which provide perfect*

security and recover source symbols with delay T, must satisfy:

$$r = \begin{cases} i & t = M + (i-2) \text{ and } T = B \\ 1 & t = M - 1 \text{ and } T > B \\ \frac{T}{B} + 1 & t = M \text{ and } T > B. \end{cases}$$

Proof. According to Lemma 6.5, we have $(M-r)B = M(T+B) - (k+\mu) = (M-1)(T+B) - k$, with $\mu = T + B$. We choose $k = tT$ to ensure that for any burst pattern the source symbols injected at time i, that is $s[i]$, can be recovered with delay T using $G_{\mathcal{T}}$ (see (5.2)) in each link. Thus, $r = \frac{T}{B}(t - M) + \frac{T}{B} + 1$. ∎

Note that $F_i = t$ for $i = 1, \ldots, T$. Also note that in Proposition 6.6 we do not make a statement about the existence of a code for our channel model.

Corollary 6.7. *For $T > B$, a code as described in Proposition 6.6 is able to recover at most $(M-1)B$-burst erasures.*

We use the result of Proposition 6.6 for the case where $r = 1$ in the following theorem.

Theorem 6.8. *For $Z = M-1$, $k = tT$ with $t = M-1$ and for a suitable, small field size q, there exist M-link codes for the channel model described in Section 6.3 that satisfy the properties in Theorem 6.2.*

Proof. In the following, the source symbols injected at time $i = 1, \ldots, T$ are specified as $s[i] = (s_i, s_{i+T}, \ldots, s_{i+(t-1)T})$ and the encoder symbols injected at time $i = 1, \ldots, T + B$ as $e[i] = e_i$.

For a code that is able to recover $M-1$ bursts of length B (each occurring on a separate link) within delay T, and that provides perfect security if an eavesdropper is observing noiselessly a link of his choice, we first select a proper field size $q = p^m$, where $m \geqslant 1$ is an integer. For odd $M > 2$, we choose $p = 2$ and for even $M > 2$, we choose any prime $p > 2$, with integer $m \geqslant 1$, so that a $[T + B, T]_q$ linear code with systematic generator matrix $G_{\mathcal{T}}$ (as in (5.2)) can be constructed. For odd M, define $a = 1$. For even M we choose any $a \in \mathbb{F}_q \backslash \{0\}$. We require $(M-1) \mod p \neq 1$ and $(M-2) \mod p \neq 0$. The former is required to ensure the linear independence of the columns at non-erased positions in G^{M-link}. We will deal with the case $M = 2$ later. For any $q = p^m$, we can always choose $G_{\mathcal{T}}$ for a code over $\mathbb{F}_p$ while maintaining its properties (i.e. $G_{\mathcal{T}}$ is still a generator matrix for a delay-optimal B-burst-erasure correcting code).

Encoding: The encoding matrix of code C is

$$G^{M-link} = \begin{pmatrix} G^* \\ G' \end{pmatrix} = \begin{pmatrix} U \otimes G_{\mathcal{T}} \\ L \otimes I_\mu \end{pmatrix}, \tag{6.2}$$

where $U = (I_{(M-1)} \; a^{M-1})$ is an $(M-1) \times M$ matrix, a^{M-1} is a column vector of length $M-1$ and L is an M-length row vector with ones as its entries, except the last one, which

is a (see Example 6.9). Note that the constructed code is causal by observing the source symbols $s[i]$ and the encoder symbols $e[i]$ injected at time i and the corresponding matrix $G_j = \begin{pmatrix} G_j^* \\ G_j' \end{pmatrix}$, $j = 1 \ldots, M$.

For the case where $M = 2$, constructions for a delay-optimal code that meets the conditions in Lemma 6.5 exist for $T = B$. The generator matrix for the code over $\mathbb{F}_q$ with $q \geqslant 2$ is

$$G^{2-link} = \begin{pmatrix} G_{\_B} & \mathbf{0}_{B\times 2B} \\ I_{2B} & I_{2B} \end{pmatrix}.$$

Decoding: Now we want to show that for $M > 2$ and any admissible erasure pattern $\tilde{\mathscr{B}}$ we can recover any source symbols $s[i]$ with a decoding delay not exceeding T. For simplicity of description we divide the source vector s^k into equal sized sub-vectors $s_j^T, j = 1, \ldots, M-1$. Define $\mathscr{Z} \subset \mathscr{M} = \{1, \ldots, M\}$ as the set of links where the erasure bursts have occurred. With a slight abuse of notation, denote the noiseless link as $\mathscr{M}\backslash\mathscr{Z}$. Let $\mathscr{B}_M \subset \{1, \ldots, T+B\}$ be the erasure pattern in the last link. In the case where $\mathscr{M}\backslash\mathscr{Z} \neq \{M\}$, we first subtract the correctly received vector $y_{\mathscr{M}\backslash\mathscr{Z}}^{T+B} = x_{\mathscr{M}\backslash\mathscr{Z}}^{T+B}$ (after multiplying it by a) from the vector transmitted on the last link. Then erasures on $\mathscr{B}_M$ in the resulting vector $c_M^{T+B} := y_M^{T+B} - a x_{\mathscr{M}\backslash\mathscr{Z}}^{T+B}$ can be reconstructed by $aG_{\_T}$ (within delay T). Note that the sum of $M-2$ codewords encoded by $aG_{\_T}$ results again in a codeword of the linear code with generator matrix $aG_{\_T}$. Thus we can determine $x_M^{T+B} = c_M^{T+B} + a x_{\mathscr{M}\backslash\mathscr{Z}}^{T+B}$. Hereafter, $a y_j^{T+B}$ with $j \in \mathscr{Z}\backslash\{M\}$ can be subtracted sequentially from the last link to obtain x_j^{T+B}. Finally, when M is odd, we only need to reconstruct $e^{T+B} = x_M^{T+B} - (x_1^{T+B} + \ldots + x_{M-1}^{T+B})$ to get $s_j^T = x_j^{T+B} - e^{T+B}$. When M is even, we obtain $T+B$ (of which we only need T) linear systems of equations with $M-1$ equations and $M-1$ unknowns, each. If we convert the $M-1$ equations to matrix form then we get the following $(M-1)\times(M-1)$ matrix $A = \begin{pmatrix} 0 & a & \cdots & \cdots & a \\ a & 0 & a & \cdots & a \\ \vdots & a & \ddots & \ddots & \vdots \\ a & a & \cdots & \cdots & 0 \end{pmatrix}$, where the entries $m_{ij} = 0$ if $i = j$ and a otherwise. To show the full rank of the matrix we can calculate the determinate of the matrix using the Leibniz formula. We see that $\prod_{i=1}^{M-1} m_{i\sigma(i)} = 0$ if there is a permutation with a fixed point, where $\sigma(i)$ is the function value of the permutation σ at the point i and the set of all such permutations, the so-called symmetric group, is denoted by $\mathcal{S}_{M-1}$. Thus we get

$$\begin{aligned} \det(A) = a^{M-1}(&|\{\sigma \in \mathcal{S}_{M-1}\colon \text{ fixpointfree, } sgn(\sigma) = 1\}| \\ &- |\{\sigma \in \mathcal{S}_{M-1}\colon \text{ fixpointfree, } sgn(\sigma) = -1\}|). \end{aligned}$$

According to [50], for even M (the number of even derangements - the number of odd derangements) $= M-2$ so that $det(A) \neq 0$ if $p \nmid (M-2)$. Thus we obtain unique solutions due to the condition $(M-2) \mod p \neq 0$. In the case where $\mathscr{M}\backslash\mathscr{Z} = \{M\}$ we

can immediately start by determining x_j^{T+B}, with $j \in \mathscr{Z}$. If $M-1 \bmod p = 0$, we can recover the source symbols as is the case for odd M.

Note that each $s[i]$ can be recovered within delay T, since $G_{\_T}$ is the generator matrix of a B-burst-erasure correcting code with decoding delay T. Also note that G_j is cyclically good, that is every T cyclically consecutive columns of the matrix are linearly independent, $j = 1, \ldots, M$.

The decoding process for the case where $M = 2$ is obvious.

Security: Note that each matrix G'_j is cyclically good. According to Lemma 6.5(c), the code provides perfect security in the case where the eavesdropper observes any link of his choice, that is $\mu = T + B$. Moreover, perfect security holds, even if he switches from $x[j,i]$ to $x[j', i+1]$, where $j' \in \{0, \ldots, M\} \setminus \{j\}$.

In summary, we obtain an $[M(T+B), MT+B]_q$ code that satisfies Lemma 6.5 and additionally, can reconstruct the source symbols within delay T, despite bursts occurring that are arbitrarily positioned in the arbitrary $M-1$ links. ∎

Observe that for $T < B$ and $Z = M-1$ we get $m = \mu$, so that Lemma 6.5(a) is not fulfilled and therefore no code exists for the channel model described in Section III with a positive secrecy rate.

We give an example for $T > B$ and $r = 1$.

Example 6.9. *For $T = 3$, $B = 2$, $M = 4$ and $Z = 3$ we select $q = 3$ and $a = 2$. Then we have*

$$G^{4-link} = \begin{pmatrix} G_{\_3} & \mathbf{0}_{3\times 5} & \mathbf{0}_{3\times 5} & 2G_{\_3} \\ \mathbf{0}_{3\times 5} & G_{\_3} & \mathbf{0}_{3\times 5} & 2G_{\_3} \\ \mathbf{0}_{3\times 5} & \mathbf{0}_{3\times 5} & G_{\_3} & 2G_{\_3} \\ I_5 & I_5 & I_5 & 2I_5 \end{pmatrix}, \tag{6.3}$$

where $G_{\_3} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{pmatrix}$ *and* $L = \begin{pmatrix} 1 & 1 & 1 & 2 \end{pmatrix}$. *Suppose that the bursts of erasures have occurred on links 1,2,3 and suppose that on link 3 the first 2 positions are erased, such that* $y_3^5 = (?, ?, c_3, c_4, c_5)$. *Let* $c^5_{j,\hat{j}} := s_j^3 2G_{\_3} + s_{\hat{j}}^3 2G_{\_3} = x_4^5 - 2x_l^5$, *where* $l \in \{1, \ldots, 3\} \setminus \left\{\hat{j}, j\right\}$ *and* $j, \hat{j} \in \{1, \ldots, 3\}$ *with* $j \neq \hat{j}$. *First we determine* $c^5_{1,2} = s_1^3 2G_{\_3} + s_2^3 2G_{\_3}$ *by* $c^5_{1,2} = x_4^5 - 2x_3^5$. *Since the output transmitted over the last link passes the channel noiselessly (that is* $y_4^5 = x_4^5$*) and the output on link 3 has erasures in the first two positions, we obtain* $(?, ?, c_3, c_4, c_5) = x_4^5 - 2y_3^5$. *Now, according to Theorem 5.5 we can recover the erasures in* $(?, ?, c_3, c_4, c_5)$ *to obtain* $c^5_{1,2}$ *(and thus* x_3^5*) by the code with generator matrix* $2G_{\_3}$ *within* $T = 3$. *Similarly, we determine* $c^5_{1,3} = s_1^3 2G_{\_3} + s_3^3 2G_{\_3}$ *and* $c^5_{2,3} = s_2^3 2G_{\_3} + s_3^3 2G_{\_3}$ *by calculating* $x_4^5 - 2y_2^5$ *and* $x_4^5 - 2y_1^5$, *respectively, and by the code with generator matrix* $2G_{\_3}$. *Then we can recover the source symbols as described for the odd case, since* $3 \bmod 3 = 0$, *that is by calculating* $s_j^T = x_j^{T+B} - e^{T+B}$. *Thus*

the source symbols injected at time i can be recovered within delay 3. Observe that $s[1] = (s_1, s_4, s_7)$, $s[2] = (s_2, s_5, s_8)$ and $e[1] = e_1$, with $F_i = 3$ up to $i = T$ and $L_i = 1$ up to $i = n$.

6.3.2. The Converse

We provide the converse result of Theorem 6.2. We proceed as in [44] where we prove the upper bound on the secrecy rate by contradiction.

In the following, random variable $X^n \backslash Y_i, \ldots, Y_j$ with $1 \leqslant i \leqslant j \leqslant n$ corresponds to realization $(x_1, \ldots, x_{i-1}, x_{j+1}, \ldots, x_n)$. $x[a, \ldots, b]$ is denoted by $x[a : b]$. We assume that the source symbols are i.i.d. uniform distributed over $\mathbb{F}_q$ and thus $H(S) = \log_q |\mathbb{F}_q|$. We start with a useful lemma.

Lemma 6.10 ([44]). *Let $X^n = (X_1, \ldots, X_n)$, with $n \geqslant 2$. If $H(S) > 0$ and $H(S|X^n \backslash X_i) = 0$, $i = 1, \ldots, n$, then $H(X^n) < \sum_{i=1}^{n} H(X_i)$.*

Next we derive the upper bound for k, that is the number of source symbols that we can securely convey over the M-link channel model described in Section 6.3. Let X^{Mn} be the random variable of the realization $(x_1^n \,||\cdots||\, x_M^n)$.

Proposition 6.11. *In an M-link channel model, where on any of the Z links, respectively, any B-burst erasure can occur and where an eavesdropper can choose μ symbols to observe, from $\mathbb{F}_q$, one can convey securely and with zero-error decoding, at most*

$$k = Mn - ZB - \mu \tag{6.4}$$

symbols, on the condition that $Mn - ZB \geqslant \mu$.

Proof. For ease of description, let $\mathscr{X}^{Mn} = \mathbb{F}_q^{Mn}$ and $\mathscr{Z}^{\mu} = \mathbb{F}_q^{\mu}$. (6.4) follows from the following two conditions:

$$H(S^k|X^{Mn} \backslash X_{\tilde{\mathscr{B}}}) = 0, \qquad \text{(perfect reliability)} \tag{6.5}$$

$$H(S^k|Z^{\mu}) = k, \qquad \text{(perfect security)} \tag{6.6}$$

where $X^{Mn} \backslash X_{\tilde{\mathscr{B}}} \in \mathscr{X}^{Mn} \backslash \tilde{\mathscr{B}}$ and $\mathscr{Z}^{\mu} \subset \mathscr{X}^{Mn}$ are the random variables of the revealed symbols at the legitimate receiver and at the eavesdropper, respectively.

Let $\mathscr{I} = \{1, \ldots, Mn\}$. If $\mu \geqslant Mn - ZB$, and the set of positions $\mathscr{L} = \mathscr{I} \backslash \tilde{\mathscr{B}}$ of the revealed symbols at the legitimate receiver and the set of positions of the revealed symbols $\mathscr{E} \subset \mathscr{I}$ at the eavesdropper are chosen so that $\mathscr{L} \subset \mathscr{E}$, we have $H(S^k|X^{Mn} \backslash X_{\tilde{\mathscr{B}}}) \geqslant H(S^k|Z^{\mu})$. Hence, conditions (6.5) and (6.6) do not hold for this case. Suppose now $\mu \leqslant Mn - ZB$ and $\mathscr{E}, \mathscr{L}$ are chosen such that $\mathscr{E} \subset \mathscr{L}$. Then we have the following necessary condition for achieving (6.5) and (6.6): $k = H(S^k|Z^{\mu}) - H(S^k|X^{Mn} \backslash X_{\tilde{\mathscr{B}}}) \leqslant (Mn - ZB) - \mu$, which is Proposition 6.11. ∎

Time	1	2	3	4	5
Link 1	×	×			
Link 2	×	×			
Link 3	×	×			
Link 4					

Time	1	2	3	4	5
Link 1	×	×			
Link 2	×	×			
Link 3					
Link 4				×	×

Figure 6.2.: On the left, the table illustrates $\tilde{\mathscr{B}}_1 = \{I_t\}_{t=1}^{3}$ and on the right $\tilde{\mathscr{B}}_0 = \{I_t\}_{t=1}^{2} \bigcup ((4,4) : (4,5))$ for $T = 3$, $B = 2$, $n = \mu = 5$, $M = 4$ and $Z = 3$. Crossed squares illustrate erasures.

For the case $T \geqslant B$, we want to show that,

$$T \geqslant n - B. \tag{6.7}$$

Suppose for a contradiction that $T = n-B-1$. Then $s[1]$ must be reconstructed from time $n-B$ at the latest. Let S_1 be the random variable of $s[1]$. We divide each x_j^n, $j = 1, \ldots, M$ into distinct segments of length B, except the last positions, to obtain $x[1 : M, 1 : (n-B)]$. Let $X_{M\times(n-B)}$ be the random variable of realization $x\,[1 : M, 1 : (n-B)]$. Furthermore, let $N_B := \lceil \frac{n-B}{B} \rceil$ be the number of segments in each x_j^n. Let $I_{j,1} = \{(j,1), \ldots, (j,B)\}\,, I_{j,2} = \{(j, B+1), \ldots, (j, 2B)\}\,, \ldots, I_{j,N_B} = \{(j, N_B B - B + 1), \ldots, (j, n-B)\}$, $j = 1, \ldots, M$ be the index sets of the segments. Overall there are MN_B segments and we index them as $I_1 = I_{1,1}$, $I_2 = I_{2,1}, \ldots$, $I_{M+1} = I_{1,2}, \ldots$, $I_{MN_B} = I_{M,N_B}$.

We consider the following erasure patterns $\tilde{\mathscr{B}}_i = \{I_t\}_{t=1}^{Z-1} \bigcup I_{Z-1+i}$, $i = 1, \ldots, MN_B - Z + 1$. We refer to Fig. 6.2 for an example. We assume that S_1 can be reconstructed at $n - B$, which implies that $H(S_1 | X_{M\times(n-B)} \backslash X_{\tilde{\mathscr{B}}_i}) = 0$. According to Lemma 6.10, where $S = S_1$ and $X_i = X_{I_{Z-1+i}}$, $i = 1, \ldots, MN_B - Z + 1$, we have

$$\begin{aligned} H(X_{M\times(n-B)} \backslash X_{\{I_t\}_{t=1}^{Z-1}}) &< \sum\nolimits_{i=1}^{MN_B - Z + 1} H(X_{I_{Z-1+i}}) \\ &= k + \mu - (M-1)B, \end{aligned} \tag{6.8}$$

where (6.8) follows from (6.4).

Now suppose the bursts of erasures occur on position $\tilde{\mathscr{B}}_0 = \{I_t\}_{t=1}^{Z-1} \bigcup \{(M, n-B+1), \ldots, (M, n)\}$, as illustrated on the right side of Fig. 6.2. Let $X_{(M-1)\times B}$ be the random variable of $x\,[1 : (M-1), (n-B+1) : n]$. Furthermore, let us denote $X_{M\times(n-B)} \backslash X_{\{I_t\}_{t=1}^{Z-1}}$ as the random variable of realization $x\,[Z : M, 1 : (n-B)]$ together with $x\,[1 : (Z-1), (B+1) : (n-B)]$. For $\mu = \mu' + \mu''$, let $Z_{\mu'}$ and $Z_{\mu''}$ be the random variables of $z_{\mu'}$ and $z_{\mu''}$, where $z_{\mu'} = x\,[M-1, 1 : (n-B)]$ and $z_{\mu''} = x\,[M-1, (n-B+1) : n]$. Let $X_{M\times(n-B)} \backslash (X_{\{I_t\}_{t=1}^{Z-1}}, Z_{\mu'})$ and $X_{(M-1)\times B} \backslash Z_{\mu''}$ be the random variables of symbols revealed to the legitimate receiver, except the symbols observed by the eavesdropper. The source symbols S^k must be reconstructed from $X_{M\times(n-B)} \backslash X_{\{I_t\}_{t=1}^{Z-1}}$ and $X_{(M-1)\times B}$, hence,

$H(S^k|X_{M\times(n-B)}\backslash X_{\{I_t\}_{t=1}^{Z-1}}, X_{(M-1)\times B}) = 0$. But

$$\begin{aligned}
0 &= H(S^k|X_{M\times(n-B)}\backslash X_{\{I_t\}_{t=1}^{Z-1}}, X_{(M-1)\times B}) \\
&\geqslant H(S^k) + H(X_{M\times(n-B)}\backslash X_{\{I_t\}_{t=1}^{Z-1}}, X_{(M-1)\times B}|S^k) \\
&- H(X_{M\times(n-B)}\backslash(X_{\{I_t\}_{t=1}^{Z-1}}, Z_{\mu'})) - H(X_{(M-1)\times B}\backslash Z_{\mu''}) \\
&- H(Z_\mu) \overset{a)}{\geqslant} H(S^k) - H(X_{M\times(n-B)}\backslash(X_{\{I_t\}_{t=1}^{Z-1}}, Z_{\mu'})) \\
&- H(X_{(M-1)\times B}\backslash Z_{\mu''}) \\
&\overset{b)}{>} k - (k+\mu-(M-1)B-\mu') - ((M-1)B-\mu'') = 0,
\end{aligned} \tag{6.9}$$

where a) follows from (6.4), where

$$\begin{aligned}
&H(X_{M\times(n-B)}\backslash X_{\{I_t\}_{t=1}^{Z-1}}, X_{(M-1)\times B}|S^k) \\
&\geqslant H(X_{M\times(n-B)}\backslash X_{\{I_t\}_{t=1}^{Z-1}}, X_{(M-1)\times B}) - H(S^k) \\
&= (Mn - ZB) - k = \mu
\end{aligned}$$

and $H(X_{M\times(n-B)}\backslash X_{\{I_t\}_{t=1}^{Z-1}}, X_{(M-1)\times B}|S^k) - H(Z_\mu) \geqslant 0$. b) follows from (6.8) and $H(X_{M\times(n-B)}\backslash(X_{\{I_t\}_{t=1}^{Z-1}}, Z_{\mu'})) = H(X_{M\times(n-B)}\backslash X_{\{I_t\}_{t=1}^{Z-1}}) - H(Z_{\mu'})$, $H(X_{(M-1)\times B}\backslash Z_{\mu''}) = H(X_{(M-1)\times B}) - H(Z_{\mu''})$. As a consequence, we obtain by (6.4) and (6.7),

$$R_s = \frac{k}{n} \leqslant M - \frac{ZB+\mu}{T+B}, \tag{6.10}$$

which matches the secrecy rate in Theorem 6.2 for $\mu = T + B$ and $Z = M - 1$.

6.4. The Secure Streaming Codes for the M-Link Channel

In this section we briefly discuss the mapping from M-link block codes to M-link streaming codes for the channel model given in Section 6.2. We use diagonal interleaving to obtain M-link streaming codes from M-link block codes described in Theorem 6.2. A detailed description of the mapping for the case where $M = 1$ is given in Subsection 5.5.1. The extension to the case $M > 1$ is straightforward.

We define the mapping of an M-link convolutional code that has encoder memory ϖ, from a causal and nested M-link block code with generator matrix $G^{M-link} = \begin{pmatrix} G^* \\ G' \end{pmatrix}$.

Definition 6.3. *An $(n, k, \mu, \varpi, T)_q$ M-link convolutional code with encoder memory ϖ is an $(n, k, \mu, T)_q$ streaming code for an M-link channel with an eavesdropper, constructed as follows: For any $i \geqslant 0$ and $j = 1, \ldots, M$, we obtain the packet*

$$x[i,j] = \sum_{l=0}^{\varpi}(s\,[i-l]\,G_{j,l}^{*conv} + e\,[i-l]\,G_{j,l}^{\prime conv}), \tag{6.11}$$

where $G_{j,l}^{*conv}$ *is a* $k \times n$ *matrix so that*

$$G_j^* = \sum_{l=0}^{n-1} G_{j,l}^{*conv} \text{ with } G^* = (G_1^* \,||\cdots||\, G_M^*) \tag{6.12}$$

and $G_{j,l}^{'conv}$ *is a* $\mu \times n$ *matrix so that*

$$G_j^{'} = \sum_{l=0}^{n-1} G_{j,l}^{'conv} \text{ with } G^{'} = (G_1^{'} \,||\cdots||\, G_M^{'}). \tag{6.13}$$

By convention, we choose $s[-1], \ldots, s[-\varpi] = \mathbf{0}_{1\times k}$ *and* $e[-1], \ldots, e[-\varpi]$*, which correspond to i.i.d random variables over* $\mathbb{F}_q^{\mu}$.

Here $G_{j,0}^{'conv} = G_j^{'}$, and $G_{j,l}^{*conv}$ corresponds to the l-th diagonal (starting from column l) of $G_{\_T}$ in the corresponding link j, and zeros elsewhere.

6.4.1. Achievability

We provide $(n, k, \mu, T)_q$ streaming codes for an M-link channel, where $Z = M - 1$ burst erasures of length $B \leqslant T$ can occur, each on a separate link, and where the eavesdropper can observe any link noiselessly.

Theorem 6.12. *For the admissible parameters* T, B, μ, M, k, q*, there exists a* $(T, B, \mu, M-1; T+1)_q$ *M-link streaming code as given in Definition 6.1, with secrecy rate*

$$R_s = \frac{k}{n} = \frac{(M-1)T}{T+B}, \tag{6.14}$$

obtained by diagonal interleaving causal M-link block codes described in Theorem 6.2.

Proof. In each link, each diagonal in the resulting M-link streaming code is an output x_j^{T+B} of the M-link block code, where $j = 1, \ldots, M$. Thus, all M-tuples of diagonals are independent random variables taking values from $\mathbb{F}_q^{M(T+B)}$. In Fig. 6.4, an M-tuple of diagonals is underlined. The obtained M-link code allows us to reconstruct the source packets induced at time i with delay of at most T. This follows from the property of the corresponding M-link block code, since each source symbol in $(M-1)B$ erased packets can be recovered with delay T, using the corresponding tuple of M diagonals.

Moreover, in Theorem 6.2 we see that due to the form of the generator matrix G^{M-link} it follows that $G_{j,l}^{*conv} = \mathbf{0}_{T\times(T+B)}, G_{j,l}^{'conv} = \mathbf{0}_{(T+B)\times(T+B)}$ for any $l \geqslant T+1$, thus $W = T+1$, and we can recover any burst of B erasures in link j as long as they are separated by at least T packets.

It remains to be shown that the code also provides perfect security. Let S^k, $X_1^{T+B}, \ldots, X_M^{T+B}$ and Z^{μ} respectively, be the random variables corresponding to the source symbols, the output of the M-link block code with realization

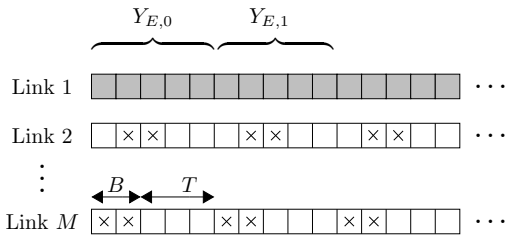


Figure 6.3.: The M-link channel used in proving the upper bound for $T \geqslant B$, with indication of which packets are observed by the eavesdropper $Y_{E,i}$ (gray squares). $Y_{(L\setminus E),i}$ is indicated by white squares. Crossed squares are erasures of length B.

$(x_1^{T+B} \;||\; \cdots \;||\; x_M^{T+B})$ and the observation at the eavesdropper, where $Z^\mu \in \mathbb{F}_q^\mu$. Recall that the eavesdropper can observe any link j, that is $\mu = T + B$. According to Lemma 6.5 (c) (see also [4]), related to our model, perfect security is achieved iff the matrix G'_j for each $j = 1, \ldots, M$ is a $(T+B) \times (T+B)$ c-good matrix. The latter implies that each coset of C' has the same number of vectors from which Z^{T+B} can be obtained by $(M-1)(T+B)$ erasures. This means, by construction of the M-link block code that we have perfect security, that is $H(S^k|Z^{T+B}) = k$. Now suppose the eavesdropper observes any link j. Consider every M-tuples of diagonals of length $T + B$ separately. Let S^k be the random variable that corresponds to the source symbols of the M-link block code which appears along the M-tuple diagonals, and Z^{T+B} be the random variable that corresponds to the eavesdropper's channel output of the M-link block code which appears along one of the M-tuple diagonals. By construction of the M-link block code, we have that $H(S^k|Z^{T+B}) = k$. Thus, the mutual information $I(S^k; Z^{T+B}) = 0$ and this holds for every M-tuple of diagonals. Furthermore, note that for any time unit i, the symbols in packets $x[i, 1], \ldots, x[i, M]$ are equiprobable and any $T + B$ consecutive packets in any link are mutually independent. This follows as the codewords are i.i.d. vectors.

Thus any $T + B$ consecutive packets observed by the eavesdropper in any link j do not reveal any information about the source symbols in those packets, as well as in the remaining $(M - 1) \times (T + B)$ packets, and in previously observed packets. This is also true when the eavesdropper changes the link it is observing.

Since $k = M(T + B) - (M - 1)B - (T + B)$, we get the result. ∎

6.4.2. Converse

Let $I_j = \{1, \ldots, P\}$ and $I = \{I_j\}_{j=1}^M$, $j = 1, \ldots, M$ respectively, be the index set for the packets in one period of length $P = T + B$ in link j and in link $1, \ldots, M$. Let $L \subset I$ and $E = I_j$, $j = 1, \ldots, M$ respectively, be the index sets of the revealed packets at the legitimate receiver and at the eavesdropper. In the i-th period, $Y_{E,i}$ and $Y_{L,i}$ are, respectively, the observations at the eavesdropper and the legitimate receiver. Fig. 6.3 shows the time slots and the size of Y_E for the case when $T = 3$ and $B = 2$. We assume

that $W = T + 1$, that is, in each of the $M - 1$ links the erasure bursts of length B are separated by T non-erased packets.

For any integer $h \geqslant 1$, we use $Y_{L,0}^{h-1}$ to denote $Y_{L,0}, Y_{L,1}, \ldots, Y_{L,h-1}$. We require a coding scheme that provides perfect reliability with delay T and perfect security for each $h \geqslant 1$, that is,

$$H(S_0^{h-1}|Y_{L,0}^{h-1}, U_{L,h}) = 0, \quad \text{(perfect reliability)} \tag{6.15}$$

$$H(S_0^{h-1}|Y_{E,0}^{h-1}, U_{E,h}) = H(S_0^{h-1}), \text{(perfect security)} \tag{6.16}$$

where $U_{L,h}$ and $U_{E,h}$ are, respectively, observations of the receiver and the eavesdropper in the interval of MT and T successive outcome packets within the h-th period, that is $|U_{L,h}| = MT$ and $|U_{E,h}| = T$. Furthermore, S_i is a random variable representing the sequence of messages produced by the source in the i-th period, thus $S_i \in \mathbb{F}_q^{P \cdot k}$ and $H(S_i) = P \cdot k$. We assume that all source packets have the same entropy. Denote $W_{L,h} = Y_{L,0}^{h-1}, U_{L,h}$ and $W_{E,h} = Y_{E,0}^{h-1}, U_{E,h}$. Hence, for achieving (6.15) and (6.16) we have the following necessary condition:

$$\begin{aligned} h \cdot P \cdot k &= H(S_0^{h-1}|W_{E,h}) \leqslant H(S_0^{h-1}, W_{L,h}|W_{E,h}) \\ &= H(W_{(L\backslash E),h}|W_{E,h}) + H(S_0^{h-1}|W_{L,h}, W_{E,h}) \\ &\leqslant H(W_{(L\backslash E),h}) = H(Y_{(L\backslash E),0}^{h-1}, U_{(L\backslash E),h}) \\ &\leqslant hH(Y_{(L\backslash E),0}) + n(MT - T). \end{aligned}$$

The latter implies that

$$R_s = \frac{k}{n} \leqslant \frac{ha + MT - T}{Ph} \overset{h\to\infty}{\to} \frac{a}{P}, \tag{6.17}$$

where $a = |(L\backslash E), 0|$, that is $H(Y_{(L\backslash E),0}) \leqslant na$.

For the case $T \geqslant B$, we have $|L| = M(T + B) - (M - 1)B$ and $|E| = T + B$ in each period. We obtain the following theorem.

Theorem 6.13. *For $T \geqslant B$, $P = T + B$, we obtain $H(Y_{(L\backslash E),0}) \leqslant [M(T + B) - (M - 1)B - (T + B)] \cdot n$, and hence $R_s \leqslant M - \frac{(M-1)B}{T+B} - 1$, which matches the secrecy rate in Theorem 6.12.*

6.5. Delay-Optimal Parallel Link Channel with an Active Eavesdropper and $Z = 1$

Here, we consider two similar models as in Section 6.2. In the first model $Z = 1$, that is, (in any sliding window) a burst erasure of length B can occur in any single link. In

Link 1

$x[i,1] =$	$\underline{s_1[i]+e_1[i]}$	$s_2[i]+e_2[i]$	$s_3[i]+e_3[i]$	$s_1[i-3]+s_3[i-1]+e_4[i]$	$s_2[i-3]+s_3[i-2]+e_5[i]$
$x[i+1,1] =$	$s_1[i+1]+e_1[i+1]$	$\underline{s_2[i+1]+e_2[i+1]}$	$s_3[i+1]+e_3[i+1]$	$s_1[i-2]+s_3[i]+e_4[i+1]$	$s_2[i-2]+s_3[i-1]+e_5[i+1]$
$x[i+2,1] =$	$s_1[i+2]+e_1[i+2]$	$s_2[i+2]+e_2[i+2]$	$\underline{s_3[i+2]+e_3[i+2]}$	$s_1[i-1]+s_3[i+1]+e_4[i+2]$	$s_2[i-1]+s_3[i]+e_5[i+2]$
$x[i+3,1] =$	$s_1[i+3]+e_1[i+3]$	$s_2[i+3]+e_2[i+3]$	$s_3[i+3]+e_3[i+3]$	$\underline{s_1[i]+s_3[i+2]+e_4[i+3]}$	$s_2[i]+s_3[i+1]+e_5[i+3]$
$x[i+4,1] =$	$s_1[i+4]+e_1[i+4]$	$s_2[i+4]+e_2[i+4]$	$s_3[i+4]+e_3[i+4]$	$s_1[i+1]+s_3[i+3]+e_4[i+4]$	$\underline{s_2[i+1]+s_3[i+2]+e_5[i+4]}$

Link 2

$x[i,2] =$	$\underline{s_4[i]+e_1[i]}$	$s_5[i]+e_2[i]$	$s_6[i]+e_3[i]$	$s_4[i-3]+s_6[i-1]+e_4[i]$	$s_5[i-3]+s_6[i-2]+e_5[i]$
$x[i+1,2] =$	$s_4[i+1]+e_1[i+1]$	$\underline{s_5[i+1]+e_2[i+1]}$	$s_6[i+1]+e_3[i+1]$	$s_4[i-2]+s_6[i]+e_4[i+1]$	$s_5[i-2]+s_6[i-1]+e_5[i+1]$
$x[i+2,2] =$	$s_4[i+2]+e_1[i+2]$	$s_5[i+2]+e_2[i+2]$	$\underline{s_6[i+2]+e_3[i+2]}$	$s_4[i-1]+s_6[i+1]+e_4[i+2]$	$s_5[i-1]+s_6[i]+e_5[i+2]$
$x[i+3,2] =$	$s_4[i+3]+e_1[i+3]$	$s_5[i+3]+e_2[i+3]$	$s_6[i+3]+e_3[i+3]$	$\underline{s_4[i]+s_6[i+2]+e_4[i+3]}$	$s_5[i]+s_6[i+1]+e_5[i+3]$
$x[i+4,2] =$	$s_4[i+4]+e_1[i+4]$	$s_5[i+4]+e_2[i+4]$	$s_6[i+4]+e_3[i+4]$	$s_4[i+1]+s_6[i+3]+e_4[i+4]$	$\underline{s_5[i+1]+s_6[i+2]+e_5[i+4]}$

.
.
.

Link 4

$x[i,4] =$	$\underline{2(s_1[i]+s_4[i]+s_7[i]+e_1[i])}$	$2(s_2[i]+s_5[i]+s_8[i]+e_2[i])$	$2(s_3[i]+s_6[i]+s_9[i]+e_3[i])$	$2(s_1[i-3]+s_4[i-3]+s_7[i-3]+s_3[i-1]+s_6[i-1]+s_9[i-1]+e_4[i])$	$2(s_2[i-3]+s_5[i-3]+s_8[i-3]+s_3[i-2]+s_6[i-2]+s_9[i-2]+e_5[i])$
$x[i+1,4] =$	$2(s_1[i+1]+s_4[i+1]+s_7[i+1]+e_1[i+1])$	$\underline{2(s_2[i+1]+s_5[i+1]+s_8[i+1]+e_2[i+1])}$	$2(s_3[i+1]+s_6[i+1]+s_9[i+1]+e_3[i+1])$	$2(s_1[i-2]+s_4[i-2]+s_7[i-2]+s_3[i]+s_6[i]+s_9[i]+e_4[i+1])$	$2(s_2[i-2]+s_5[i-2]+s_8[i-2]+s_3[i-1]+s_6[i-1]+s_9[i-1]+e_5[i+1])$
$x[i+2,4] =$	$2(s_1[i+2]+s_4[i+2]+s_7[i+2]+e_1[i+2])$	$2(s_2[i+2]+s_5[i+2]+s_8[i+2]+e_2[i+2])$	$\underline{2(s_3[i+2]+s_6[i+2]+s_9[i+2]+e_3[i+2])}$	$2(s_1[i-1]+s_4[i-1]+s_7[i-1]+s_3[i+1]+s_6[i+1]+s_9[i+1]+e_4[i+2])$	$2(s_2[i-1]+s_5[i-1]+s_8[i-1]+s_3[i]+s_6[i]+s_9[i]+e_5[i+2])$
$x[i+3,4] =$	$2(s_1[i+3]+s_4[i+3]+s_7[i+3]+e_1[i+3])$	$2(s_2[i+3]+s_5[i+3]+s_8[i+3]+e_2[i+3])$	$2(s_3[i+3]+s_6[i+3]+s_9[i+3]+e_3[i+3])$	$\underline{2(s_1[i]+s_4[i]+s_7[i]+s_3[i+2]+s_6[i+2]+s_9[i+2]+e_4[i+3])}$	$2(s_2[i]+s_5[i]+s_8[i]+s_3[i+1]+s_6[i+1]+s_9[i+1]+e_5[i+3])$
$x[i+4,4] =$	$2(s_1[i+4]+s_4[i+4]+s_7[i+4]+e_1[i+4])$	$2(s_2[i+4]+s_5[i+3]+s_8[i+4]+e_2[i+4]$	$2(s_3[i+4]+s_6[i+4]+s_9[i+4]+e_3[i+4])$	$2(s_1[i+1]+s_4[i+1]+s_7[i+1]+s_3[i+3]+s_6[i+3]+s_9[i+3]+e_4[i+4])$	$\underline{2(s_2[i+1]+s_5[i+1]+s_8[i+1]+s_3[i+2]+s_6[i+2]+s_9[i+2]+e_5[i+4])}$

Figure 6.4.: A $(5, 9, 5, 3, 3)_3$ 4-link convolutional code constructed by diagonally interleaving the 4-link block code with generator matrix G^{4-link} given in Example 6.9, where $Z = 3$, $T = 3$, $B = 2$.

addition, the eavesdropper is able to observe any interval of length μ in any link of his choice. In the second model, a complete link may fail and the eavesdropper is able to observe any complete link. We assume that the eavesdropper causes the erasures in any link of the legitimate receiver. Furthermore, we consider the case where the eavesdropper is able to change the link in which it causes erasures.

Again first we construct multi-link block codes and diagonally interleave them to multi-link streaming codes. The multi-link block codes can be used for a wiretap channel II model with delay constraint where the eavesdropper can choose any link and observe noiselessly any interval of μ symbols (also end-around) from n symbols transmitted to the legitimate receiver. Additionally, the eavesdropper can cause any burst of erasures of length B on any chosen link, that is $Z = 1$. We give explicit constructions of *optimal* multi-link block codes that achieve maximum secrecy rate, provide perfect security (i.e.

the eavesdropper can obtain no information about the secret message) and provide zero-error decoding with minimum possible delay. Moreover, we consider a multi-link wiretap channel II model with an active eavesdropper for streams of encoded packets. For $T \geqslant B$, we propose constructions of multi-link streaming codes for the case where the eavesdropper can cause in any link and in any sliding window of size W_1 an interval of erasures of at most B packets and observe any interval of at most μ packets in any sliding window of size W_2 in the same link or in any other. The multi-link streaming codes provide perfect security, zero-error decoding with delay T and have the maximum secrecy rate. In Subsection 6.5.2, we discuss the case where the eavesdropper is able to switch between the links, but it costs him δ time units.

When the channel packet $x[i, j^*]$ at time i is transmitted over link j^* then the legitimate receiver observes either $y[i, j^*] = ?$ if the channel packet at time i on link j^* is erased (caused by the eavesdropper) or $y[i, j^*] = x[i, j^*]$ if the channel packet is passed to the receiver noiselessly. Correspondingly, the channel packets observed by the eavesdropper on link j' are either passed noiselessly, i.e. $z[i, j'] = x[i, j']$ or are erased, i.e, $z[i, j'] = ?$.

The encoding and decoding procedure is the same as in Section 6.2. However, for $T \geqslant B$, we introduce an additional parameter V, which determines the length of an interval of erased packets separating the observed packets of the eavesdropper. In Subsection 6.5.2, we specify V, δ and the window size W_2, which are fully characterized by the parameters n, k, μ, T. When $T \geqslant B$, we call a code that fulfills the above requirements a $(T, B, \mu, Z = 1; W_1, W_2)_2$ M-link streaming code.

For the case where $T < B$, we construct binary codes for a slightly different model, where the eavesdropper can erase and observe a complete link of its choice. In this case, $V, W_1, W_2, \delta = 0$. The code for the channel model where the eavesdropper can erase and observe a complete link of its choice and that provides zero-error decoding within delay T and provides perfect security, we denote as a $(T, B, \mu, Z = 1)_2$ M-link streaming code

6.5.1. Construction of the Multi-Link Block Codes

We define a *secure delay-optimal B-burst-erasure correcting M-link block code* that can be converted to an M-link convolutional code for the channel model introduced in Section 6.5.

Definition 6.4. *We call an M-link binary block code C with generator matrix $G^{M-link} = \begin{pmatrix} G^* \\ G' \end{pmatrix}$ secure delay-optimal B-burst-erasure correcting, or $\mu - (Mn, m)$ secure delay-optimal B-burst-erasure correcting if*

1) C is an optimal burst-erasure correcting code, that is $m = Mn - B$.

2) The transmitter can convey $k = m - \mu$ symbols with perfect security and the code

has maximum secrecy rate

$$R_s = \frac{k}{n} = \frac{m-\mu}{n} = \begin{cases} M - \frac{B+\mu}{T+B} & T \geqslant B \\ M - 1 - \frac{\mu}{B} & T < B, \end{cases} \tag{6.18}$$

where $n = T + B$ *for the case where* $T \geqslant B$.

3) Every source symbol can be reconstructed with delay of at most T

Definition 6.5. *We call* $G^{M-link} = \begin{pmatrix} G^* \\ G' \end{pmatrix}$ *a systematic matrix if* G^* *and* G'_j*, the submatrix of* G'*, are of a systematic form, respectively, with* $j = 1, \ldots, M$.

The next Lemma follows from Lemma 4.18.

Lemma 6.14. *Let* A *be an* $r \times n$ *c-good matrix. Then* $A \otimes I_m$ *is an* $rm \times nm$ *c-good matrix, where "*$\otimes$*" is the Kronecker product.*

In the following we present systematic binary secure delay-optimal B-burst-erasure correcting M-link block codes for the M-link channel model described above.

Theorem 6.15. *There exist explicit constructions of systematic* $\mu - (Mn, m)$ *secure delay-optimal* B*-burst-erasure correcting* M*-link block codes that achieve the maximum secrecy rate* R_s *given in* (6.18) *for the following cases:*

1. *For* $B < \mu = T$ *where* $T = tB$, $n = T + B$ *and integer* $t \geqslant 2$.
2. *For* $B = \mu \leqslant T$ *where* $T = tB$, $n = T + B$, *with integer* $t \geqslant 1$, *and* $M \cdot \frac{T+B}{B}$ *is even.*
3. *For* $B < \mu < T$ *where,* $T = tB$, $n = T + B$ *and* $\mu = n/2$, *with* $t = 3$.
4. *For* $B = n = \mu$ *with* $T = 0$ *and even* M.

Proof. We divide the proof into two parts. The achievability we prove by the construction given in the sequel and the converse is given below.

The Achievability

In the first three cases, the main idea is to construct for admissible integers t, M an $(M(t+1)-1) \times M(t+1)$ systematic binary generator matrix $\acute{G} = \begin{pmatrix} \acute{G}^* \\ \acute{G}' \end{pmatrix} = \left(\begin{array}{c|c|c} \acute{G}^*_1 & \cdots & \acute{G}^*_M \\ \acute{G}'_1 & \cdots & \acute{G}'_M \end{array} \right)$, $j = 1, \ldots, M$ for a secure single erasure correcting M-link code. This, in fact, is the case when $\acute{G}$ is c-good and the $\mu' \times (t+1)$ submatrices $\acute{G}'_j$ of $\acute{G}$ for each j are c-good. (In this case the eavesdropper is able to observe μ' consecutive symbols from $t+1$, including cyclic intervals of length μ' in any link.) Then we apply the Kronecker product to $\acute{G}$ and I_B.

The resulting $(M(t+1)-1)B \times M(t+1)B$ matrix $G^{M-link} = (G_1 \,||\, \cdots \,||\, G_M)$ is c-good according to Lemma 6.14. Correspondingly, the $\mu \times (t+1)B$ submatrices G'_j of G^{M-link} are c-good, where $\mu = \mu' B$. According to Lemma 6.5 we have an optimal B-burst-erasure correcting M-link block code C that provides perfect security even if the eavesdropper is able to observe cyclic intervals of length μ in the chosen link and that achieves the secrecy rate given in (6.18). Then it remains to be shown that the code can reconstruct arbitrary source symbols with delay of at most T in each link.

1. For any integer $t \geqslant 2$, we construct an $M(t+1)-1 \times M(t+1)$ matrix $\acute{G} = (\acute{G}_1 \,||\, \cdots \,||\, \acute{G}_M)$, with

$$\acute{G}_j = \begin{pmatrix} \mathbf{0}_{(j-1)(t+1)\times(t+1)} \\ I_{t+1} \\ \mathbf{0}_{(M-1-j)(t+1)\times(t+1)} \\ I_t \;\; \mathbb{1}^t \end{pmatrix}, \qquad \acute{G}_M = \begin{pmatrix} \mathbf{0}_{(M-1)(t+1)\times t} \;\; \mathbb{1}^{(M-1)(t+1)} \\ I_t \;\; \mathbb{1}^t \end{pmatrix},$$

where $j = 1, \ldots, M-1$ and $\mathbb{1}^r = (1, 1, \ldots, 1)^\top$ of length r. Note that $\acute{G}$ is c-good, which contains c-good $t \times (t+1)$ submatrices $\acute{G}'_j$, $j = 1, \ldots, M$.

Now we take $\acute{G} \otimes I_B$ to obtain the binary systematic $M(T+B) - B \times M(T+B)$ generator matrix

$$G^{M-link} = (G_1 \,||\, G_2 \,||\, \cdots \,||\, G_M), \tag{6.19}$$

with

$$G_j = \begin{pmatrix} \mathbf{0}_{(j-1)(T+B)\times(T+B)} \\ I_{T+B} \\ \mathbf{0}_{(M-1-j)(T+B)\times(T+B)} \\ I_T \;\; A \end{pmatrix}, \qquad G_M = \begin{pmatrix} \mathbf{0}_{(M-1)(T+B)\times T} \;\; A' \\ I_T \;\; A \end{pmatrix}$$

for a secure B-burst-erasure correcting M-link block code, where $j = 1, \ldots, M-1$. Observe that $A = \mathbb{1}^t \otimes I_B$, $A' = \mathbb{1}^{(M-1)(t+1)} \otimes I_B$ and that $G'_j = (I_T \; A)$, $j = 1, \ldots, M$, is a $T \times (T+B)$ c-good matrix as well as G'.

To show that the M-link code can recover source symbols within delay T if a B-burst erasure occurs, consider the output $(x_1^{(t+1)B} || \cdots || x_M^{(t+1)B})$, where $tB = T$ and $(t+1)B = T + B$.

Note that the urgent source symbols are the first B symbols in x_j^{T+B}, $j = 1, \ldots, M$ since they must be reconstructed with delay of at most T. W.l.o.g, suppose a B-burst erasure occurs on the last link, that is, the burst erasure affects $x_M^{(t+1)B}$, then $(x_1^{(t+1)B} || \cdots || x_{M-1}^{(t+1)B})$ are received correctly. Let $(x_{1,i}^{t+1} || \cdots || x_{M,i}^{t+1})$ be, respectively, the subvectors of $(x_1^{(t+1)B} || \cdots || x_M^{(t+1)B})$, where $x_{j,i}^{t+1} = (x(j,i), x(j,i+B), \ldots, x(j,i+tB))$, with $i = 1, \ldots, B$ and $j = 1, \ldots, M$. Note that $(x_{1,i}^{t+1} || \cdots || x_{M,i}^{t+1})$, where $i = 1, \ldots, B$ is the output of the t-delay single erasure correcting M-link code with generator matrix $\acute{G}$. In subvector $x_{M,i}^{t+1} = (x(M,i), x(M,i+B), \ldots, x(M,i+tB))$ of $x_M^{(t+1)B}$, there is at most one erased symbol. Thus the source symbols can be reconstructed with delay of at most $tB = T$.

2. For any integer t, consider the $M(t+1)-1 \times M(t+1)$ matrix $\acute{G} = (\acute{G}_1 \,||\, \cdots \,||\, \acute{G}_M)$

of the form

$$\acute{G}_j = \begin{pmatrix} \mathbf{0}_{(j-1)(t+1)\times(t+1)} \\ I_{t+1} \\ \mathbf{0}_{(M-1-j)(t+1)+t-1\times(t+1)} \\ \mathbb{1}_{t+1} \end{pmatrix}, \qquad \acute{G}_M = \begin{pmatrix} \mathbf{0}_{(M-1)(t+1)\times t} & \mathbb{1}^{(M-1)(t+1)} \\ I_{t-1} & \mathbf{0}_{(t-1)\times 1} \; \mathbb{1}^{t-1} \\ & \mathbb{1}_{t+1} \end{pmatrix},$$

where $j = 1, \ldots, M-1$ and $\mathbb{1}_r = (1, 1, \ldots, 1)$ of length r. Observe that for even $M(t+1)$, $\acute{G}$ is a c-good matrix which contains c-good $1 \times (t+1)$ submatrices $\acute{G}'_j$, $j = 1, \ldots, M$. Then the Kronecker product of $\acute{G}$ and I_B yields a c-good systematic matrix G^{M-link}, which is the generator matrix for an optimal B-burst-erasure correcting M-link code with delay $T = tB$. This we can show by arguing in the same way as in case 1. For even $M(t+1)$ we obtain that $M \cdot \frac{T+B}{B}$ is even since $M(t+1)B = M(T+B)$. Moreover, G^{M-link} contains a $\mu \times (T+B)M$ matrix $G' = (G'_1 \,||\ldots||\, G'_M) = (\underbrace{I_B \cdots I_B}_{(t+1) \text{ times}} \,||\ldots||\, I_B \cdots I_B)$, which is c-good as well as G'_j, $j = 1, \ldots, M$.

3. For $t = 3$, consider the c-good $4M - 1 \times 4M$ matrix $\acute{G} = (\acute{G}_1 \,||\, \cdots \,||\, \acute{G}_M)$ of the form

$$\acute{G}_j = \begin{pmatrix} \mathbf{0}_{(j-1)4\times 4} \\ I_4 \\ \mathbf{0}_{(M-1-j)4+1\times 4} \\ I_2 \; I_2 \end{pmatrix}, \qquad \acute{G}_M = \begin{pmatrix} \mathbf{0}_{(M-1)4\times t} & \mathbb{1}^{(M-1)4} \\ 1 & \mathbf{0}_{4\times 1} \; 1 \\ & I_2 \; I_2 \end{pmatrix},$$

where $j = 1, \ldots, M-1$. The Kronecker product of $\acute{G}$ and I_B yields a c-good systematic $4MB - B \times 4MB$ matrix G^{M-link}, which contains M c-good $\mu \times B4$ submatrices with $\mu = B(4/2)$. Obviously, we can recover each source symbol with delay of at most $T = 3B$, as in cases 1 and 2.

Thus we obtain a generator matrix for an optimal B-burst-erasure correcting M-link code.

4. For any $n = \mu \geqslant 1$ and even $M > 2$, let $G^{M-link} = (G_1 \,||\cdots||\, G_M)$ with

$$G_j = \begin{pmatrix} G^*_j \\ G'_j \end{pmatrix} = \begin{pmatrix} \mathbf{0}_{(j-1)n\times n} \\ I_n \\ \mathbf{0}_{(M-2-j)(n)\times n} \\ I_n \end{pmatrix}, \quad G_{M-1} = \begin{pmatrix} \mathbf{0}_{(M-2)(n)\times n} \\ I_n \end{pmatrix}, \qquad G_M = \begin{pmatrix} I_n^{(1)} \\ \vdots \\ I_n^{(M-1)} \end{pmatrix}$$

being an $(M-1)n \times Mn$ systematic binary c-good matrix that contains an $n \times Mn$ c-good submatrix $G' = \underbrace{(I_n \,||\cdots||\, I_n)}_{\text{M times}}$, where $G'_j = I_n$, $j = 1, \ldots, M$. Note that if a link outage occurs on link j^*, that is all the symbols transmitted over link j^* are erased, the source symbols $s_{i+(j-1)n}$, $i = 1, \ldots, n$ and $j = 1, \ldots, M-2$ can be reconstructed immediately from symbols $x[\hat{j}, i]$, $\hat{j} \in J \backslash j^*$, where $J = \{1, \ldots, M\}$. Thus delay $T = 0$. Note that the code also can correct n erasures which can be distributed over several links but must not occur at the same time, e.g. one deletion in the first link at time $i = 1$ and $n-1$ deletions in the second link at time $i = 2, ..., n$.

According to Lemma 6.5 and since $T = 0$, we have a secure delay-optimal n-burst-

erasure correcting M-link code, which is secured against an eavesdropper who is able to observe a complete link of his choice and which achieves the secrecy rate as given in (6.18). Note also that for odd M, G^{M-link} is not c-good and the condition $M > 2$ is necessary according to Lemma 6.5 (a). The latter applies since in the case where $M = 2$, the generator matrix G^{M-link} for an n-burst-erasure correcting code is of dimension $n \times 2n$, which implies that $\mu = n = m$.

Example 6.16. *As an example of a code construction in Theorem 6.15 for case 1 where* $M = 2$, $T = \mu = 4$ *and* $B = 2$, *consider the following systematic generator matrix* $\acute{G} = (\acute{G}_1 \mid\mid \acute{G}_2) = \left(\begin{array}{ccc||ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{array}\right)$ *for a binary secure single erasure correcting 2-link block code, where* $t = 2$. *Then take* $\acute{G} \otimes I_B$ *to obtain the generator matrix*

$$G^{2-link} = \left(\begin{array}{cccccc||cccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{array}\right)$$

for a 4-$(12, 10)$ *secure delay-optimal* 2*-burst-erasure correcting* 2*-link block code with secrecy rate* $R_s = 1$.

Converse

For $T \geqslant B$, the converse can be proven as in Subsection 6.3.2 for the case $Z = 1$.

For the case $T < B$, similar to [44], we want to show that

$$n \leqslant B. \tag{6.20}$$

Suppose for a contradiction that $n = B + 1$.

Consider the erasure pattern $\mathscr{B}_0 = \{(1, 1), \ldots, (1, B)\}$. Let $X_{(M-1)\times(T+1)}$ be the random variable corresponding to $x[2 : M, 1 : T + 1]$, where $T + 1 \leqslant B$. Since S_1 (all source symbols injected at time 1) is recovered at $T + 1$, we have $H(S_1, E_1|X_{(M-1)\times(T+1)}) = 0$. We assume that the encoder symbols $E_1, \ldots, E_\mu$ are i.i.d. uniform distributed over $\mathbb{F}_q$.

Next, we consider the erasure pattern $\mathscr{B}_1 = \{(1, 2), \ldots, (1, n)\}$. The source symbols S^k must be reconstructed from $X_{M\times 1}$, which corresponds to realization $x[1 : M, 1]$ and

$X_{(M-1)\times(n-1)}$, which corresponds to realization $x[2 : M, 2 : n]$. Hence, $H(S^k|X_{M\times 1}, X_{(M-1)\times(n-1)}) = 0$.

But

$$\begin{aligned}
0 &= H(S^k|X_{M\times 1}, X_{(M-1)\times(n-1)}) \\
&= H(S^k, X_{M\times 1}, X_{(M-1)\times(n-1)}) - H(X_{M\times 1}, X_{(M-1)\times(n-1)}) \\
&\overset{a)}{\geqslant} H(S^k, X_{M\times 1}, X_{(M-1)\times(n-1)}) - H(S_1, E_1, X_{(M-1)\times n}) \\
&\overset{b)}{=} H(S^k, X_{M\times 1}, X_{(M-1)\times(n-1)}) - H(f(X_{(M-1)\times(T+1)}), X_{(M-1)\times n}) \\
&\overset{c)}{\geqslant} H(S^k, X_{M\times 1}, X_{(M-1)\times(n-1)}) - H(X_{(M-1)\times n}, X_{(M-1)\times n}) \\
&\geqslant H(S^k) + H(X_{M\times 1}, X_{(M-1)\times(n-1)}|S^k) - H(X_{(M-1)\times n}\backslash Z_\mu) - H(Z_\mu) \\
&\geqslant H(S^k) - H(X_{(M-1)\times n}\backslash Z_\mu) \\
&\overset{d)}{=} k - ((M-1)n - \mu) = 1,
\end{aligned}$$

where $a)$ follows the causality of the code, thus X_1 must be a function of S_1 and E_1, $b)$ follows $S_1, E_1 = f(X_{(M-1)\times(T+1)})$, $c)$ follows that $T + 1 < n$, d) follows from (6.4), where $H(X_{M\times 1}, X_{(M-1)\times(n-1)}|S^k) \geqslant H(X_{M\times 1}, X_{(M-1)\times(n-1)}) - H(S^k) = (Mn - B - k) = \mu$ and $H(X_{M\times 1}, X_{(M-1)\times(n-1)}|S^k) - H(Z_\mu) \geqslant 0$. Thus, (6.20) together with (6.4) imply

$$R_s = \frac{k}{n} \leqslant M - 1 - \frac{\mu}{B}, \tag{6.21}$$

which matches the secrecy rate in Definition 6.4 and in Theorem 6.15.

6.5.2. The Multi-Link Streaming Codes

We use diagonal interleaving to obtain M-link streaming codes from M-link block codes given in Theorem 6.15. For the case where $M = 1$, a detailed description of the mapping is given in Subsection 5.5.1, and the extension of the case where $M > 1$ is straightforward.

We obtain the following result.

Theorem 6.17. *For the admissible parameters T, B, μ, M, there exist for $T \geqslant B$, $(T, B, \mu, 1; T + 1, W_2)_2$ M-link streaming codes, and for $T < B$, $(T, B, \mu, 1)_2$ M-link streaming codes with secrecy rate*

$$R_s = \frac{k}{n} = \frac{m-\mu}{n} = \begin{cases} M - \frac{B+\mu}{T+B} & T \geqslant B \\ M - 1 - \frac{\mu}{B} & T < B, \end{cases} \tag{6.22}$$

obtained by diagonal interleaving systematic binary $\mu - (Mn, m)$ secure delay-optimal B-burst-erasure correcting M-link block codes described in Theorem 6.15. W_2 is case dependent so that

Link 1

$x[i,1] =$	$\underline{s_1[i] + e_1[i]}$	$s_2[i] + e_2[i]$	$s_3[i] + e_3[i]$	$s_4[i] + e_4[i]$	$s_5[i] + e_1[i-4] + e_3[i-2]$	$s_6[i] + e_2[i-4] + e_4[i-2]$
$x[i+1,1] =$	$s_1[i+1] + e_1[i+1]$	$\underline{s_2[i+1] + e_2[i+1]}$	$s_3[i+1] + e_3[i+1]$	$s_4[i+1] + e_4[i+1]$	$s_5[i+1] + e_1[i-3] + e_3[i-1]$	$s_6[i+1] + e_2[i-3] + e_4[i-1]$
$x[i+2,1] =$	$s_1[i+2] + e_1[i+2]$	$s_2[i+2] + e_2[i+2]$	$\underline{s_3[i+2] + e_3[i+2]}$	$s_4[i+2] + e_4[i+2]$	$s_5[i+2] + e_1[i-2] + e_3[i]$	$s_6[i+2] + e_2[i-2] + e_4[i]$
$x[i+3,1] =$	$s_1[i+3] + e_1[i+3]$	$s_2[i+3] + e_2[i+3]$	$s_3[i+3] + e_3[i+3]$	$\underline{s_4[i+3] + e_4[i+3]}$	$s_5[i+3] + e_1[i-1] + e_3[i+1]$	$s_6[i+3] + e_2[i-1] + e_4[i+1]$
$x[i+4,1] =$	$s_1[i+4] + e_1[i+4]$	$s_2[i+4] + e_2[i+4]$	$s_3[i+4] + e_3[i+4]$	$s_4[i+4] + e_4[i+4]$	$\underline{s_5[i+4] + e_1[i] + e_3[i+2]}$	$s_6[i+4] + e_2[i] + e_4[i+2]$
$x[i+5,1] =$	$s_1[i+5] + e_1[i+5]$	$s_2[i+5] + e_2[i+5]$	$s_3[i+5] + e_3[i+5]$	$s_4[i+5] + e_4[i+5]$	$s_5[i+5] + e_1[i+1] + e_3[i+3]$	$\underline{s_6[i+5] + e_2[i+1] + e_4[i+3]}$

Link 2

$x[i,2] =$	$\underline{e_1[i]}$	$e_2[i]$	$e_3[i]$	$e_4[i]$	$s_1[i-4] + s_3[i-2] + s_5[i]$ $e_1[i-4] + e_3[i-2]$	$s_2[i-4] + s_4[i-2] + s_6[i]$ $e_2[i-4] + e_4[i-2]$
$x[i+1,2] =$	$e_1[i+1]$	$\underline{e_2[i+1]}$	$e_3[i+1]$	$e_4[i+1]$	$s_1[i-3] + s_3[i-1] + s_5[i+1]$ $e_1[i-3] + e_3[i-1]$	$s_2[i-3] + s_4[i-1] + s_6[i+1]$ $e_2[i-3] + e_4[i-1]$
$x[i+2,2] =$	$e_1[i+2]$	$e_2[i+2]$	$\underline{e_3[i+2]}$	$e_4[i+2]$	$s_1[i-2] + s_3[i] + s_5[i+2]$ $e_1[i-2] + e_3[i]$	$s_2[i-2] + s_4[i] + s_6[i+2]$ $e_2[i-2] + e_4[i]$
$x[i+3,2] =$	$e_1[i+3]$	$e_2[i+3]$	$e_3[i+3]$	$\underline{e_4[i+3]}$	$s_1[i-1] + s_3[i+1] + s_5[i+3]$ $e_1[i-1] + e_3[i+1]$	$s_2[i-1] + s_4[i+1] + s_6[i+3]$ $e_2[i-1] + e_4[i+1]$
$x[i+4,2] =$	$e_1[i+4]$	$e_2[i+4]$	$e_3[i+4]$	$e_4[i+4]$	$\underline{s_1[i] + s_3[i+2] + s_5[i+4]}$ $\underline{e_1[i] + e_3[i+2]}$	$s_2[i] + s_4[i+2] + s_6[i+4]$ $e_2[i] + e_4[i+2]$
$x[i+5,2] =$	$e_1[i+5]$	$e_1[i+5]$	$e_3[i+5]$	$e_4[i+5]$	$s_1[i+1] + s_3[i+3] + s_5[i+5]$ $e_1[i+1] + e_3[i+3]$	$\underline{s_2[i+1] + s_4[i+3] + s_6[i+5]}$ $\underline{e_2[i+1] + e_4[i+3]}$

Figure 6.5.: A secrecy rate-1 code constructed by diagonally interleaving the $4-(12,10)$ delay-optimal secure two-link block code (see Example 6.16).

$$W_2 = \begin{cases} T+1 & \text{if } V = T \\ \mu+1 & \text{if } V < T \end{cases}. \tag{6.23}$$

In the latter case, the eavesdropper can observe either an interval of length of at most μ or at most $W_2 - V$ packets separated by V erased packets in any sliding window W_2.

Proof. For both the case where $T \geqslant B$ and $T < B$, Lemma 5.17 can be extenden for the M-link case in a straightforward way. Thus, the streaming codes obtained by diagonal interleaving systematic binary $\mu-(Mn,m)$ secure delay-optimal B-burst-erasure correcting M-link block codes as described in Theorem 6.15 are able to recover source packets with delay T when a B-burst-erasure occurs. Due to the form of the generator matrix G^{M-link} for the $\mu-(Mn,m)$ secure delay-optimal B-burst-erasure correcting M-link block code, it follows that for $\mu \leqslant T$ and for each $j = 1,\ldots,M$, we have that $G_{j,l}^{*conv}$ and $G_{j,l}^{\prime conv}$ are zero-matrices for any $l \geqslant T+1$, thus $\varpi = T$.

It remains to be shown that the M-link streaming code obtained by applying diagonal interleaving to the M-link nested block code provides perfect security when the eavesdropper observes in any link j either μ consecutive packets in any sliding window of size W_2 or the complete link, that is $\mu = n$.

Let S^k, $X_1^n,\ldots,X_M^n$ and Z^{Mn}, respectively, be the random variables corresponding to the source symbols, the output of the block code $(x_1^n \,||\, \cdots \,||\, x_M^n)$ and the observation at

the eavesdropper on any link $j = 1, \ldots, M$, where $Z^{Mn} \in (\mathbb{F}_q \cup \{?\})^{Mn}$. In the interest of simplification, we consider Z^{Mn} instead of $Z^{\mu} \in (\mathbb{F}_q \cup \{?\})^{\mu}$ and complete the remaining random variables of Z^{Mn} with "?". Recall that for $T \geqslant B$, we consider an eavesdropper who is also able to observe any cyclic interval of length μ in x_j^n. In the case where $T < B$, all n symbols of the chosen link j can be observed, that is $\mu = n$. According to Lemma 6.5 (c), related to our model, perfect security is achieved iff the submatrix G'_j of G for each $j = 1, \ldots, M$ is a $\mu \times n$ c-good matrix. The latter implies that each coset of C' has the same number of vectors from which Z^{Mn} can be obtained by $Mn - \mu$ erasures. This means, by construction of the secure nested M-link block code, we have perfect security, that is, $H(S^k|Z^{Mn}) = k$. Note that for $T \geqslant B$, the eavesdropper may observe any μ consecutive codeword symbols (also wrap-around) noiselessly from x_j^n, which implies that $n - \mu$ symbols of x_j^n are erased. Now suppose the eavesdropper observes packets in link j, e.g., $x[i, j], \ldots, x[i + \mu - 1, j]$, which is a $\mu \times n$ matrix (see Fig. 6.5). Consider every M-tuples of diagonals of length n containing entries of this matrix. In Fig. 6.5, an M-tuple of diagonals is underlined. Let S^k and Z^{Mn}, respectively, be the random variables that correspond to the source symbols and the eavesdropper's channel output of the block code which appears along the M-tuple diagonals. By construction of the block code we have that $H(S^k|Z^{Mn}) = k$, and this holds for every M-tuple of diagonals. Furthermore, note that for any time slot i and link j the symbols in packet $x[i, j]$ are equiprobable and any μ consecutive packets are mutually independent. This follows as the codewords are i.i.d. vectors.

For $i < 0$, by convention we choose $s[-1], \ldots, s[-T] = \mathbf{0}_{1\times k}$ and $e[-1], \ldots, e[-T]$, that correspond to i.i.d random variables over $\mathbb{F}_q^{\mu}$. The latter is necessary to provide perfect reliability and perfect security even if the eavesdropper observes any link j at time i.

W.l.o.g, we consider link j. For $T \geqslant B$ with $\mu < T + B \leqslant k$, the subvector $(x_{\mu+1}[0, j], \ldots, x_n[0, j])$ of $x[0, j]$ consists of the linear combination of the source symbols $(s_{\mu+1}[0, j], \ldots, s_g[0, j])$ with $g \leqslant k$ and the symbols of $e[-1], \ldots, e[-T]$. Suppose, the eavesdropper was observing a burst erasure of length $V \geqslant T$ before observing $x[0, j]$, then the encoder packets $e[-1], \ldots, e[-V]$ are unknown to him. By construction of the block codes described in Theorem 6.15, the source symbols $(s_{\mu+1}[0, j], \ldots, s_g[0, j])$ are secured by $e[-1], \ldots, e[-T]$. For the case where $V < T$, by construction of the block codes, the source symbols $(s_{\mu+1}[0, j], \ldots, s_g[0, j])$ are secured by $e[-1], \ldots, e[-V]$. Otherwise, the last $\mu - (n - T)$ rows at column-positions $\{\mu + 1, \ldots, n\}$ of G'_j would be zero, which would imply that the systematic matrix G'_j is not c-good.

Thus any μ consecutive packets observed by the eavesdropper in any link j do not reveal any information about the source symbols in those packets as well as in other observed packets, as long as they are separated by an interval of at least $V = n - \mu$ erased packets.

The latter implies that for $T \geqslant B$,

$$W_2 = \begin{cases} T+1 & \text{if } V = T \\ \mu+1 & \text{if } V < T \end{cases}. \tag{6.24}$$

■

For $T < B$ with $\mu = n$, each $x[i,j]$ is a linear combination of source packet $s[i]$ and encoder packet $e[i]$. Thus for $i < 0$, we can set any packet $e[i] = 0$ without violating the security condition.

Discussion 1. *Suppose the eavesdropper decides to change the link he observes. We assume that this operation costs the eavesdropper δ time units. To provide perfect security we have to choose $\delta = V$, since the eavesdropper could noiselessly observe the last μ packets before he jumps to the next link. When we allow the eavesdropper to switch between the links where he causes bursts of erasures, we have to assume that the lost time by changing the link is at least T, since we have to assume that the last packet has been erased before he jumps to another link. Thus we have to choose $\delta = \max\{T, V\}$ to be able to communicate with perfect reliability and perfect security; however there is loss in terms of the maximum secrecy rate R_s. For $T = \mu$ we have $\delta = V = T$, which implies that we can communicate with the maximum secrecy rate. Also note, that for the case $B = \mu$ we have $W_1 = W_2$. Note that for the case where $T < B$, due to the block code construction 4 in the achievability part, $\delta = 0$.*

6.5.3. Converse for the secrecy rate for streaming codes

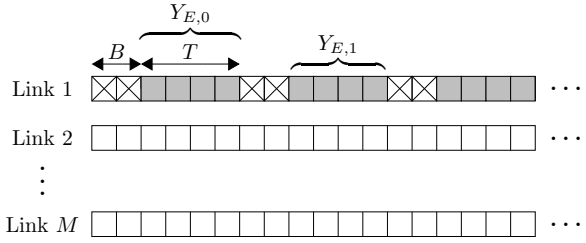


Figure 6.6.: The M-link channel used in proving the upper bound for $T \geqslant B$, with indication of which packets are observed by the eavesdropper $Y_{E,i}$ (gray squares). $Y_{(L\setminus E),i}$ is indicated by white squares. Crossed squares are erasures of length B.

In this section we provide the converse of Theorem 6.17. For $T \geqslant B$, we consider a periodic erasure channel in link 1 in the presence of an eavesdropper who can noiselessly observe μ packets in any sliding window of size

$$W_2 = \begin{cases} T+1 & \text{if } V = T \\ \mu+1 & \text{if } V < T \end{cases}, \tag{6.25}$$

where $V = n - \mu$, with $n = T + B$.

Let $I = \{1, \ldots, P\}_{j=1}^{M}$ and $I_j = \{1, \ldots, P\}$, $j = 1, \ldots, M$, respectively, be the index set for the packets in one period of length $P = lcm(V + \mu, T + B)$ and the index set for the packets in one period of length P in link j. Let $L \subset I$ and $E \subset I_j$, $j = 1, \ldots, M$ be the index sets of the revealed packets in the M-link channel and the eavesdropper's link, respectively. In the i-th, period $Y_{E,i}$ and $Y_{L,i}$ are, respectively, the observations at the eavesdropper and the legitimate receiver. Fig. 6.6 shows the time slots and the size of Y_E for the case when $V < T$ and $T = \mu$. Each erasure burst of length B is separated by T non-erased packets and the eavesdropper's observation of μ consecutive packets is separated by V erased packets.

For any integer $h \geqslant 1$ we use $Y_{L,0}^{h-1}$ to denote $Y_{L,0}, Y_{L,1}, \ldots, Y_{L,h-1}$. We require a coding scheme that provides perfect reliability with delay T and perfect security for each $h \geqslant 1$, that is,

$$H(S_0^{h-1}|Y_{L,0}^{h-1}, U_{L,h}) = 0, \quad \text{(perfect reliability)} \tag{6.26}$$

$$H(S_0^{h-1}|Y_{E,0}^{h-1}, U_{E,h}) = H(S_0^{h-1}), \quad \text{(perfect security)} \tag{6.27}$$

where $U_{L,h}$ and $U_{E,h}$ are, respectively, observations of the receiver and the eavesdropper in the interval of MT and T successive outcome packets within the h-th period, that is $|U_{L,h}| = MT$ and $|U_{E,h}| = T$. $S_i \in \mathbb{F}_q^{P \cdot k}$ and $H(S) = P \cdot k$. We assume that all source packets have the same entropy. Denote $W_{L,h} = Y_{L,0}^{h-1}, U_{L,h}$ and $W_{E,h} = Y_{E,0}^{h-1}, U_{E,h}$. Hence, for achieving (6.26) and (6.27), we have the following necessary condition:

$$\begin{aligned}
h \cdot P \cdot k &= H(S_0^{h-1}|W_{E,h}) - H(S_0^{h-1}|W_{L,h}) \\
&\leqslant H(S_0^{h-1}, W_{L,h}|W_{E,h}) - H(S_0^{h-1}|W_{L,h}) \\
&= H(W_{(L\backslash E),h}|W_{E,h}) + H(W_{E,h}|W_{E,h}, W_{(L\backslash E),h}) \\
&+ H(S_0^{h-1}|W_{L,h}, W_{E,h}) - H(S_0^{h-1}|W_{L,h}) \leqslant H(W_{(L\backslash E),h}) \\
&= H(Y_{(L\backslash E),0}^{h-1}, U_{(L\backslash E),h}) \leqslant hH(Y_{(L\backslash E),0}) + n(MT - \mu).
\end{aligned}$$

The latter implies that

$$R_s = \frac{k}{n} \leqslant \frac{ha + MT - \mu}{Ph} \overset{h\to\infty}{\to} \frac{a}{P}, \tag{6.28}$$

where $a = |(L\backslash E), 0|$, that is $H(Y_{(L\backslash E),0}) \leqslant na$.

For the case $T \geqslant B$, we have $|L| = (M-1)P + T$ and $|E| = \mu$ in one period.

Theorem 6.18. *For $T \geqslant B$, $V = T + B - \mu$ such that $P = T + B$, we obtain $H(Y_{(L\backslash E),0}) \leqslant [M(T + B) - B - \mu] \cdot n$, and hence $R_s \leqslant M - \frac{B+\mu}{T+B}$, which matches the secrecy rate in Theorem 6.17.*

For the case $T < B$, according to [35], if a code can decode all bursts of length B with delay $T < B$, then it can decode when any link is completely erased. Thus, when we choose $P = B = n$ we obtain $|L| = (M-1)B$ and $|E| = \mu$, where $\mu \leqslant n$.

Theorem 6.19. *For $T < B$, $V = B - \mu$ such that $P = B$, we obtain $H(Y_{(L\setminus E),0}) \leqslant [(M-1)B - \mu] \cdot n$, and hence $R_s \leqslant M - \frac{B+\mu}{B}$, which matches the secrecy rate in Theorem 6.17.*

6.6. Conclusion

For admissible parameters T, B, μ, M, Z, we constructed M-link codes over a small finite field $\mathbb{F}_q$ that can perfectly recover Z erasure bursts of length B in any sliding window of size $T+1$, each occurring on a separate link with minimum possible delay. In addition, the codes provide perfect security while the eavesdropper is observing an interval of at most μ packets in any sliding window of size W_2 (W_2 is case-dependent) or a copy of any link, i.e. $\mu = T + B$. The codes achieve the maximum secrecy rate for the channel models.

For $Z > 1$, it is worth mentioning that code constructions exist for a wider class of code parameters for the channel model if the positions of the bursts are the same in the corresponding Z links.

For future work it would be interesting to construct codes for the channel model where $Z = M$ and $Z < M - 1$ (perhaps using other methods of code construction). It would also be interesting to consider parallel burst-erasure wiretap channels, where on each link an eavesdropper is able to observe parts of the communication noiselessly.

Part III.

Modular Codes for the Wiretap Channel in the Finite Blocklength Regime

7. The Seeded Modular Code

7.1. Introduction

We consider a seeded modular code for the additive white Gaussian noise (AWGN) wiretap channel consisting of a security layer, an error-correction layer and a modulation layer. For reliable transmission, we use any forward error-correction (FEC) code and modulation method. In the security layer, a universal hash function (UHF) is used, which depends on a randomly chosen seed s. We consider three communication scenarios in which the advantage (the security measure) at the eavesdropper is measured in different ways. In the first two scenarios, the message distribution may be arbitrary, so these setups would be variants of "semantic security" in common terminology. In the third scenario, the advantage is measured under the assumption of a uniformly distributed message. This is usually referred to as "strong security". The eavesdropper uses the maximum likelihood (ML) test as an attack strategy. To assess the security performance, we derive the operational meaning of the advantages in terms of the error probability.

Contribution

We consider three communication scenarios in which the advantage at the eavesdropper is measured in different ways. Among them we consider the advantage at Eve under *distinguishing security* [1] [13]. The difference between semantic security and distinguishing security is that distinguishing security considers only the subclass of message distributions, whose support is a set of two equally probable messages. In [13] it is shown that distinguishing security is equivalent to semantic security asymptotically, but distinguishing security is easier to handle. That is, Eve observes a random vector Z^c for any message pair from the message set $\mathcal{M}$, and tries to identify to which message Z^c belongs. We analyze Eve's optimal attack strategy which is the maximum likelihood (ML) test. In the first two scenarios we interpret Eve as active in the sense that she can choose the message pair to be transmitted.

In preparation for the simulations, we consider different security metrics in each communication scenario and show some relevant relationships between the advantages of the three scenarios. Among them we extend the proof of Bellare et al. [20], which shows that strong security implies semantic security, to the AWGN case with BPSK or QPSK

[1] An early instance of distinguishability is used in [51].

input. Furthermore, we discuss the operational meanings of the distinguishing security and the strong security. They amount to Eve performing an optimal ML test for every message pair. For this purpose, we introduce a new strong security metric to compare with the distinguishing security. Furthermore, we have made a working hypothesis; increasing the average Hamming distance of a coset pair improves Eve's performance and thus increases her advantage. In Chapter 8 the simulations indicate that our hypothesis might be correct.

Related Work

In information theory, UHF's were first studied by Bennett et al. [52]. Hayashi [53] proposed using the UHF as a technique for wiretap coding. In [54] and in [13], it is shown that with a modular UHF scheme a variant of semantic security - where Eve can choose the message distribution after getting to know the seed - can be achieved if the wiretap channel is discrete, degraded and symmetric. Furthermore, it is shown that the modular UHF scheme achieves secrecy capacity under semantic security in this case. In the case of an additive white Gaussian noise (AWGN) channel, it is shown in [55] that strong secrecy capacity can be achieved. In [56], a special UHF is used as the security component, by which a variant of semantic security is achievable - where the message distribution is arbitrary but independent of the seed - for arbitrary discrete memoryless wiretap channels. Similar to [56], it can be shown that the modular UHF scheme is semantically secure for the Gaussian channel. The seed it requires is longer than that needed by, e.g., the function in [55] and in [13]. In [57], a novel type of functions called biregular irreducible (BRI) functions is introduced and applied as security components (instead of, e.g., universal hash functions) in seeded modular wiretap coding schemes. In [57] it is shown that semantic security can be achieved for a discrete and Gaussian wiretap channel by using BRI functions. During the preparation of this work, efficiently computable BRI functions were constructed [58].

Previous works have already implemented and analyzed codes for the wiretap channel, such as in [59], [60]. In [59], the performance of LDPC codes for the Gaussian wiretap channel under strong security was analyzed. In [60], additional inner coding layers were created that generate a discrete memoryless channel (DMC) for Eve and Bob so that the outer wiretap code already available in [61], [46] can be used. However, Eve is required to process the channel output before security is evaluated. Moreover, strong security is shown heuristically for the three layer coding scheme, and the migration effort in existing systems is high compared to the proposed seeded modular coding scheme. In [61] and in [62], alternative concepts to the modular scheme are presented.

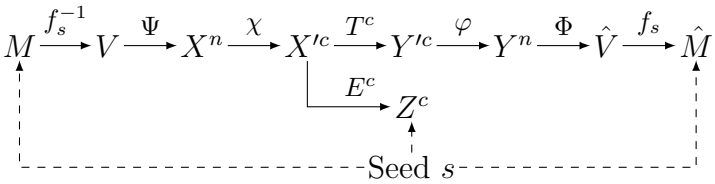


Figure 7.1.: A seeded modular code for the wiretap channel (T^c, E^c), where n is the codewordlength of the FEC code and c is the blocklength.

Outline

Section 7.2 introduces notation and provides the preliminary background about the AWGN wiretap channel, as well as the AWGN wiretap code. Furthermore, we define the security metrics if an unseeded modular scheme is used. In Section 7.3, we describe the seeded modular UHF coding scheme for the AWGN wiretap channel and the explicit construction of the code. We introduce three communication scenarios and define the corresponding security metrics. Then, we consider the relationships between the security metrics. Section 7.4 gives the operational meaning of the advantage at Eve under distinguishing security and under strong security. Then we discuss how to maximize Eve's performance to simulate a worst case scenario. Section 7.5 concludes with discussion and open problems.

7.2. Preliminary

7.2.1. Notations

Throughout the paper, we write $X \sim unif(\mathcal{X})$ to denote that X is a uniform random variable over some discrete set $\mathcal{X}$. The logarithm log and the exponential function exp will always be taken to base 2. We denote by 0^k a zero vector of length k. The operation $[\cdot]_k$ selects the k most significant bits and $(\cdot||\cdot)$ denotes the concatenation of two vectors. The statistical difference between X_1 and X_2 is defined by $\|P_{X_1} - P_{X_2}\| = \frac{1}{2}\int_{\mathcal{X}} |p_{X_1}(x) - p_{X_2}(x)|\, dx$.

7.2.2. The AWGN Wiretap Channel

The goal of Alice is to communicate a message $M \in \mathcal{M}$ of length $\log|\mathcal{M}|$ to Bob (w.l.o.g. we assume that M is a binary sequence), which is distributed according to P_M. Alice performs this task by encoding M to a vector $X'^c \in \mathcal{X}'^c$ of length c and transmitting X'^c. For the case where the channels are AWGN, we have

$$Y' = X' + N_T, \qquad Z = X' + N_E, \tag{7.1}$$

where $X' \in \mathcal{X}' \subset \mathbb{C}$ and N_T and N_E are Gaussian noises of Bob's channel T and Eve's channel E, respectively. We assume that N_T and N_E are circularly symmetric accord-

ing to $\mathcal{CN}(0, 2\sigma_T^2)$ and $\mathcal{CN}(0, 2\sigma_E^2)$, respectively. Since the channel is memoryless, we respectively obtain after c channel uses $Y'^c = X'^c + N_T^c$ and $Z^c = X'^c + N_E^c$, where N_T^c and N_E^c are white Gaussian noises, respectively. Alice is subject to a transmission power restriction P. The encoding of a message M by Alice should be such that Bob is able to decode M *reliably*, and using the appropriate security metric, Z^c should give Eve as little advantage as possible about M.

7.2.3. The AWGN Wiretap Code and the Security Metrics

Definition 7.1. *An (ξ, ζ)-Code $\mathcal{C}_c$ for the AWGN wiretap channel (T^c, E^c) consists of a stochastic encoder at the transmitter*

$$\xi : \mathcal{M} \to \mathcal{X}'^c, \tag{7.2}$$

and a decoder at the legitimate receiver

$$\zeta : \mathcal{Y}'^c \to \mathcal{M}. \tag{7.3}$$

The maximum probability that the decoding fails is

$$\hat{P}_e(\mathcal{C}_c) = \max_{m \in \mathcal{M}} Pr\left((\zeta \circ T^c \circ \xi)(m) \neq m\right), \tag{7.4}$$

where $\zeta \circ T^c \circ \xi$ denotes the concatenation of ζ, the channels T^c and ξ.

We wish $\hat{P}_e(\mathcal{C}_c)$ to be small, then the transmission of messages through T^c applying the wiretap code $\mathcal{C}_c$ is close to noiseless.

At the same time, Eve observing the output of E^c should learn as little as possible about the message M, that is, we require the advantage to be close to zero. Let $E^c(\xi(M))$ be the channel output at Eve when message M was sent and encoded with ξ, so that $Z^c(M) = E^c(\xi(M))$.

We adopt the definitions (7.5) - (7.8) from [13]. The advantage at the eavesdropper under semantic security (SS) is defined as follows:

Let h be a function defined on the message set $\mathcal{M}$, so that $h(M)$ is the image, then

$$Adv^{SS}(\xi; E^c) = \max_{h,M}\big(\max_{\mathcal{A}} Pr(\mathcal{A}(Z^c(M)) = h(M)) - \max_{\mathcal{G}} Pr(\mathcal{G}(k) = h(M))\big), \tag{7.5}$$

where $\mathcal{A}$ is the attack strategy of the eavesdropper, $\mathcal{G}$ is any simulator that has knowledge of the length of the message k and the implicit knowledge of h and M.

The advantage at the eavesdropper under distinguishing security (DS) is defined as

follows:

$$\begin{aligned} Adv^{DS}(\xi; E^c) &= \max_{\mathcal{A},m_1,m_2} && 2Pr(\mathcal{A}(m_1, m_2, Z^c(m_B)) = B) - 1 \\ &= \max_{m_1,m_2} && \frac{1}{2}\int_{\mathcal{Z}^c} |p(z^c|m_1) - p(z^c|m_2)|\, dz^c, \end{aligned} \tag{7.6}$$

where B is uniformly distributed over $\{1, 2\}$ and the maximum is over all messages m_1, m_2 and all $\{1, 2\}$-valued eavesdropper strategies $\mathcal{A}$.

Furthermore, we consider the advantage under the mutual information security (MIS):

$$Adv^{MIS}(\xi; E^c) = \max_{M} I(M; Z^c(M)), \tag{7.7}$$

where the maximum is over all random variables $M \in \mathcal{M}$ and M is assumed to be distributed arbitrarily.

In information theory the more common security metric is mutual information security for random messages ($MIS - R$) also known as strong security:

$$Adv^{MIS-R}(\xi; E^c) = I(M; Z^c(M)), \tag{7.8}$$

where $M \sim unif\{\mathcal{M}\}$.

In addition, we define the advantage under average distinguishing security ($DS - R$):

$$Adv^{DS-R}(\xi; E^c) = \frac{1}{|\mathcal{M}|^2} \sum_{m_1,m_2 \in \mathcal{M}} \frac{1}{2}\int_{\mathcal{Z}^c} |p(z^c|m_1) - p(z^c|m_2)| dz^c. \tag{7.9}$$

A common interpretation of Adv^{xs} is that the channel of Eve has d_{xs} bits of xs-security if $Adv^{xs} \leqslant 2^{-d_{xs}}$. We call d_{xs} the security level.

Let $\mathcal{C} = \left\{\mathcal{C}_{c(g)}\right\}_{g\in\mathbb{N}}$ be a sequence of AWGN wiretap codes for Bob's channel with blocklength $c(g)$, where c is a monotonically increasing function of g. We assume that the channel input fulfills the average power constraint, that is, $\frac{1}{c(g)}\sum_{i=1}^{c(g)} |x_i'|^2 \leqslant P$. Note that we consider the case where the modulation alphabet can change with the blocklength.

We call $\mathcal{C}$ an AWGN wiretap coding scheme, and xs-secure if the scheme fulfills the properties of the following definition.

Definition 7.2 (Achievable Asymptotic Secrecy Rate). *A non-negative real number R_{sec} is called an achievable asymptotic secrecy rate under xs-security if there exists a strictly increasing sequence $\{c(g)\}_{g\geqslant 1}$ and a sequence $\left\{\mathcal{C}_{c(g)}\right\}_{g\in\mathbb{N}} = (\left\{\xi_{c(g)}\right\}_{g\in\mathbb{N}}, \left\{\zeta_{c(g)}\right\}_{g\in\mathbb{N}})$*

of wiretap codes, where $\mathcal{C}_{c(g)}$ complies with the average transmit power P, such that

$$\lim_{g\to\infty} \frac{1}{c(g)} \log \left|\mathcal{M}_{c(g)}\right| \geqslant R_{sec},$$
$$\lim_{g\to\infty} Adv^{xs}(\xi_{c(g)}; E^{c(g)}) = 0,$$
$$\lim_{g\to\infty} P_e(\mathcal{C}_{c(g)}) = 0.$$

The supremum of all achievable asymptotic secrecy rates under xs-security is called the xs secrecy capacity of the AWGN wiretap channel. The xs secrecy capacity is given as follows.

Proposition 7.1. *The xs secrecy capacity of the AWGN wiretap channel is for all $xs \in \{SS, DS, MIS, DS-R, MIS-R\}$*

$$C_s(\sigma_T^2, \sigma_E^2, P) = \begin{cases} C_T(\sigma_T^2, P) - C_E(\sigma_E^2, P) & \sigma_T^2 \leqslant \sigma_E^2 \\ 0 & otherwise, \end{cases} \tag{7.10}$$

where $C_T(\sigma_T^2, P) = \log(1 + \frac{P}{2\sigma_T^2})$ and $C_E(\sigma_E^2, P) = \log(1 + \frac{P}{2\sigma_E^2})$.

This was shown in [63] for the case of strong security, i.e., for the case where $xs \in \{MIS-R, DS-R\}$. In [64], it is shown that the secrecy capacity is given by $C_s(\sigma_T^2, \sigma_E^2, P)$ if the message may have an arbitrary distribution, i.e., for the case where $xs \in \{MIS, SS, DS\}$.

Remark 7.2. *According to Definition 7.2, it is possible that the sequence of codes $\mathcal{C}$ is defined for a subsequence of the set of blocklengths, and analysis of the converse proofs shows that the achievable secrecy rate is not increased compared with the secrecy rates of the common definition, where the sequence of codes attains all blocklengths.*

Next, we consider the equivalences[1] of the security metrics given above.

The following relationship between SS-security and DS-security for the discrete wiretap channel setup is given in [13, Theorem 4.1], and can be extended to the Gaussian setup.

$$Adv^{SS}(\xi; E^c) \leqslant Adv^{DS}(\xi; E^c) \leqslant 2Adv^{SS}(\xi; E^c). \tag{7.11}$$

Thus, distinguishing security is equivalent to semantic security asymptotically.

Furthermore, according to [62, Proposition 1] we have

$$Adv^{MIS}(\xi; E^c) \leqslant 2Adv^{DS}(\xi; E^c) \log \frac{|\mathcal{M}|}{2Adv^{DS}(\xi; E^c)},$$
$$Adv^{DS}(\xi; E^c) \leqslant 2\sqrt{2Adv^{MIS}(\xi; E^c)}. \tag{7.12}$$

[1] We call two security measures equivalent if one security measure approaches 0 if and only if the other approaches 0.

If $Adv^{DS}(\xi; E^c)$ decreases exponentially with c, then DS-security implies MIS-security.

The relationship between strong security and average distinguishing security is given in the following proposition.

Proposition 7.3.

$$Adv^{MIS-R}(\xi; E^c) \leqslant 2Adv^{DS-R}(\xi; E^c) \log \frac{|\mathcal{M}|}{2Adv^{DS-R}(\xi; E^c)}. \tag{7.13}$$

Proof. We use the upper bound by [62, Appendix I], so that

$$Adv^{MIS-R}(\xi; E^c) \leqslant 2 \int_{z^c \in \mathcal{Z}^c} \gamma \log |\mathcal{M}| d\mu - \big(2 \int_{z^c \in \mathcal{Z}^c} \gamma d\mu\big) \log \big(2 \int_{z^c \in \mathcal{Z}^c} \gamma d\mu\big),$$

where $\gamma = \frac{1}{2} \sum_{m \in \mathcal{M}} \big|P_M(m) - P_{M|z^c}(m)\big|$ and $d\mu = p(z^c)dz^c$ is the probability measure associated to Z^c. Observe that,

$$\|P_{Z^c,M} - P_{Z^c} P_M\| = \int_{z^c \in \mathcal{Z}^c} \gamma d\mu, \tag{7.14}$$

where

$$\begin{aligned}
&\|P_{Z^c,M} - P_{Z^c} P_M\| \\
&= \sum_{m \in \mathcal{M}} P_M(m) \frac{1}{2} \int_{\mathcal{Z}^c} |p(z^c|m) - p(z^c)| dz^c \\
&= \frac{1}{|\mathcal{M}|} \sum_{m_1 \in \mathcal{M}} \frac{1}{2} \int_{\mathcal{Z}^c} |p(z^c|m_1) - \frac{1}{|\mathcal{M}|} \sum_{m_2 \in \mathcal{M}} p(z^c|m_2)| dz^c \\
&= \frac{1}{|\mathcal{M}|} \sum_{m_1 \in \mathcal{M}} \frac{1}{2} \int_{\mathcal{Z}^c} |\frac{1}{|\mathcal{M}|} \sum_{m_2 \in \mathcal{M}} \big(p(z^c|m_1) - p(z^c|m_2)\big)| dz^c \\
&\leqslant \frac{1}{|\mathcal{M}|^2} \sum_{m_1 \in \mathcal{M}} \sum_{m_2 \in \mathcal{M}} \frac{1}{2} \int_{\mathcal{Z}^c} |p(z^c|m_1) - p(z^c|m_2)| dz^c \\
&= Adv^{DS-R}(\xi; E^c).
\end{aligned} \tag{7.15}$$

■

7.3. The Seeded Modular UHF Code

Fig. 7.1 shows a seeded modular wiretap code for the wiretap channel. We suppose, that both, the channel of Bob and the channel of Eve are AWGN. We assume that Eve's SNR is smaller than Bob's SNR and that all participants have seed s, e.g., because of access to sufficient common randomness. A possible scenario is when Alice transmits the seed and the message in succession. Furthermore, the seed is chosen according to a random variable S which is uniformly distributed over a finite set $\mathcal{S}$. We consider the seeded modular UHF code (ξ, ζ) for the AWGN wiretap channel (T^c, E^c) that consists of a stochastic seeded encoder at Alice $\xi : \mathcal{S} \times \mathcal{M} \to \mathcal{X}'^c$, and a seeded decoder at Bob $\zeta : \mathcal{S} \times \mathcal{Y}'^c \to \mathcal{M}$. The

code consists of the following three layers; the modulation layer (χ, φ), the error-correction layer (Ψ, Φ) and the security layer (f_s^{-1}, f_s), so that $\xi : \chi \circ \Psi \circ f_s^{-1}$ and $\zeta = f_s \circ \Phi \circ \varphi$, where χ, φ are the modulation and demodulation functions, respectively, and Ψ, Φ the encoding and decoding functions of the linear FEC code, respectively. Functions f_s and f_s^{-1} are defined below. Bob's goal is to decode the message m correctly. First, he demodulates the noisy version $y'^c = x'^c + N_T^c$ of the modulated codeword x'^c by $\varphi : y'^c \mapsto y^n$ and then decodes $\hat{v} = \Phi(y^n)$. Finally, Bob decodes the message $\hat{m}$ as $\hat{m} = f_s(\hat{v})$.

In the following, we consider binary FEC codes, but the code is not limited to a binary alphabet. Denote by $\mathbb{F}_{2^l}$ the finite field with 2^l elements, $\mathbb{F}_{2^l}^* = \mathbb{F}_{2^l} \setminus \{0\}^l$, and let $*$ and $\oplus$ denote multiplication and addition in $\mathbb{F}_{2^l}$, respectively.

Security Layer: For two sets of $\mathcal{V} = \{0,1\}^l$ and $\mathcal{M} = \{0,1\}^k$, we use a family of UHF's $\mathcal{F} = \left\{ f_s : \{0,1\}^l \to \{0,1\}^k \,|s \in \mathcal{S} \right\}$, so by definition

$$|\{s \in \mathcal{S} | f_s(v_1) = f_s(v_2)\}| \leqslant \frac{|\mathcal{S}|}{2^k} \tag{7.16}$$

for every $v_1 \neq v_2 \in \{0,1\}^l$. Alice encodes the message $m \in \{0,1\}^k$ by using the randomized inverse $f_s^{-1}(m)$, which uniformly at random picks an element v of the set $\{v' : f_s(v') = m\}$.

We consider the following two families of UHF's:

1)

$$\mathcal{F}_1 = \left\{ f_{a,t} : \{0,1\}^l \to \{0,1\}^k \,|a \in \mathbb{F}_{2^l}^*, t \in \mathbb{F}_{2^l} \right\}, \tag{7.17}$$

where $f_{a,t}(v) = [(a * v) \oplus t]_k$ and $s = (a,t) \in \mathcal{S}$.

Accordingly, for some random vector $R \sim unif(\{0,1\}^{l-k})$, the randomized inverse is

$$f_s^{-1}(m) = a^{-1} * ((m||R) \oplus t), \tag{7.18}$$

where $((m||R) \oplus t) \in \{0,1\}^l$.

Next, we show that $\mathcal{F}_1$ is a family of UHF's, and thus that for every $v_1 \neq v_2 \in \{0,1\}^l$, (7.16) is fulfilled. Wegman and Carter [65] proved Proposition 7.4 for the case of finite fields $\mathbb{F}_p$ with p prime.

Proposition 7.4. *$\mathcal{F}_1$ defined in (7.17) is a family of UHF's.*

Proof. For any given $v_1 \neq v_2 \in \{0,1\}^l$ we have to count how many seeds satisfy $f_s(v_1) = f_s(v_2)$ and thus satisfy $[(a * v_1) + t]_k = [(a * v_2) + t]_k$. We can reformulate the equation to $[(a * v_1)]_k + [t]_k = [(a * v_2)]_k + [t]_k$, so that it remains to count how many a satisfy $0^k = [a * v_1]_k + [a * v_2]_k = [a * v']_k = [a * (m||r)]_k$, where $v' = v_1 + v_2 = (m||r)$. Since $a \neq 0^l$, there is a unique value $a * v'$. If we fix m, then there are $2^{l-k} - 1$ choices of a to obtain $0^k = [a * (m||r)]_k$, and we have 2^l choices for t. We obtain $2^{-k}2^l(2^l - 2^k)$ choices for (a,t), where $2^{-k}2^l(2^l - 2^k) \leqslant \frac{|\mathcal{S}|}{2^k}$. ∎

2)

$$\mathcal{F}_2 = \left\{ f_s : \{0,1\}^l \to \{0,1\}^k \,|s \in \mathbb{F}_{2^l}^* \right\}, \tag{7.19}$$

where $f_s(v) = [s * v]_k$, and the randomized inverse is

$$f_s^{-1}(m) = s^{-1} * (m||R). \tag{7.20}$$

Remark 7.5. *Proposition 7.4 is also valid for (7.19). Furthermore, for a restricted message set, the functions are BRI functions [57].*

Error-Correction Layer: In the error-correction layer, Alice encodes v using some FEC code (Ψ, Φ) of rate $R_{FEC} = l/n$, so that

$$\Psi(v) = x^n = vG, \tag{7.21}$$

where G is the $l \times n$ generator matrix of the FEC code.

Modulation Layer: We consider BPSK and QAM. We denote the corresponding symbol alphabet by $\mathcal{X}' \subset \mathbb{C}$. It has size $2^{R_{\text{mod}}}$, where R_{mod} denotes the number of bits per symbol. In the modulation layer, Alice modulates x^n to x'^c using a modulation scheme (χ, φ), where $c = n/R_{mod}$. In order to satisfy the transmit power constraint, we choose it in such a way that

$$P_{av} = \frac{1}{2^{R_{mod}}} \sum_{x' \in \mathcal{X}'} |x'|^2. \tag{7.22}$$

The product of R_{FEC} and R_{mod} gives the effective rate

$$R_{eff} = l/c, \tag{7.23}$$

which is the rate of Bob's channel. The secrecy rate R_{sec} is defined as

$$\frac{k}{c} = \frac{k}{l} R_{eff}. \tag{7.24}$$

Note that the blocklength is c.

7.3.1. Communication Scenarios

In the following, we introduce three communication scenarios where in analogy to Section 7.2.3, we define the decoding error probability of Bob and the corresponding seeded advantages at Eve under xs-security, where S is taken as additional knowledge of Eve. For a better overview, we provide the security metrics with a number that indicates the communication scenario.

The achievable asymptotic secrecy rates R_{sec} under xs-security of the seeded modular UHF coding scheme for the AWGN wiretap channels can be defined analogous to Section 7.2.3 and are given in (7.10). For the case with seed, no converse is known.

Communication scenario 1: This is a communication scenario as considered by Bellare and Tessaro in [66], where Eve can actively affect the selection of messages. In this scenario, $xs \in \{DS_1, MIS_1\}$. If $xs = DS_1$, first a randomly chosen seed $S \in \mathcal{S}$ is given, and then Eve chooses a message pair (m_1, m_2), where $m_1, m_2 \in \{0,1\}^k$, so that the choice of (m_1, m_2) depends on the choice of the seed. If $xs = MIS_1$, the process is the same except that Eve chooses the conditional message distribution $P_{M|S}$.

The error probability of the seeded modular code (ξ, ζ) is then defined as

$$P_e(\xi, T^c, \zeta) = \mathsf{E}_S \max_m Pr\left(\zeta(Y'^c(S,m))) \neq m\right). \tag{7.25}$$

Since $f_s(f_s^{-1}(m)) = m$, the error depends only on the FEC code, the modulation mapping and the channel.

We choose the FEC code and the modulation scheme so that $P_e(\xi, T^c, \zeta)$ is sufficiently small (e.g. $P_e(\xi, T^c, \zeta) < 10^{-4}$). For the following scenarios, the error probabilities are adjusted to Bob analogously to the advantages.

The seeded advantage at Eve under DS_1-security is defined in [66] as follows:

$$\begin{aligned} Adv^{DS_1}(\xi; E^c; \mathcal{S}) &= \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \max_{\mathcal{A}, m_1, m_2} 2Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_B, s)) = B) - 1 \\ &= \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \max_{m_1, m_2} \frac{1}{2} \int_{\mathcal{Z}^c} |p(z^c|m_1, s) - p(z^c|m_2, s)| \, dz^c \\ &= \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \max_{m_1, m_2} ||P_{Z^c|M=m_1, S=s} - P_{Z^c|M=m_2, S=s}||, \end{aligned} \tag{7.26}$$

where B is uniformly distributed over $\{1,2\}$ and the maximum is over all $\{1,2\}$-valued eavesdropper strategies $\mathcal{A}$ and k-bit messages m_1, m_2. Note, if the advantage is small, then the probability of a seed appearing that is favorable for Eve is also small.

The seeded advantage at Eve under MIS_1-security:

$$Adv^{MIS_1}(\xi; E^c; \mathcal{S}) = \max_{P_{M|S}} I(M; Z^c(M,S)|S). \tag{7.27}$$

Since Eve knows $P_{M|S}$ and thus $I(M;S)$, the advantage depends on the right term of (7.27) only.

Communication scenario 1a: A possible similar scenario that we consider in Section 8.2 is the case where for a fixed channel and given code parameters, we choose a specific seed from the seed set $\mathcal{S}$. This corresponds to the case of unseeded encryption. The advantage

under $DS_{1a)}$-security given a specific seed s is

$$\begin{aligned} Adv^{DS_{1a)}}(\xi; E^c; s) &= \max_{\mathcal{A},m_1,m_2} 2Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_B, s)) = B) - 1 \\ &= \max_{m_1,m_2} \frac{1}{2} \int_{\mathcal{Z}^c} |p(z^c|m_1, s) - p(z^c|m_2, s)| \, dz^c \\ &= \max_{m_1,m_2} ||P_{Z^c|M=m_1,S=s} - P_{Z^c|M=m_2,S=s}||. \end{aligned} \tag{7.28}$$

We are interested in seeds that are unfavorable for Eve. In Section 8.2 we are looking for such seeds. This scenario would drastically reduce the complexity of the modular UHF scheme because then the seed must only be made public once.

Remark 7.6. *The security measures (7.5) - (7.9) and the equivalences from 7.2.3 for the case without seed immediately apply to scenario 1a).*

Communication Scenario 2: Depending on whether DS_2 or MIS_2 is considered, Eve first chooses a message pair (m_1, m_2) or P_M that is beneficial for her, and only then is $S \in \mathcal{S}$ randomly chosen, so that the message and the seed are independent.

The seeded advantage at Eve under DS_2-security:

$$\begin{aligned} Adv^{DS_2}(\xi; E^c; \mathcal{S}) &= \max_{\mathcal{A},m_1,m_2} 2Pr(\mathcal{A}(S, m_1, m_2, Z^c(m_B, S)) = B) - 1 \\ &= \max_{m_1,m_2} \frac{1}{2} \int_{\mathcal{Z}^c} \frac{1}{|\mathcal{S}|} \sum_{s\in\mathcal{S}} |p(z^c|m_1, s) - p(z^c|m_2, s)| \, dz^c \\ &= \max_{m_1,m_2} \frac{1}{|\mathcal{S}|} ||P_{Z^c|M=m_1,S=s} - P_{Z^c|M=m_2,S=s}||, \end{aligned} \tag{7.29}$$

where B, m_1, m_2 and $\mathcal{A}$ are as in (7.26).

The seeded advantage at Eve under MIS_2-security:

$$Adv^{MIS_2}(\xi; E^c; \mathcal{S}) = \max_{P_M} I(M; Z^c(M, S), S). \tag{7.30}$$

Next, we will see that when the message is chosen independently of the seed, and using (7.17), proposed by Hayashi [56], as the security component in the seeded modular scheme, MIS_2-security can be achieved with positive secrecy rate, i.e. $Adv^{MIS_2}(\xi; E^c; \mathcal{S})$ tends to 0. See also the next subsection. The following upper bound of the advantage at Eve under MIS_2-security has been first proven in a more general context in [56], in terms of the conditional Renyi entropy. In [57] the result was extended, and is given in terms of the smooth Renyi divergence.

Proposition 7.7. *Given the family of UHF's $\{f_s : s \in \mathcal{S}\}$ as proposed in (7.17), [56, Lemma 21] implies*

$$Adv^{MIS_2}(\xi; E^c; \mathcal{S}) \leqslant \frac{1}{\ln 2} 2^{-c(R_{eff} - R_{sec} - (C_E(\sigma_E^2, P) + \delta))} + \epsilon(\delta, c) c R_{sec}, \tag{7.31}$$

where $\epsilon(\delta, c)c \to 0$ *for* $\delta \to 0$ *and* $c \to \infty$, *and* R_{sec} *and* R_{eff} *are given in* (7.24) *and* (7.23), *respectively.*

Proof. This follows immediately from [56, Lemma 21] and the discussion in [57]. ∎

As a consequence of (7.31) we obtain the following corollary. Let $(\cdot)^+ = \max(\cdot, 0)$.

Corollary 7.8. *Using an FEC code and a modulation scheme, and using the family of UHF's given in* (7.17), *the seeded modular UHF wiretap scheme* $\mathcal{C}$ *can achieve all secrecy rates* R_{sec} *satisfying*

$$R_{sec} < (R_{eff} - (C_E(\sigma_E^2, P) + \delta))^+ \tag{7.32}$$

with MIS_2*-security.*

Proof. For $c \to \infty$ the right hand side of (7.31) should tend to zero. Therefore, we require $\lim_{c\to\infty}(R_{eff} - R_{sec} - (C_E(\sigma_E^2, P) + \delta)) > 0$. Furthermore, for any δ, one can choose $\epsilon(\delta, c)$ such that $\epsilon(\delta, c)c \to 0$ as $c \to \infty$, then $\lim_{c\to\infty} \epsilon(\delta, c)cR_{sec} = 0$. Since δ can be chosen arbitrarily small (but constant), one achieves any rate smaller than the right side of (7.32). ∎

This implies that if R_{eff} can be arbitrarily close to the channel capacity $C_T(\sigma_T^2, P)$, the seeded modular UHF wiretap scheme can achieve the secrecy capacity under MIS_2-security. Here we refer to Proposition 7.1.

The inequality (7.31) says that $Adv^{MIS_2}(\xi; E^c; \mathcal{S}) \leqslant 2^{-d_{MIS_2}}$ if

$$R_{sec} \approx R_{eff} - \log\left(1 + \frac{P}{2\sigma_E^2}\right) - \frac{d_{MIS_2}}{c}. \tag{7.33}$$

To achieve a certain security level d_{MIS_2} at a given R_{sec}, R_{eff} and $P = 2\sigma_T^2 SNR_B$, the following approximately applies

$$d_{MIS_2} \geqslant l - k - c\log\left(1 + \frac{P}{2\sigma_E^2}\right). \tag{7.34}$$

Under the same conditions as in Proposition 7.7, we can derive a bound for DS_2-security.

Proposition 7.9. *Given the family of UHF's* $\{f_s : s \in \mathcal{S}\}$ *as proposed in* (7.17), *the upper bound of* $Adv^{DS_2}(\xi; E^c; S)$ *is*

$$Adv^{DS_2}(\xi; E^c; \mathcal{S}) \leqslant 4\sqrt{2^{-c(R_{eff} - R_{sec} - (C_E(\sigma_E^2, P) + \delta))}} + 2\epsilon(\delta, c), \tag{7.35}$$

where $\epsilon(\delta, c) \to 0$ *for* $\delta \to 0$ *and* $c \to \infty$.

Corollary 7.8 for DS_2-security also applies here.

The inequality (7.35) says that $Adv^{DS_2}(\xi; E^c; \mathcal{S}) \leqslant 2^{-d_{DS_2}}$ if

$$R_{sec}(d_{DS_2}) \approx R_{eff} - C_E(\sigma_E^2, P) - \frac{2d_{DS_2} + 4}{c}. \tag{7.36}$$

To achieve a certain security level d_{DS_2} at a given R'_{sec}, R_{eff} and $P = 2\sigma_T^2 SNR_B$, we approximately obtain

$$d_{DS_2} \geqslant \frac{1}{2}\left(l - k - c\log\left(1 + \frac{P}{2\sigma_E^2}\right)\right) - 2. \tag{7.37}$$

Remark 7.10. *Since in scenario 2 Eve has to choose the message pair to be independent of the seed* $Adv^{DS_2}(\xi; E^c; \mathcal{S}) \leqslant Adv^{DS_1}(\xi; E^c; \mathcal{S})$*, and thus we can upper bound* d_{DS_1} *by* d_{DS_2}*. In Section 8.2, we use the right-hand side of (7.37) to determine the amount of encoding randomness.*

Remark 7.11. *The estimations in Proposition 7.7 and 7.9 are valid for sufficiently large* c *but we use it because of its simple form. Furthermore, the actual security parameters are found by simulations anyway.*

Communication scenario 3: We consider a communication scenario where the seed $S \in \mathcal{S}$ and the message $M \sim unif(\{0,1\}^k)$ are randomly chosen independently.

We introduce the seeded advantage at Eve under DS_3-security:

$$\begin{aligned} Adv^{DS_3}(\xi; E^c; \mathcal{S}) &= \frac{1}{|\mathcal{S}|\, 2^{2k}} \sum_{s\in\mathcal{S}} \sum_{\substack{m_1,m_2 \\ \in\{0,1\}^k}} \frac{1}{2} \int_{\mathcal{Z}^c} |p(z^c|m_1, s) - p(z^c|m_2, s)| dz^c \\ &= \frac{1}{|\mathcal{S}|\, 2^{2k}} \sum_{s\in\mathcal{S}} \sum_{\substack{m_1,m_2 \\ \in\{0,1\}^k}} ||P_{Z^c|M=m_1,S=s} - P_{Z^c|M=m_2,S=s}|| \\ &= \frac{2}{|\mathcal{S}|\, 2^{2k}} \sum_{s\in\mathcal{S}} \sum_{\substack{m_1,m_2 \\ \in\{0,1\}^k}} \max_{\mathcal{A}} Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_B, s)) = B) - 1, \end{aligned} \tag{7.38}$$

where B and $\mathcal{A}$ are as in (7.26). The uniform distribution of the message in this scenario is reflected in the averaging over message pairs in (7.38). Adv^{DS_3} is a measure of strong security that we consider in Section 8.2. A similar form already appears in [18].

The seeded advantage at Eve under MIS_3-security is given as follows:

$$Adv^{MIS_3}(\xi; E^c; S) = I(M; Z^c(M, S), S). \tag{7.39}$$

Since the communication scenario 2 is difficult to simulate (because the first thing to do is to find a message pair that maximizes the advantage), Section 8.2 considers mainly the communication scenarios 1, 1a) and 3, namely the DS_1-security, $DS_{1a)}$-security and the DS_3-security.

7.3.2. Relations between the Security Metrics

MIS_3-Security Implies MIS_2-Security

To determine $Adv^{DS_2}(\xi; E^c; \mathcal{S})$ and $Adv^{DS_1}(\xi; E^c; \mathcal{S})$ it is necessary to maximize over a message pair, which raises problems. We therefore also consider $Adv^{MIS_3}(\xi; E^c; \mathcal{S})$ and extend the proof of Bellare et al. [20] to the AWGN case. This shows that when the channel is a binary input AWGN channel, the FEC code is linear and when $f_s^{-1}(m)$ is given as in (7.20) then $Adv^{MIS_2}(\xi; E^c; \mathcal{S})$ decreases if $Adv^{MIS_3}(\xi; E^c; \mathcal{S})$ decreases. We want to mention that for the unseeded case, an alternative proof exists in [67] that shows the relationship when the channel is symmetric and the universal hash function is linear.

Consider an AWGN channel with zero mean and variance σ_E^2. Suppose we use BPSK with $\mathcal{X}' = \{-a, a\}$ and a uniform quantizer that maps Z to the nearest value in the set $\hat{\mathcal{Z}} = \left\{-L + \frac{1}{2^L}, -L + \frac{2}{2^L}, \ldots, L - \frac{2}{2^L}, L - \frac{1}{2^L}\right\}$, so that $\hat{Z} \in \hat{\mathcal{Z}}$. Furthermore, L can become arbitrarily large, so that the Gaussian density function can be arbitrarily approximated due to its smoothness, and thus the advantage as well. Note that we can partition the AWGN channel outputs $\hat{\mathcal{Z}}$ in such a way that for each subset, the matrix of transition probabilities has the property that each row is a permutation of each other row and each column is a permutation of each other column. Thus, the channel is symmetric according to [68]. More precisely, the set of outputs of the $\mathcal{X}' - to - \hat{\mathcal{Z}}$ channel can be partitioned into subsets, so that in terms of transition probability matrices of the subsets (using inputs as rows and outputs of the subset as columns), with $\hat{\mathcal{Z}} = \bigcup_{i=1}^{\lceil L/2 \rceil} \hat{\mathcal{Z}}_i$, we have for all $\hat{z}, \hat{z}^* \in \hat{\mathcal{Z}}_i$ that the list of probabilities of $W[a, \cdot]$ and $W[-a, \cdot]$ and of $W[\cdot, \hat{z}]$ and $W[\cdot, \hat{z}_i^*]$ is the same, respectively. Furthermore, the same applies to the AWGN channel if we use QPSK, because this corresponds to two BPSK.

Consider the group $(\mathcal{X}', \oplus)$, where a is the identity and where $0 \mapsto a$ and $1 \mapsto -a$. Furthermore, let $\xi : \mathcal{S} \times \{0,1\}^k \to \mathcal{X}'^c$ be a random function which can be realized by a deterministic function $\xi : \mathcal{S} \times \{0,1\}^{l-k} \times \{0,1\}^k \to \mathcal{X}'^c$, that has an additional uniformly at random input vector.

According to [13][Theorem 4.12], if a random function $\xi : \mathcal{S} \times \{0,1\}^k \to \mathcal{X}'^c$ is separable and message-linear and the channel $E' : \mathcal{X}' \to \hat{\mathcal{Z}}$ is symmetric, then

$$Adv^{MIS_2}(\xi; E^c; \mathcal{S}) \leqslant Adv^{MIS_3}(\xi; E^c; \mathcal{S}), \tag{7.40}$$

where ξ is separable if

$$\xi(s, r, m) = \xi(s, r, 0^k) \oplus \xi(s, 0^{l-k}, m) \tag{7.41}$$

for all $s \in \mathcal{S}$, $r \in \{0,1\}^{l-k}$ and $m \in \{0,1\}^k$, and message linear if

$$\xi(s, 0^{l-k}, m + m') = \xi(s, 0^{l-k}, m) \oplus \xi(s, 0^{l-k}, m') \tag{7.42}$$

for all $s \in \mathcal{S}$, $m, m' \in \{0,1\}^k$. Obviously, for the case where f_s^{-1} is given as in (7.20), $\xi = BPSK \circ \Psi \circ f_s^{-1}$ is separable and message-linear.

Since we can approximate the Gaussian density function arbitrarily close with the appropriate partition of $\hat{\mathcal{Z}}$, (7.40) also holds for the continuous channel.

We summarize the result as follows.

Theorem 7.12. *Let $f_s^{-1}(m)$ be as in (7.20) and let the randomized function $\xi = BPSK \circ \Psi \circ f_s^{-1} : \mathcal{S} \times \{0,1\}^k \to \mathcal{X}'^c$ be a separable and message-linear encryption function, and $E : \mathcal{X}' \to \mathcal{Z}$ a symmetric AWGN channel, then (7.40) is true.*

DS_3-Security Implies DS_1-Security

First, we review [66, Lemma 5.8].

Lemma 7.13. *For the case where the channel $E : \mathcal{X}' \to \mathcal{Z}$ is symmetric and $\xi : \mathcal{S} \times \{0,1\}^k \to \mathcal{X}'^c$ is a separable and message-linear function, $||P_{Z^c|m,s} - P_{Z^c|s}||$ is the same regardless of the choice of the input $m \in \{0,1\}^k$.*

Theorem 7.14. *Let $f_s^{-1}(m)$ be as in (7.20). For the symmetric AWGN channel $E : \mathcal{X}' \to \mathcal{Z}$, if $\xi = BPSK \circ \Psi \circ f_s^{-1} : \mathcal{S} \times \{0,1\}^k \to \mathcal{X}'^c$ is a separable and message-linear randomized function, then*

$$Adv^{DS_1}(\xi; E^c; \mathcal{S}) \leqslant 2Adv^{DS_3}(\xi; E^c; \mathcal{S}). \tag{7.43}$$

Proof. Let $M \sim unif(\{0,1\}^k)$.

$$\begin{aligned} Adv^{DS_3}(\xi; E^c; \mathcal{S}) &= \frac{1}{|\mathcal{S}|\, 2^{2k}} \sum_{s \in \mathcal{S}} \sum_{\substack{m_1, m_2 \\ \in \{0,1\}^k}} \frac{1}{2} \int_{\mathcal{Z}^c} |p(\hat{z}^c|m_1, s) - p(\hat{z}^c|m_2, s)| dz^c \\ &\geqslant \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \frac{1}{2^k} \sum_{\substack{m \\ \in \{0,1\}^k}} \frac{1}{2} \int_{\mathcal{Z}^c} |p(\hat{z}^c|m, s) - p(\hat{z}^c|s)| dz^c \\ &= \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \frac{1}{2} \int_{\mathcal{Z}^c} |p(\hat{z}^c|m, s) - p(\hat{z}^c|s)| dz^c, \end{aligned}$$

where the inequality follows from the triangle inequality and the last equality from Lemma 7.13. Moreover,

$$\begin{aligned} Adv^{DS_1}(\xi; E^c; \mathcal{S}) &= \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \max_{m_1, m_2} \frac{1}{2} \int_{\mathcal{Z}^c} |p(\hat{z}^c|m_1, s) - p(\hat{z}^c|m_2, s)|\, dz^c \\ &\leqslant \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \max_{m_1, m_2} \Big[\frac{1}{2} \int_{\mathcal{Z}^c} |p(\hat{z}^c|m_1, s) - p(\hat{z}^c|s)| dz^c \\ &+ \frac{1}{2} \int_{\mathcal{Z}^c} |p(\hat{z}^c|s) - p(\hat{z}^c|m_2, s)| dz^c \Big] \\ &= 2 \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \frac{1}{2} \int_{\mathcal{Z}^c} |p(\hat{z}^c|m, s) - p(\hat{z}^c|s)| dz^c, \end{aligned} \tag{7.44}$$

where the inequality follows from the triangle inequality and the last equality follows from Lemma 7.13. ∎

Note that a similar security measure to $Adv^{DS_3}(\xi; E^c; \mathcal{S})$ is defined in [66]:

$$Adv^{RDS}(\xi; E^c; \mathcal{S}) = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \frac{1}{2^k} \sum_{\substack{m \\ \in \{0,1\}^k}} \frac{1}{2} \int_{\mathcal{Z}^c} |p(\hat{z}^c|m, s) - p(\hat{z}^c|s)| dz^c. \tag{7.45}$$

They have the following relationship:

$$Adv^{DS_3}(\xi; E^c; \mathcal{S}) \leqslant 2Adv^{RDS}(\xi; E^c; \mathcal{S}) \leqslant 2Adv^{DS_3}(\xi; E^c; \mathcal{S}). \tag{7.46}$$

DS_3-Security and MIS_3-Security

The empirical study of mutual information is difficult and therefore we can only approach the analysis of MIS_3-security theoretically.

Proposition 7.15. *Let $\xi : \mathcal{S} \times \{0,1\}^k \to \mathcal{X}'^c$ be a stochastic encoder and E^c the channel of Eve. Then,*

$$Adv^{MIS_3}(\xi; E^c; \mathcal{S}) \leqslant 2Adv^{DS_3}(\xi; E^c; \mathcal{S}) \log \frac{2^k}{2Adv^{DS_3}(\xi; E^c; \mathcal{S})}. \tag{7.47}$$

Proof. For the AWGN channel (and the DMC channel) we can use the upper bound proposed by [62, Appendix I]:

$$Adv^{MIS_3}(\xi; E^c; \mathcal{S}) \leqslant 2 \left\| P_{Z^c,M|S=s} - P_{Z^c|S=s} P_M \right\| \log \frac{2^k}{2 \left\| P_{Z^c,M|S=s} - P_{Z^c|S=s} P_M \right\|}. \tag{7.48}$$

Observe that,

$$\begin{aligned}
&Adv^{MIS_3}(\xi; E^c; S) \\
&\overset{a)}{\leqslant} \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} 2 \left\| P_{Z^c,M|S=s} - P_{Z^c|S=s} P_M \right\| \log \frac{2^k}{2 \left\| P_{Z^c,M|S=s} - P_{Z^c|S=s} P_M \right\|} \\
&\overset{b)}{\leqslant} \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} 2 \left\| P_{Z^c,M|S=s} - P_{Z^c|S=s} P_M \right\| \log \frac{2^k}{\frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} 2 \left\| P_{Z^c,M|S=s} - P_{Z^c|S=s} P_M \right\|},
\end{aligned}$$

where a) follows from (7.48), and b) as $-x \log x$ is concave.

Furthermore,

$$
\begin{aligned}
&\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}}\left\|P_{Z^c,M|S=s}-P_{Z^c|S=s}P_M\right\| \\
&=\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}}\sum_{m\in\{0,1\}^k}P_M(m)\frac{1}{2}\int_{\mathcal{Z}^c}|p(z^c|m,s)-p(z^c|s)|dz^c \\
&=\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}}\frac{1}{2^k}\sum_{m_1\in\{0,1\}^k}\frac{1}{2}\int_{\mathcal{Z}^c}|p(z^c|m_1,s)-\frac{1}{2^k}\sum_{m_2\in\{0,1\}^k}p(z^c|m_2,s)|dz^c \\
&=\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}}\frac{1}{2^k}\sum_{m_1\in\{0,1\}^k}\frac{1}{2}\int_{\mathcal{Z}^c}|\frac{1}{2^k}\sum_{m_2\in\{0,1\}^k}\big(p(z^c|m_1,s)-p(z^c|m_2,s)\big)|dz^c \\
&\leqslant\tfrac{1}{|\mathcal{S}|2^{2k}}\sum_{s\in\mathcal{S}}\sum_{\substack{m_1,m_2\\\in\{0,1\}^k}}\frac{1}{2}\int_{\mathcal{Z}^c}|p(z^c|m_1,s)-p(z^c|m_2,s)|dz^c \\
&=Adv^{DS_3}(\xi;E^c;\mathcal{S}). \qquad (7.49)
\end{aligned}
$$

■

If $Adv^{DS_3}(\xi;E^c;\mathcal{S})$ decreases exponentially with c, then DS_3-security implies MIS_3-security.

DS_3-Security and MIS_2-Security

From (7.40) and (7.47) follows:

Proposition 7.16. *Let $\xi:\mathcal{S}\times\{0,1\}^k\to\mathcal{X}'^c$ be a separable and message-linear stochastic encoder and E^c a symmetric AWGN channel of Eve. Then,*

$$
Adv^{MIS_2}(\xi;E^c;\mathcal{S})\leqslant 2Adv^{DS_3}(\xi;E^c;\mathcal{S})\log\frac{2^k}{2Adv^{DS_3}(\xi;E^c;\mathcal{S})}. \qquad (7.50)
$$

This is also true for $f_s^{-1}(m)$ given in (7.20). If $Adv^{DS_3}(\xi;E^c;\mathcal{S})$ decreases exponentially with c, then DS_3-security implies MIS_2-security.

7.4. Measurement of security by simulation

7.4.1. Operational Meaning

We consider the operational meaning of the DS_1-security and DS_3-security to be able to evaluate the simulation results in Section 8. For simulations when considering distinguishing security and strong security, we can use the ML decoder at Eve which is an optimal attack strategy, because only message distributions with equally probable message pairs are considered.

Distinguishing Security of Scenarios 1 and 1a)

We can reformulate the first equation of (7.26) to obtain

$$\begin{aligned}
&Adv^{DS_1}(\xi; E^c; \mathcal{S}) \\
&= \frac{1}{|\mathcal{S}|} \sum_{s\in\mathcal{S}} \max_{\mathcal{A},m_1,m_2} Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_1, s)) = 1) - Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_2, s)) = 1) \\
&= \frac{1}{|\mathcal{S}|} \sum_{s\in\mathcal{S}} \max_{\mathcal{A},m_1,m_2} 1 - Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_1, s)) = 2) \\
&- Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_2, s)) = 1),
\end{aligned} \tag{7.51}$$

where $Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_1, s)) = 2)$ is the probability of error of the first kind, and $Pr(\mathcal{A}(s, m_1, m_2, Z^c(m_2, s)) = 1)$ the probability of error of the second kind.

Every $\{1, 2\}$-valued eavesdropper strategy $z^c \mapsto \mathcal{A}(s, m_1, m_2, z^c)$ is a hypothesis test for distinguishing m_1 and m_2. Thus for fixed s, m_1, m_2, the maximum over $\mathcal{A}$ in (7.51) is attained by an ML test with threshold η, as given in Subsection 7.4.2. Then (7.51) becomes

$$Adv^{DS_1}(\xi; E^c; \mathcal{S}) = \frac{1}{|\mathcal{S}|} \sum_{s\in\mathcal{S}} \max_{\eta,m_1,m_2} (1 - \lambda_1(s, m_1, m_2, \eta) - \lambda_2(s, m_1, m_2, \eta)), \tag{7.52}$$

where $\lambda_1(s, m_1, m_2, \eta)$ and $\lambda_2(s, m_1, m_2, \eta)$ are the probability of error of the first kind and the second kind, respectively.

If we define the *distinguishing error rate* at Eve:

$$DER_{E_1}(\xi; E^c; \mathcal{S}) = \frac{1}{|\mathcal{S}|} \sum_{s\in\mathcal{S}} \min_{\eta,m_1,m_2} \frac{(\lambda_1(s, m_1, m_2, \eta) + \lambda_2(s, m_1, m_2, \eta))}{2}, \tag{7.53}$$

then we can see that $Adv^{DS_1}(\xi; E^c; \mathcal{S})$ and $DER_{E_1}(\xi; E^c; \mathcal{S})$ are two different representations of Eve's performance and therefore can be translated into each other by

$$Adv^{DS_1}(\xi; E^c; \mathcal{S}) = 1 - 2 \cdot DER_{E_1}(\xi; E^c; \mathcal{S}).$$

In all communication scenarios, we can use DER_E as a benchmark value for the advantage, where DER_E close to $1/2$ means "high security".

Accordingly, (7.28) can be reformulated so that

$$Adv^{DS_{1a)}}(\xi; E^c; s) = \max_{\eta,m_1,m_2} (1 - \lambda_1(s, m_1, m_2, \eta) - \lambda_2(s, m_1, m_2, \eta)). \tag{7.54}$$

The distinguishing error rate at Eve for given η, m_1, m_2 and s is then,

$$DER_{E_{1a)}}(s, m_1, m_2, \eta) = \frac{\lambda_1(s, m_1, m_2, \eta) + \lambda_2(s, m_1, m_2, \eta)}{2}.$$

We define $DER_{E_{1a)}}(s, m_1, m_2, \eta)$ for each message pair because we apply them in the simulation.

The maximization over the message pair is discussed in Section 7.4.3.

Distinguishing Security of Scenario 3

Here, we obtain $Adv^{DS_3}(\xi; E^c; \mathcal{S})$ in terms of the probability of error of the first and the second kind.

Proposition 7.17. *Let $\xi : \mathcal{S} \times \{0,1\}^k \to \mathcal{X}'^c$ be a stochastic encoder and E^c the channel of Eve. Then,*

$$Adv^{DS_3}(\xi; E^c; \mathcal{S}) \quad = \quad 1 - \bar{\lambda}(\xi; E^c; \mathcal{S}), \tag{7.55}$$

where $\bar{\lambda}(\xi; E^c; \mathcal{S}) = \frac{1}{2^{2k}|\mathcal{S}|} \sum_{s\in\mathcal{S}} \sum_{m_1,m_2\in\{0,1\}^k} \min_\eta(\lambda_1(s, m_1, m_2, \eta) + \lambda_2(s, m_1, m_2, \eta))$.

Proof. As in (7.52) and (7.54), we can replace the arbitrary distinguishing strategies $\mathcal{A}$ by ML tests. Thus

$$\begin{aligned} Adv^{DS_3}(\xi; E^c; \mathcal{S}) &= \frac{1}{2^{2k}\,|\mathcal{S}|} \sum_{s\in\mathcal{S}} \sum_{m_1,m_2\in\{0,1\}^k} \max_\eta(1 - \lambda_1(s, m_1, m_2, \eta) - \lambda_2(s, m_1, m_2, \eta)) \\ &= \frac{1}{2^{2k}|\mathcal{S}|} \sum_{s\in\mathcal{S}} \sum_{m_1,m_2\in\{0,1\}^k} (1 - \min_\eta(\lambda_1(s, m_1, m_2, \eta) + \lambda_2(s, m_1, m_2, \eta))) \\ &= 1 - \bar{\lambda}(\xi; E^c; \mathcal{S}), \end{aligned} \tag{7.56}$$

with $\bar{\lambda}(\xi; E^c; \mathcal{S})$ as defined in the statement. ∎

The corresponding distinguishing error rate at Eve is defined as

$$DER_{E_3}(\xi; E^c; \mathcal{S}) \quad = \quad \frac{\bar{\lambda}(\xi; E^c; \mathcal{S})}{2}. \tag{7.57}$$

7.4.2. Attack strategy of the eavesdropper

Eve applies the ML test. Consider the log likelihood ratio with threshold η,

$$LLR(z^c|s, m_1, m_2) = \log\left(\frac{p(z^c|s, m_1)}{p(z^c|s, m_2)}\right) \begin{matrix} \geqslant^{\hat{M}=m_1} \\ <_{\hat{M}=m_2} \end{matrix} \log(\eta),$$

where Eve decides for m_1 if $LLR(z^c|s, m_1, m_2)$ is greater than or equal to $\log\eta$ and for m_2 otherwise.

For the case where the channel is AWGN, we use the following conditional probability densities when message m_b with $b \in \{1, 2\}$ was sent,

$$p(z^c|s, m_b) = \frac{1}{2^{l-k}} \sum_{v:f_s(v)=m_b} \prod_{i=1}^{c} w(z_i(m_b, s)|\chi(\Psi(v))_i), \tag{7.58}$$

where $\chi(\Psi(v))_i$ denotes the i-th symbol in the length-c channel input $\chi(\Psi(v)) \in \mathbb{C}^c$, and, for any channel input $x' \in \mathbb{C}$ and output $z \in \mathbb{C}$,

$$w(z|x') = \frac{SNR_E}{\pi} \exp(-|z - x'|^2 SNR_E). \tag{7.59}$$

7.4.3. Determining the Best Performance of Eve under DS_1- Security

Now we focus on how to maximize (7.52) over the pair of messages out of the message set $\{0, 1\}^k$. We consider the problem from the coding point of view and analyze the code structure.

We want to analyze the relationship between the advantage at Eve under distinguishing security and the Hamming distance of the chosen codeword pairs of the associated message pair (m_1,m_2). If a message pair (m_1, m_2) is chosen for which the codeword pairs have the maximum Hamming distance, intuitively this should be close to optimal for Eve.

Recall that the randomized inverse of the UHF of (7.17) provides a stochastic mapping from messages to the FEC inputs, so that as the message m is chosen, the encoder chooses uniformly at random a vector v from the set $\{v' : f_s(v') = m\}$, and encodes it to a codeword x^n via the FEC code with generator matrix G. Therefore we do not search for a single codeword pair but for a *coset* pair of codewords that have the maximum average Hamming distance. Let us denote the set $C'(m, s) := \{v' : f_s(v') = m\} G$ that corresponds to a certain message m as coset, where $C'(m, s)$ is a subset of the codeword set of the seeded modular UHF wiretap code $\mathcal{C}_n \subset \{0, 1\}^n$, with $|C'(m, s)| = 2^{l-k}$ and $|\mathcal{C}_n| = 2^l$. $C'(m, s)$ is in fact a coset since the UHF is affine-linear and $\{0, 1\}^l$ corresponds to the elements of $GF(2^l)$. Furthermore, $\dot{\bigcup}_{h=1,\ldots,2^k} C'(m_h, s) = \mathcal{C}_n$.

Thus, we consider the following working hypothesis. If a message pair is chosen for which the cosets $C'(m_1, s)$ and $C'(m_2, s)$ have the maximum average Hamming distance among all pairs of such sets, this should intuitively be close to optimal for Eve. For any linear FEC code with $l \times n$ generator matrix G, we first look for two cosets $C'(m_1, s) = \{v' : f_s(v') = m_1\} G$, $C'(m_2, s) = \{v'' : f_s(v'') = m_2\} G$, whose codewords have in average the maximum Hamming distance to each other, where $m_1 \neq m_2 \in \{0, 1\}^k$. The maximum average Hamming distance of a coset pair $(C'(m_1, s), C'(m_2, s))$ is defined as follows:

$$d_{max}(s) = \max_{m_1, m_2} d_H(C'(m_1, s), C'(m_2, s)),$$

where

$$d_H(C'(m_1,s),C'(m_2,s)) := \frac{1}{2^{(l-k)2}} \sum_{x_1^n \in C'(m_1,s)} \sum_{x_2^n \in C'(m_2,s)} d_H(x_1^n, x_2^n)$$

is the average Hamming distance between $C'(m_1,s)$ and $C'(m_2,s)$. We also define the minimum such distance,

$$d_{min}(s) = \min_{m_1,m_2} d_H(C'(m_1,s),C'(m_2,s)).$$

Remark 7.18. *If we arrange the Hamming distances of all codeword pairs from the coset pair in a matrix - whose columns number the codewords from coset 1 and whose rows number the codewords from coset 2 - we see that the matrix is bisymmetric and additionally that the diagonal and prediagonal each have uniform values. This insight saves computational power.*

To find a message pair or the corresponding coset pair of maximum average Hamming distance, we analyze the code as follows. W.l.o.g. we can set $t = 0$, because both for $t = 0$ and $t \neq 0$, the average Hamming distance between two cosets remains unchanged. Let $x^n(m,r) \in C'(m,s)$, where m and r specifies the codeword. Since $(m||r) = (m||0^{l-k}) + (0^k||r)$ and the distributive law holds, we can write (7.18) as follows.

$$s^{-1} * (m||r) = s^{-1} * \left[(m||0^{l-k}) + (0^k||r)\right] = s^{-1} * (m||0^{l-k}) + s^{-1} * (0^k||r).$$

Then (7.21) becomes

$$x^n(m,r) = (s^{-1} * (m||0^{l-k}))G + (s^{-1} * (0^k||r))G. \tag{7.60}$$

The set of n-bit vectors $\left\{s^{-1} * (0^k||r)\right\}_{\forall r \in \{0,1\}^{l-k}} G$ forms the coset $C'(0^k,s)$, and the cosets $C'(m_i,s)$ are given by $b_i + C'$ with $b_i = (s^{-1} * (m_i||0^{l-k}))G$, which correspond to message m_i, with $i \in \left\{1,\ldots,2^k\right\}$. If vectors b_i and b_j, $i \neq j \in \left\{1,\ldots,2^k\right\}$ have the maximum Hamming distance then for any given r the corresponding codewords $x^n(m_i,r) \in C'(m_i,s)$ and $x^n(m_j,r) \in C'(m_j,s)$ have the maximum Hamming distance, too. However, this gives no information about the coset pairs which have the maximum average Hamming distance. Since the codeword pairs are randomly selected from the coset pairs, their Hamming distances are not known in advance. Furthermore, we could not theoretically show a correlation between Adv^{DS_1} and the Hamming distance of the codeword pairs. This is where the simulations come into action, provided in Section 8.2. We gain helpful insights into the interaction of seed and Eve's advantage or the other code parameters, e.g., l,k.

Note that for higher order modulation, i.e. for $j \geqslant 4$, the performance of Eve depends not only on the code but also on the modulation. Here it would be interesting for future work to investigate how the interaction of code and modulation affects the performance

of Eve by additionally analyzing the Euclidean distance for the modulated words. Our assumption is that the impact of the modulation decreases as the length of the random vector r increases.

It is also interesting to see how the mean over the seed set affects the choice of message pair. Let's take a closer look at equation (7.52).

$$\begin{aligned}
&Adv^{DS_1}(\xi; E^c; \mathcal{S}) \\
&= \frac{1}{|\mathcal{S}|} \sum_{s\in\mathcal{S}} \max_{\eta} \max_{m_1,m_2} \left[1 - Pr\left(\frac{\sum_{v:f_s(v)=m_1} \prod_{i=1}^{c} w(z_i(m_1,s)|\chi(\Psi(v))_i)}{\sum_{v:f_s(v)=m_2} \prod_{i=1}^{c} w(z_i(m_1,s)|\chi(\Psi(v))_i)} < \eta \right) \right. \\
&\left. - Pr\left(\frac{\sum_{v:f_s(v)=m_1} \prod_{i=1}^{c} w(z_i(m_2,s)|\chi(\Psi(v))_i)}{\sum_{v:f_s(v)=m_2} \prod_{i=1}^{c} w(z_i(m_2,s)|\chi(\Psi(v))_i)} \geqslant \eta \right) \right],
\end{aligned} \tag{7.61}$$

where $w(z|x')$ follows (7.59).

The probability ratio contains in both the denominator and the numerator Gaussian mixture densities, which makes the analysis difficult. Since in the case of DS_1-security, Eve selects the message pair that maximizes the advantage after the selection of the seed, the selection of the message pair plays an important role. However, according to Proposition 7.19 in the case where the message pair and the seed are chosen independently, if our working hypothesis is correct, $Adv^{DS_2}(\xi; E^c; \mathcal{S})$ is independent of the choice of the message pairs.

Proposition 7.19. *The averaged Hamming distance of any coset pair, averaged over all seeds $s = (a, t)$ with $a \in \mathbb{F}_{2^l}^*$, $t \in \mathbb{F}_{2^l}$, is independent of the choice of the message pair.*

Proof. Let w be the Hamming weight, $r \in \{0,1\}^{l-k}$ the random vector $v(r) \in f_s^{-1}(m, r)$ and $p = (a^{-1} * t)G$. Consider for any message pair m_1, m_2, and $m = m_1 + m_2$ the following

$$\begin{aligned}
&\sum_{r_1,r_2,s} d_H((a^{-1} * (m_1||r_1))G + p, (a^{-1} * (m_2||r_2))G + p) \\
&= \sum_{r_1,r_2,a} w((a^{-1} * (m_1 + m_2||r_1 + r_2))G) \\
&= \sum_{a} 2^{l-k} \sum_{r} w(a^{-1} * (m||r)G) \\
&= \sum_{v\in\{0,1\}^l\setminus\{0\}^l} 2^{l-k} \sum_{r} w(v(r)G),
\end{aligned}$$

where the sum is over all $a \in \mathbb{F}_{2^l}^*$, all $r_1, r_2 \in \{0,1\}^{l-k}$. Moreover, the last equation holds since for given m, r with $(m||r) \neq 0^l$ we have

$$\left\{(a^{-1} * (m||r))G : a^{-1} \in \mathbb{F}_{2^l}^*\right\} = \left\{vG : v \in \mathbb{F}_{2^l}^*\right\}. \qquad \blacksquare$$

If our working hypothesis is true, then $Adv^{DS_3}(\xi; E^c; S) \approx Adv^{DS_2}(\xi; E^c; S)$.

7.5. Conclusion

We considered a seeded modular UHF code for the AWGN wiretap channel. Furthermore, we introduced three communication scenarios, each reflecting the operational meaning of different security measures and different assumptions about Eve's strengths. We showed some relevant relationships between the advantages of the three scenarios. We introduced a new strong security metric to compare with the distinguishing security. We derived the operational meanings of the distinguishing security and the strong security in the three communication scenarios. We have made a working hypothesis that increasing the average Hamming distance of a coset pair improves Eve's performance and thus increases her advantage. But we could not prove the correlation. Also, we could not find the coset pairs with maximum average Hamming distance by analyzing the code structure. The reason the Hamming distances of the codeword pairs are not known in advance is that they are randomly selected from the coset pairs.

In a finite blocklength regime, security is more quantitative, given by a certain number of secure bits. It is difficult to classify systems as secure or insecure, and they can be application specific. Therefore, we derived a security level d in terms of code parameters, which specifies how many bits are secure. We can use d to estimate the necessary amount of encoding randomness at given code parameters and channel parameters.For future work, it would be interesting to extend the modular coding scheme to the fading and Multiple Input Multiple Output (MIMO) case.

8. Simulations and Results on the Seeded Modular Code

8.1. Introduction

We experimentally verify the information-theoretic security of a seeded modular code for the Additive White Gaussian noise (AWGN) wiretap channel consisting of a security layer, an error-correction layer and a modulation layer. Depending upon the communication scenario and the operating SNR of the eavesdropper's channel (SNR_E), we determine the advantage at Eve under distinguishing security. We use the advantage in terms of the error probability derived in Section 7.4 to be able to assess the security performance. We gain helpful insights from the simulation results, e.g., the impact of code parameters and seed choice on security. For BPSK and QAM, we find that for small blocklengths, the advantage under the required security metric is close to 0 for suitable code parameters. We also verify that our simulation results support the theoretical results. In addition, we compare the achievable secrecy rates with the simulated secrecy rates in terms of the advantage under a given security metric. Finally, we compare the decoding performance of the attack strategy proposed in Section 7.4.2 with other attack strategies with lower computational effort.

8.1.1. Contribution

To the best of our knowledge, we are the first to verify the distinguishing security [13] of a seeded modular code for the AWGN wiretap channel by simulations. We use the FEC codes and the modulation schemes proposed by the 3GPP standards [69] and [70], respectively, which fulfill a desired decoding probability ($< 10^{-4}$) at a certain signal-to-noise ratio SNR_B of Bob's channel. As the security component, we use the UHF proposed in [56]. Our simulation results show that for given SNR_E and a suitable, positive *d-secure rate* (see Section 8.2), Eve's advantage under distinguishing security is close to zero, even for small blocklengths. The *security level* d specifies the sufficient amount of randomness. We analyze the advantage at Eve in terms of the security level with different modulation alphabets. In order to see how the seed choice affects the advantage, we analyze the correlation between the average Hamming distance of the codeword sets corresponding to distinct message pairs, and the advantage at Eve under distinguishing security. We

observe a positive correlation. We observe that the seed set can be divided into two subsets. In the one subset, all seeds ensure a consistent advantage at Eve. The other subset consists of *dispersing* seeds that can affect the code in such a way, so that the advantage at Eve under distinguishing security can be increased. This means that we want the probability of occurrence of *dispersing* seeds to be as small as possible. We find that the cardinality of the subset of *dispersing* seeds decreases exponentially with the length of the random vector.

8.1.2. Related Work

We refer to the related work in Section 7.1.

8.1.3. Outline

Section 8.2 contains the simulations, the simulation results and the insights we gain through the simulations. Section 8.3 we present other attack strategies and compare them with the ML test from Section 7.4.2. Section 8.4 concludes the chapter.

8.2. Simulations and Results

In our simulations we vary either k or SNR_E, while SNR_B and all other code parameters are fixed so that $P_e(\xi, T^c, \zeta) < 10^{-4}$. The average transmit power P_{av} is set to 1 in our simulations, so that only the noise of the respective channel changes with the SNR. Note that P_{av} is different from P used in the asymptotic analysis in Section 7.2.3 for QAM modulations with $R_{mod} > 2$. It is possible that individual codewords occur which require $P > 1$. In addition, we can improve the bound in (7.37) by replacing the capacity of the Gaussian channel $C_E(\sigma_E^2, P)$ with $C_{unif_E}(P_{av} - \alpha)$. According to [71], $C_{unif_i}(P_{av} - \alpha)$ with $i \in \{T, E\}$ is the capacity for the AWGN channel of $2^{R_{mod}}$-QAM in the limit of asymptotically large code blocklength, where $\alpha > 0$ is a small constant power margin. In the case of BPSK and QPSK $\alpha = 0$, otherwise $\alpha = 0.05$. $C_{unif_i}(P_{av})$ is given in [72] as

$$C_{unif_i}(P_{av}) := R_{mod} - \frac{1}{2^{R_{mod}}} \sum_{k=1}^{2^{R_{mod}}} \mathsf{E}_{N_i}[\log(\sum_{l=1}^{2^{R_{mod}}} e^{\frac{|N_i|^2 - |X'_k + N_i - X'_l|^2}{2\sigma_i^2}})], \tag{8.1}$$

where X'_k denotes the k-th modulation symbol uniformly distributed over $2^{R_{mod}}$ modulation symbols with $k \in \{1, 2, ..., 2^{R_{mod}}\}$, and N_i with $i \in \{T, E\}$ denotes the Gaussian noise of channel i.

From now on we use (7.37) to choose the code parameter, where we replace $C_E(\sigma_E^2, P)$ by $C_{unif_E}(P_{av} - \alpha)$:

$$d := \frac{1}{2}\Big(l - k - cC_{unif_E}(P_{av} - \alpha)\Big) - 2. \tag{8.2}$$

In the following we will see that not only d_{DS_2} but even d_{DS_1} can be larger than d (see Remark 7.10). Furthermore, we define the d-secure rate

$$R_{sec}(d) := \frac{k}{c} = \frac{k}{l} R_{eff}, \tag{8.3}$$

with

$$Adv^{DS_2}(\xi; E^c; \mathcal{S}) \leqslant 2^{-d}. \tag{8.4}$$

For a given SNR_E, the code parameters l, k and c determine d. We use d to approximate d_{DS_3} and $d_{DS_{1a)}}(s)$ for specific seed s. In Subsections 8.2.1 and 8.2.2, respectively, we will see why the approximations are admissible. Later we empirically determine $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s)$ by substituting the error probabilities which we obtain from the simulation for the corresponding parameters l, k, c and SNR_E into (7.54). Then we get the empirical $\hat{d}_{DS_{1a)}}(s) := -\log(\overline{Adv}^{DS_{1a)}}(\xi; E^c; s))$ and compare it with the theoretical security level d. Similarly, we estimate $\overline{Adv}^{DS_3}(\xi; E^c; \mathcal{S})$ using (7.55), and $\hat{d}_{DS_3}$.

For our simulations, we use the Matlab 5G Toolbox. Furthermore, we use the seed and the UHF given in (7.17) as security component. The implementation of the UHF and its inverse was done by using cyclotomic polynomials for a faster computation. This restricts the choice of l in our simulations because $l + 1$ has to be a prime number.

For the error-correction and the modulation layer we use polar codes in the uplink scenario as proposed in the 5G new radio standard. Authors in [73] give a report to the channel coding 5G new radio and show complete coding chains for the NR polar codes. The core components of the FEC encoder are the cyclic-redundancy-check (CRC) of length $n_{crc} = 6$ or 11 bits, the polar encoding kernel, and the rate matcher. In addition, in the uplink scenario a segmentation is performed before the CRC encoder. Furthermore, a parity check encoder is applied before the polar encoding kernel. The rate matcher, which contains a subblock interleaver is followed by a channel interleaver. All functions of the coding chain are linear and therefore we let G be the concatenation of the linear functions. Note that n_{crc} is not included in the value l and therefore has no influence on R_{eff}. At Bob, the soft-demodulated channel outputs are transformed into log likelihood ratios (LLR), which are rate recovered and then decoded with a CRC-aided successive cancellation list decoder of list size $L_{polar} = 8$. In the modulation layer, the following modulation schemes are supported: BPSK, $\pi/2$-BPSK, QPSK, 16QAM, 64QAM and 256QAM. The modulation scheme follows a Gray coding.

Eve's attack strategy is implemented as shown in Section 7.4.2, where we have chosen $\eta = 1$ to optimize Eve's performance. That is, by $\eta = 1$ we maximize (7.52), (7.54), (7.55). Recall that Eve knows the selected message pair and its distribution, the seed, the coding scheme, and the channel. The drawback of Eve's attack strategy is that the computational cost grows exponentially with $l - k$, since the cosets grow exponentially with $l - k$. For example, if we use nodes with 28 cores that have a nominal frequency of

2.6 GHz and a DDR4 memory of 64 GB per node, the simulations for a single SNR value require 22 hours for $l-k=22$ and about 91 hours for $l-k=24$. To obtain $d>0$, we can see in (8.2) that for small values of $l-k$, the SNR_E has to be chosen small or negative.

8.2.1. Distinguishing Security - Scenario 1, 1a)

Correlation Between the Average Hamming Distance and Eve's Seeded Advantage

For comparison we consider Eve's distinguishing performance as a function of SNR_E for a seed s and message pairs that provide all average Hamming distances between $d_{max}(s)$ and $d_{min}(s)$. In the simulations, first the seed $s=(a,t)$ is chosen, then the source needs to select between two uniformly probable messages m_1, m_2 to be transmitted that provide $d_{max}(s)$, $d_{min}(s)$ or any distance in between. Eve receives $z^c \in \mathbb{C}^c$ and has to choose between m_1 and m_2. The message pair and the seed are fixed for 10^3 iterations. In each iteration, the source uniformly chooses one of the two messages to be transmitted.

Fig. 8.1 shows the distinguishing error rate of Eve $DER_{E_{1a)}}(s,m_1,m_2)$ and Fig. 8.2 shows $\lambda_1(s,m_1,m_2)$ and $\lambda_2(s,m_1,m_2)$. Both figures illustrate two curves of $DER_{E_{1a)}}$, $\lambda_1(s,m_1,m_2)$ or $\lambda_2(s,m_1,m_2)$ for each modulation alphabet. The solid curves belong to a message pair with $d_{max}(s)=19$ and the dashed curves belong to a message pair with $d_{min}(s)=14$. The solid curves are always strictly smaller than the dashed curves for the same modulation scheme. For BPSK and QPSK, $DER_{E_{1a)}}(s,m_1,m_2)$ decreases monotonously in $d_H(C'(m_1,s),C'(m_2,s))$. For higher order modulations, if $d_H(C'(m_1,s),C'(m_2,s))$ is not much larger than $d_H(C'(m_3,s),C'(m_4,s))$, the reverse behavior is possible, i.e., $DER_{E_{1a)}}(s,m_1,m_2) > DER_{E_{1a)}}(s,m_3,m_4)$. For example, we have observed that for the 16QAM case the $DER_{E_{1a)}}(s,m_1,m_2)$ with a specific coset pair having $d_H(C'(m_1,s),C'(m_2,s))=14$ in Fig. 8.1 is higher than the $DER_{E_{1a)}}(s,m_1,m_2)$ with a specific coset pair having $d_H(C'(m_3,s),C'(m_4,s))=15$. However, this phenomenon gets less frequent as $l-k$ increases, i.e., with increasing coset size. We suspect that this smoothes out the influence of higher order modulation due to the increasing number of codewords in the coset pair. Consequently, we can assume that Eve's advantage, with few exceptions, correlates positively with the average Hamming distance.

In Fig. 8.2, we can observe that $\lambda_1(s,m_1,m_2)$ and $\lambda_2(s,m_1,m_2)$ are very similar to each other which implies an optimal choice of η, since we have a binary hypothesis test with two uniformly distributed messages and a symmetric channel.

Dispersing vs. Non-Dispersing Seeds

Since we are interested in seeds that maximize the advantage at Eve, we analyze how the seeds affect the partitioning of the code $\mathcal{C}_n$ and the average Hamming distance. For our analyses of the average Hamming distance, we set $t=0^l$ in (7.17), because t does not affect the average Hamming distance. For given code parameters l,k,n, and for all $s \in \mathcal{S}$, we de-

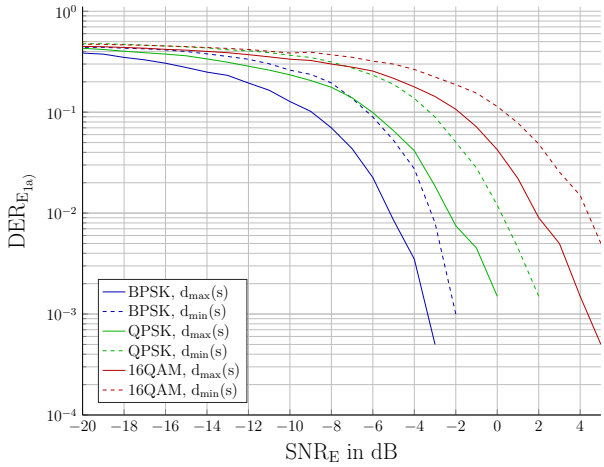


Figure 8.1.: Eve's distinguishing performance for $n = 32, l = 12, k = 6, n_{crc} = 6$ and for a fixed seed. For BPSK we choose $SNR_B = 2.5$dB, for QPSK SNR_B= 5dB and for 16QAM $SNR_B = 10.5$dB. The solid curves show the distinguishing performance at Eve for a coset pair with $d_{max}(s) = 19$ and the dashed curves show the performance for a coset pair with $d_{min}(s) = 14$.

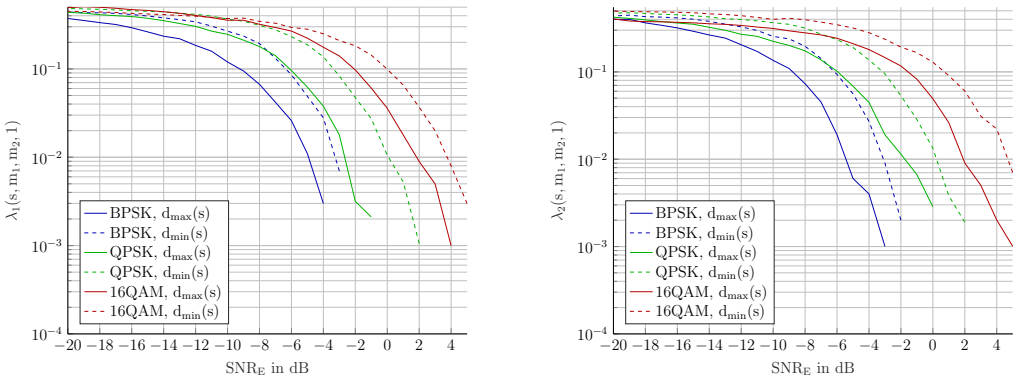


Figure 8.2.: $\lambda_1(s, m_1, m_2, 1)$ and $\lambda_2(s, m_1, m_2, 1)$ for the same parameters as in Fig. 8.1

termine the average Hamming distance of all coset pairs $\{(C'(m_i, s), C'(m_j, s))\}_{i \neq j \in \{1,\ldots,2^k\}}$.

We observed that the seed set can be divided into two subsets. One subset consists of "non-dispersing" seeds, or $s_{nd} \in \mathcal{S}_{nd} \subset \mathcal{S}$ for short. Non-dispersing seeds provide coset pairs that have all the same average Hamming distance. In the case where $s_{nd} \in \mathcal{S}_{nd}$ is used, the choice of the message pair and the non-dispersing seed does not affect Eve's performance. The other subset consists of "dispersing" seeds, or $s_d \in \mathcal{S}_d \subset \mathcal{S}$ for short, which partition the code $\mathcal{C}_n$ in such a way that some coset pairs have a larger average Hamming distance than that of a non-dispersing seed, thus improving Eve's performance. For any fixed seed s_{nd}, the mean value of the average Hamming distances of all possible coset pairs is $\frac{n}{2}$, whereas for a given s_d the mean value may differ slightly from $\frac{n}{2}$. Note that the seed chosen for the analyses in Section 8.2.1 necessarily has to be a dispersing seed, since $d_{min}(s_d) < d_{max}(s_d)$. The density distribution of the number of message pairs as a function of $d_H(C'(m_1, s), C'(m_2, s))$ for any $s = s_{nd} \in \mathcal{S}_{nd}$ is shown in Fig. 8.3 in the

left image and for a certain $s = s_d \in \mathcal{S}_d$ with $d_{max}(s_d) = 19$ and $d_{min}(s_d) = 14$ in the right image.

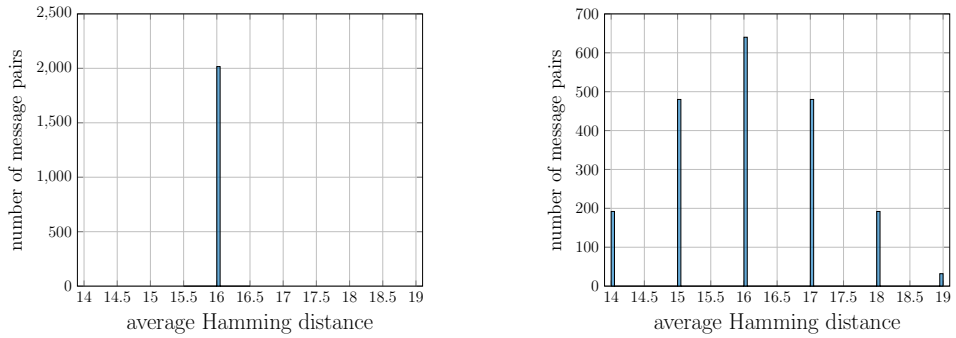


Figure 8.3.: Density distribution of the number of message pairs as a function of the average Hamming distance for a non-dispersing seed on the left and for a dispersing seed on the right for $n = 32, l = 12, k = 6$ and $n_{crc} = 6$.

We want to know the occurrence probability of the dispersing seeds. For this purpose, we analyze the average Hamming distance of all possible coset pairs for $n = 32$, $l = 12$, $n_{crc} = 6$ and $k = \{2, 3, 4, 5, 6\}$ for all possible seeds. For each seed we consider a coset pair with $d_{max}(s)$. For different values of k, the number of seeds that belongs to the respective average maximum Hamming distance $d_{max}(s_d)$ is listed in Table 8.1 on the left and the number of seeds that belongs to the average minimum Hamming distance $d_{min}(s_d)$ is listed in Table 8.1 on the right. We observed that for $k = 6$, 36.73% of all seeds are dispersing. If we increase $l - k$ so that, $k = 3$, the number of all dispersing seeds decreases to 4.93%. Altogether, we observed that $|\mathcal{S}_d|$ diminishes with $\mathcal{O}(2^{-(l-k)})$.

Table 8.1.: The average Hamming distance of all possible coset pairs for $n = 32$, $l = 12$, $n_{crc} = 6$ and $k = \{2, 3, 4, 5, 6\}$ for all possible dispersing seeds.

$d_{max}(s_d)$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
16.5	82	182	370	675	1114
17	7	18	43	122	296
17.5	0	2	8	19	73
18	0	0	0	4	16
18.5	0	0	0	0	3
19	0	0	0	0	2
$\sum$	89	202	421	820	1504

$d_{min}(s_d)$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
15.5	82	184	370	675	1114
15	0	18	43	122	296
14.5	0	0	8	19	73
14	0	0	0	4	18
13.5	0	0	0	0	3
14	0	0	0	0	0
$\sum$	82	202	421	820	1504

It is also interesting to consider the variance of the density distribution of the number of message pairs as a function of the average Hamming distance for dispersing seeds. For this we have varied k for fixed l and observed that the variance decreases as $l-k$ increases.

The Effect of Seed Choice on the Advantage

Next, we compare $Adv^{DS_1}(\xi; E^c; \mathcal{S})$ from (7.52) with $Adv^{DS_{1a)}}(\xi; E^c; s)$ from (7.54), where s is either dispersing or non-dispersing.

Since $d_H(C'(m_1, s_{nd}), C'(m_2, s_{nd}))$ does not change with m_1, m_2 as long as $m_1 \neq m_2$, we assume by our working hypothesis that the choice of message pair does not influence

Eve's advantage and that this advantage is the same for all non-dispersing seeds. In other words,

$$Adv^{DS_{1a)}}(\xi; E^c; s_{nd}) = max_{\mathcal{A}} 2Pr(\mathcal{A}(s_{nd}, m_1, m_2, Z^c(m_B, s_{nd})) = B) - 1.$$

In contrast, if $s_d \in \mathcal{S}_d$ is chosen, then $d_{max}(s_d) > d_{max}(s_{nd}) = \frac{n}{2}$, where s_{nd} is arbitrary. Table 8.2 summarizes the performance comparison and $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd})$ for different modulation schemes and security levels d for any non-dispersing seed. In addition, for different modulation schemes and values of d, Table 8.3 shows $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_d)$ where all coset pairs have $d_H(C'(m_1, s_d), C'(m_2, s_d)) = 17.5$, which was the largest Hamming distance we have found for the parameters. Again by our working hypothesis,

$$Adv^{DS_{1a)}}(\xi; E^c; s_d) \geqslant Adv^{DS_{1a)}}(\xi; E^c; s_{nd}) \tag{8.5}$$

for any $s_d \in \mathcal{S}_d$ and $s_{nd} \in \mathcal{S}_{nd}$, which is also supported by the comparison of Tables 8.2 and 8.3. Since

$$Adv^{DS_1}(\xi; E^c; \mathcal{S}) = \frac{1}{|\mathcal{S}|}\Big[\sum_{s_{nd} \in \mathcal{S}_{nd}} Adv^{DS_{1a)}}(\xi; E^c; s_{nd}) + \sum_{s_d \in \mathcal{S}_d} Adv^{DS_{1a)}}(\xi; E^c; s_d)\Big],$$

together with (8.5) we can infer that there exists an $s_d \in \mathcal{S}_d$ such that for all $s_{nd} \in \mathcal{S}_{nd}$,

$$Adv^{DS_{1a)}}(\xi; E^c; s_d) \geqslant Adv^{DS_1}(\xi; E^c; \mathcal{S}) \geqslant Adv^{DS_{1a)}}(\xi; E^c; s_{nd}),$$

which is supported by Fig. 8.4. Additionally, recall that we observed that $|\mathcal{S}_d|$ becomes very small with increasing $l - k$, by which the gap between $Adv^{DS_1}(\xi; E^c; \mathcal{S})$ and $Adv^{DS_{1a)}}(\xi; E^c; s_{nd})$ decreases. Note that since $Adv^{DS_{1a)}}(\xi; E^c; s_{nd}) \leqslant Adv^{DS_1}(\xi; E^c; \mathcal{S})$, we can assume that $\hat{d}_{DS_{1a)}}(s_{nd}) \geqslant \hat{d}_{DS_1}$. Furthermore, we can observe in Table 8.2 that the empirical security level is higher than the theoretical security level, that is, $\hat{d}_{DS_{1a)}}(s_{nd}) > d$. However, this does not apply to $\hat{d}_{DS_{1a)}}(s_d)$.

For completeness, we plotted $\overline{Adv}^{DS_2}(\xi; E^c; \mathcal{S})$ in Fig. 8.4 and we see that as expected, $\overline{Adv}^{DS_1}(\xi; E^c; \mathcal{S}) \geqslant \overline{Adv}^{DS_2}(\xi; E^c; \mathcal{S})$ (see Remark 7.10). According to Fig. 8.4, we can assume that $\hat{d}_{DS_{1a)}}(s_{nd}) \approx \hat{d}_{DS_1} \approx \hat{d}_{DS_2} \geqslant d$. Since the communication scenarios 1 and 2 are difficult to simulate, especially for bigger parameters for l and k, we consider communication scenarios 1a) and 3 in the remaining part.

For QPSK, Fig. 8.5 shows $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_d)$ and $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd})$ as a function of SNR_E. Fig. 8.6 shows the performance comparison for different values of $l - k$, for the same parameters, and the same s_{nd} and s_d as in Table 8.2 and 8.3.

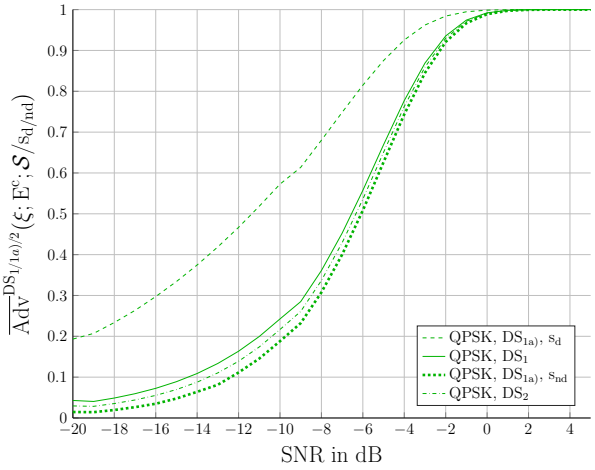


Figure 8.4.: $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_d)$, $\overline{Adv}^{DS_1}(\xi; E^c; \mathcal{S})$, $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd})$ and $\overline{Adv}^{DS_2}(\xi; E^c; \mathcal{S})$ for QPSK, $n = 32$, $l = 12$, $k = 6$ and $n_{crc} = 6$. In the simulation, $d_H(C'(m_1, s_d), C'(m_2, s_d)) = 19$, which is the worst case.

Table 8.2.: The performance comparison with $n = 32$, $l = 18$, $n_{crc} = 6$, $SNR_E = -5dB$ for any non-dispersing seed. For BPSK we choose $SNR_B = 5dB$, for QPSK $SNR_B = 7.5dB$ and for 16QAM $SNR_B = 13dB$.

Modulation	d	$R_{sec}(d)$	$l-k$	$\hat{d}_{DS_{1a)}}(s_{nd})$	$\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd})$	$DER_{E_{1a)}}(s_{nd}, m_1, m_2)$	$\lambda_1(s_{nd}, m_1, m_2)$	$\lambda_2(s_{nd}, m_1, m_2)$
BPSK	0.4080	0.0625	16	1.9467	0.2594	0.3702	0.3753	0.3654
QPSK	0.8360	0.3750	12	2.8059	0.1430	0.4285	0.4485	0.4086
QPSK	2.8360	0.1250	16	4.3688	0.0484	0.4758	0.4753	0.4764
16QAM	0.5524	1.2500	8	2.7583	0.1478	0.4265	0.4111	0.4411
16QAM	2.5524	0.7500	12	6.3688	0.0121	0.4940	0.4951	0.4928

Rate Comparison

In Table 8.4, for non-dispersing seeds and $SNR_E = -5dB$ we compare $R_{sec}(d)$ given in (8.3) with secrecy capacity $C_s(\sigma_T^2, \sigma_E^2, P_{av})$ given in (7.10) and two theoretical secrecy rates $R^*_{sec}(P_{av} - \alpha)$, $R^{**}_{sec}(P_{av} - \alpha)$ with limited modulation alphabet whose secure levels are for asymptotically perfect security.

The secrecy rates are defined as follows.

$$R^*_{sec}(P_{av} - \alpha) := R_{eff} - C_{unif_E}(P_{av} - \alpha) \tag{8.6}$$

and

$$R^{**}_{sec}(P_{av} - \alpha) := R^*_{eff}(c, P_e(\xi, T^c, \zeta)) - C_{unif_E}(P_{av} - \alpha), \tag{8.7}$$

where R_{eff} is given in (7.23) and $R^*_{eff}(c, P_e(\xi, T^c, \zeta))$ is the maximum achievable coding rate for the AWGN channel with $2^{R_{mod}}$-QAM input and average power constraint P_{av} with block error probability $P_e(\xi, T^c, \zeta)$ and blocklength c. More precisely, $R^*_{eff}(c, P_e(\xi, T^c, \zeta))$

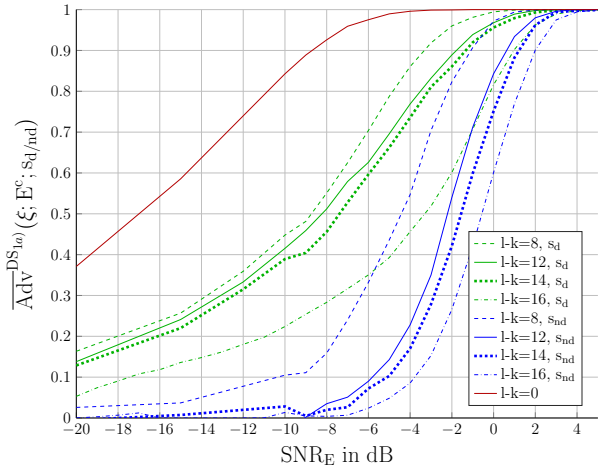


Figure 8.5.: $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_d)$ and $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd})$ for QPSK and the same parameters as in Tables 8.2 and 8.3. When s_d is used, the largest $d_H(C'(m_1, s_d), C'(m_2, s_d))$ we could find is 17.5 for $l - k = \{8, 12, 14\}$ and 16.5 for $l - k = 16$.

Table 8.3.: The performance comparison for $n = 32$, $l = 18$, $n_{crc} = 6$, $SNR_E = -5dB$ and for a certain dispersing seed. The message pairs correspond to the coset pairs with $d_{max}(s_d) = 17.5$. For the case where $l - k = 16$, $d_{max}(s_d) = 16.5$. For BPSK we choose $SNR_B = 5dB$, for QPSK $SNR_B = 7.5dB$ and for 16QAM $SNR_B = 13dB$.

Modulation	d	$R_{sec}(d)$	$l-k$	$\hat{d}_{DS_{1a)}}(s_d)$	$\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_d)$	$DER_{E_{1a)}}(s_d, m_1, m_2)$	$\lambda_1(s_d, m_1, m_2)$	$\lambda_2(s_d, m_1, m_2)$
BPSK	0.4080	0.0625	16	0.7456	0.5964	0.2018	0.2019	0.2018
QPSK	0.8360	0.3750	12	0.5230	0.6959	0.1520	0.1489	0.1551
QPSK	2.8360	0.1250	16	1.2351	0.4248	0.2876	0.2915	0.2837
16QAM	0.5524	1.2500	8	0.7520	0.5938	0.2030	0.2066	0.1996
16QAM	2.5524	0.7500	12	0.8267	0.5638	0.2185	0.2475	0.1886

is given in [71] by

$$R^*_{eff}(c, P_e(\xi, T^c, \zeta)) := C_{unif_T}(P_{av} - \alpha) - \sqrt{\frac{U_{unif}(P_{av} - \alpha)}{c}} Q^{-1}(P_e(\xi, T^c, \zeta)) + \mathcal{O}(\frac{1}{c}), \tag{8.8}$$

where $Q(.)$ denotes the Gaussian complementary CDF and $U_{unif}(P_{av})$ is defined as

$$\begin{aligned} U_{unif}(P_{av}) := \frac{1}{2^{R_{mod}}} \sum_{k=1}^{2^{R_{mod}}} \mathrm{Var}_{N_i}[\log(\sum_{l=1}^{2^{R_{mod}}} e^{\frac{||N_i||_2^2 - ||X'_k + N_i - X'_l||_2^2}{2\sigma_i^2}})] \\ + \mathrm{Var}_{X'}[\mathsf{E}_{N_i}[\log(\sum_{l=1}^{2^{R_{mod}}} e^{\frac{||N_i||_2^2 - ||X' + N_i - X'_l||_2^2}{2\sigma_i^2}})|X']]. \end{aligned} \tag{8.9}$$

Note, that the comparison between $R^*_{sec}(P_{av} - \alpha)$ and $R^{**}_{sec}(P_{av} - \alpha)$ reveals the difference between R_{eff} and $R^*_{eff}(c, P_e(\xi, T^c, \zeta))$. By simulations we measure $R_{sec}(d)$ with different

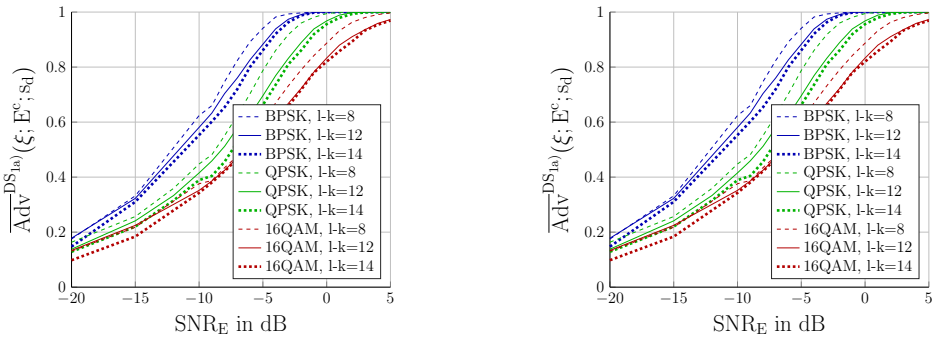


Figure 8.6.: Eve's distinguishing performance in scenario 1a) with s_d where $d_{max}(s_d) = 17.5$ on the left and with any s_{nd} on the right for $n = 32, l = 18, k = \{4, 6, 10\}, n_{crc} = 6$. For BPSK we choose $SNR_B = 5$dB, for QPSK SNR_B= 7.5dB and for 16QAM $SNR_B = 13$dB.

modulation schemes and can observe that as expected $R_{sec}(d) < R^*_{sec}(P_{av}-\alpha) < R^{**}_{sec}(P_{av}-\alpha) < C_s(\sigma^2_T, \sigma^2_E, P_{av})$. For BPSK and QPSK, R_{eff} is much smaller than R^*_{eff} and R_{eff} approaches R^*_{eff} as blocklength increases. We observe that the determination of $R_{sec}(d)$ for a chosen security level d by (8.2) is suboptimal. Note that our focus is primarily on the security aspect and not on the maximum achievable secrecy rates. The $DER_{E_{1a)}}$ that belong to Table 8.4 can be found in Table 8.2.

Table 8.4.: Rate comparison under distinguishing security in communication scenario 1a) for $s = s_{nd}$. $R_{sec}(d)$ is specified by $l = 18$, $n = 32$, $n_{crc} = 6$ and $SNR_E = -5dB$. $P_{av} = 1$, and for BPSK and QPSK $\alpha = 0$ and for 16QAM $\alpha = 0.95$.

	BPSK	QPSK	16QAM
SNR_B	$5dB$	$7.5dB$	$13dB$
C_s	1.661	2.331	3.993
$R^{**}_{sec}(P_{av} - \alpha)$	0.454	1.117	2.146
$R^*_{sec}(P_{av} - \alpha)$	0.213	0.730	1.888
$l - k = 16$			
$R_{sec}(d)$	0.063	0.125	0.250
$l - k = 14$			
$R_{sec}(d)$		0.250	0.500
$l - k = 12$			
$R_{sec}(d)$		0.375	0.750
$l - k = 8$			
$R_{sec}(d)$			1.250

8.2.2. Comparison of Scenarios 1a) and 3

We evaluate the seeded advantage at the eavesdropper under DS_3 security. Since we average over message pairs in contrast to the maximization of scenarios 1 and 1a), we

obtain

$$Adv^{DS_3}(\xi; E^c; \mathcal{S}) \leqslant Adv^{DS_1}(\xi; E^c; \mathcal{S}),$$

so that $d_{DS_3} \geqslant d_{DS_1}$. Since we do not have to maximize the seeded advantage at Eve over the message pair, we save the computational complexity of determining the coset pairs with maximum average Hamming distance. Moreover, we can choose l and k larger with the constraint that $l - k < 22$, due to computation time issues as mentioned above.

In the following, the simulations for DS_3-security were performed in such a way that a random seed S and a random message pair (M_1,M_2) were chosen independently and fixed for 10^3 iterations. The simulations were repeated for 10^3 randomly selected message pairs and randomly selected seeds. To perform the $DER_{E_3}(\xi; E^c; \mathcal{S})$ calculation for DS_3 security, we averaged $\lambda_1(s, m_1, m_2)$ and $\lambda_2(s, m_1, m_2)$ over the number of message pairs transmitted and the seeds selected, so that we get $\bar{\lambda}(\xi; E^c; \mathcal{S})$. We can insert $\bar{\lambda}(\xi; E^c; \mathcal{S})$ into (7.55) to obtain $\overline{Adv}^{DS_3}(\xi; E^c; \mathcal{S})$.

For a randomly chosen s_{nd} and a certain dispersing seed s_d, we compare $DER_{E_3}(\xi; E^c; \mathcal{S})$ with $DER_{E_{1a)}}(s, m_1, m_2)$ in Fig. 8.7 on the left and $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd})$ with $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_d)$ and $\overline{Adv}^{DS_3}(\xi; E^c; \mathcal{S})$ in Fig. 8.7 on the right, for QPSK, 16QAM and 64QAM. In Fig. 8.7 we can see in the SNR range from -20 to 5 dB that $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_d)$ is larger than $\overline{Adv}^{DS_3}(\xi; E^c; \mathcal{S})$, whereas $\overline{Adv}^{DS_3}(\xi; E^c; \mathcal{S}) \approx \overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd})$, and therefore $\hat{d}_{DS_3} \approx \hat{d}_{DS_{1a)}(s_{nd})} > d$. For example for 16QAM and $SNR_E = -5dB$ we get

$$\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd}) = 0.0079,$$

$$\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_d) = 0.4229$$

and

$$\overline{Adv}^{DS_3}(\xi; E^c; \mathcal{S}) = 0.0065.$$

It is not surprising that $\overline{Adv}^{DS_3}(\xi; E^c; \mathcal{S})$ and $\overline{Adv}^{DS_{1a)}}(\xi; E^c; s_{nd})$ are similar, since the mean of the average Hamming distance of all coset pairs for each seed is approximately $\frac{n}{2}$. Thus if an element of $\mathcal{S}_{nd}$ can be identified, communication scenario 1a) should be chosen. The advantage of this scenario is that the seed can remain fixed once it has been identified and thus the complexity of the seeded modular UHF code ca be reduced. Moreover, $Adv^{DS_{1a)}}(\xi; E^c; s_{nd}) \approx Adv^{DS_3}(\xi; E^c; \mathcal{S})$ means that the performance is close to that of the DS_3 security. Unfortunately, if we have a fading wiretap channel instead of an AWGN channel, then the set $\mathcal{S}_{nd}$ for Eve will in general depend on the channel state. Note that because of $Adv^{DS_{1a)}}(\xi; E^c; s_{nd}) \approx Adv^{DS_3}(\xi; E^c; \mathcal{S})$, respectively the advantages and d-secure rates listed in Table 8.2 and 8.4 are similar to $\overline{Adv}^{DS_3}(\xi; E^c; S)$. Furthermore,

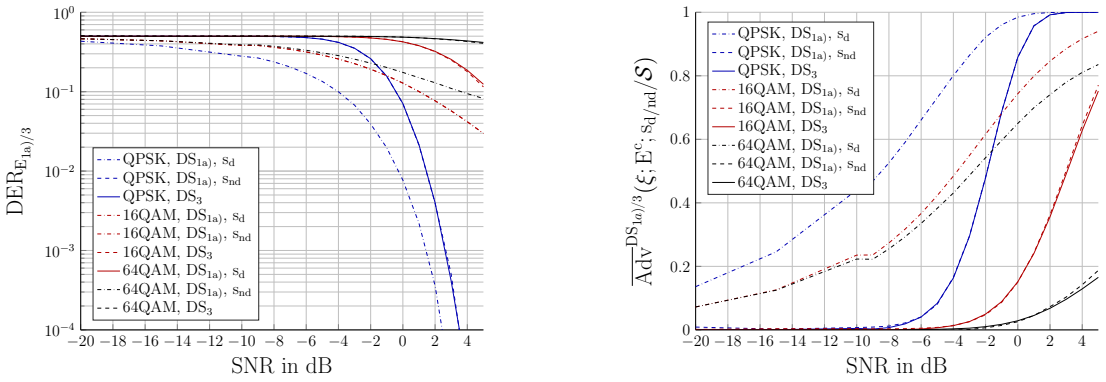


Figure 8.7.: Eve's distinguishing performance in scenarios 1a) and 3 for $n = 48, l = 36, k = 17$ and $n_{crc} = 11$. For QPSK we choose $SNR_B = 10dB$, for 16 QAM $SNR_B = 16dB$ and for 64 QAM $SNR_B = 22dB$. When $s = s_d$, $d_H(C'(m_1, s_d), C'(m_2, s_d)) = 25.5$, which is the largest average Hamming distance we could find.

according to Fig. 8.4 and Fig. 8.7 we can assume that $\hat{d}_{DS_3} \approx \hat{d}_{DS_{1a)}}(s_{nd}) \approx \hat{d}_{DS_1} \approx \hat{d}_{DS_2} \geqslant d$.

8.3. Comparison of Different Attack Strategies of the Eavesdropper

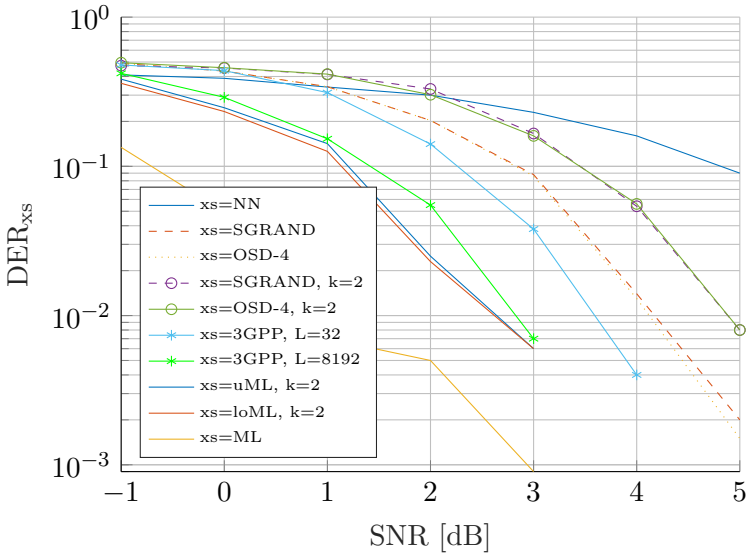


Figure 8.8.: The distinguishing error rate for different attack strategies for $n = 64$, $l = 36$, $k = 12$. We used a non-dispersing seed.

Since the attack strategy proposed in Section 7.4.2 has a high level of complexity, for code parameters $n = 64$, $l = 36$, $k = 12$ we compare different attack strategies by simulation, where some of them approximate the maximum likelihood test for linear block codes and for suitable code parameters. The modulation we choose is QPSK, so that after

modulation the channel input consists of c symbols. Fig.8.8 illustrates the distinguishing performance of the decoding strategies described below. We consider the Soft Guessing Random Additive Noise Decoder (SGRAND) [74] and the Ordered Statistics Decoder of order 4 (OSD-4) [75] modified for our purposes. The SGRAND is an ML decoder for arbitrary additive memoryless channels and tries to identify the noise/error vector that has corrupted the codeword. The algorithm queries error vectors in a specific order and iterates until a condition is met, whereas in the OSD-u algorithm all error vectors starting with Hamming weight 0 up to u are queried. Only then is a condition checked. Several works exist that focus particularly on reducing the complexity of the OSD [76], [77], [78]. It was shown in [75] that for binary transmission over an AWGN channel, reprocessing order equal to $\lceil d(C)/4 - 1 \rceil$ achieves practically optimum ML decoding performance for a block code C of minimum Hamming distance $d(C)$.

Modified SGRAND: The pseudocode for the modified SGRAND is given in Algorithm 1. The modified SGRAND differs in that after the possible estimated codeword x^n has been checked to see if it is an element of the codebook, i.e. $H(x^n)^T = 0$, where $x^n = \Theta(z^c) - e^n$ and Θ is the demodulator, it is additionally checked for another criterion in lines 10-13 and that is whether $f_s(f(x^n)) = m_b$, with $b = 1, 2$. Function f extracts l information bits from x^n to obtain the output vector of the security layer v^l at Alice. If v^l does not belong to any of the message pair then the next error vector is queried until the criteria are met or the maximum number of queries b has been reached. If no admissible x^n could be found in b queries then the decoder randomly chooses one of the two messages m_1, m_2. No erasures are declared in the modified SGRAND as is the case in SGRANDAB [74]. We denote the distinguishing error rate of the modified SGRAND by $DER_{SGRAND}(s, m_1, m_2)$. The worst case complexity of the modified SGRAND is $\mathcal{O}(bn^2)$.

Modified OSD: We have used the code for the OSD from [30] and modified it as given in Algorithm 2. The decoder is adapted for the polar code as proposed for 3GPP standard. The procedure from line 1 to 12 is the same as in [75], with the exception that h is a concatenation of two functions. The first function soft demodulates the channel output observed by Eve $z^c \in \mathbb{C}^c$ and the second recovers the rate as proposed in the 3GPP standard. We obtain a vector of reliabilities $y^n \in \mathbb{R}^n$. In line 13 the OSD-u searches for the error vector e that affects the positions of the hard-demodulated codeword u^n_{HD}, such that the absolute values of the log-likelihood ratios of r^n at the respective positions are small in sum (called min_value in the algorithm), and searches for the error vector that yields the smallest min_value.

Once we have found the error vector, the OSD-u determines the estimated codeword by permuting u^n_{new} using the inverse permutations $\lambda_1^{-1}\lambda_2^{-1}$, i.e. $x^n = \lambda_1^{-1}(\lambda_2^{-1}(u^n_{new}))$. The modified algorithm checks in line 16 whether x^n belongs to one of the messages. Therefore,

in line 14 we determine the output vector of the security layer v^l at Alice, where

$$F = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \tag{8.10}$$

and A is known as the bit-reversal matrix. Note that the $(l + n_{crc} + n_{pc}) \times n$ matrix G is a submatrix of the $n \times n$ matrix B, where n_{crc} is the bit-length of the CRC and n_{pc} is the bit-length of additional parity check bits which are appended to the information bits for certain parameters. We obtain G after omitting the rows on the frozen bit-positions from B. We denote the distinguishing error rate by $DER_{OSD-4}(s, m_1, m_2)$. The decoding complexity of an order-4 OSD can be as high as $\mathcal{O}((l + n_{crc} + n_{pc})^4)$.

Neural Network Decoder: The third attack strategy we consider is a neural network (NN) decoder that we implemented with the help of the deep learning Matlab toolbox. The theory of deep learning is described in [79]. An NN consists of many connected neurons. In such a neuron, all of its weighted inputs are added up, a bias is optionally added, and the result is propagated through a nonlinear activation function, e.g., a rectified linear unit (ReLU) as in our case, which is defined as

$$g_{ReLu}(z) = \max\{0, z\}. \tag{8.11}$$

The NN decoder consists of an input layer, an output layer and so-called hidden layers. Each layer consists of neurons which are connected to neurons of other layers without feedback connections. Each layer i with c_i inputs and k_i outputs performs the mapping $f(i) : \mathbb{R}^{c_i} \rightarrow \mathbb{R}^{k_i}$ with the weights and biases of the neurons as parameters. When y^c is the input of the NN and the output is denoted as m^k, then the mapping is defined as

$$m^k = f(y^c; \Theta) = f^{(L-1)}(f^{(L-2)}(\ldots(f^{(0)}(y^c)))), \tag{8.12}$$

where Θ denotes the weights of the NN and L the number of layers. The weights of the NN which minimize the loss function over the training set can be found by the use of gradient descent optimization methods and the backpropagation algorithm. When training the network we use "Adam", which is a method of stochastic gradient descent optimization [80]. We design an NN decoder that consists of four hidden fully connected layers of respectively 64, 128, 128, 64 neurons. Since the task of our network is to distinguish two messages from each other the input and output layers consist of $2c = 64$ and two neurons, respectively. Note that since the input vector is complex valued, we have split the vector into an imaginary part and a real part, so that each variable corresponds to one feature. We obtain a total vector length of $2c$.

Since the output layer represents which message was sent, that is m_1 or m_2, a softmax function forces the output neurons to be between zero and one. Thus, the output of the softmax function can be used to represent a probability distribution over 2 different

Algorithm 1 High-level description of modified SGRAND

INPUT: b, H, z^c, s, t, m_1, m_2
OUTPUT: $\hat{m}_{out} \in \{m_1, m_2\}$

$g = 0$ {g counts queries performed}
$\mathcal{S} = \{0^n\}$ {$\mathcal{S}$ contains candidate error vectors e^n}
$i = (i_1, ..., i_n)$ = ordered error indices vector {Based on z^n}
while $g \leqslant b$ **do**
  $e^n = \arg\max_{v^n \in \mathcal{S}} p(z^c | \Theta(z^c) - v^n)$
  $\mathcal{S} = \mathcal{S} \backslash \{e^n\}$
  $g = g + 1$
  if $H(\Theta(z^c) - e^n) = 0^{n-k}$ **then**
    $x^n = \Theta(z^c) - e^n$
    $\hat{v}^l = f(x^n)$ {f extracts l information bits}
    $\hat{m} = [(s * \hat{v}^l) \oplus t]_k$
    if $\hat{m} \in \{m_1, m_2\}$ **then**
      $\hat{m}_{out} = \hat{m}$
      return
    end if
  else
    if $e^n = 0^n$ **then**
      $j* = 0$
    else
      $j* = \max\{j : e_{i_j} \neq 0\}$
    end if
    if $j* < n$ **then**
      $e_{i_{j*+1}} = 1$
      $\mathcal{S} = \mathcal{S} \cup \{e^n\}$
      if $j* > 0$ **then**
        $e_{i_{j*}} = 0$
        $\mathcal{S} = \mathcal{S} \cup \{e^n\}$
      end if
    end if
  end if
end while
$\hat{m}_{out} = rand(m_1, m_2)$ {randomly choose $\hat{m}_{out}$}
return

Algorithm 2 High-level description of modified OSD-u for polar decoding

INPUT: u, G, $B = AF^{\otimes \log(n)}$, z^c, s, t, m_1, m_2
OUTPUT: $\hat{m}_{out} \in \{m_1, m_2\}$

$y^n = h(z^c)$
$\hat{y}^n = \lambda_1(y^n)$ with $|\hat{y}_1| \geqslant |\hat{y}_2| \geqslant \cdots \geqslant |\hat{y}_n|$
$G' = \lambda_1(G)$
$G'' = \lambda_2(G') = \lambda_2(\lambda_1(G))$
$r^n = \lambda_2(\hat{y}^n)$ with $|r_1| \geqslant |r_2| \geqslant \cdots \geqslant |r_{l+n_{crc}+n_{pc}}|$
and $|r_{l+n_{crc}+n_{pc}+1}| \geqslant |r_{l+n_{crc}+n_{pc}+2}| \geqslant \cdots \geqslant |r_n|$
$G_{sys} \overset{(.)}{\leftarrow} (G'')$ {. = rowoperation}
u^n_{HD} =HardDecision(r^n)
$min\_value = \infty$
for $1 \leqslant i \leqslant u$ **do**
  for $0 \leqslant k \leqslant |\mathcal{E}_i|$ **do**
    $u^n_{new} = (u^{(l+n_{crc}+n_{pc})}_{HD} \oplus e^{(l+n_{crc}+n_{pc})}_k) \cdot G_{sys}$ $\{u^{(l+n_{crc}+n_{pc})}_{HD} = u_{HD_1}, \ldots, u_{HD_{(l+n_{crc}+n_{pc})}},\ e^{(l+n_{crc}+n_{pc})}_k \in \mathcal{E}_i$, where $e^{(l+n_{crc}+n_{pc})}$ is the error vector of length $(l + n_{crc} + n_{pc})$ with $wt(e^{(l+n_{crc}+n_{pc})}_k) = i$ and $|\mathcal{E}_i| = \binom{l+n_{crc}+n_{pc}}{i}\}$
    $value = \sum_{j:u^n_{new_j} \neq u^n_{HD_j}} |r_j|$
    if $value < min\_value$ **then**
      $\hat{v}^l = w(\lambda_1^{-1}(\lambda_2^{-1}(u^n_{new}))B^{-1})$ {w extracts l information bits}
      $\hat{m} = [(s * \hat{v}^l) \oplus t]_k$
      if $\hat{m} \in \{m_1, m_2\}$ **then**
        $\hat{m}_{out} = \hat{m}$
        $min\_value = value$
      end if
    end if
  end for
end for

possible events. If the probability suggests the label, the loss (e.g., the mean squared error) should be increased only slightly, while large errors should result in a very large loss.

We have collected 700,000 labeled codewords (data) for training the NN and 150,000 independent labeled data, each for validation and testing. The data were trained at an SNR of 10dB, which turned out to be a good SNR. Smaller SNRs led to overfitting. We trained the NN in epochs, where in each epoch the gradient of the loss function is calculated over the training set. The mini batch size is a term that refers to the number of training examples utilized in one iteration. The mini batch size is smaller than the training set and specifies how many iterations complete one epoch.

Since our seeded modular UHF wiretap code consists of a randomized encoder during training the NN is more difficult to generalize to codewords that it has never seen than in the case of structured codes [81]. Generalization means that after training an NN, it is able to find the correct outputs that correspond to new inputs. We have made similar insights as the authors in the paper [81]. The larger we made the randomness, i.e. the larger $l-k$ became, the more examples were necessary to train the NN, or they had to be trained over more epochs. This is due to the fact that the code looses structure with increasing $l-k$. It should be noted that the NN decodes without knowledge of the SNR. In addition, it learns the channel distribution itself with the help of the labeled data in comparison to the ML decoder, where the knowledge of the channel distribution is assumed. Another advantage is that, after training the NN decoder for fixed code parameters, the decoding effort is small compared to the ML decoder. The distinguishing error rate of an NN decoder is

$$DER_{NN}(s, m_1, m_2) = 1 - (accuracy/100), \tag{8.13}$$

where $accuracy$:= (number of correct predicted messages/total number of messages to be predicted). The total number of messages to be predicted is equivalent to the dataset size.

Modified Polar Decoder: Furthermore, we compare the above attack strategies with the modified CRC-aided successive cancellation list (SCL) decoder. We use the CRC-aided SCL decoder as proposed by the 3GPP standard. Since Eve has to decide between a message pair (m_1, m_2), we have modified the decoder to additionally calculate the Hamming distance between the decoded message m (which can be any message from the message space $\mathcal{M}$) and message m_1 and m_2, respectively, and output the message that has a smaller Hamming distance to the original decoded message. If both pairs have the same Hamming distance, that is $d_H(m, m_1) = d_H(m, m_2)$, then the decoder randomly chooses m_1 or m_2. We denote the distinguishing error rate by $DER_{3GPP}(s, m_1, m_2)$ and report the results with list sizes $L = 32$ and $L = 8192$ (see Fig. 8.8).

To evaluate the distinguishing performance of the considered attack strategies for $l-k >$

Table 8.5.: Key parameters of Adam

Parameter	
miniBatchSize	500
InitialLearnRate	0.002
MaxEpochs	1000
GradientDecayFactor	0.9

24 we have determined an upper bound for the ML test numerically. Given the channel output z^c, the standard polar decoder (SCL decoder) generates a list set $\mathcal{L}$ with $L = |\mathcal{L}|$ most probable codewords. We reduce the list set to the codewords from coset $C'(m_1)$ and $C'(m_2)$ and obtain the list set $\mathcal{L}'_b = \mathcal{L} \bigcap C'(m_b)$, $b = 1, 2$. For the case where the channel is AWGN, instead of (7.58) we obtain the following conditional probability densities when message m_b with $b \in \{1, 2\}$ was sent

$$p_{approx}(z^c|s, m_b) = \frac{1}{|\mathcal{L}'_b|} \sum_{v:\mathcal{L}'_b} \prod_{i=1}^{c} w(z_i(m_b, s)|\chi(\Psi(v))_i), \tag{8.14}$$

where $\chi(\Psi(v))_i$ denotes the i-th symbol in the length-c channel input $\chi(\Psi(v)) \in \mathbb{C}^c$.

Since the list set $\mathcal{L}'_1 \cup \mathcal{L}'_2$ is smaller than $\mathcal{L}$, a higher decoding error probability is expected and as soon as $L \to 2^l$, DER_{uML} approaches the DER_{ML} of the ML test from Section 7.4.2. Therefore, we use this decoder to determine the upper bound of the DER_{ML}. In addition, an error is declared if $\mathcal{L}'_1 = \mathcal{L}'_2 = \varnothing$. We observed that as the list size L increases, the upper bound approaches the ML test. It is worth noting that here the coset pairs are not generated, instead the messages corresponding to the codewords from the list $\mathcal{L}$ are computed and then the codewords are partitioned into $\mathcal{L}'_1$ and $\mathcal{L}'_2$. We determine the list size L using a lower bound on the DER_{ML}. The calculation of the lower bound DER_{loML} differs from the calculation of the upper bound in that the set $\mathcal{L}$ is extended by the transmitted codeword x^n if it does not already appear in the list. Thus the decoder makes an error as soon as it finds a codeword that corresponds to the wrong coset, which leads to a greater probability of the channel output conditioned on this codeword than conditioned on the actually sent x^n. With increasing L, the DER_{loML} converges to the DER_{ML} from below. The complexity of the calculation of the lower and upper bound is $\mathcal{O}(Ln \log n)$, while the complexity of the calculation of the ML test is $\mathcal{O}(n2^{(l-k)})$. In Fig. 8.9 for $n = 32$, $l = 18$, $k = 2$ and for $L = 512, L = 1024, L = 2048$, DER_{loML}, DER_{uML} and DER_{ML} are illustrated. In the SNR_E range of interest to Alice (that is where $DER \approx 0.5$) we observe that the larger $\mathcal{L}$ is, the closer DER_{uML} and DER_{loML} to DER_{ML} are and the more accurately we can estimate DER_{ML}. In Fig. 8.8, $L = 8192$ was used to determine the upper and lower bound for DER_{ML}.

Using the lower bound, for given n, k, SNR_E, we determined the list size L for any security level d given in (8.2) and the admissible $l - k$ that provide $DER_{loML} \approx 0.5$.

We started with $d = 3$. If the lower bound does not converge to $DER = 0.5$ with increasing list size up to a certain value, l is increased up to the greatest prime $< n$, where $l = \text{prime} - 1$ and thus d as well. Then L is again increased for the new code parameters. The process is repeated until a list size with code parameters is chosen so that the lower bound converges to $DER = 0.5$, or until l reaches the closest value prime-1 to n. If no l could be found with $DER_{loML} \approx 0.5$, then the security criteria is not satisfied for the given n. For complexity reasons we have limited the list size to $L = 32768$. Some list sizes and the corresponding code parameters are listed in Table 8.6.

For small blocklengths, we can observe that R_{eff} must be high to transmit at a positive d-secure rate $R_{sec}(d)$. For fixed SNR_E, and with increasing blocklength and fixed rate R_{eff}, $R_{sec}(d)$ increases.

Furthermore, we have observed that for given n, SNR_E, and modulation, the decoding performance of Eve depends only on $l - k$ and not on how large k and l are in detail, i.e. with $l = 100$ and $k = 20$ Eve decodes with the same error as with $l = 82$ and $k = 2$.

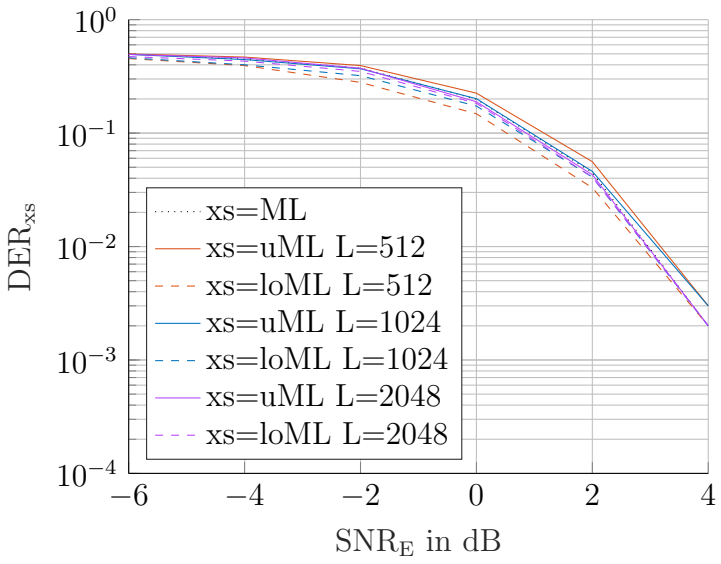


Figure 8.9.: DER_{uML}, DER_{loML} and DER_{ML} for QPSK, $n = 32$, $l = 18$, $k = 2$ and $L = 512, 1024, 2048$.

In Fig. 8.8 we can observe that the modified SGRAND and the modified OSD-4 perform equally. From 1dB, the modified SGRAND performs worse than the modified polar decoder with $L = 32$. The NN decoder performs better below 1.5dB than the modified SGRAND and Polar decoder with $L = 32$. Above 1.5dB, however, the NN decoder performs worse. Unfortunately, compared to our ML test, the modified SGRAND with $b = 10^7$ and the modified OSD-4 perform much worse. In Fig. 8.8, we observe that the performance of the modified SGRAND and the modified OSD-4 depends on $l - k$, where $l = 36$ and k vary. Everywhere where $k = 2$ is selected is labeled in the legend. It is evident that for an increasing random vector at the encoder, i.e. with increasing $l - k$ where l is fixed and k decreases, the DER_{ML} approaches the DER of the alternative

Table 8.6.: Code parameters that provide $DER_{Eve} \approx 0.5$

SNR_E [dB]	n	$l-k$	L	d
1	150	110	8192	10.79
2	150	124	8192	11.84
2	512	376	32768	21.60
3	150	136	16384	11.94
3	512	418	32768	22.49

decoders. This is because the distinguishing performance of the modified polar decoder and the modified SGRAND or the modified OSD-4 depends on $R_{FEC} = l/n$ and, respectively, is hardly and little affected by the size k, while the performance of the ML test with varying k changes drastically as shown in Fig. 8.8. The DER_{3GPP} approaches the DER_{ML} as the list size increases. Our main observation is that the best attack strategy for Eve is to decode as we do, to obtain the upper bound on the DER_{ML}.

8.4. Conclusion

The main objective of this work was to calculate the seeded advantage at the eavesdropper Eve under distinguishing security and strong security for small blocklengths, using a seeded modular UHF code for an AWGN channel. We measured the seeded advantage as a function of security level d. We used d to estimate the necessary amount of encoding randomness at given code parameters and channel parameters. We then have simulated Eve's advantage in two communication scenarios, each reflecting the operational meaning of different security measures and different assumptions about Eve's capability. With the help of the simulations we have gained important insights. For admissible code parameters, we have observed that the advantage at Eve is close to zero even for small blocklengths. We went even deeper by analyzing the impact of seeds on Eve's distinguishing performance. We observed that for given code parameters, some seeds increase Eve's advantage under distinguishing security. We made the important observation that the selection probability of such seeds decreases exponentially with the length of the random vector $l-k$. This means that a selection criterion can be met by the seed set at the transmitter for a given code to avoid a worst case scenario. There are several reasons why it is desirable to identify and fix a non-dispersing seed for application in the security layer. We already mentioned reducing the complexity of the coding scheme. Another advantage of having a fixed seed is that if the seed were to be chosen anew for every transmission, this seed would have to be made known to Alice and Bob before every message transmission. This would cause a dramatic rate loss and should be avoided. Furthermore, for BPSK and QPSK we could observe in simulations that the average Hamming distance of a coset pair, corresponding to a certain message pair correlates positively with Eve's seeded ad-

vantage when the sample size is large enough. This simulatively confirms our working hypothesis that increasing the average Hamming distance of a coset pair improves Eve's performance and thus increases her advantage. A further problem is the computational complexity, so that we only were able to simulate the performance of Eve for $l - k < 22$, so that we had to choose small and negative SNR_E to achieve acceptable results (e.g. advantage at Eve ≈ 0). For future work, the search for a universal attack strategy for Eve with similar performance as the ML test but lower complexity is important for further analysis. Simulations with other FEC codes are of interest as well. We have observed that the security, e.g., measured in bits, is higher than theoretically estimated. This may be due to the loose upper bound, so that in future work other upper bounds could be used.

9. Experimental Evaluation of a Modular UHF Code

9.1. Introduction

Physical layer security (PLS) can provide provable information-theoretic security, in contrast to public key encryption that relies on the unproven assumption that certain mathematical operations, e.g., factorization of large prime numbers, are computationally hard to invert. Additionally, PLS systems do not require a secret key exchange. Therefore, we consider a modular UHF scheme, which is given in Chapter 7, for PLS consisting of three layers; a modulation layer, an error-correction layer and a security layer. Under this approach, an existing forward error-correction (FEC) code is used, preceded by a pre-processing step responsible for the security. This scheme has the advantage that well-researched FEC's can still be used, easing PLS integration in deployed systems.

Contribution

We experimentally evaluate the modular universal hash function (UHF) code that is given in Chapter 7, where software defined radios (SDR's) represent Alice, Bob, and Eve. In order to avoid external radio conditions from affecting our experiments, the SDR's are connected via coaxial cables. The wiretap setup is implemented using splitters and combiners: Alice's transmit signal is split in two channels, where two independent Gaussian noise sources are connected using combiners. Two noise generators with different power levels are used for this purpose. This realizes the different channel statistics required by our model. Before the corresponding signals are fed to Bob and Eve, they are attenuated 30 dB to keep a link budget below the saturation level of the analog-to-digital converters (ADC's) at the receiving Universal Software Radio Peripherals (USRP), thus avoiding signal clipping. We use a distinguishing security metric that can be evaluated experimentally (to asses Eve's performance) and is independent of the message distribution. In real signal transmission, we measured the performance of Eve and compared it with the simulation results. We observed that the experimental results are close to the simulation results. This means that the synchronization and signal processing algorithms implemented for our experimental setup do not contribute to the degradation of the communication at the bit level. To the best of our knowledge, this is the first time that PLS in wiretap channels

using a modular scheme is evaluated experimentally.

Related Work

In the literature, novel codes that achieve PLS and provide both reliability and security simultaneously have been investigated. For instance in [59], the secrecy performance of LDPC codes is studied for a uniform message distribution. Despite their practical approach, the integration of such codes into existing systems would demand major design changes. A different approach is presented in [82], where three-layer wiretap codes for the AWGN wiretap channel are evaluated using a restricted security analysis based on mutual information that does not allow the eavesdropper to perform arbitrary operations on the received data. Furthermore, neither [59] nor [82] has been experimentally validated yet.

Outline

In the next section, we briefly reproduce the modular UHF code. In Section 9.3, we give the security measure and its operational meaning in terms of the error probability to evaluate the performance of Eve. In Section 9.4, we present the experimental setup of the communication system at the signal processing level and set the key parameters. In Section 9.5, we compare the experimental results with the simulation results provided by Matlab and conclude the paper with Section 9.6.

9.2. The Seeded Modular Code for the Wiretap Channel

We briefly reproduce the seeded modular code from Chapter 7 and sum up some useful information needed for the experimental setup.

9.2.1. Security Layer

Recall that we use a UHF that was proposed by Hayashi and Matsumoto [56]. We assume that all participants have knowledge of a seed $s = (a, t)$. The two components of the seed, a and t, are bit strings of length l, and randomly chosen from $\{0,1\}^l \backslash \{0\}^l$ and $\{0,1\}^l$, respectively. Messages come from the set $\mathcal{M} = \{0,1\}^k$. For a message m and a bit string $r \in \{0,1\}^{l-k}$, $k < l$, we define the mapping $f_s^{-1} \colon \{0,1\}^k \times \{0,1\}^{l-k} \to \{0,1\}^l$ according to

$$f_s^{-1}(m, r) = a^{-1} * ((m||r) \oplus t), \tag{9.1}$$

where $m||r$ denotes the concatenation of the bit strings m and r, a^{-1} is the inverse, $*$ the multiplication, and $\oplus$ the addition in the corresponding field $\mathbb{F}_{2^l}$. At the security layer on Alice's side, for a given message $m \in \mathcal{M}$, we randomly choose a bit string r and compute $v = f_s^{-1}(m, r)$. The bit string v is then further processed by the error-correction layer.

At the security layer on Bob's side we have to reverse the action of f_s^{-1}. To this end, we use the mapping $f_s\colon \{0,1\}^l \to \{0,1\}^k$, defined by

$$f_s(\hat{v}) = \left[(a * \hat{v}) \oplus t\right]_k , \tag{9.2}$$

where $[x]_k$ denotes the operation of selecting the first k bits of x. f_s is applied on the output $\hat{v}$ of the coding layer, which results in $\hat{m} = f_s(\hat{v})$. If the transmission over the channel T has been error free, i.e. if $\hat{v} = v$, then we have $\hat{m} = m$ because $f_s(f_s^{-1}(m, r)) = m$ for all s, m, and r.

Although the function f_s^{-1}, which was given in (9.1), is a function of two arguments - the message m and a randomly chosen bit string r - it will be convenient to interpret it as mapping with only one argument m. Recall that the randomized inverse of the UHF of $f_s^{-1}(m)$ provides a stochastic mapping from messages to the FEC inputs, so that as the message m is chosen, the encoder chooses uniformly at random a vector v from the set $\{v' : f_s(v') = m\}$, and encodes it to a codeword x^n via the FEC code.

9.2.2. System Integration of the Security Layer

In this section, we describe how the security layer integrates into the communication system. The used mappings and variables are displayed in Fig. 9.1.

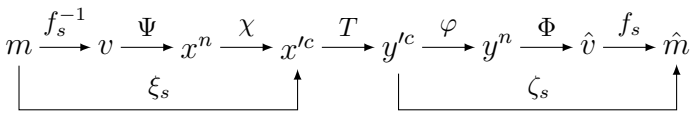


Figure 9.1.: Modular coding scheme from Alice to Bob.

Error-Correction Layer: Alice encodes the output of the security layer v, using some forward error-correction (FEC) code with encoder-decoder pair (Ψ, Φ) of rate $R_{\text{FEC}} = l/n$. Hence, we have $x^n = \Psi(v) \in \{0,1\}^n$. For this layer, we use polar codes, which are also used in the 5G New Radio standard [73]. The core components of the FEC encoder are the cyclic-redundancy-check (CRC) encoder with CRC lengths of 6 and 11 bits, the polar encoding kernel, and the rate matcher. At Bob, the soft-demodulated channel outputs are transformed into log likelihood ratios (LLR), which are rate recovered and then decoded with a CRC-aided successive cancellation list decoder of list size $L = 8$. We use the implementation of the polar encoder and decoder in the 5G Toolbox in Matlab.

Modulation Layer: For modulation, we consider quadrature phase shift keying (QPSK). We denote the corresponding symbol alphabet by $\mathcal{X}' \subset \mathbb{C}$. It has size $2^{R_{\text{mod}}}$, where $R_{\text{mod}} = 2$ denotes the number of bits per symbol. By χ we denote the constellation mapper and by φ the constellation demapper. Alice modulates x according to $x'^c = \chi(x^n) \in \mathcal{X}'^c$, where $c = n/R_{\text{mod}}$. The modulation scheme follows a Gray encoding. The product of R_{FEC} and R_{mod} gives the effective rate $R_{\text{eff}} = l/c$.

We assume that the channel to Bob and the channel to Eve are both complex additive white Gaussian noise (AWGN) channels. Since the channel is memoryless, we obtain after c channel uses $y'^c = x'^c + w_T$ and $z^c = x'^c + w_E$, where $x' \in \mathcal{X}'^c$ is the channel input, while $y'^c \in \mathbb{C}^c$ and $z^c \in \mathbb{C}^c$ are the channel outputs at Bob and Eve, respectively. w_T and w_E are complex circularly-symmetric Gaussian random vectors.

Combining the security layer and the traditional coding layer, we obtain the total encoding function $\xi_s = \chi \circ \Psi \circ f_s^{-1}$ and the total decoding function $\zeta_s = f_s \circ \Phi \circ \varphi$. The mapping $\xi_s \colon \mathcal{M} \to \mathcal{X}$ is the stochastic seeded encoder and $\zeta_s \colon \mathcal{Y} \to \mathcal{M}$ the seeded decoder. We call $(\xi, \zeta) = (\{\xi_s\}_{s\in\mathcal{S}}, \{\zeta_s\}_{s\in\mathcal{S}})$ a seeded modular code. Its secrecy rate is $R_{\text{sec}} = \frac{k}{l} R_{\text{eff}} = k/c$.

9.2.3. Decoding at Bob

Bob receives a noisy version $y'^c = T(x'^c)$ of the channel input $x'^c = \xi_s(m)$, and his goal is to decode the message m correctly. To this end, Bob computes $\hat{m} = \zeta_s(y'^c)$. Since we use a polar code for the traditional coding layer, we use, in fact, a soft demodulation at Bob.

The error probability of the seeded modular code (ξ, ζ) and channel T is given by

$$P_{\text{e}}(\xi, \zeta, T) = \max_{s\in\mathcal{S}} \max_{m\in\mathcal{M}} \left[\Pr\left(\zeta_s(T(\xi_s(m))) \neq m\right)\right]. \tag{9.3}$$

Note that the error probability as defined in (9.3) is a worst case error probability, where we maximize over all possible messages $m \in \mathcal{M}$ and seeds $s \in \mathcal{S}$. The error depends only on the FEC code, the modulation mapping and the channel.

9.2.4. Information-Theoretic Security

Eve should learn as little as possible about the message m when observing the output of the channel E. Traditionally, information-theoretic security is measured in terms of entropy or mutual information. Since we are interested in experimentally measuring Eve's "advantage" of learning the message, we need a metric with an immediate operational meaning. Thus, we employ the distinguishing security (DS) metric

$$\begin{aligned} &Adv^{DS_{1a)}}(\xi, \mathcal{S}, E) \\ &= \frac{1}{|\mathcal{S}|} \sum_{s\in\mathcal{S}} \max_{\mathcal{A}, m_1, m_2} 2\Pr[\mathcal{A}(s, m_1, m_2, E(\xi_s(m_B))) = B] - 1, \end{aligned} \tag{9.4}$$

which has been introduced in Chapter 7. B is a uniformly distributed random variable over $\{1, 2\}$ and can be seen as a random challenge bit. In (9.4) we maximize all messages m_1, m_2 and all adversary strategies $\mathcal{A}$.

The closer $Adv^{DS_{1a)}}$ is to zero, the more secure a security system is. For further information about security metrics, see Chapter 7 and [13].

9.3. Performance Evaluation

We take non-dispersing seeds, for which Eve's performance does not depend on the message pair selected for transmission. Non-dispersing seeds provide a stronger level of security than the others. We use such seeds in our experiments and denote them by $s = (a, t)$ in the following.

When evaluating (9.4) for a fixed seed s, we obtain

$$\max_{\mathcal{A}} 2\Pr[\mathcal{A}(s, m_1, m_2, E(\xi_s(m_\Theta))) = B] - 1. \tag{9.5}$$

The maximum over m_1, m_2 can be omitted due to the choice of the seed, and thus m_1, m_2 can be fixed arbitrarily. Alice randomly chooses $B \in \{1, 2\}$ and transmits the message m_B. Eve receives $E(\xi_s(m_B))$, and, based on that information, has to decide whether m_1 or m_2 was sent. The attack strategy of Eve in the experiments is to use a maximum likelihood (ML) decoder, which is given in Section 7.4.2.

The distinguishing error probability $\mathrm{DER_E}$, i.e., the probability that Eve decides incorrectly, is given by

$$\mathrm{DER_E} = \Pr[\hat{m}_{\mathrm{Eve}} \neq m_B]. \tag{9.6}$$

Then (9.5) is equal to $1 - 2\mathrm{DER_E}$, and $\mathrm{DER_E}$ close to $1/2$ means "high security".

Note that Eve has to decide which one out of two given messages was sent. In contrast, Bob decodes ordinarily without this additional information and tries to determine which message out of the set of all possible messages $\mathcal{M}$ was sent. Hence, the decoding task of Bob in our setting is more intricate than the decoding task of Eve. Bob's block error rate is given by

$$\mathrm{BLER_B} = \Pr\left(\zeta_s(T(\xi_s(m))) \neq m_B\right).$$

Performing the maximum likelihood decoding as in (7.58) requires the computation of all words in the set $\{v' : f_s(v') = m_B\}$. The size of this set, and consequently the time needed to evaluate (7.58) grows exponentially in $l - k$. Thus, the ML decoding at Eve is computationally feasible only up to $l - k < 22$.

The pseudo code given in Algorithm 3 summarizes how $\mathrm{BLER_B}$ and $\mathrm{DER_E}$ are determined.

9.4. Experimental Setup

9.4.1. Hardware Setup

The experimental setup consists of three NI USRP-2954R software defined radios (SDRs) representing Alice, Bob, and Eve. In order to have reproducible conditions, the SDRs are connected via coaxial cables, as indicated in Fig. 9.2. Alice's transmit signal is split, and white Gaussian noise is added from an R&S SMW200A signal generator, which includes

Algorithm 3 High-level description of the performance evaluation of the seeded modular coding scheme

INPUT: $\mathrm{SNR_B}, \mathrm{SNR_E}, c, l, k, \mathrm{num\_codewords}, s, m_1, m_2$
OUTPUT: $\mathrm{BLER_B}, \mathrm{DER_E}$

codeword_errors_Bob := 0;
codeword_errors_Eve := 0;
for $j = 1$ to num_codewords **do**
  choose $m \in \{m_1, m_2\}$ randomly
  choose v randomly from $f_s^{-1}(m)$
  $x'^c = \chi(\Psi(v))$
  Decoder of Bob:
  $y'^c = T(x'^c)$
  $\hat{m} = \zeta_s(y'^c)$
  if $\hat{m} \neq m$ **then**
    codeword_errors_Bob := codeword_errors_Bob + 1;
  end if
  Decoder of Eve:
  $z^c = E(x'^c)$
  $\mathrm{LLR}(z^c|s, m_1, m_2) \gtrless_{\hat{m}_{\mathrm{Eve}}=m_2}^{\hat{m}_{\mathrm{Eve}}=m_1} \log(1)$
  if $\hat{m}_{\mathrm{Eve}} \neq m$ **then**
    codeword_errors_Eve := codeword_errors_Eve + 1;
  end if
end for
$\mathrm{BLER_B}$ = codeword_errors_Bob/num_codewords
$\mathrm{DER_E}$ = codeword_errors_Eve/num_codewords

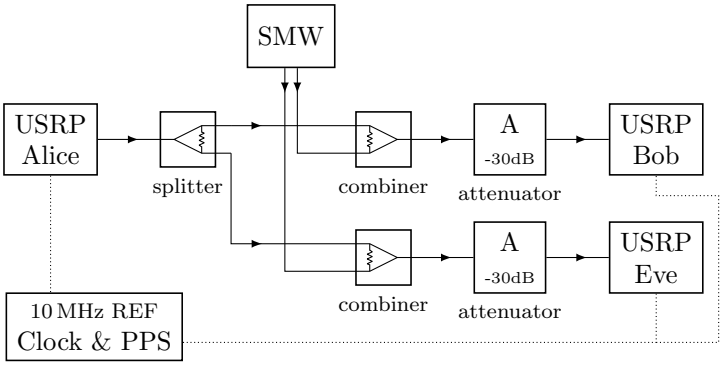


Figure 9.2.: Hardware setup for the experiments.

two separate SMW-K62 noise generators. The resulting signals are attenuated by 30 dB and fed into Bob and Eve.

The USRP gains and SMW noise power have been selected after a careful budget link planning in order to avoid signal clipping and to minimize quantization errors at the receivers.

All USRPs are synchronized using a clock distribution system CDA-2990. Both clock and pulse-per-second (PPS) signals are generated by a GPS disciplined clock with an accuracy of 5 parts per billion.

9.4.2. Communication Scheme

We have deployed a single-carrier transmission communication protocol. Alice sends messages to Bob in the form of periodic bursts, e.g., 32.768 ms for $n = 28$, with sampling rate f_r and using the frame structure shown in Fig. 9.3. Since we are concerned with experimentally demonstrating information-theoretic security, the frame structure and digital signal-processing (DSP) algorithms have been chosen to minimize DSP-related errors. For frame synchronization we employ a Barker sequence of length N_{sync}, which is repeated twice [83]. For phase ambiguity resolution we employ a Gold sequence of length N_{pilot} [84]. Both of these sequences are used for phase offset estimation. The padding sequence of length N_{padding} separates the noise-free signal on its left from the noisy signal on its right, as discussed in Section 9.4.4. The SNR is estimated using the second Gold sequence after countering the channel effects on it. The entire preamble is modulated using BPSK.

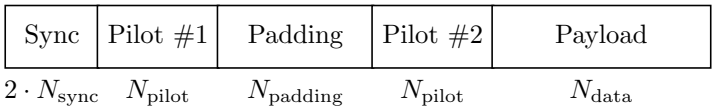


Figure 9.3.: Frame structure of the transmission scheme.

The transmitted payload is $N_{\text{data}} = N'_{\text{data}}/\log_2(M_{\text{data}})$ symbols long, where N'_{data} is the length of the bit sequence and $M_{\text{data}} = 2^{R_{\text{mod}}}$ is the modulation order used for the payload.

9.4.3. Signal Processing Implementation

The DSP steps performed after encoding the message using the modular scheme are illustrated in Fig. 9.4. The generated codewords are encapsulated into protocol data units (PDU) by either segmenting the bitstream in chunks of N'_{data} bits or adding padding bits, in case the codeword is smaller than the PDU. In our experiments, the PDU length has been chosen in order to send one frame per transmission only. The PDU is then converted into I/Q symbols using a Gray-encoded constellation mapper. Next, the preamble is appended, which contains the synchronization, padding, and pilot symbols. Finally, this stream of I/Q symbols is pulse-shaped, using a square-root-raised-cosine finite impulse

response (FIR) filter and upsampled, using a factor F before being transferred to USRP-Alice. The SDR transforms the digital baseband signal to the analog domain using a 16-bit digital-to-analog converter (DAC) with a sampling rate of f_r, before its RF front-end upconverts it using carrier frequency f_c, amplifies it with gain G_{tx} and sends it over the channel.

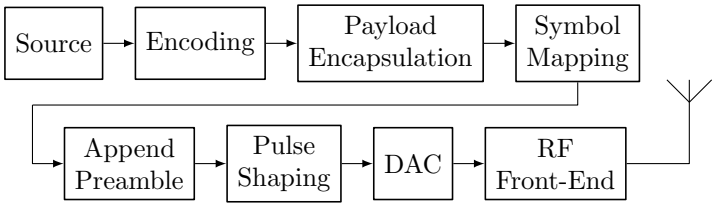


Figure 9.4.: Transmit signal processing.

At the receiver end, the analog bandpass signal is filtered, amplified using the receiver gain G_{rx}, and downconverted to baseband at the RF-frontend before being digitized by a 14-bit analog-to-digital converter (ADC). The carrier phase offset is compensated using a phase-locked loop (PLL) synchronizer [85, p. 333] before matched filtering (MF) takes place. The SNR maximization feature of MF is exploited to perform timing recovery via the output power maximization algorithm [86, p. 261] followed by downsampling. The start of the frame is then identified using a cross-correlation algorithm, which exploits the good autocorrelation property of the Barker sequences. Phase ambiguity is finally resolved by using the known pilots [85, p. 366]. These steps are depicted in Fig. 9.5.

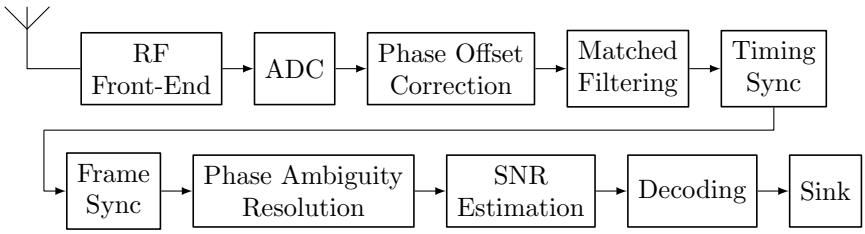


Figure 9.5.: Receive signal processing.

Real-time operation is realized through *time-based synchronization* via PPS signals, a common reference clock, and parallel modular processes with dedicated CPU affinities and hierarchical priorities. Data exchange among threads is managed via signaling notifications and queuing buffers.

Table 9.1 summarizes key implementation parameters.

9.4.4. Guaranteeing Correct Low SNR Measurements

In order to achieve the SNRs required at Eve while ensuring correct signal processing, the SMW delays the initiation of the AWGN noise generation until the time interval assigned for padding symbols is reached. This is selected long enough to account for a 1–2 ms jitter

Table 9.1.: Key Parameters of the Experimental Setup

Parameter	Variable	Value
Barker sequence length	N_{sync}	13 symbols
Pilot sequence length	N_{pilot}	128 symbols
Padding sequence length	N_{padding}	256 symbols
PDU length	N'_{data}	512 , 2048 bits
Up- / downsampling factor	F	16
Modulation (preamble)	M_{preamble}	BPSK
Modulation (payload)	M_{data}	QPSK
Carrier frequency	f_c	2.437 GHz
USRP bandwidth	B	50 kHz
I/Q sampling rate	f_r	390 625 Sps
USRP transmit, receive gain	G_{tx}, G_{rx}	28 dB, 18 dB
SMW noise bandwidth	B_{noise}	100 kHz

Table 9.2.: Key parameters of the experiments

Parameter	Exp. 1	Exp. 2	Exp. 3	Exp. 4
n	28	28	128	28
l	18	18	72,78,88,96	18
k	4	4	16	4,10,18
CRC length	6	6	11	6

in the SMW's response time. Such process is repeated periodically for each transmitted frame, i.e. disabling the SMW shortly after a frame has been received.

9.5. Results

In our first experiment, we use the two messages $m_1 = (0,0,0,0)^{\text{T}}$ and $m_2 = (0,0,0,1)^{\text{T}}$, as well as the seed $s = (a,t)$ with $a = (0,0,1,1,1,1,0,1,0,1,1,1,1,1,1,1,1,1)^{\text{T}}$ and t identically zero. This seed has the properties discussed in Section 9.3. Further system parameters are listed in Table 9.2. During the experiment, the noise power in the channel to Bob is fixed, and the one in the channel to Eve is changed in 0.25 dB steps. For each setting, we repeat the measurement at least 484 times, corresponding to at least 13,568 codewords. The SNRs at Bob SNR_{B} and Eve SNR_{E} are estimated based on the noisy pilot signal using a data-aided ML estimator [87].

We measured $\text{SNR}_{\text{B}} = 8.1\,\text{dB}$. Having a fixed SNR_{B}, we virtually obtain the same error rate $\text{BLER}_{\text{B}} = 0.00086$ for all measurements from one batch. For Eve, we determine the distinguishing error rate DER_{E} as well as the BLER_{E}, the block error rate when Eve applies the same decoder as Bob. Note that the BLER_{E} is not considered to be a security metric in the strict sense. The values of DER_{E} and BLER_{E} for different SNR_{E} values are shown in Fig. 9.6. The lower the SNR_{E}, the better the security level of the scheme is,

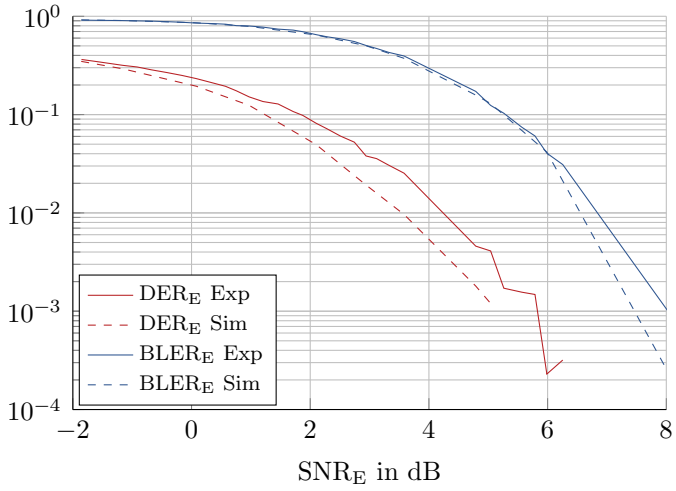


Figure 9.6.: $\mathrm{DER_E}$ and $\mathrm{BLER_E}$ as a function of $\mathrm{SNR_E}$ (n=28; l=18; k=4).

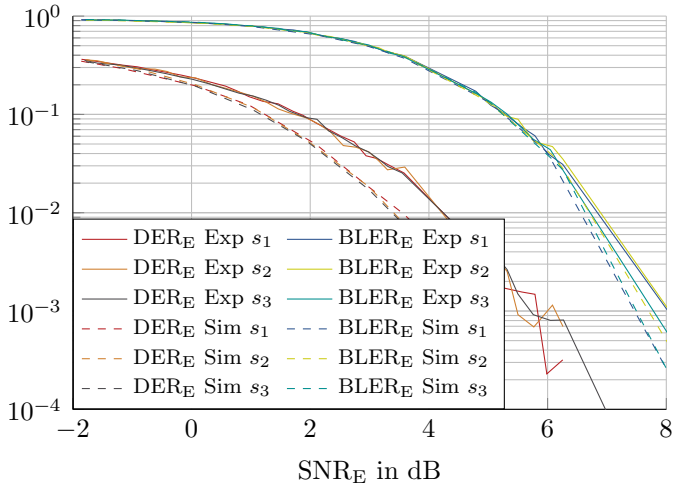


Figure 9.7.: Comparison of 3 different seeds (n=28; l=18; k=4).

as measured by the distinguishing error rate $\mathrm{DER_E}$. At an $\mathrm{SNR_E}$ of $-1.9\,\mathrm{dB}$ we have a $\mathrm{DER_E}$ of 0.36. For comparison, at this $\mathrm{SNR_E}$, the $\mathrm{BLER_E}$ is 0.91. If we compare this $\mathrm{SNR_E}$ to the $\mathrm{SNR_B}$, we observe that there is quite a large difference of 10 dB. In our next experiments we will see that this SNR margin depends on the blocklength and the difference $l - k$. In addition, the simulated values generated with $5,000$ codewords are plotted for comparison. We see that the experimental results (solid curves) are close to the simulation results (dashed curves).

In our second experiment, we use the same messages m_1 and m_2, and compare three different seeds s_1, s_2, s_3, each with the properties discussed in Section 9.3. As before, t is chosen to be identically zero for all three seeds, and a is given by

$a_1 = (0,0,1,1,1,1,0,1,0,1,1,1,1,1,1,1,1,1)^{\mathrm{T}}$,

$a_2 = (0,0,1,1,1,1,0,1,0,1,1,1,1,1,1,1,1,0)^{\mathrm{T}}$,

$a_3 = (1,0,1,0,1,1,0,1,0,1,1,0,1,1,1,1,0,0)^{\mathrm{T}}$.

In Fig. 9.7 we see that the distinguishing error rate $\mathrm{DER_E}$, and hence the security level,

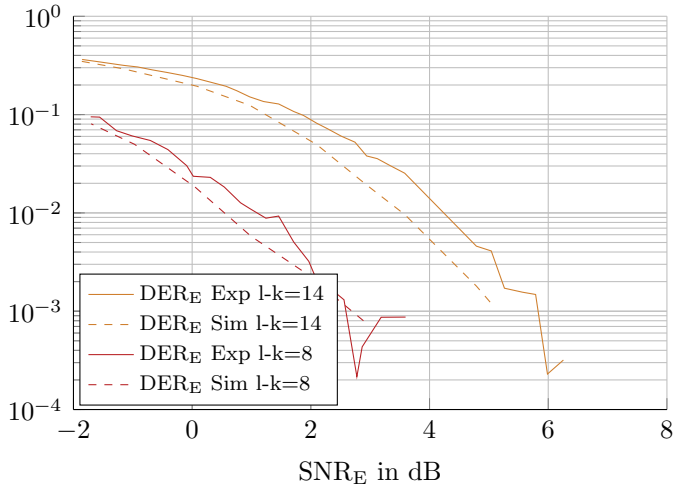


Figure 9.8.: Effect of the difference $l-k$ on the $\mathrm{DER_E}$ error rate (n=28; l=18; k=4,10,18). For $k = 18$, i.e., $l - k = 0$, we have a measured $\mathrm{DER_E}$ of zero, which cannot be displayed.

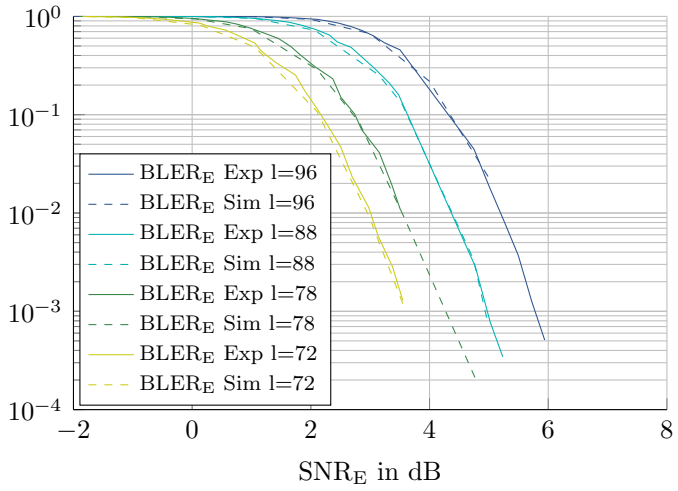


Figure 9.9.: $\mathrm{BLER_E}$ for larger blocklengths (n=128; l=78, 88, 96; k=16).

has almost the same behavior for all seeds. Simulations have shown that the set of good seeds that provide a high security level is rather large. This property is important for practical implementations.

In our third experiment, we use longer, more realistic blocklengths. In Fig. 9.9 we see the $\mathrm{BLER_E}$ for $n = 128$, $k = 16$ and $l = 72, 78, 88, 96$. Due to the large difference $l - k$, it is computationally infeasible to determine the $\mathrm{DER_E}$ in this scenario. As expected from the asymptotic theory of channel capacity, the transition from the $\mathrm{SNR_E}$ region with high $\mathrm{BLER_E}$ to the region where $\mathrm{BLER_E}$ decreases is sharper than in Fig. 9.7, where the blocklength is smaller. Moreover, the decrease is faster in Fig. 9.9.

In our fourth experiment, whose results are displayed in Fig. 9.8, we vary the difference $l - k$ by holding $l = 18$ fixed and choosing $k = 4, 10, 18$. For $k = 18$ we have $l - k = 0$, which gives a $\mathrm{DER_E}$ of zero for all $\mathrm{SNR_E}$. It can be clearly seen that increasing $l - k$ for

a fixed $\mathrm{SNR_E}$ leads to a larger $\mathrm{DER_E}$, i.e., a higher security level. To put it differently, a fixed security level can be sustained even for higher $\mathrm{SNR_E}$ if $l-k$ is increased. This effect can also be observed in Fig. 9.9.

In addition, we have taken screenshots of signal measurements for the testbed setup. The configurations of the individual communication participants as well as the network configurations can be seen in Fig. 9.10 and 9.11. For $n = 32$, $l = 18$ and $k = 4$ the distinguishing performance of Eve and $BLER_B$ of Bob were recorded when the security layer is switched on. In the lower right of the figures, DER_E is plotted as a function of time for three different attack strategies, and $BLER_B$ is plotted to the left. The SNR values at the given time are shown to the left of the $BLER_B$ curve. In the legend, the NN decoder is labeled "deep learning" and the modified polar decoder with list size $L = 8$ is labeled "Polar SCL" (see Section 8.3). The ML test was introduced in Section 7.4.2. Bob decodes according to the 3GPP standard. In Fig. 9.11 we consider the same scenario when the security layer is disabled, i.e. the random vector has length $l-k = 0$. We can see that in the second image, when the security layer is disabled, Eve can decode messages over the entire SNR range with $DER_E \approx 0$. Thus, without the security layer and thus without the randomized encoding, no security can be provided (in our case distinguishing security), with the exception of the modified polar decoder. We can see that the security layer has no influence on the modified polar decoder. Therefore, the modified decoder continues to decode with errors after deactivating the security layer. Thus Eve would not choose the modified polar decoder as her attack strategy.

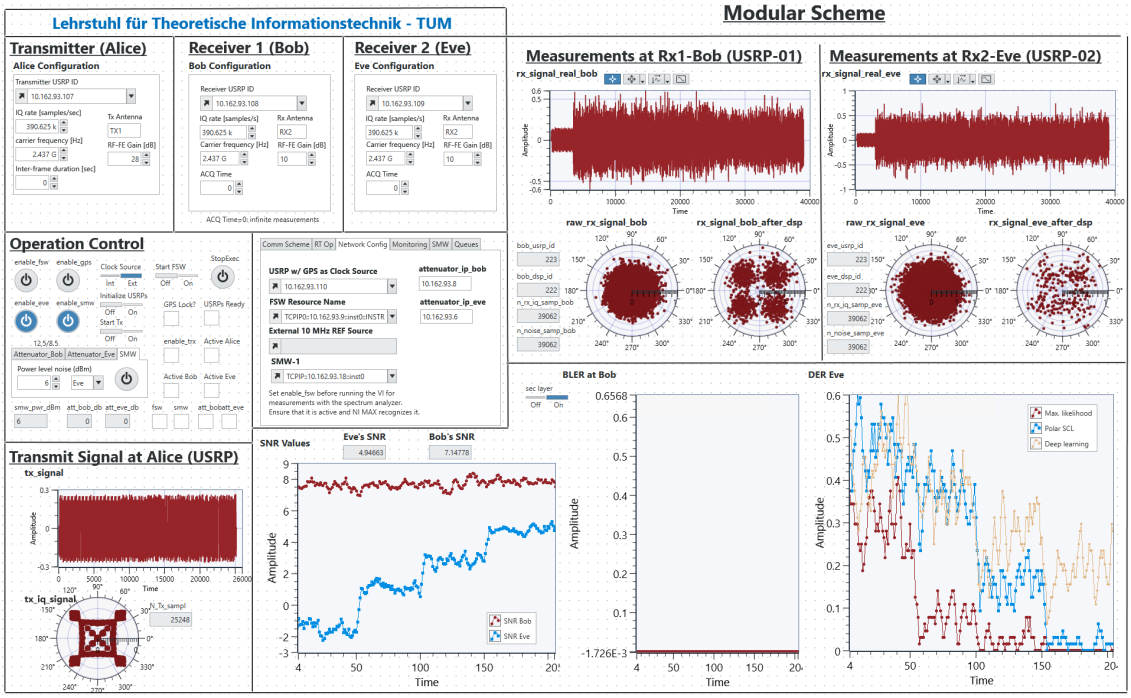


Figure 9.10.: Communication scheme with activated security layer for $n = 32$, $l = 18$ and $k = 4$.

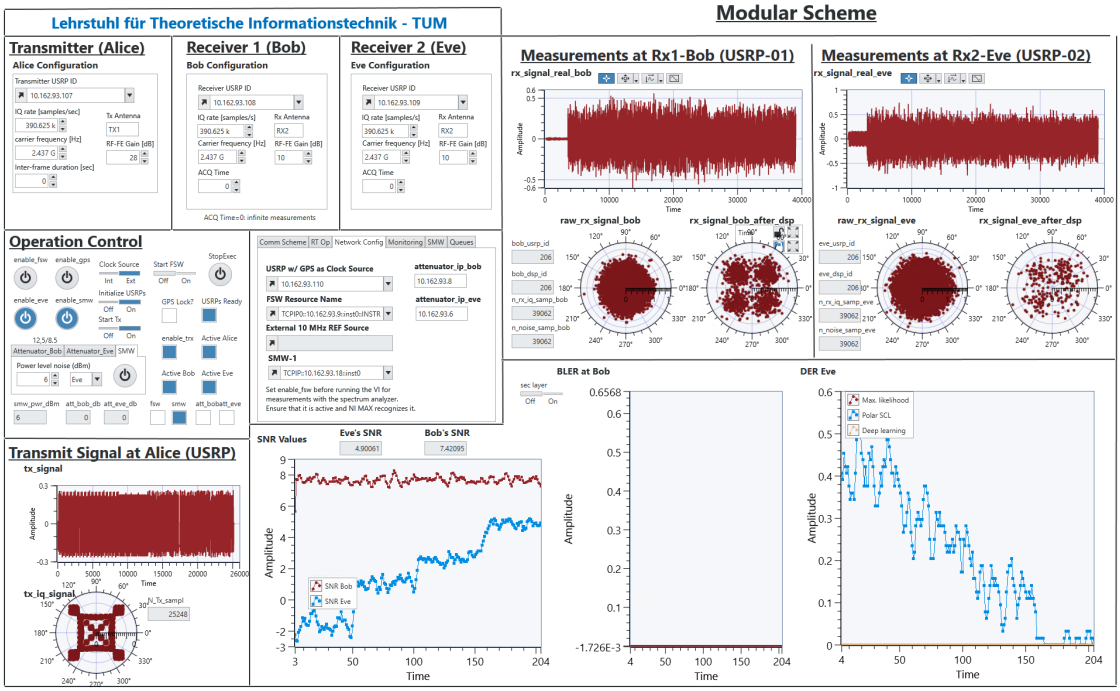


Figure 9.11.: Communication scheme with deactivated security layer for $n = 32$, $l = 18$ and $k = 4$.

9.6. Conclusion

We experimentally evaluated a seeded modular physical layer security scheme using software defined radios. To the best of our knowledge, this is the first time such a demonstration has been done using real signal transmission. The used blocklengths are rather short, given that the computational load of Eve's ML decoder needed to assess the security level via the DS metric would otherwise be too big. We observed that the experimental results are close to the simulation results. A relevant future research direction is to find other security metrics with an operational meaning that do not require this costly operation.

Bibliography

[1] C. E. Shannon, "Communication theory of secrecy systems," *The Bell system technical journal*, vol. 28, no. 4, pp. 656–715, 1949.

[2] A. D. Wyner, "The wire-tap channel," *Bell Laboratories technical journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[3] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[4] L. H. Ozarow and A. D. Wyner, "Wire-tap channel II," *Bell Laboratories technical journal*, vol. 63, no. 10, pp. 2135–2157, 1984.

[5] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC codes to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2933–2945, 2007.

[6] R. Liu, H. V. Poor, P. Spasojevic, and Y. Liang, "Nested codes for secure transmission," in *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–5, IEEE, 2008.

[7] E. Martinian and C.-E. Sundberg, "Burst erasure correction codes with low decoding delay," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2494–2502, 2004.

[8] E. Martinian and M. Trott, "Delay-optimal burst erasure code construction," in *2007 IEEE International Symposium on Information Theory*, pp. 1006–1010, IEEE, 2007.

[9] A. Badr, A. Khisti, W.-t. Tan, and J. Apostolopoulos, "Streaming codes with partial recovery over channels with burst and isolated erasures," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 3, pp. 501–516, 2015.

[10] S. L. Fong, A. Khisti, B. Li, W.-T. Tan, X. Zhu, and J. Apostolopoulos, "Optimal streaming codes for channels with burst and arbitrary erasures," *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4274–4292, 2019.

[11] L. Lai, H. El Gamal, and H. V. Poor, "Authentication over noisy channels," *IEEE Transactions on Information Theory*, vol. 55, no. 2, pp. 906–916, 2009.

[12] V. Aggarwal, L. Lai, A. R. Calderbank, and H. V. Poor, "Wiretap channel type II with an active eavesdropper," in *2009 IEEE International Symposium on Information Theory*, pp. 1944–1948, IEEE, 2009.

[13] M. Bellare, S. Tessaro, and A. Vardy, "A cryptographic treatment of the wiretap channel," *arXiv preprint arXiv:1201.2205*, 2012.

[14] G. P. Fettweis and H. Boche, "6g: The personal tactile internet-and open questions for information theory," *IEEE BITS the Information Theory Magazine*, 2021.

[15] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[16] T. M. Cover and J. A. Thomas, "Elements of information theory second edition solutions to problems," *Internet Access*, pp. 19–20, 2006.

[17] F. Jessie-Macwilliams and N. J. A. Sloane, "The theory of error correcting codes," 1981.

[18] I. Csiszár, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol. 32, no. 1, pp. 40–47, 1996.

[19] U. M. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," *in Advances in Cryptology - Eurocrypt 2000, Lecture Notes in Computer Science*, vol. B. Preneel, p. 351, 2000.

[20] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," in *Annual Cryptology Conference*, pp. 294–311, Springer, 2012.

[21] S. Goldwasser and S. Micali, "Probabilistic encryption," *Journal of Computer and System Sciences*, vol. 28, no. 2, pp. 270–299, 1984.

[22] M. Bloch and J. Barros, "Physical-layer security: from information theory to security engineering," 2011.

[23] M. Van Dijk, "On a special class of broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 712–714, 1997.

[24] Y. Liang, H. V. Poor, and S. Shamai, "Information theoretic security," 2009.

[25] H. Boche and R. F. Schaefer, "Capacity results and super-activation for wiretap channels with active wiretappers," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 9, pp. 1482–1496, 2013.

[26] P. Wang and R. Safavi-Naini, "A model for adversarial wiretap channels," pp. 40–44, 2014.

[27] Y. E. Rouayheb, E. Soljanin, and A. Sprintson, "Secure network coding for wiretap networks of type II," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1361–1370, 2012.

[28] S. Jaggi and M. Langberg, "Secure network coding: Bounds and algorithms for secret and reliable communications," in *Network Coding*, pp. 183–215, Elsevier, 2012.

[29] T. Moon, *Error correcting coding: Mathematical methods and algorithms*. John Wiley & Sons, 2005.

[30] R. H. Morelos-Zaragoza, *The art of error correcting coding*. John Wiley & Sons, 2006.

[31] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1250–1276, 2002.

[32] H. D. Hollmann and L. M. Tolhuizen, "Optimal codes for correcting a single (wrap-around) burst of erasures," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4361–4364, 2008.

[33] C. Di, D. Proietti, I. E. Telatar, T. J. Richardson, and R. L. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1570–1579, 2002.

[34] M. Fossorier, "Universal burst error correction," in *2006 IEEE International Symposium on Information Theory*, pp. 1969–1973, IEEE, 2006.

[35] E. Martinian, "Dynamic information and constraints in source and channel coding," 2006.

[36] Z. Li, A. Khisti, and B. Girod, "Correcting erasure bursts with minimum decoding delay," pp. 33–39, 2011.

[37] A. Badr, D. Lui, and A. Khisti, "Streaming codes for multicast over burst erasure channels," *IEEE Transactions on Information Theory*, vol. 61, no. 8, pp. 4181–4208, 2015.

[38] A. Khisti and J. P. Singh, "On multicasting with streaming burst-erasure codes," pp. 2887–2891, 2009.

[39] N. Adler and Y. Cassuto, "Burst-erasure correcting codes with optimal average delay," *IEEE Transactions on Information Theory*, vol. 63, no. 5, pp. 2848–2865, 2017.

[40] A. Badr, A. Khisti, and E. Martinian, "Diversity embedded streaming erasure codes (DE-SCo): Constructions and optimality," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 5, pp. 1042–1054, 2011.

[41] A. Badr, A. Khisti, W.-T. Tan, and J. Apostolopoulos, "Robust streaming erasure codes using mds constituent codes," pp. 158–163, 2013.

[42] A. Badr, P. Patil, A. Khisti, W.-T. Tan, and J. Apostolopoulos, "Layered constructions for low-delay streaming codes," *IEEE Transactions on Information Theory*, vol. 63, no. 1, pp. 111–141, 2016.

[43] D. Leong, A. Qureshi, and T. Ho, "On coding for real-time streaming under packet erasures," pp. 1012–1016, 2013.

[44] Z. Li, A. C. Begen, and B. Girod, "Delay-optimal burst erasure codes for parallel links," pp. 1–6, 2011.

[45] Z. Li, A. Khisti, and B. Girod, "Forward error protection for low-delay packet video," pp. 1–8, 2010.

[46] A. Frank, H. Aydinian, and H. Boche, "Type II wiretap channel with an active eavesdropper in finite blocklength regime," pp. 258–263, 2016.

[47] G. Forney, "Burst-correcting codes for the classic bursty channel," *IEEE Transactions on Communication Technology*, vol. 19, no. 5, pp. 772–781, 1971.

[48] X. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross, "A measurement study of a large-scale P2P IPTV system," *IEEE Transactions on Multimedia*, vol. 9, no. 8, pp. 1672–1687, 2007.

[49] T. Silverston and O. Fourmaux, "Measuring P2P IPTV systems," in *Proceedings of NOSSDAV*, vol. 7, p. 2, 2007.

[50] R. Chapman, "An involution on derangements," *Discrete Mathematics*, vol. 231, no. 1, pp. 121–122, 2001.

[51] M. Hayashi, "Upper bounds of eavesdropper's performances in finite-length code with the decoy method," *Physical Review A*, vol. 76, no. 1, p. 012329, 2007.

[52] C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM journal on Computing*, vol. 17, no. 2, pp. 210–229, 1988.

[53] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3989–4001, 2011.

[54] I. Tal and A. Vardy, "Channel upgrading for semantically-secure encryption on wiretap channels," in *2013 IEEE International Symposium on Information Theory*, pp. 1561–1565, IEEE, 2013.

[55] H. Tyagi and A. Vardy, "Universal hashing for information-theoretic security," *Proceedings of the IEEE*, vol. 103, no. 10, pp. 1781–1795, 2015.

[56] M. Hayashi and R. Matsumoto, "Secure multiplex coding with dependent and non-uniform multiple messages," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2355–2409, 2016.

[57] M. Wiese and H. Boche, "Semantic security via seeded modular coding schemes and ramanujan graphs," *IEEE Transactions on Information Theory*, vol. 67, no. 1, pp. 52–80, 2020.

[58] M. Wiese and H. Boche, "Mosaics of combinatorial designs for information-theoretic security," *Designs, Codes and Cryptography*, pp. 1–40, 2022.

[59] M. Baldi, G. Ricciutelli, N. Maturo, and F. Chiaraluce, "Performance assessment and design of finite length LDPC codes for the Gaussian wiretap channel," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 435–440, IEEE, 2015.

[60] W. K. Harrison, T. Fernandes, M. A. Gomes, and J. P. Vilela, "Generating a binary symmetric channel for wiretap codes," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 2128–2138, 2019.

[61] M. Bloch, M. Hayashi, and A. Thangaraj, "Error-control coding for physical-layer secrecy," *Proceedings of the IEEE*, vol. 103, no. 10, pp. 1725–1746, 2015.

[62] C. Ling, L. Luzzi, J.-C. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the gaussian wiretap channel," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6399–6416, 2014.

[63] J. Barros and M. Bloch, "Strong secrecy for wireless channels (invited talk)," in *International Conference on Information Theoretic Security*, pp. 40–53, Springer, 2008.

[64] L. Liu, Y. Yan, and C. Ling, "Achieving secrecy capacity of the Gaussian wiretap channel with polar lattices," *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1647–1665, 2018.

[65] M. N. Wegman and J. L. Carter, "New hash functions and their use in authentication and set equality," *Journal of computer and system sciences*, vol. 22, no. 3, pp. 265–279, 1981.

[66] M. Bellare and S. Tessaro, "Polynomial-time, semantically-secure encryption achieving the secrecy capacity," *arXiv preprint arXiv:1201.3160*, 2012.

[67] M. Hayashi and M. Hayashi, "Quantum wiretap channel with non-uniform random number and its exponent and equivocation rate of leaked information," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5595–5622, 2015.

[68] R. G. Gallager, *Information theory and reliable communication*, vol. 2. Springer, 1968.

[69] "3GPP TS 38.211 version 15.2.0 release 15", 5G; NR; "Physical channels and modulation," *Technical Specification Group Radio Access Network*, Jul 2018.

[70] "3GPP TS 38.212 version 15.2.0 release 15", 5G; NR; "Multiplexing and channel coding," *3rd Generation Partnership Project (3GPP)*, July 2018.

[71] E. M. Jazi and J. N. Laneman, "Coded modulation for gaussian channels: Dispersion- and entropy-limited regimes," in *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 528–533, IEEE, 2015.

[72] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55–67, 1982.

[73] D. Hui, S. Sandberg, Y. Blankenship, M. Andersson, and L. Grosjean, "Channel coding in 5G new radio: A tutorial overview and performance comparison with 4G LTE," *IEEE Vehicular Technology Magazine*, vol. 13, no. 4, pp. 60–69, 2018.

[74] A. Solomon, K. R. Duffy, and M. Médard, "Soft maximum likelihood decoding using GRAND," in *2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2020.

[75] M. P. Fossorier and S. Lin, "Soft-decision decoding of linear block codes based on ordered statistics," *IEEE Transactions on Information Theory*, vol. 41, no. 5, pp. 1379–1396, 1995.

[76] C. Choi and J. Jeong, "Fast and scalable soft decision decoding of linear block codes," *IEEE Communications Letters*, vol. 23, no. 10, pp. 1753–1756, 2019.

[77] C. Yue, M. Shirvanimoghaddam, B. Vucetic, and Y. Li, "A revisit to ordered statistics decoding: Distance distribution and decoding rules," *IEEE Transactions on Information Theory*, 2021.

[78] S. E. Alnawayseh and P. Loskot, "Ordered statistics-based list decoding techniques for linear binary block codes," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–12, 2012.

[79] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[80] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," 2017.

[81] T. Gruber, S. Cammerer, J. Hoydis, and S. ten Brink, "On deep learning-based channel decoding," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, IEEE, 2017.

[82] W. K. Harrison, T. Fernandes, M. A. Gomes, and J. P. Vilela, "Generating a binary symmetric channel for wiretap codes," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 2128–2138, 2019.

[83] R. Barker, "Group sysnchronizing of binary digital systems," *Communication Theory*, pp. 273–287, 1953.

[84] R. Gold, "Optimal binary sequences for spread spectrum multiplexing (corresp.)," *IEEE Transactions on Information Theory*, vol. 13, no. 4, pp. 619–621, 1967.

[85] M. Rice, *Digital communications: a discrete-time approach*. Prentice Hall, 2009.

[86] C. R. Johnson Jr, W. A. Sethares, and A. G. Klein, *Software receiver design: Build your own digital communication system in five easy steps*. Cambridge University Press, 2011.

[87] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Transactions on Communications*, vol. 48, pp. 1681–1691, Oct. 2000.