UNIVERSITÀ POLITECNICA DELLE MARCHE
SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE DELL'INGEGNERIA
CURRICULUM IN INGEGNERIA ELETTRONICA, ELETTROTECNICA E DELLE
TELECOMUNICAZIONI

Information-theoretic security techniques for data communications and storage

Ph.D. Dissertation of:
Linda Senigagliesi

Advisor:
Prof. Luca Spalazzi

Coadvisor:
Dr. Marco Baldi

Curriculum Supervisor:
Prof. Francesco Piazza

XVII edition - new series

UNIVERSITÀ POLITECNICA DELLE MARCHE
SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE DELL'INGEGNERIA
CURRICULUM IN INGEGNERIA ELETTRONICA, ELETTROTECNICA E DELLE
TELECOMUNICAZIONI

Information-theoretic security techniques for data communications and storage

Ph.D. Dissertation of:
Linda Senigagliesi

Advisor:
Prof. Luca Spalazzi

Coadvisor:
Dr. Marco Baldi

Curriculum Supervisor:
Prof. Francesco Piazza

XVII edition - new series

UNIVERSITÀ POLITECNICA DELLE MARCHE
SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE DELL'INGEGNERIA
FACOLTÀ DI INGEGNERIA
Via Brecce Bianche – 60131 Ancona (AN), Italy

To my family

Abstract

The last years have seen a growing need of security and privacy in many aspects of communications, together with the technological progress. Most of the implemented security solutions are based on the notion of computational security, and must be kept continuously updated to face new attacks and technology advancements. To meet the more and more strict requirements, solutions based on the information-theoretic paradigm are gaining interest to support pure cryptographic techniques, thanks to their capacity to achieve security independently on the attacker's computing resources, also known as unconditional security. In this work we investigate how information-theoretic security can be applied to practical systems in order to ensure data security and privacy.

Information-theoretic metrics to assess the secrecy performance of realistic wireless communication settings under practical conditions are defined, together with a protocol that mixes coding techniques for physical layer security and cryptographic solutions. This scheme is able to achieve some level of semantic security at the presence of a passive attacker. At the same time, multiple scenarios are considered: security analysis for parallel relay channels is provided, thus finding an optimal resource allocation that maximizes the secrecy rate. Then, by exploiting a probabilistic model checker, we define the parameters for heterogeneous distributed storage systems that permit us to achieve perfect secrecy in practical conditions. For privacy purposes, we propose a scheme which guarantees private information retrieval of files for caching at the wireless edge against multiple spy nodes. We find the optimal content placement that minimizes the backhaul usage, thus reducing the communication cost of the system.

Foreword

During my period at Dipartimento di Ingegneria dell'Informazione of Università Politecnica delle Marche as Ph.D. student, I had the pleasure to work with the research groups leaded by Prof. Luca Spalazzi, Dr. Marco Baldi and Franco Chiaraluce.

In our work we have studied information-theoretic techniques to improve security and privacy in wireless networks and distributed storage systems. In particular, we focused on schemes able to guarantee security and privacy regardless of the attacker's computing power, considering realistic conditions for the channel.

During these years I had the possibility to collaborate with other important European universities as the Chalmers University of Technology of Gothenburg, Sweden. In particular, I spent almost five months at the Department of Electrical Engineering, where part of this thesis was developed with the Communication Systems group leaded by Prof. Alexandre Graell i Amat.

The contents of this thesis have been partially included in the following publications.

- S. Kumar, E. Rosnes, A. Graell i Amat and **L. Senigagliesi**, "Private Information Retrieval From a Cellular Network with Caching at the Edge", IEEE Transactions on Communications, accepted with major revision.

- **L. Senigagliesi**, M. Baldi and S. Tomasin, "Resource Allocation for Secure Gaussian Parallel Relay Channels with Finite-Length Coding and Discrete Constellations", Jul. 2018, arXiv:1807.06448 [cs.IT]. [Online].
  Available: https://arxiv.org/abs/1807.06448.

- **L. Senigagliesi**, M. Baldi and F. Chiaraluce, "Semantic Security with Practical Transmission Schemes over Fading Wiretap Channels", MDPI Entropy, Sept. 2017, DOI: 10.3390/e19090491.

- M. Baldi, E. Bartocci, F. Chiaraluce, A. Cucchiarelli, **L. Senigagliesi**, L. Spalazzi and F. Spegni, "A Probabilistic Small Model Theorem to Assess Confidentiality of Dispersed Cloud Storage", in Proc. International Conference on Quantitative Evaluation of SysTems (QEST), pp. 123-139, Berlin, Germany, Sept. 2017, DOI: 10.1007/978-3-319-66335-7-8.

- M. Baldi, F. Chiaraluce, **L. Senigagliesi**, L. Spalazzi and F. Spegni, "Security in heterogeneous distributed storage systems: a practically achievable information-

theoretic approach", in Proc. IEEE Symposium on Computers and Communications (ISCC), pp. 1021-1028, Heraklion, Crete, Jul. 2017, DOI: 10.1109/ISCC.2017.8024659.

- M. Baldi, **L. Senigagliesi** and F. Chiaraluce, "On the security of transmissions over fading wiretap channels in realistic conditions", in Proc. IEEE International Conference on Communications (ICC), pp. 1-6, Paris, France, May 2017, DOI: 10.1109/ICC.2017.7996662.

Contents

List of Figures

List of Tables

List of abbreviations

ACK acknowledge.

AES advanced encryption standard.

AF amplify and forward.

AONT all-or-nothing transform.

ARQ automatic repeat request.

AWGN additive white Gaussian noise.

BCC broadcast channel with confidential messages.

BER bit error rate.

BPSK binary phase shift keying.

CDF cumulative distribution function.

CER codeword error rate.

CSI channel state information.

CTS clear to send.

DEC decoder.

DF decode and forward.

DMC discrete memoryless channel.

DSS distributed storage system.

ENC encoder.

FK fake packet.

GRS generalized Reed-Solomon.

HARQ hybrid automatic repeat request.

i.i.d. independent and identically distributed.

l.h.s. left-hand-side.

LDPC low density parity check.

LT Luby transform.

MBS macro-cell base station.

MDP Markov decision process.

MDS maximum distance separable.

MIMO multiple input multiple output.

MIS *mutual information security*.

MRC maximal ratio combining.

NA not available.

NACK not acknowledge.

OFDM orthogonal frequency division multiplexing.

OOT *on–off transmission*.

OOT-FP *on–off transmission with fake packets*.

p.d.f. probability density function.

PIR private information retrieval.

PLS physical layer security.

PMC probabilistic model checking.

PMD probability mass distribution.

PMF probability mass function.

PPP Poisson point process.

PRNG pseudo-random numbers generator.

QAM quadrature amplitude modulation.

r.h.s. right-hand-side.

RS Reed-Solomon.

RTS request to send.

SBS small-cell base station.

SFC slow fading channel.

SNR signal to noise ratio.

SP service provider.

SS *semantic security*.

Chapter 1

Introduction

Classical practical solutions for data security are based on the concept of computational security. The level of secrecy guaranteed by a cryptosystem is strictly dependent on the power of the resources available to the attacker. In other words, the security is measured with the average number of attempts required by an opponent to break the system. This number obviously changes in time: due to the exponential growth of computing resources, it is necessary to continuously update the algorithms and the key sizes recommended by the security standards. Anyway, computational security provides a strong notion of secrecy in practical systems, since the absence of any information leakage is guaranteed for a reasonably long time under the hypothesis of a resource-constrained attacker.

Modern algorithms with suitable key lengths (e.g. AES-128 [1], RSA-2048, etc.) are not susceptible to brute force attack, even with massive amounts of computing power, and they would take centuries or, in some cases, even longer than the lifetime of the universe to break. However, many other attacks exist, which are continuously updated and improved. Moreover, with the advent of quantum computing, the time of attack is dramatically reduced, and many of the current encryption algorithms will be rendered essentially useless once quantum computers reach a certain scale.

On the one hand, many efforts are focused on developing new post-quantum cryptographic algorithms. On the other hand, information-theoretic security is gaining a growing interest, thanks to its ability to reach the so-called *unconditional security*. This notion of security is independent on the attacker's computing power.

This, jointly with the low algorithmic complexity of information-theoretic security techniques, has made this area of research of great relevance for some scenarios like that of resource-constrained wireless devices. In fact, the diffusion of the Internet of Things (IoT) and a growing number of devices deployed over networks have forced security schemes to meet new requirements, such as low latency and delay, limited storage and energy constraints. Traditional cryptographic methods, however, are characterized by a complexity that barely fits these requirements, while information-theoretic techniques help in reducing the burden of heavy computational calculations and guarantee a level of security which does not depend on the attacker's capacity.

Nevertheless, the idea of information-theoretic security is often related to ideal con-

ditions almost impossible to achieve in practice. Users are differentiated only on the basis of the intrinsic randomness and uniqueness of the transmission channel, without the need of any shared secret, and for this reason such a paradigm is also often referred as physical layer security [2]. Unfortunately, some of its characteristics make it difficult to be accepted as the sole security mechanism in a practical setting. One of such limitations is the inability to achieve perfect secrecy in practical systems, which forces to rely on some weaker notion of secrecy implying to cope with some information leakage in non-asymptotic regimes.

In this work we study whether unconditional security is achievable also in practical systems, and under which bounds and conditions. We start considering a wireless channel under realistic assumptions: the channel is affected by fading, messages are encoded using short codes and transmitted with finite constellation signaling (we also measured the secrecy performances of coding schemes that are already included in standards, and implemented in commercial transceivers, such that no new code design is required to implement the proposed protocol and existing hardware and software can easily be reused.). As channel model we consider the wiretap channel [3], where the attacker experiences a noisier version of the main channel. We define and compute suitable information theoretic metrics to assess the secrecy performance while taking into account the constraints due to the realistic wireless channel models, referring to the notion of *mutual information security* and *semantic security*.

A promising research direction is that of using physical layer security and computational security jointly, in such a way as to overcome their limitations and exploit their benefits to the utmost. Motivated by these observations, we propose a joint physical layer / computational security protocol aimed at achieving a given level of computational security without the need of any secret key, either pre-shared or distilled from the channel. The main elements on which our protocol is built are coding and all-or-nothing transforms. In order to provide a significant estimate of the security level, we consider the notion of semantic security, which is a well-known concept used in cryptography. Moreover, such a notion of security is gaining a prominent role also within the context of information theoretic security for the finite block length regime [4].

We then consider a model which includes a set of cooperating relays and where each link comprises a set of parallel channels, modeling for example an orthogonal frequency division multiplexing transmission, still taking into account the impact of discrete constellations and finite-length coding. We propose a power and channel allocation algorithm that maximizes the achievable secrecy rate by resorting to two coupled Gale-Shapley algorithms for stable matching problem. We consider the scenarios of both full and partial channel state information at Alice. In the latter case, we only guarantee an outage secrecy rate, i.e., the rate of a message that remains secret with a given probability. In the case of fading channels, we provide results in terms of average outage secrecy rate, showing that practical schemes achieve a performance quite close to that of ideal ones.

Successively, we extend our analysis to distributed storage systems and caching systems. We exploit a version of our protocol adapted to these scenarios to reach a twofold goal: estimating the levels of information-theoretic security and defining a practical scheme able to achieve them. Proper dimensioning of the scheme's parameters, in order to optimize the security metrics, is achieved through the application of an effective probabilistic model checking, thus removing most of the limitations related to more conventional methods, which are not able to deal with heterogeneous networks, where nodes are grouped in classes with different properties.

Moreover, we are also interested in evaluating privacy in the mentioned systems. In particular, we consider private information retrieval (PIR) of content from a library of files, i.e., the user wishes to download a file and does not want the network to learn any information about which file she is interested in. While there are existing solutions that guarantee PIR in a distributed storage system, we consider the problem of downloading content from a cellular network where content is cached at the wireless edge while achieving privacy. To reduce the backhaul usage, content is cached at the wireless edge in a number of small-cell base stations (SBSs) using maximum distance separable codes. We propose a PIR scheme for this scenario that achieves privacy against a number of spy SBSs that (possibly) collaborate.

1.1 Outline of the thesis

The document is organized as follows.

Chapter 2

In Ch. 2 the main notions about information-theoretic security are given. These notions are useful for the following chapters.

Chapter 3

In Ch. 3 we show how the level of security at the physical layer can be assessed from the information theoretic standpoint while taking into account the constraints of practical transmissions over realistic wireless wiretap channels. For this purpose, we consider the notion of mutual information security. Moreover, the chapter contains the proposal of an on–off protocol for communication over wireless wiretap channels with security at the physical layer. The proposed method does not require either pre-shared secret keys or public keys. It also leverages the noisy and fading nature of the channel and exploits coding and all-or-nothing transforms to achieve the desired level of semantic security. We show that the use of fake packets in place of skipped transmissions during low channel quality periods yields significant advantages in terms of time needed to complete transmission of a secret message.

Chapter 4

Chapter 4 studies the problem of resource allocation for confidential communications

over the Gaussian parallel relay channels. The achievable secrecy rate is derived in an ideal case and under practical constraints, and it is used to find the optimal power allocation for point-to-point secure transmission. Then we exploit power and rate adaptation algorithms coupling two Gale and Shapley algorithms to allocate resources over the parallel relay channels.

Chapter 5
Ch. 5 contains the proposal of a mixed cryptographic/information-theoretic approach to assess security of complex scenarios with heterogeneous network topologies and a passive attacker eavesdropping the channel between user and storage nodes. The presented methodology helps to determine the optimal values of the design parameters needed to achieve a given level of security and also to evaluate which level of security can be guaranteed by a dispersed cloud storage, given a certain configuration.

Chapter 6
Chapter 6 contains the proposal of a private information retrieval scheme for caching at the edge. The private retrieval of content from a library of files that have different popularities is considered, and the backhaul rate for the PIR case as a function of the content placement is derived. We prove that uniform content placement, i.e., all files that are cached are encoded with the same code rate, is optimal.

Chapter 7
Finally, Ch. 7 concludes the thesis.

1.2 Main contributions of the thesis

In the following list the main contributions of this thesis are reported.
Chapter 3

- Suitable information-theoretic metrics to assess the secrecy performance while taking into account the constraints due to practical coding and modulation schemes and realistic wireless channel models are defined.

- A protocol based on coding for physical layer security and cryptographic solutions able to achieve some level of semantic security without need of secret keys is proposed.

Chapter 4

- A formulation of the secrecy rate under practical constraints is provided and compared with the achievable rate in ideal conditions, considering both perfect and partial channel state information under outage constraints.

- An approximated expression for the secrecy rate under practical conditions is proposed to optimize the link-level parallel channel power allocation generalizing the solution of the ideal transmission scenario.
- Extending the approach in [5], the secrecy rate is maximized by resorting to an iterative algorithm based on the Gale and Shapley theory for the stable matching problem.

Chapter 5

- A framework to design and assess heterogeneous distributed storage systems with a prefixed level of security is proposed.
- An approach that joins information-theory and formal verification is considered in order to achieve perfect secrecy.
- The ratio between the message original size and the total size is optimized.

Chapter 6

- A private information retrieval scheme against a number of spy small-cell base stations possibly colluding in a wireless caching scenario is proposed.
- An expression of the backhaul rate for the PIR case as a function of the content placement is derived.
- A minimization of a weighted sum of the backhaul rate and the communication rate from the small-cell base stations is investigated.

Chapter 2

Preliminary concepts

In this chapter we outline the main concepts that will help to understand the theoretical background of this dissertation. We first introduce the notions which constitute the base of information theory, with a particular consideration for the secrecy aspects, and physical layer security. Then we present the secrecy notions of mutual information security and semantic security. These concepts help us to give a valid estimate of which level of secrecy can be achieved in practical communication systems, starting from the secrecy at the physical layer in **Chapter 3** and **Chapter 4** to the analysis of security (**Chapter 5**) and privacy (**Chapter 6**) obtainable in more complex networks and systems.

2.1 Information-theoretic security

Information theory is a branch of mathematics which was originally proposed by Claude E. Shannon in his milestone paper *A mathematical theory of communication* in 1948 [6]. His work found application in many topics, including data compression, channel coding and cryptography.

2.1.1 Channel model

In [6] Shannon defines a communication system consisting of 5 parts as shown schematically in Fig. 2.1. These parts are

- An information source which produces a message or sequence of messages to be communicated to the receiving terminal;
- A transmitter which operates on the message in some way to produce a signal suitable for transmission over the channel;
- The channel, which is the medium used to transmit the signal from transmitter to receiver;
- The receiver, that performs the inverse operation of that done by the transmitter, reconstructing the message from the signal;

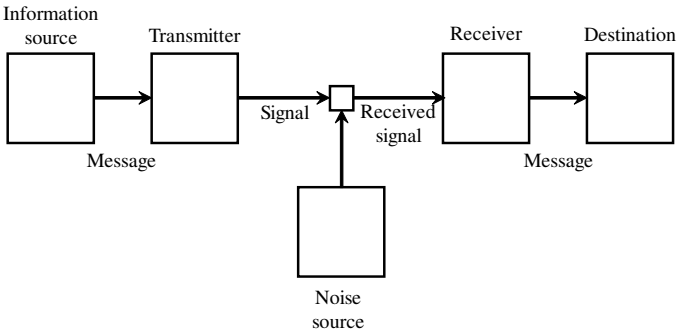


Figure 2.1: Schematic diagram of a general communication system.

- The destination, i.e. the person (or thing) for whom the message is intended.

This channel scheme is considered as a reference model for the rest of this section.

2.1.2 Information-theoretic metrics

One of the key concepts of Shannon's work is the definition of *entropy* as a measure of information, choice and uncertainty. The entropy $\mathbb{H}(X)$ of a discrete random variable X with possible values $\{x_1, \cdots, x_n\}$ and probability mass function (PMF) $p(x)$ is defined as[1]

$$\mathbb{H}(X) = -\sum_x p(x) \log(p(x)).$$

$\mathbb{H}(X)$ is approximately equal to how much information it is possible to learn on average from one instance of the random variable X.

It is now useful to introduce two other quantities, the joint entropy and the conditional entropy. Let X and Y be two discrete random variables, with PMFs equal to, respectively, $p(x)$ and $p(y)$, and joint PMF $p(x, y)$, the joint entropy is defined as the measure of the uncertainty associated with a set of variables. In formulas

$$\mathbb{H}(X, Y) = -\sum_x \sum_y p(x, y) \log(p(x, y)).$$

The conditional entropy (or *equivocation*) quantifies the amount of information needed to describe the outcome of a random variable X given that the value of another random variable Y is known, i.e.

$$\mathbb{H}(X|Y) = -\sum_x \sum_y p(x, y) \log(p(x|y)),$$

[1] From now on in this chapter logarithms are intended base 2.

where $p(x|y)$ represents the conditional probability of X given the knowledge of Y. The conditional entropy is a measure of how much uncertainty remains about the random variable X when we know the value of Y.

Closely related to conditional entropy is the concept of mutual information. The mutual information between two discrete random variables X, Y jointly distributed according to $p(x, y)$ is given by

$$\mathbb{I}(X, Y) = \sum_x \sum_y p(x, y) \log \frac{p(x, y)}{p(x)p(y)},$$

which corresponds to

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

The mutual information represents the reduction on uncertainty of X due to the knowledge of Y.

The last definition of this section concerns channel capacity. The channel capacity is defined as the tight upper bound on the rate at which information can be reliably transmitted over a communication channel, corresponding to the maximum mutual information on every possible input distribution. In formulas

$$C = \max \mathbb{I}(X, Y).$$

2.1.3 Perfect secrecy

The strongest existing notion of information-theoretic security is known as *perfect secrecy* [7]. Considering messages and codewords as random variables M and C, respectively, a cryptosystem achieves perfect secrecy when C yields no information about M, or, equivalently, if Shannon uncertainty of the message after observing the codeword is equal to the *a priori* uncertainty of the message. In formulas

$$\mathbb{H}(M|C) = \mathbb{H}(M),$$

This notion of secrecy is often called *unconditional security*, since it does not depend on the computing resources available to the attacker. Unfortunately, there exist few practical methods satisfying the above criterion, illustrated by Shannon in the so-called *one-time pad*, which include the generation of a new perfectly random key of at least the same length of the message M for each exchange session.

This channel model, however, does not take into account the physical reality of communication channels. Especially, it does not consider the degradation of signals because of noise. A more realistic communication model is introduced in the next section.

2.2 Physical layer security

According to the paradigm of information theoretic security, when two legitimate communicators exchange some coded data on a noisy channel, a part of such data, depending on the channel conditions, is secret to the attacker, independently of her computational power [8]. For this reason, information theoretic security is often denoted as unconditional security, in that it does not rely on any limitation on Eve's computing resources. Moreover, users differentiation is only based on the intrinsic randomness and uniqueness of the transmission channel, without the need of any shared secret, and for this reason such a paradigm is also often referred to as physical layer security [2]. The channel model considered, where noise in the main channel and eavesdropper's channel is explicitly introduced, is known as wiretap channel.

2.2.1 The wiretap channel

The wiretap channel model was introduced by Wyner in 1975 [3]. The main actors are the legitimate sender, called Alice, whose aim is to communicate in a secure way with the intended receiver, called Bob, in the presence of a passive eavesdropper, Eve. From now on in this work, we will refer to them with these names. A scheme of the wiretap channel in depicted in Fig. 2.2.

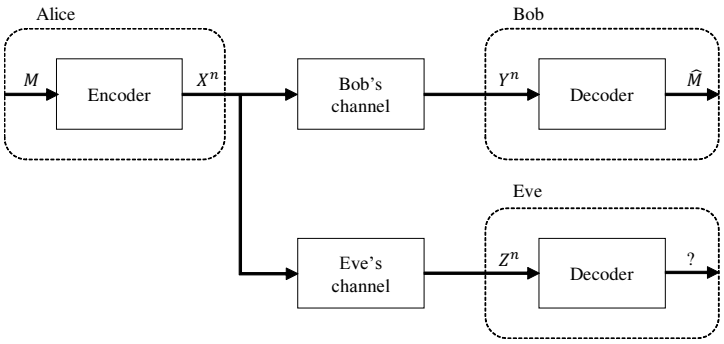


Figure 2.2: The wiretap channel.

Alice wishes to send a secret message M of k bits to Bob. First she encodes M into a codeword X^n of length n and transmits it through the channel which, in Wyner's original work, is supposed to be a discrete memoryless channel (DMC). Bob receives a version Y^n of the message, while Eve observes the transmission through the eavesdropper's channel, and intercept a noisy estimation Z^n of M, which is supposed to be more degraded than Bob's one. Eve knows perfectly the statistics of the channel between Alice and Bob and she is also aware of the techniques used to process the message (encoding, puncturing, scrambling, etc). Moreover, she has unlimited computing resources. The reliability target is satisfied when Bob is able to obtain the

original message without errors, while the system is considered secure if the message intercepted by Eve is different enough from the original to render it unintelligible for the eavesdropper.

In physical layer security, a system is able to achieve *weak secrecy* when the mutual information between Eve' observation Z^n and the codeword X^n tends to zero when the code length n tends to infinity, i.e.

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{I}(Z^n; X^n) = 0\,.$$

While we have *strong secrecy* when

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

The *secrecy capacity* is the maximum rate achievable over the main channel under the secrecy condition for the wiretapper's channel. Interestingly, both strong and weak secrecy requirements result in the same secrecy capacity [9], [10].

These notions of security however are suitable to describe secrecy in ideal systems, which are very far from reality. When we consider practical assumptions, we need to resort to more strong definitions for security, which are described in the next section.

2.3 Other security definitions

As already pointed out, Shannon's perfect secrecy is a strong notion of security, but very difficult to achieve in practice. In the following chapter we refer to other security definitions more suitable for the channel model and the settings we consider: *mutual information security* (MIS) and *semantic security* (SS).

The first definition of *semantic security* was given in [11], and it is usually considered as the computational analogue to Shannon's perfect secrecy. information about the plaintext of a given ciphertext. A cryptosystem is defined semantically secure if only negligible information about the secret message can be feasibly extracted from the ciphertext. Goldwasser et al. also demonstrated that semantic security is equivalent to another definition of security called ciphertext indistinguishability under chosen-plaintext attack. This latter definition is more common than the original definition of semantic security, that does not suggest any method for security evaluation of practical cryptosystems.

Using the same definitions of [12], a system is mis-r (mutual-information secure for random messages) if the maximum "advantage" of an adversary in breaking the scheme is equal to the mutual information between the secret message M and the information that the attacker can extract from her channel. However, due to the hypothesis of random messages uniformly distributed over $\{0,1\}^m$, this metric becomes very weak in a cryptographic perspective, since real messages are not random. The

authors then introduced the definition of *mutual information security* as the maximum mutual information between M and the attacker's observation, where the maximum is over all random variables M over $\{0,1\}^m$, regardless of their distribution.

Moreover, in [12] equivalence between MIS and SS under asymptotic regime is proven, meaning that an encryption scheme is MIS-secure if and only if it is SS-secure.

Unlike classical approaches to physical layer security (PLS), where asymptotic conditions like infinite block length and continuous modulations are assumed, we need security metrics working in the finite block length and discrete modulation regime. As in our analysis we do not make any assumption on the message distribution and consider the maximum of Eve's channel mutual information over all possible distributions, the security notion we will use in the next chapter for PLS is MIS.

First, we propose some tools to study the levels of MIS achievable by practical transmission schemes over Nakagami-m fading channels. Then, starting from these tools, we propose a protocol able to reach some level of SS, where the estimated level of MIS is used as a substrate to achieve some level of SS through cryptographic tools like AONT. In fact, if we only resort to PLS stemming from the communication channel and measured through MIS, it is still possible for Eve to discover some part of the message with less attempts than those estimated through her total equivocation. To avoid this, we propose to pre-process the secret data through an AONT prior to transmission.

Chapter 3

Semantic security with practical transmission schemes over fading wiretap channels

The wiretap coding problem is a well known information theoretic problem introduced in Wyner's seminal work [3]. Such a problem consists of finding coded transmission schemes able to achieve reliability towards a legitimate receiver (Bob) while ensuring secrecy against a wiretapper (Eve) without leveraging any pre-shared secret (like secret keys used in computational security techniques), but only exploiting the differences between Bob's and Eve's channels. The wiretap channel model considered in this chapter is shown in Fig. 3.1.

According to the classical formulation of the problem, information theoretic metrics are used for both reliability and security, codes with length $n \to \infty$ are considered and: i) the reliability target is to achieve error free transmission towards Bob, while ii) the secrecy target is to achieve vanishing mutual information between the secret message and its noisy encoded version received by Eve. Coding schemes which are able to meet these targets in the asymptotic regime have been devised, like LDPC [13], polar [14] and lattice codes [15]. These analyses, however, rely on assumptions which are rather far from practical wireless communications, like infinite code lengths, discrete channels or continuous channels with Gaussian signaling. Instead, it would be interesting to know which reliability and secrecy performance is achievable by using practical codes and modulations, like those imposed by some standard for wireless communications, over realistic channel models. A previous solution to this problem comes from the use of the bit error rate (BER) as a performance metric for both Bob and Eve. In fact, the BER can be estimated (through analytical approaches or numerical simulations) by taking into account all the constraints imposed by practical coding and modulation schemes. This has been done, for example, in [16–18].

However, while Bob's BER is a well accepted reliability metric, measuring secrecy through Eve's BER is less robust than using information theoretic metrics, like Eve's equivocation about the secret message. More precisely, we can say that a high BER at Eve's is a necessary but not a sufficient condition to achieve physical layer security,

as it is easily seen in the case of correlated errors.

The aim of this chapter is to define and compute suitable information-theoretic metrics to assess the secrecy performance while taking into account the constraints due to the use of practical coding and modulation schemes and realistic wireless channel models. For this purpose, we resort to the notions of *mutual information security* (MIS) and *semantic security* (SS), which are proved to be equivalent in [12]. Moreover, we propose a protocol based on coding for physical layer security and cryptographic solutions, such as the AONT, able to achieve some level of semantic security.

3.1 Related works

Many previous works have been devoted to the study of PLS over fading wiretap channels (see, for example, [19–22] and references therein). However, they are mostly focused on asymptotic secrecy targets and consider ideal conditions (like capacity achieving codes and continuous modulations). Moreover, suitably designed coding schemes for the wiretap channel are commonly considered [4], while our aim is to exploit classical coding schemes to achieve some level of PLS in practical conditions. Recent works are focused on finite block-length physical-layer security such as [23], where metrics for security analysis of short blocklength codes are proposed. Those metrics are then further detailed and analyzed in a more recent publication [24].

Some literature also exists on protocols in which the transmission of each packet depends on the occurrence of a certain channel condition, known as *on–off schemes*. Examples can be found in some recent papers [25–27]. The use of fake packets has already been introduced in [28], where the author considers the deliberate transmission of random message blocks when Bob's channel quality is poor. The use of dummy messages has also been considered in previous works concerning hybrid automatic repeat request (HARQ) protocols for secure transmissions [29, 30]. All these approaches, however, consider asymptotic notions of secrecy (like the secrecy capacity, the secrecy throughput or the ergodic secrecy rate) with an underlying notion of weak secrecy, and do not take into account either practical constraints or SS as a metric. Moreover, they focus on optimal codes specifically designed for the wiretap channel and do not consider practical modulation formats, thus being still far from practical applications. On the contrary, our goal is to propose and assess an on–off scheme able to provide some level of SS at the physical layer by exploiting practical and simple coding and modulation schemes.

3.2 Security metrics

We focus on practical, already implemented coding and modulation schemes, hence we consider classical deterministic coding instead of random coding or coset coding,

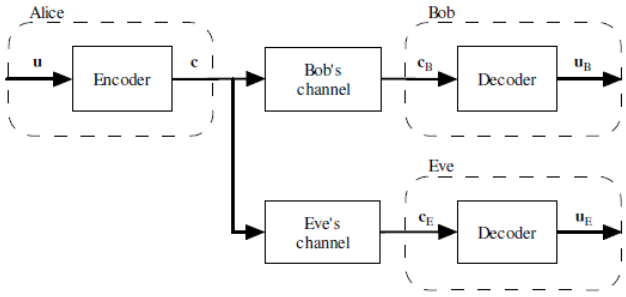


Figure 3.1: Wiretap channel model with legitimate (Bob) and malicious (Eve) receivers.

which are often invoked in the literature for this kind of systems. Therefore, each k-bit block of data x (subscript is omitted for the sake of simplicity) is univocally mapped into a codeword c, differently from coding schemes specifically designed for the wiretap channel, which usually exploit random binning based on coset coding. For this reason we refer to Fig. 3.1, where we denote by c_E the noise corrupted vector received by Eve upon transmission of the codeword c. Then, noting by $\mathbb{H}(\cdot)$ the entropy function, $s = \mathbb{H}(c|c_E) = \mathbb{H}(x|c_E)$ is Eve's total equivocation about the transmitted codeword, that is, the conditional entropy of c given c_E. As in general the input messages are not independent and identically distributed (i.i.d.), we have $\mathbb{H}(c) = \mathbb{H}(x) = k' \leq k$. When $s = k'$, we have perfect secrecy in Wyner's sense [3]. Instead, if $0 < s < k'$, perfect secrecy is not achieved; however, Eve still has to perform 2^s attempts on average in order to correctly decode c from c_E. The source entropy rate is $R_h = \frac{k'}{n} \leq \frac{k}{n} = R_c$, being R_c the code rate. Eve's equivocation can be expressed as

$$s = \mathbb{H}(c|c_E) = \mathbb{H}(c) - \mathbb{I}(c, c_E),$$

where $\mathbb{I}(c, c_E)$ denotes the mutual information between c and c_E. The dependence of $\mathbb{I}(c, c_E)$ on the distribution of x and c can be removed by resorting to the upper bound $\mathbb{I}(c, c_E) \leq \frac{n}{q} C_E$, where C_E is Eve's channel capacity and q is the number of bits per transmitted symbol. As we consider generally distributed messages and take the maximum of Eve's channel mutual information over all message distributions, we are under a MIS notion according to [12].

Let us denote by $R_e = \frac{s}{n}$ Eve's equivocation rate. By using the above upper bound on $\mathbb{I}(c; c_E)$ and taking into account that Eve's equivocation cannot be negative, we have

$$R_e \geq \max\left\{0, \frac{1}{n}\left[k' - \frac{n}{q} C_E\right]\right\} = \left[R_h - \frac{C_E}{q}\right]^+.$$

Then, normalizing to the code rate, we obtain

$$\bar{R_e} \geq \frac{\left[R_h - \frac{C_E}{q}\right]^+}{R_c} = \widetilde{R_e}, \tag{3.1}$$

and, finally,

$$s = nR_e = k\bar{R_e} \geq n\left[R_h - \frac{C_E}{q}\right]^+ = \widetilde{s}. \tag{3.2}$$

In order to apply this notion of secrecy to coded transmissions over the wiretap channel, we use the wiretapper's equivocation rate as a metric. In fact, measuring Eve's equivocation allows to obtain a lower bound on the size of a list that she can reliably limit the message to [31]. By using this metric, we show that some prefixed secrecy level can be reached in practical conditions, which however is bounded away from the ultimate limits found in ideal conditions.

3.2.1 Approximate Input-Constrained Capacity

Based on (3.2), we have a simple tool to assess the minimum of Eve's total equivocation on each transmitted block. However, in order to compute it, we need to know the value of Eve's channel capacity C_E, which depends on the modulation order $M = 2^q$ and on Eve's channel SNR. Actually, C_E can be computed through classical formulations of the input-constrained channel capacity. However, though providing exact estimates, such formulations are not in closed form, and therefore not amenable to manipulate. For this reason, we exploit the following approximation based on simple logarithmic functions

$$C(\gamma) \approx \begin{cases} \alpha_1 \log_2\left(\frac{1+\alpha_2\gamma}{1+\alpha_3\gamma}\right), & \text{for } \gamma \leq \gamma_{\max}, \\ q, & \text{for } \gamma > \gamma_{\max}, \end{cases} \tag{3.3}$$

where C is the input-constrained capacity and γ is the channel SNR. The values of the parameters α_1, α_2, α_3 and $\gamma_{\max}$ for the cases of BPSK, 4-QAM and 16-QAM have been found through a least-squares fitting procedure and are reported in Table 3.1. The corresponding curves are shown in Figure 3.2, as functions of γ, and compared with their exact counterparts. From the figure, we observe that, for these cases, the approximation provided by (3.3) is very tight.

3.2.2 Channel model

In order to model the fading nature of the channels, we consider the Nakagami-m distribution [32], which is a consolidated mathematical model for small-scale fading in high-frequency radio wave propagation. We consider this type of distribution for our channel model since it is a versatile statistical distribution that can be used to model

Table 3.1: Parameters used in (3.3) to compute the approximate input-constrained capacity for BPSK, 4-QAM and 16-QAM.

Modulation	α_1	α_2	α_3	$\gamma_{\max}$ **[dB]**
BPSK	10.4798	1.4095	1.3007	5.36
4-QAM	15.1700	0.7823	0.7034	9.17
16-QAM	11.7634	0.3150	0.2441	16.43

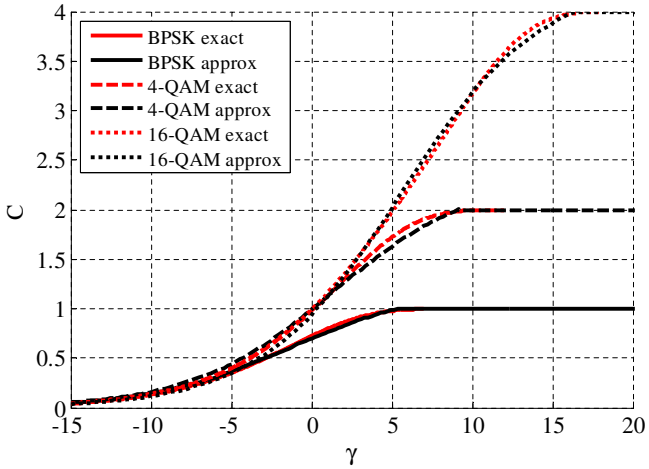


Figure 3.2: Exact and approximate input-constrained capacity for BPSK, 4-QAM and 16-QAM, as a function of the SNR γ.

a variety of fading environments, including one sided Gaussian fading (for $m = 1/2$) and Rayleigh fading (for $m = 1$). It also gives an approximation of the Rician fading that is amenable for mathematical manipulations.

According to the Nakagami-m distribution, the probability density function (p.d.f.) of the signal-to-noise ratio γ can be written as

$$p_\Gamma(\gamma) = \begin{cases} \frac{1}{\Gamma(m)} \left(\frac{m}{\bar{\gamma}}\right)^m \gamma^{m-1} e^{-\frac{m}{\bar{\gamma}}\gamma}, & \text{for } \gamma \geq 0, \\ 0, & \text{for } \gamma < 0, \end{cases} \tag{3.4}$$

where $\bar{\gamma}$ is the average SNR and $\Gamma(\cdot)$ is the Gamma function.

The parameter m represents the "shape factor" of the distribution, and it controls the severity, or intensity, of fading. Values of m lower than 1 correspond to a fading more severe than Rayleigh fading, while values higher than 1 lead to a less severe fading.

For $m = 1$ the p.d.f. of γ corresponds to the Rayleigh's model,, in which the channel gain α is the modulus of a Rayleigh complex random variable h, having two

Gaussian random variables with mean 0 and variance $1/2$ as real and imaginary parts. In this case the p.d.f. is given by

$$p_\gamma(\gamma) = \begin{cases} \frac{1}{\bar{\gamma}} e^{-\frac{\gamma}{\bar{\gamma}}}, & \gamma \geq 0, \\ 0, & \gamma < 0. \end{cases}$$

It is useful to calculate a primitive of the p.d.f. in (3.4), i.e., a solution of the integral

$$I = \int \frac{1}{\Gamma(m)} \left(\frac{m}{\bar{\gamma}}\right)^m \gamma^{m-1} e^{-\frac{m}{\bar{\gamma}}\gamma} d\gamma.$$

Indeed, this integral can be solved by exploiting the properties of the Gamma function, thus obtaining

$$I = -\frac{\Gamma\left(m, \frac{m}{\bar{\gamma}}\gamma\right)}{\Gamma(m)} + a, \tag{3.5}$$

where a is a constant and $\Gamma(\cdot,\cdot)$ is the incomplete Gamma function.

The availability of (3.3) allows us to characterize the p.d.f. of the approximate input-constrained capacity through a classical random variable analysis. More precisely, starting from (3.4) and (3.3), the following closed form expression can be easily obtained

$$p_C(C) = \begin{cases} \beta\gamma_f(C)^{m-1} e^{-\frac{m}{\bar{\gamma}}\gamma_f(C)} + \theta\delta(C-q), & 0 \leq C \leq q, \\ 0, & \text{otherwise}, \end{cases} \tag{3.6}$$

where

$$\gamma_f(C) = \frac{2^{C/\alpha_1} - 1}{\alpha_2 - \alpha_3 2^{C/\alpha_1}}$$

is the inverse of the right-hand-side (r.h.s.) of (3.3) for $\gamma \leq \gamma_{\max}$, $\theta = \frac{1}{\Gamma(m)}\Gamma\left(m, \frac{m}{\bar{\gamma}}\gamma_{\max}\right)$, $\beta = \frac{1}{\Gamma(m)}\left(\frac{m}{\bar{\gamma}}\right)^m \frac{\ln(2)(1+\alpha_2\gamma_f(C))(1+\alpha_3\gamma_f(C))}{\alpha_1(\alpha_2-\alpha_3)}$ and $\delta(x)$ denotes the Dirac delta function in $x = 0$.

For $m = 1$, the Nakagami-m model reduces to a Rayleigh fading model, and (3.6) becomes

$$p_C(C) = \begin{cases} \xi e^{-\frac{\gamma_f(C)}{\bar{\gamma}}} + e^{-\frac{\gamma_{\max}}{\bar{\gamma}}}\delta(C-q), & 0 \leq C \leq q, \\ 0, & \text{otherwise}, \end{cases} \tag{3.7}$$

where $\xi = \frac{\ln(2)(1+\alpha_2\gamma_f(C))(1+\alpha_3\gamma_f(C))}{\bar{\gamma}\alpha_1(\alpha_2-\alpha_3)}$.

From now on we consider the assumption of a SFC. This means that the channel gain α may vary between codewords, but it does not vary during transmission of a codeword. This hypothesis is realistic because, in practical wireless communication systems like those we consider, the transmission rates are usually high with respect to the channel variations and the codeword lengths are typically short.

In order to consider more realistic channel models, we can extend (3.1) to the case of the SFC, for which the statistics of Eve's channel capacity C_E can be described through the p.d.f. $p_C(C)$ in (3.7). By taking this into account in the expression (3.1) of $\widetilde{R_e}$, we obtain

$$p_{\widetilde{R_e}}(\widetilde{R_e}) = \begin{cases} qR_s p_{C_E}(\xi) + \theta\delta(\widetilde{R_e}), & 0 \leq \widetilde{R_e} \leq \frac{k'}{k}, \\ 0, & \text{otherwise}, \end{cases} \tag{3.8}$$

where $\xi = q(R_s - \widetilde{R_e}R_c)$ and

$$\theta = \Pr\{C_E > qR_s\} = \int_{qR_s}^{q} p_{C_E}(C_E)dC_E. \tag{3.9}$$

Analogously we obtain the p.d.f. of the wiretapper's equivocation $\widetilde{s}_e$ as

$$p_{\widetilde{s_e}}(\widetilde{s}_e) = \begin{cases} \frac{q}{n} p_{C_E}(\tau) + \phi\delta(\widetilde{s}_e), & 0 \leq \widetilde{s}_e \leq k', \\ 0, & \text{otherwise}, \end{cases}$$

where $\tau = q\left(R_h - \frac{\widetilde{s}_e}{n}\right)$ and $\phi = \Pr\{C_E > qR_h\} = \int_{qR_h}^{q} p_{C_E}(C_E)dC_E$.

Since for any $\rho \leq q$, from (3.7) we have:

$$\int_{\rho}^{q} p_C(C)dC = -e^{-\frac{\gamma_f(C)}{\bar{\gamma}}}\Big|_{\rho}^{q} + e^{-\frac{\gamma_{\max}}{\bar{\gamma}}},$$

the value of θ in (3.9) can be computed in closed form, *i.e.*,

$$\theta = -e^{-\frac{\gamma_f(q)}{\bar{\gamma}_E}} + e^{-\frac{\gamma_f(qR_s)}{\bar{\gamma}_E}} + e^{-\frac{\gamma_{\max}}{\bar{\gamma}_E}},$$

where $\bar{\gamma}_E$ is Eve's average SNR.

Concerning Bob's channel, we suppose that Alice has full channel state information (CSI) about it. Therefore, Bob's instantaneous SNR γ_B is known by Alice and she can decide to transmit only when it overcomes some threshold to ensure that Bob reliably receives the transmitted data.

3.3 Performance with discrete modulations and optimal codes

In this section, we estimate the performance achievable under the constraint of discrete modulation formats, but with optimal codes, that is, under the hypothesis that the coding scheme used by Alice permits to achieve the input-constrained capacity of Bob's channel. We start from static AWGN wiretap channels and then consider SFCs.

3.3.1 Optimal codes over AWGN channels

Let us denote by S_g the SNR gap between Bob and Eve, *i.e.*, the ratio between Bob's and Eve's channel SNRs, that is

$$S_g = \frac{\gamma_B}{\gamma_E}. \tag{3.10}$$

This quantity was used with the name of *security gap* in [33] and has been often used as a performance indicator for wiretap coding schemes [17, 34, 35].

In optimal conditions, Alice uses a code achieving Bob's channel capacity C_B and hence having rate $\bar{R}_c = C_B/q$. Under the hypothesis that $\gamma_B \geq \gamma_E$, we have $C_B \geq C_E$ and, from (3.1),

$$\widetilde{R_e} = \frac{C_B - C_E}{C_B}. \tag{3.11}$$

Therefore, for some given values of γ_B and γ_E (or, equivalently, S_g), based on (3.11) it is straightforward to compute the level of mutual information security $\widetilde{R_e}$ achievable over AWGN channels with optimal coding.

3.3.2 Optimal codes over Rayleigh fading channels under outage constraints

Let us consider an SFC model with full CSI about Bob's channel. In this case, we are interested to know the average value of Eve's channel SNR $\bar{\gamma}_E$ for which the probability that one realization of Eve's channel has $\gamma \geq \gamma_E = \gamma_B/S_g$ falls below some threshold ω. For this purpose, we can estimate the outage probability as

$$P_o = \int_{\gamma_E}^{+\infty} p_\Gamma(\gamma) d\gamma = \exp\left(-\frac{\gamma_E}{\bar{\gamma}_E}\right). \tag{3.12}$$

Hence, by setting $P_o = \omega$ we have $\bar{\gamma}_E = \frac{\gamma_E}{\ln\left(\frac{1}{\omega}\right)}$. For the case with fading, let us define the SNR gap as $\bar{S}_g = \frac{\gamma_B}{\bar{\gamma}_E}$. From (3.10) and (3.12) we have

$$\bar{S}_g = S_g \ln\left(\frac{1}{\omega}\right).$$

3.3.3 Optimal codes with finite length and decoding errors

In the previous subsections, we have supposed that the code used by Alice permits to achieve Bob's channel capacity. Under this hypothesis, error free transmission is reached on Bob's channel through ideally infinite length codewords. In order to consider more realistic scenarios, we can take into account the effect of a finite codeword length and in the occurrence of decoding errors by Bob.

For this purpose, we can follow the approach in [36, 37]. According to [37, eq. (37)], if we fix the code length (n) and the decoding error probability (P_e), the highest transmission rate that Bob can achieve is equal to

$$\bar{R} \approx C - \sqrt{\frac{\left(\log_2 e\right)^2}{2n}} Q^{-1}\left(P_e\right).$$

By setting $\epsilon = \sqrt{\frac{\left(\log_2 e\right)^2}{2n}} Q^{-1}\left(P_e\right)$ and $\bar{R} = q\bar{R}_c$, we obtain

$$\bar{R}_c \approx \frac{C - \epsilon}{q}. \tag{3.13}$$

3.4 Performance of practical codes

In this section, we take into account the constraints imposed by the use of practical and widespread codes. For this purpose, as a significant case study we consider the state-of-the-art LDPC codes included in the WiMax standard [38]. Therefore, the values of the code length n and the code rate R_c are chosen among those recommended by the standard. Concerning Bob, we fix an error rate target and find the corresponding SNR γ_B. Eve's SNR must instead be estimated through the metrics introduced in the previous sections, starting from the desired mutual information security level. We model Eve's channel as an SFC with average SNR $\bar{\gamma}_E$, while we assume that full CSI is available about Bob's channel.

Although perfect secrecy ($R_e/R_s = 1$) is not achievable in practice, we can define a lower threshold $\widetilde{R_e}_{\min}$ and impose that $\widetilde{R_e} \geq \widetilde{R_e}_{\min}$ unless some outage probability ω. This obviously means that $\widetilde{R_e} \geq \widetilde{R_e}_{\min}$ with probability $1 - \omega$ or higher. Using (3.8), we have

$$\int_{\widetilde{R_e}_{\min}}^{\frac{k'}{k}} p_{\widetilde{R_e}}(\widetilde{R_e}) d\widetilde{R_e} = \int_0^{q\left(R_s - \widetilde{R_e}_{\min} R_c\right)} p_{C_E}(C_E) dC_E = 1 - \omega. \tag{3.14}$$

By solving (3.14) we can find the maximum value $\bar{\gamma}_E$ of Eve's channel average SNR that is required to achieve some desired level $\widetilde{R_e}_{\min}$ of mutual information security, that is

$$\bar{\gamma}_E = \frac{2^{\frac{q\left(R_s - \widetilde{R_e}_{\min} R_c\right)}{\alpha_1}} - 1}{\alpha_2 - \alpha_3 2^{\frac{q\left(R_s - \widetilde{R_e}_{\min} R_c\right)}{\alpha_1}}} \cdot \frac{1}{\ln\left(\frac{1}{\omega}\right)}. \tag{3.15}$$

We now focus on LDPC codes with length $n = 2304$, which is the maximum length defined in the WiMax standard, considering four different code rates $(1/2, 2/3, 3/4, 5/6)$ among those supported by the system. The performance of these codes with several modulation formats has been assessed in [39], from which we can obtain the minimum value of Bob's SNR γ_B that is required to achieve a certain decoding error probability

P_e; in particular, we consider the case of $P_e \leq 10^{-4}$ for our model. Starting from these values, and fixing ω and $\widetilde{R}_{e\min}$, we can compute $\bar{\gamma}_E$ according to (3.15), as well as the corresponding $\bar{S}_g$.

For the sake of comparison, we can consider the theoretical case of an optimal code with the same length and decoding error probability, having the code rate resulting from (3.13). Such a value of $\bar{R}_c$ can be used in (3.15), solving for $\widetilde{R}_{e\min}$. This way, we obtain that the minimum level of mutual information security $\widetilde{R}_{e\min}^{(\mathrm{opt})}$ achievable in optimal coding conditions is

$$\widetilde{R}_{e\min}^{(\mathrm{opt})} = 1 - \frac{\alpha_1}{q\bar{R}_c} \log_2 \frac{1 - \bar{\gamma}_E \alpha_2 \ln(\omega)}{1 - \bar{\gamma}_E \alpha_3 \ln(\omega)}.$$

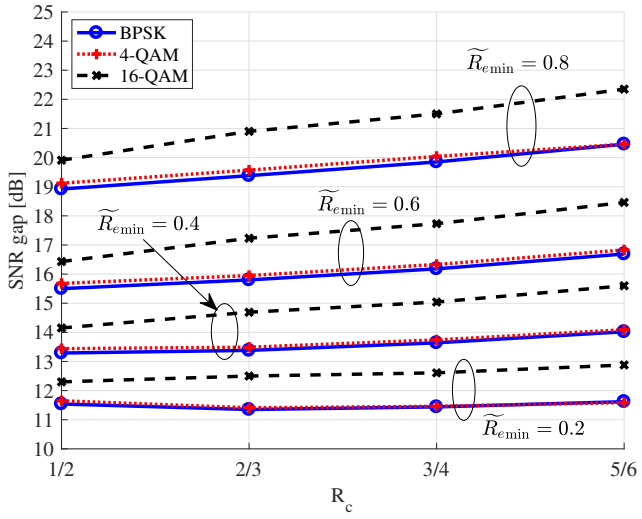
As an example, let us consider transmission over an SFC with several coding and modulation schemes compliant with the WiMax standard, under an outage constraint $\omega = 10^{-3}$. Let us focus on four target levels of mutual information security, namely, $\widetilde{R}_{e\min} = 0.2, 0.4, 0.6, 0.8$ and compute Bob's and Eve's channels threshold SNRs according to the previous analysis. The results obtained are reported in Table 3.2. As a comparison, in the table we also report the values of the code rate $\bar{R}_c$ resulting from the use of optimal codes, for the same values of Bob's channel SNR, and the corresponding MIS levels $\widetilde{R}_{e\min}^{(\mathrm{opt})}$.

From the table we observe that using standard coding schemes forces to work with smaller code rates with respect to the optimal values. This reflects into a degraded secrecy performance with respect to optimal conditions and to perfect secrecy, since we always have $\widetilde{R}_{e\min} < \widetilde{R}_{e\min}^{(\mathrm{opt})} < 1$. However, we also observe that the gap in terms of mutual information security level between standard and optimal coding is not very large. In particular, using high rate codes and low order modulation schemes allows to achieve a value of $\widetilde{R}_{e\min}$ which is very close to $\widetilde{R}_{e\min}^{(\mathrm{opt})}$.

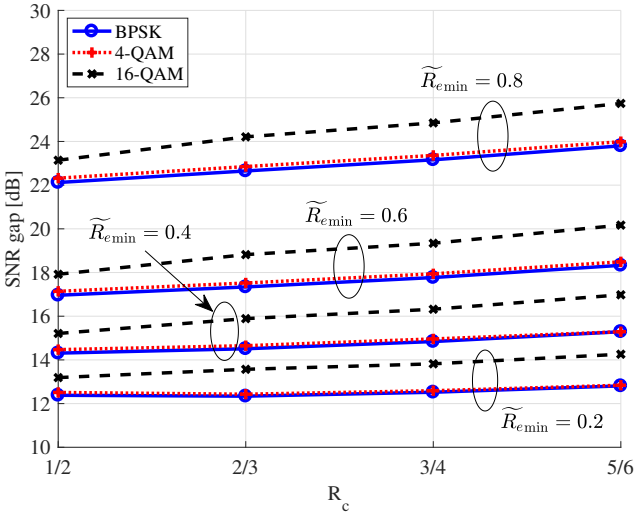
In Fig. 3.3 we report the values of the SNR gap $\bar{S}_g$ needed to achieve a level of mutual information security equal to $\widetilde{R}_{e\min}$ with the considered codes and modulation formats taken from the WiMax standard. From the figures we observe that these curves generally exhibit an increasing trend, meaning that working with low code rates is usually beneficial from the SNR gap standpoint. In addition, for a fixed $\widetilde{R}_{e\min}$, the curves corresponding to lower order modulations are below those corresponding to higher order modulations, thus yielding to the conclusion that the former outperform the latter in terms of SNR gap.

Table 3.2: Bob's and Eve's channels threshold SNRs needed to achieve a level of mutual information security equal to $\widetilde{R}_{e\min}$ with WiMax-compliant coding and modulation schemes over an SFC with $\omega = 10^{-3}$, and comparison with the MIS level $\widetilde{R}_{e\min}^{(\mathrm{opt})}$ achievable with optimal codes having rate $\bar{R}_c$, considering $k' = k$ and $k' = 0.9$ (into brackets).

				$\widetilde{R}_{e\min} = 0.2$			$\widetilde{R}_{e\min} = 0.4$			$\widetilde{R}_{e\min} = 0.8$		
R_c	Mod.	γ_B	$\bar{R}_c$	$\bar{\gamma}_E$	$\bar{S}_g$	$\widetilde{R}_{e\min}^{(\mathrm{opt})}$	$\bar{\gamma}_E$	$\bar{S}_g$	$\widetilde{R}_{e\min}^{(\mathrm{opt})}$	$\bar{\gamma}_E$	$\bar{S}_g$	$\widetilde{R}_{e\min}^{(\mathrm{opt})}$
1/2	BPSK	−1.26	0.580	−12.80	11.54	0.310	−14.55	13.29	0.483	−20.18	18.92	0.828
				(−13.64)	(12.38)	(0.396)	(−15.57)	(14.31)	(0.569)	(−23.38)	(22.12)	(0.914)
	4-QAM	1.75	0.601	−9.90	11.65	0.334	−11.69	13.44	0.501	−17.37	19.12	0.834
				(−10.76)	(12.51)	(0.418)	(−12.72)	(14.47)	(0.584)	(−20.57)	(22.32)	(0.917)
	16-QAM	7.16	0.609	−5.14	12.30	0.344	−6.99	14.15	0.508	−12.75	19.91	0.836
				(−6.03)	(13.19)	(0.426)	(−8.05)	(15.21)	(0.590)	(−15.98)	(23.14)	(0.918)
2/3	BPSK	0.58	0.723	−10.77	11.35	0.262	−12.80	13.38	0.446	−18.80	19.38	0.815
				(−11.76)	(12.34)	(0.354)	(−13.93)	(14.51)	(0.539)	(−22.07)	(22.65)	(0.908)
	4-QAM	3.59	0.744	−7.82	11.41	0.283	−9.90	13.49	0.462	−15.98	19.57	0.821
				(−8.84)	(12.43)	(0.373)	(−11.06)	(14.65)	(0.552)	(−19.26)	(22.85)	(0.910)
	16-QAM	9.55	0.754	−2.95	12.50	0.293	−5.142	14.69	0.470	−11.35	20.90	0.823
				(−4.02)	(13.57)	(0.381)	(−6.34)	(15.89)	(0.558)	(−14.66)	(24.21)	(0.912)
3/4	BPSK	1.63	0.795	−9.81	11.44	0.245	−12.01	13.64	0.434	−18.21	19.85	0.811
				(−10.89)	(12.52)	(0.339)	(−13.21)	(14.84)	(0.528)	(−21.53)	(23.16)	(0.906)
	4-QAM	4.64	0.816	−6.82	11.46	0.264	−9.10	13.74	0.448	−15.40	20.04	0.816
				(−7.95)	(12.59)	(0.356)	(−10.32)	(14.96)	(0.540)	(−18.72)	(23.36)	(0.908)
	16-QAM	10.74	0.822	−1.87	12.61	0.270	−4.30	15.04	0.452	−10.76	21.50	0.817
				(−3.08)	(13.82)	(0.361)	(−5.58)	(16.32)	(0.544)	(−14.11)	(24.85)	(0.909)
5/6	BPSK	2.76	0.858	−8.86	11.62	0.223	−11.26	14.02	0.418	−17.69	20.46	0.806
				(−10.05)	(12.81)	(0.320)	(−12.53)	(15.29)	(0.515)	(−21.04)	(23.80)	(0.903)
	4-QAM	5.77	0.880	−5.82	11.59	0.242	−8.32	14.09	0.432	−14.87	20.64	0.811
				(−7.07)	(12.84)	(0.337)	(−9.63)	(15.40)	(0.526)	(−18.22)	(23.99)	(0.905)
	16-QAM	12.12	0.889	−0.76	12.88	0.250	−3.48	15.60	0.438	−10.23	22.35	0.813
				(−2.14)	(14.26)	(0.344)	(−4.86)	(16.98)	(0.531)	(−13.62)	(23.74)	(0.906)



(a)



(b)

Figure 3.3: SNR gap (dB) between Bob's and Eve's channels needed to achieve a level of mutual information security equal to $\widetilde{R}_{e\min}$ using WiMax-compliant codes with length $n = 2304$ and rates $1/2$, $2/3$, $3/4$ and $5/6$, in conjunction with BPSK, 4-QAM and 16-QAM, considering (a) $k' = k$ and (b) $k' = 0.9k$.

3.5 OOT protocol

As mentioned above, in the model we consider, a legitimate sender (Alice) wishes to transmit some secret data to a legitimate receiver (Bob) over a fading wireless channel in the presence of a passive eavesdropper (Eve). We now propose a protocol designed to achieve this target of under reliability (towards Bob) and security (against Eve) constraints exploiting a special processing of the message prior to transmissions that relies on three main functions: encryption (with padding), slicing and encoding. Then, transmission is performed according to an OOT scheme based on channel quality estimates. All of these elements of the protocol are described next.

3.5.1 Encryption

Let us denote as M the private data that Alice wishes to securely transmit to Bob. First of all, she transforms M into X through an AONT. The concept of AONT was introduced by Rivest in [40] and can be seen as a random-like transformation that is infeasible to invert, even partially, unless the transformed data is completely available. Therefore, an AONT-processed message cannot be recovered, even in part, if some part of it is missing. In the original proposal [40], the message is divided in blocks of fixed length and a random encryption key is generated and used to encrypt each block through some symmetric block cipher. The random key is necessary to feed the symmetric cipher, but it does not represent a secret to be shared between legitimate users. Then, a cryptographic hash function is used to compute the digests of all blocks, that are XORed together and with the random key. This way, a last block is obtained that is appended to the transformed message. Therefore, the AONT can be easily inverted (by inverting the above procedure) by anyone retrieving the entire amount of transformed data, without the need of any prior knowledge of the random key (since it is embedded with the transformed data). The procedures of AONT encryption and decryption are schematically described in Figure 3.4. This first implementation of an AONT relies on cryptographic primitives, and hence follows a computational security paradigm. However, it has been shown in [41] that it is also possible to define AONTs with unconditional security.

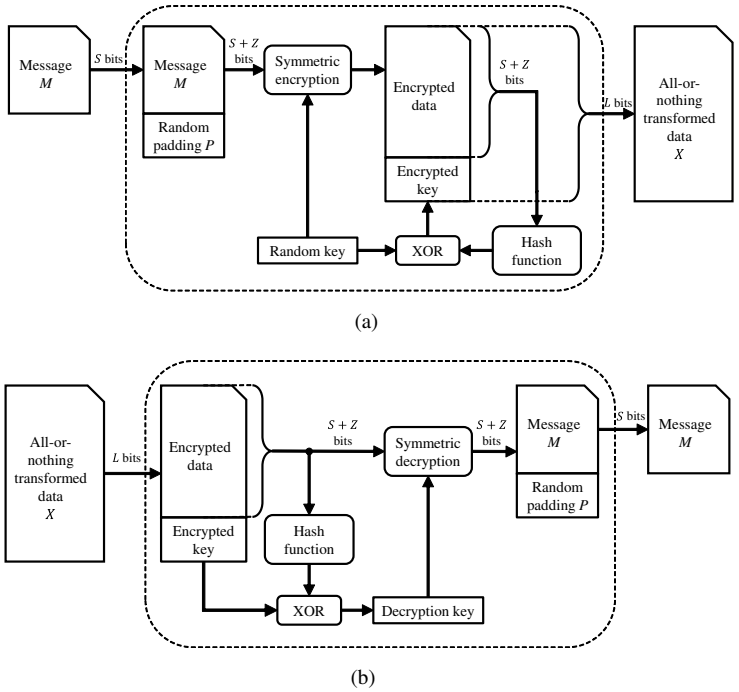


Figure 3.4: Block diagram of: (**a**) AONT encryption and (**b**) AONT decryption of a secret message M.

Coherent with the figure, let us denote the length of M in bits as S and the length of X in bits as L. Before application of the AONT, the message M is concatenated with a random padding string P having length $Z \geq 0$, i.e., $X = \mathrm{A}ONT([M|P])$, where $|$ denotes concatenation. We have $L \geq S + Z$ and the value of Z is chosen such that L is a multiple of an integer k coincident with the dimension of the binary linear block code $C_1(n, k)$ used in the encoding step (see Section 3.5.3).

3.5.2 Slicing

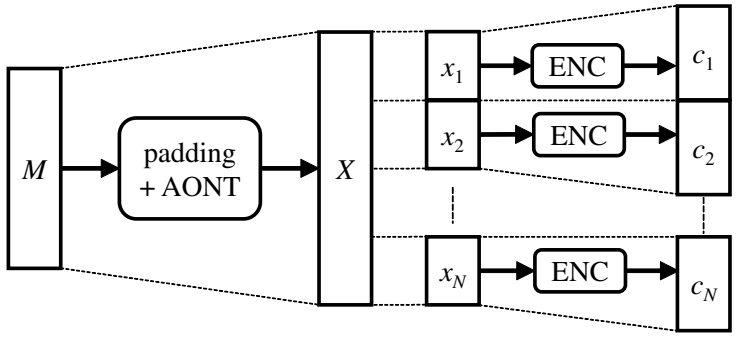
As shown in Figure 3.5a, X is split into N blocks which are then separately encoded. Each block is called slice and is k bits long. The length k coincides with the dimension of the linear block code $C_1(n, k)$ used in the subsequent encoding phase, and the number of slices is $N = L/k$. As also shown in Figure 3.5a, we denote the i-th block as x_i, $i = 1, 2, 3, \ldots, N$.

3.5.3 Encoding

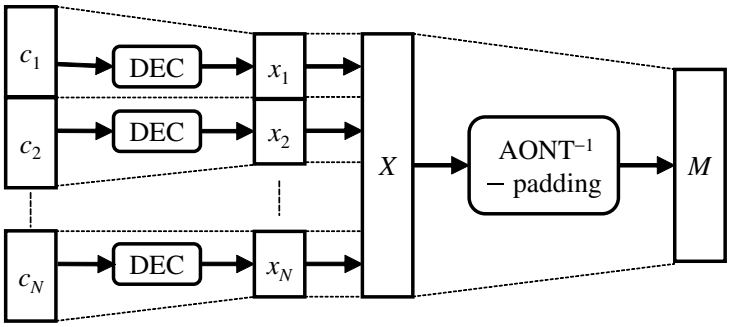
Each block is encoded through $C_1(n, k)$, where n denotes the code length and k is the code dimension. This way, Alice obtains a set of n-bit codewords c_i, $i =$

$1, 2, 3, \ldots, N$, which are then modulated and serially transmitted to Bob over the wireless channel.

When these codewords are received, they are decoded into x_i, $i = 1, 2, 3, \ldots, N$, through the decoder of $C_1(n, k)$. This way, X can be recovered and the AONT can be inverted. Then, the random padding P is discarded and the secret message M is re-obtained. The whole procedure for transforming a secret message M into a set of n-bit codewords c_i, $i = 1, 2, 3, \ldots, N$, to be transmitted through an ideal error-free channel using the proposed protocol and its inverse is schematically depicted in Figure 3.5.



(a)



(b)

Figure 3.5: Block diagram of: (**a**) the procedure for transforming a secret message M into a set of n-bit codewords c_i, $i = 1, 2, 3, \ldots, N$, to be transmitted and (**b**) its inverse.

3.5.4 Transmission

The OOT transmission scheme we consider works as follows. Transmission is synchronous, i.e., organized in time slots, each slot having the duration of n bits. This coincides with the length of a codeword c_i, $i = 1, 2, 3, \ldots, N$. Each transmission starts at the beginning of a time slot and lasts for n bits or less. Before transmitting

any data packet, the following preliminary steps are performed:

1. Alice sends a request to send (RTS) message to Bob containing some known string.

2. Based on the received string, Bob estimates the channel SNR, noted as γ_{AB}.

3. Bob sends γ_{AB} to Alice after syndrome encoding through a second binary linear block code $C_2(f, b)$, as explained next. The length of the syndrome is $z = f - b$. Bob's reply is protected through a third, ultra-reliable binary linear block code $C_3(n, z)$ such that all channel errors can be corrected with very high probability. This can be achieved by designing the system parameters in such a way that the rate of $C_3(n, z)$ is much smaller than the rate of $C_1(n, k)$, that is, $z \ll k$.

4. Alice decodes Bob's message through the decoder of $C_3(n, z)$ and recovers its original content, that is, the syndrome of γ_{AB}.

5. By comparing the decoded data with the received signal, Alice estimates the channel SNR, noted as γ_{BA}.

6. Alice reconciliates her estimate (γ_{BA}) with that of Bob (γ_{AB}) by exploiting the procedure explained next.

The aim of the reconciliation phase is to allow γ_{BA} to converge to γ_{AB}, i.e., allow Alice to obtain the same information of Bob as regards the channel SNR. Noting by γ_B this common value, if $\gamma_B \geq \gamma_B^*$, where γ_B^* is a prefixed threshold, Alice will then transmit a codeword c_i containing valid data. Otherwise, Alice will transmit a fake packet, which is recognized as such by Bob and hence discarded. The aim of the fake packet transmission is to confuse Eve. The entire procedure is summarized in Figure 3.6. In the following, this protocol will be named OOT-FP.

3.5.5 Reconciliation

As mentioned above, the main purpose of the reconciliation phase is to find a common estimate of the channel SNR, in order to allow both Alice and Bob to verify if this value overcomes the threshold γ_B^* or not. For this purpose, let us suppose that Alice represents γ_{BA} with a $[1 \times f]$ binary string s_A. Similarly, Bob represents γ_{AB} with a $[1 \times f]$ binary string s_B. Since, in general, $\gamma_{AB} \neq \gamma_{BA}$ the same will be for the two binary strings, i.e., $s_B \neq s_A$. The binary representation used by Alice and Bob, that can be the result of a quantization followed by a suitable mapping, must be chosen such that the following requirements are fulfilled:

- s_A and s_B must be two dense binary vectors, that is, their Hamming weights $w_H(s_A)$ and $w_H(s_B)$ must be on the order of $f/2$.

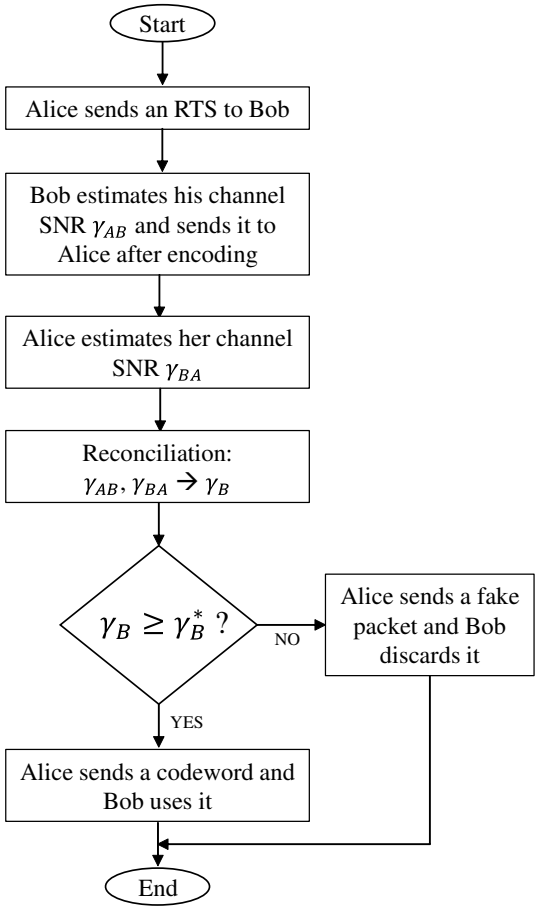


Figure 3.6: Flow chart of the transmission of a packet according to the OOT-FP protocol.

- Noted by e the binary sum of s_A and s_B (that is, $e = s_A \oplus s_B$, where $\oplus$ denotes the XOR operation), with a very high probability, it must be $w_H(e) \leq t \ll f$, being t a parameter chosen in advance. In other terms, the binary representation used for the SNR values must be chosen in such a way that small differences between γ_{AB} and γ_{BA} translate into low-weight difference vectors between s_A and s_B. Practical binary representations with this feature can be easily devised.

The reconciliation phase exploits the binary linear block code $C_2(f, b)$ having length f and rate b/f, able to correct t errors or less. Contrary to $C_1(n, k)$, that uses soft-decision decoding, $C_2(f, b)$ exploits hard-decision syndrome decoding. As an example, a classical Bose–Chaudhuri–Hocquenghem (BCH) code provided with Berlekamp–Massey hard-decision decoding [42] can be used as C_2. Its $z \times f$ parity-check matrix is denoted by H in the following. A detailed description of the operations performed during the

reconciliation phase is provided next. After having estimated γ_{AB} and converted it into the binary string s_B, Bob computes the syndrome h_B, which is a $1 \times z$ vector obtained as

$$h_B^T = H \cdot s_B^T,$$

where $\cdot$ denotes the matrix-vector multiplication and T denotes transposition. Bob then encodes h_B through the ultra-reliable low-rate binary linear block code $C_3(n, z)$ mentioned above and sends it to Alice. This ultra-reliable code should allow Alice to correct all transmission errors through decoding. The rare cases in which this does not occur are commented next. When Alice receives h_B without errors, she computes

$$h^T = h_B^T \oplus H \cdot s_A^T = H \cdot s_B^T \oplus H \cdot s_A^T = H \cdot (s_A \oplus s_B)^T = H \cdot e^T.$$

Since $w_H(e) \leq t$, Alice is able to recover e from h through syndrome decoding of $C_2(f, b)$. Finally, by knowing e, Alice can easily compute

$$s_B = s_A \oplus e.$$

At this point, both Alice and Bob know the same SNR value $\gamma_B = \gamma_{AB}$ and, by comparing it with γ_B^*, they can consistently decide whether the packet must be an information packet (if $\gamma_B \geq \gamma_B^*$) or a fake packet (if $\gamma_B < \gamma_B^*$). In the rare cases in which Alice is unable to correct all errors on h_B, she obtains $\tilde{h}_B \neq h_B$ and computes

$$\tilde{h}^T = \tilde{h}_B^T \oplus H \cdot s_A^T = H \cdot \tilde{s}_B^T \oplus H \cdot s_A^T = H \cdot (s_A \oplus \tilde{s}_B)^T = H \cdot \tilde{e}^T.$$

For the properties of syndromes, $\tilde{s}_B$ is generally significantly different from s_B. Therefore, $\tilde{e}$ has a large weight, yielding a failure in syndrome decoding that is detected by Alice. Hence, Alice becomes aware of the failure and can restart the procedure. Being a rare event, for the sake of simplicity, this fact will be not considered in the following.

Concerning Eve, after receiving Alice's RTS through her channel, she estimates a value γ_{AE} of the SNR of her channel, and represents it through a $1 \times f$ vector s_E, which, however, is different from s_B. Thus, even if we assume that Eve (who can use the same decoders of the legitimate users) can correctly recover the syndrome h_B transmitted by Bob:

- She cannot recover s_B since $w_H(s_B)$ largely exceeds the correction capability of $C_2(f, b)$ under hard-decision syndrome decoding. On the other hand, the values of the parameters are such to prevent that Eve successfully decodes s_B even with more powerful soft-decision decoders. Moreover, soft-decision decoding is practically infeasible for several families of classical codes (like BCH codes) unless their length is very short.

- She could try to exploit (3.5.5) with s_E in place of s_A, but we suppose that her channel is independent of Bob's one and different enough, so that $w_H(s_E \oplus s_B) = w_H(e') > t$, such that syndrome decoding does not permit Eve to recover e' from $h_B^T \oplus H \cdot s_E^T$. The meaning and applicability of this assumption will be further discussed in Section 3.5.6.

Therefore, Eve is not able to discover γ_B and, consequently, she has no information to decide whether the packet flowing from Alice to Bob after reconciliation is an information packet or a fake packet.

In order to measure the advantage coming from the use of fake packets, in the following we will compare the performance of the OOT-FP protocol with that of a basic version of it without fake packets, simply denoted as OOT. As in the OOT-FP protocol, in OOT Alice and Bob perform the initial estimate and reconciliation phases. After reconciliation, Alice compares γ_B with γ_B^* and, when $\gamma_B < \gamma_B^*$, instead of sending a fake packet, she simply skips transmission and restarts the procedure at the next time slot. In the OOT protocol, Bob still estimates $\gamma_B = \gamma_{AB}$ and compares it with the threshold γ_B^* but, differently from the setting we consider in OOT-FP, Bob sends to Alice a clear to send (CTS) message if $\gamma_B \geq \gamma_B^*$, or a not available (NA) message otherwise. Therefore, Alice decides whether to transmit a codeword or to skip transmission based on Bob's reply. This has the advantage of reducing complexity by avoiding computations needed for reconciliation of Alice's and Bob's estimates of the SNR. On the other hand, the broadcast transmission of CTS packets provides Eve with a clear indication of which time slots Alice is going to use to transmit valid codewords.

As a performance metric, in the following we use the number of time slots needed to transmit a message M with a given security level. In this respect, we ignore the processing times required at Bob's and Alice's premises (this is a reasonable assumption, since the processors' speed is usually orders of magnitude greater than the transmitters' speed).

3.5.6 Applicability of the Protocol

From the description given in Section 3.5.4, it immediately follows that the feasibility of the proposed protocol is conditioned on the following hypotheses:

1. Eve's channel SNR is significantly different from that of the main channel between Alice and Bob, so that its binary representation is also significantly different.

2. The main channel between Alice and Bob remains stationary during the execution of the steps required by the protocol for transmitting each packet (this means we are considering *slow fading*).

Concerning Hypothesis 1, we can rely on it because we assume that the main and eavesdropper's channels fade independently. Moreover, in the following analysis we often consider that Eve's channel is significantly degraded (with an SNR penalty in the order of 3 dB or more) with respect to the main channel. In those cases in which this does not occur, that are also of interest in practice, we have that the secret message is spread on a long sequence of (50 ore more) packets. The latter condition means that, even if in this case Eve could successfully attack the reconciliation phase of some packet transmissions, this is very unlikely to occur for all the packets of a sequence, and the use of AONTs prevents Eve from gathering any partial information about the secret message.

Hypothesis 2 instead depends on the channel and transmission characteristics. Since we focus on short packets, it is likely that the channel can be considered stationary during each execution of the protocol. For the sake of simplicity, we will denote each three-way protocol execution as a *single packet session* in the following. In general terms, the assumption of stationary channels during each single packet session requires the use of codes with short length (i.e., small values of n) and high order modulations. Therefore, when missing, such a condition can be restored by changing the data rates, the coding rates and/or the modulation order.

We can observe that Eve cannot mount active attacks, like transmitting fictitious RTS packets or impersonating Bob, since these would be detected by Alice and Bob, due to the broadcast nature of the wireless channel.

On Bob's side, the whole set of codewords must be received and decoded into the vectors x_i, $i = 1, 2, 3, \ldots, N$. Then, such vectors are used to reconstruct X and to invert the AONT to recover M. Due to the presence of the AONT, the message M can be recovered only on condition that all its corresponding codewords are correctly received and decoded, such that the AONT can be inverted. This means that, in the proposed protocol, reliability is not less important than security.

In the next sections, we show that it is possible to achieve some desired level of SS through this protocol even when the average SNR of Bob's channel is lower than the average SNR of Eve's channel. This is somehow in contrast with previous results concerning the wiretap channel, where it has been shown that in such conditions there cannot be secrecy unless some retransmission scheme is exploited [29]. Indeed, there is no contradiction with our results, since we adopt a different notion of secrecy with a different target, that is, SS against passive attackers with computational constraints. The following analyses also show that the OOT-FP protocol achieves transmission of an S-bit message by requiring a significantly smaller number of time slots with respect to the OOT protocol, though maintaining the same security level.

3.6 Security level

In the OOT-FP protocol, the total equivocation on a packet can be obtained as the sum of two contributions as follows: $\widetilde{s}_{tot} = \widetilde{s}_e + \widetilde{s}_{fk}$. The term $\widetilde{s}_e$ represents Eve's equivocation due to her channel, and it is the only contribution present in the OOT protocol. The term $\widetilde{s}_{fk}$ represents Eve's equivocation due to the presence of the fake packets, that, in turn, depends on Bob's channel.

Let us denote by p_{fk} the probability to have a fake packet, i.e., the probability that the channel SNR between Alice and Bob is below some prefixed threshold γ_B^*. Since Eve's and Bob's channels are independent, Eve's equivocation about the valid or fake nature of each packet can be written as the binary entropy following from p_{fk}, that is,

$$\widetilde{s}_{fk} = -p_{fk}\log_2 p_{fk} - (1-p_{fk})\log_2(1-p_{fk}). \tag{3.16}$$

The value of p_{fk} can be easily computed as

$$\begin{aligned} p_{fk} &= \mathrm{P}r\left\{\gamma_B < \gamma_B^*\right\} = 1 - \mathrm{P}r\left\{\gamma_B \geq \gamma_B^*\right\} \\ &= 1 - \frac{1}{\Gamma(m)}\left(\frac{m}{\bar{\gamma}_B}\right)^m \int_{\gamma_B^*}^{\infty} \gamma_B^{m-1} e^{-\frac{m}{\bar{\gamma}_B}\gamma_B} d\gamma_B. \end{aligned} \tag{3.17}$$

Let us express the average SNR experienced by Bob in terms of the threshold γ_B^* as follows:

$$\bar{\gamma}_B = \gamma_B^* \cdot \Omega, \tag{3.18}$$

that is, Ω is the ratio of Bob's channel average SNR to its threshold value. The last integral in (3.17) can be solved exploiting (3.5), thus obtaining

$$p_{fk} = 1 - \left[-\frac{\Gamma\left(m, \frac{m}{\bar{\gamma}_B}\gamma_B\right)}{\Gamma(m)}\right]_{\gamma_B^*}^{\infty} = 1 - \frac{\Gamma\left(m, \frac{m}{\Omega}\right)}{\Gamma(m)}. \tag{3.19}$$

Eve's maximum equivocation about the valid or fake nature of each packet occurs when p_{fk} is equal to 0.5. In fact, in this case, we have $\widetilde{s}_{fk} = 1$, since the occurrence of a valid or a fake packet is equally probable. The value of Ω that yields $p_{fk} = 0.5$ can be obtained from (3.19). As an example, for $m = 1$, it results in $\Omega = \frac{1}{\ln 2} = 1.44$, i.e., about 1.58 dB. However, the choice of $\bar{\gamma}_B$ is influenced also by Bob's error probability, i.e., reliability requirements. For this reason, in the following, we do not assume to be in the optimal situation with $\widetilde{s}_{fk} = 1$, but we compute the value of $\widetilde{s}_{fk}$ following from the required γ_B^* through (3.19) and (3.16). Then, we consider the same ratio Ω for all modulation formats and code rate.

We now estimate the wiretapper's equivocation $\widetilde{s}_e$ which is due to the quality of the Alice–Eve channel, expressed by the SNR value γ_E. Indeed, $\widetilde{s}_e$ is the only contribution present in the OOT protocol, while, in the case of OOT-FP, it sums up with the

contribution $\widetilde{s}_{fk}$, due to the inclusion of the fake packets, in order to obtain the total equivocation $\widetilde{s}_{tot}$.

Let P_o denote the equivocation outage probability, i.e., the probability that $\widetilde{s}_e$ falls below some lower threshold $\widetilde{s}_{\min} > 0$. Once having fixed the total equivocation $\widetilde{s}_{tot}$, according to the desired security level, the value of $\widetilde{s}_{\min}$ results as $\widetilde{s}_{\min} = \widetilde{s}_{tot} - \widetilde{s}_{fk}$. Obviously, for the OOT protocol, $\widetilde{s}_{fk} = 0$ and $\widetilde{s}_{\min} = \widetilde{s}_{tot}$. By definition, we have

$$P_o = \int_0^{\widetilde{s}_{\min}} p_{\widetilde{s}_e}(\widetilde{s}_e) d\widetilde{s}_e = 1 - \int_{\widetilde{s}_{\min}}^{k'} p_{\widetilde{s}_e}(\widetilde{s}_e) d\widetilde{s}_e.$$

Equivalently, $1/P_o$ is the number of channel realizations within which $\widetilde{s}_e \leq \widetilde{s}_{\min}$ occurs once, on average. Thus, taking into account possible outage events, we must impose $1/P_o \geq 2^{\widetilde{s}_{\min}}$ or, explicitly,

$$\log_2 \left(\frac{1}{P_o} \right) \geq \widetilde{s}_{\min}, \tag{3.20}$$

which fixes the MIS level.

Since $\widetilde{s}_{\min} > 0$, we have

$$\begin{aligned} \int_{\widetilde{s}_{\min}}^{k'} p_{\widetilde{s}_e}(\widetilde{s}_e) d\widetilde{s}_e &= \int_{\widetilde{s}_{\min}}^{k'} \frac{q}{n} p_{C_E}(\tau) d\widetilde{s}_e \\ &= \int_0^{q\left(R_h - \frac{\widetilde{s}_{\min}}{n}\right)} p_{C_E}(\tau) d\tau. \end{aligned}$$

By replacing (3.6), with $C = C_E$ and using (3.5), we obtain

$$P_o = \frac{1}{\Gamma(m)} \Gamma \left(m, \frac{m}{\bar{\gamma}_E} \eta \right), \tag{3.21}$$

where $\bar{\gamma}_E$ is the average SNR of the channel between Alice and Eve, and $\eta = \gamma_f \left(q \left(R_h - \frac{\widetilde{s}_{\min}}{n} \right) \right)$.

Through (3.20) and (3.21), we can compute the value of $\bar{\gamma}_E$ required to reach a given value of $\widetilde{s}_{\min}$. Because of the presence of the incomplete Gamma function, this value is not easily obtainable in closed form. In essence, the problem consists of calculating $\bar{\gamma}_E$ such that

$$\Gamma(m) 2^{-\widetilde{s}_{\min}} \geq \Gamma \left(m, \frac{m}{\bar{\gamma}_E} \eta \right).$$

We must note that $\widetilde{s}_{\min}$ appears on both sides of this inequality: explicitly at the left-hand-side (l.h.s.) and implicitly in the computation of η at the r.h.s.. According to our target, however, the unknown variable is the upper bound on $\bar{\gamma}_E$. Thus, we set $y = \frac{m}{\bar{\gamma}_E} \eta$ and we evaluate the inverse incomplete Gamma function. Once having

found the value of y, $\bar{\gamma}_E$ is easily computed as

$$\bar{\gamma}_E \leq \frac{m}{y}\eta = \bar{\gamma}_E^*. \tag{3.22}$$

If we consider the special case of Rayleigh fading, it is easy to verify that $y = \widetilde{s}_{\min}\ln(2)$, so that (3.22) becomes

$$\bar{\gamma}_E \leq \frac{\eta}{\widetilde{s}_{\min}\ln(2)} = \bar{\gamma}_E^*. \tag{3.23}$$

Based on (3.22) and (3.23) (for the special case), we can define the conditions under which some given level of MIS is achieved.

3.6.1 Design Criteria

Equation (3.22) defines the upper threshold $\bar{\gamma}_E^*$ that must be imposed on Eve's channel quality in order to meet the security requirements. More precisely, imposing such an upper threshold ensures that Eve must perform $2^{\widetilde{s}_{\min}}$ attempts, on average, to fully recover a transmitted codeword. This must be considered in addition to the lower threshold γ_B^* on Bob's channel quality, which instead is required to achieve some reliability target for transmission from Alice to Bob. These two thresholds can be collected into the parameter

$$S_g = \frac{\gamma_B^*}{\bar{\gamma}_E^*}, \tag{3.24}$$

which represents the minimum SNR gap between Bob's and Eve's channels that is required to achieve both the reliability and the security targets.

However, we still need to avoid that Eve is able to gather some (even small) part of a transmitted codeword with less attempts than $2^{\widetilde{s}_{\min}}$. This is achieved by preprocessing the messages through the AONT. The use of the AONT also allows for concatenating together N data blocks before transmission, which are processed together through the AONT itself. This way, Eve needs to correctly recover all the N codewords corresponding to those blocks before being able to invert the AONT. Therefore, $2^{N\widetilde{s}_{tot}}$ attempts are required on average by Eve to correctly recover the whole set of packets and invert the AONT to recover the secret bits. In other words, the use of the AONT allows us to achieve SS from MIS, while this form of concatenation allows us to tune the security level according to the desired target. Numerical examples are given in the next section.

3.7 Numerical results

In order to assess the performance of the protocols we consider, let us compute the number of time slots required to achieve a given security level, i.e., a prefixed value of

$\widetilde{s}_{\min}$, as a function of the SNR gap. This way, we can also compare the performance of the OOT protocol with that of the OOT-FP protocol. For the sake of fairness, the comparison assumes the same value of S, that is, the length of the secret message M. Instead, based on Eve's equivocation achieved by each protocol, the resulting length of the random padding Z may be different. Considering that the handshaking phase between Alice and Bob, i.e., the exchange of initial messages and the subsequent reconciliation, involves the same number of time slots in the two protocols, the following analysis will be focused on the number of single packet sessions required to successfully transmit a message under given reliability and security constraints.

In order to refer to a significant, practical example concerning wireless transmissions, we suppose that the code $C_1(n,k)$ is one of the LDPC codes included in the WiMax standard [38], and we consider modulation schemes compliant with the same standard. In WiMax, four code rates ($1/2$, $2/3$, $3/4$, $5/6$) and several code lengths for each code rate are supported. As an example, we focus on codes with length $n = 2304$, but the analysis can be obviously extended to the other code lengths. Since our protocol relies on the hypothesis that the channel does not vary during each single packet session between Alice and Bob, for a given transmission bandwidth we suppose that the modulation order and the channel symbol rate are chosen in such a way that the channel coherence time is longer than a single packet session. The chance to actually meet this constraint depends on any specific setting. In the following numerical examples, we suppose that such a condition is met with $n = 2304$ and $q = 1, 2, 4$, but the analysis could be obviously repeated by assuming shorter codes and higher order modulations.

Denoting the desired security level as SL (expressed in bits), the number of codewords that is necessary to transmit for achieving SL-bit security is easily obtained as

$$N = \frac{SL}{\widetilde{s}_{tot}},$$

where $\widetilde{s}_{tot} = \widetilde{s}_{\min} + \widetilde{s}_{fk}$ for the OOT-FP protocol and $\widetilde{s}_{tot} = \widetilde{s}_{\min}$ for the OOT protocol.

Although a direct comparison with a complete WiMax system is not possible (since we focus on one instance of the physical layer and neglect higher layer features like channel adaptivity of coding and modulation techniques), we observe that N represents an upper bound on the overhead introduced by our method with respect to plain transmission. Such an upper bound is reached when data to be transmitted are so few that they are contained in a single codeword, while all the other $N - 1$ codewords must be filled with encoded padding bits. In this case, the overhead introduced by our protocol might be large. However, we must consider that this is an upper bound, while the average overhead depends on the statistical features of the source, whose consideration is out of the scope of this work.

Once having fixed the average quality required for the main channel and the result-

ing threshold γ_B^*, the value of p_{fk} can be determined from (3.19), and then that of $\widetilde{s}_{fk}$ follows from (3.16). For a range of $\widetilde{s}_{\min}$ values, the values of $\bar{\gamma}_E^*$ can then be computed according to (3.22). The number of single packet sessions finally results as

$$N_{sps} = \frac{N}{1 - p_{fk}},$$

and can be plotted as a function of S_g.

The numerical values obtained depend on the particular choices made for the degrees of freedom of the system. Indeed, the performance of the codes we consider with several modulation formats over different channels has been assessed in [39]. In order to provide some significant examples, let us fix the reliability requirement in terms of a decoding error probability $\leq 10^{-4}$ experienced by Bob. We note that the decoding error probability, that is the complement of the probability of successful transmission, is an input of our model since the target reliability of the system is fixed beforehand. The corresponding values of γ_B^* obtained from [39] are reported in Table 3.3. Let us also consider a ratio $\Omega = 3$ dB of the average quality of the main channel to its threshold value. According to (3.18), this fixes the value of Bob's average SNR.

R_c	1/2	1/2	1/2	2/3	2/3	2/3
Mod.	BPSK	4-QAM	16-QAM	BPSK	4-QAM	16-QAM
γ_B^*	-1.26	1.75	7.16	0.58	3.59	9.55
R_c	3/4	3/4	3/4	5/6	5/6	5/6
Mod.	BPSK	4-QAM	16-QAM	BPSK	4-QAM	16-QAM
γ_B^*	1.63	4.64	10.47	2.76	5.77	12.12

Table 3.3: Values of γ_B^* (in dB) required to achieve decoding error probability $\leq 10^{-4}$ for LDPC codes with $n = 2304$ and several rates and modulation schemes compliant with WiMax.

In Figures 3.7 and 3.8, we compare the results obtained using the OOT-FP protocol (continuous line) with those achieved by the OOT protocol (dashed line), for two different modulation formats, namely BPSK and 16-QAM, and a couple of WiMax code rates. The figures report the values of N_{sps}, that is the number of single packet sessions, required to achieve a decoding error probability $\leq 10^{-4}$ (towards Bob) and a semantic security of 128 bits (against Eve). The independent variable is the ratio S_g, defined by (3.24). Moreover, as an example, we fix k' equal to $0.9k$. In order to match binary coding with non-binary modulations, we follow a pragmatic approach,

according to which coding is applied first and modulation acts downstream on groups of encoded symbols.

As it is reasonable and expected, independently of the protocol used, the values of N_{sps} become smaller and smaller for increasing S_g. In the figures we consider channels with different fading intensity: from $m = 0.5$, which is the minimum value allowed for the Nakagami-m distribution, to $m = 5$, which represents a more stationary channel. The case with $m = 1$, coinciding with the Rayleigh fading model, is considered as well.

From these figures, we observe that, interestingly, the proposed protocol is able to achieve 128-bit SS even when the average SNR of Eve's channel is comparable to or even slightly better than that of Bob's channel, although this is obviously paid in terms of an increased number of single packets sessions needed to achieve such a security level. Finally, and very important, we observe that the use of fake packets provides a significant improvement in performance, since Alice needs to transmit for a shorter time, measured in time slots, in order to reach 128-bit SS.

Also the value of m has a relevant impact, both in absolute terms and as regards the comparison between the two protocols. In particular, we must consider that strong fading, as described by low values of m, yields to a greater number of fake packet transmissions that hence dominate Eve's equivocation. On the contrary, more stationary channels yield less fake packet transmissions. This reflects on higher values of N_{sps} when Eve's channel is better (or, at least, not significantly worse) than Bob's channel, since fake packets are those responsible for Eve's equivocation in these conditions.

Figures 3.9 and 3.10, in turn, highlight the differences in terms of performance between various code rates and modulation formats, respectively. In this case, the analysis has been repeated for different values of k'. The value of the shape factor has been fixed to $m = 3$. Although the variations are almost negligible, we can observe from Figure 3.9 that using high rate codes permits us to achieve the target security level with a smaller number of time slots with respect to low rate codes. Instead, from Figure 3.10, we see that using high order modulation schemes is not beneficial from the number of time slots standpoint. 16-QAM, in particular, always requires the largest amounts of single packet sessions to achieve the target security level. It must be said that high order modulation schemes may be needed in order to ensure that the channel can be considered stationary during each single packet session. In these cases, using high order modulations still allows for achieving the desired SS level with a moderate increase in the number of required single packet sessions.

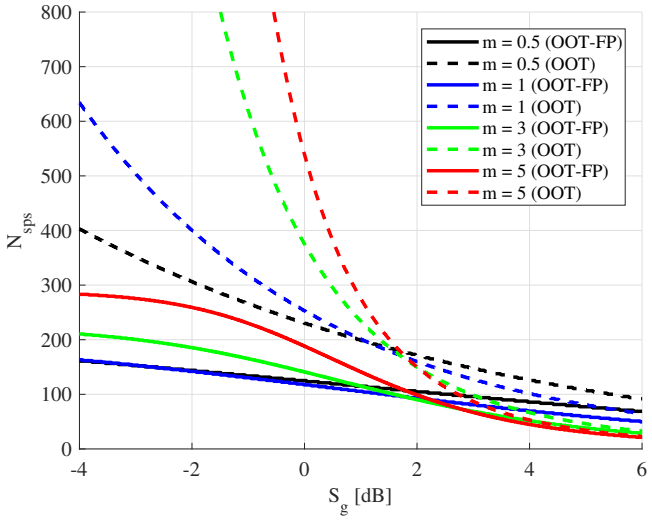


Figure 3.7: Number of single packet sessions needed to achieve 128-bit SS versus SNR gap with WiMax LDPC codes having length $n = 2304$, rate $R_c = 1/2$ and BPSK, for the case of $k' = 0.9k$.

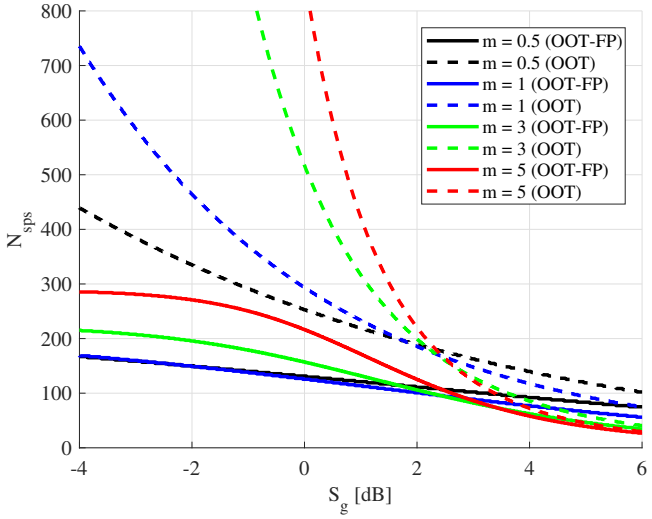
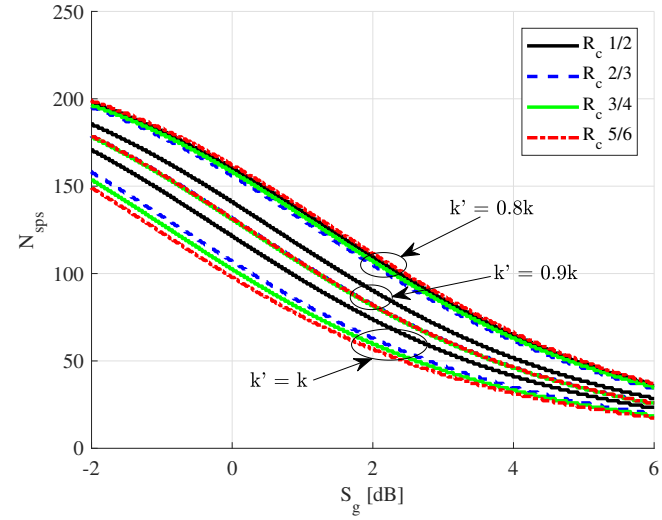
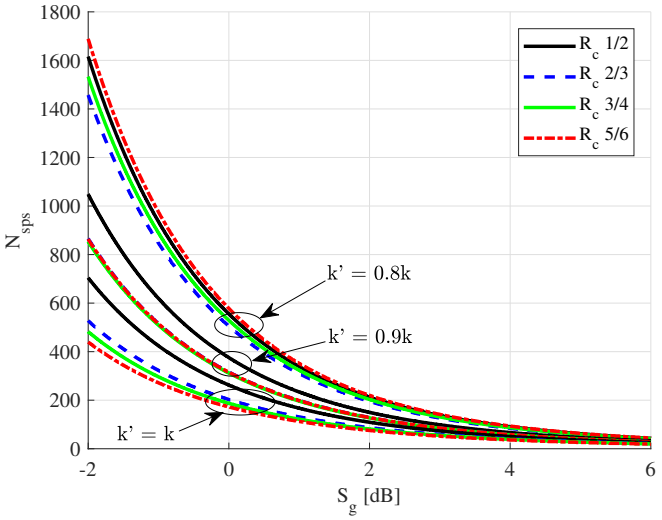


Figure 3.8: Number of single packet sessions needed to achieve 128-bit SS versus SNR gap with WiMax LDPC codes having length $n = 2304$, rate $R_c = 5/6$ and 16-QAM, for the case of $k' = 0.9k$.

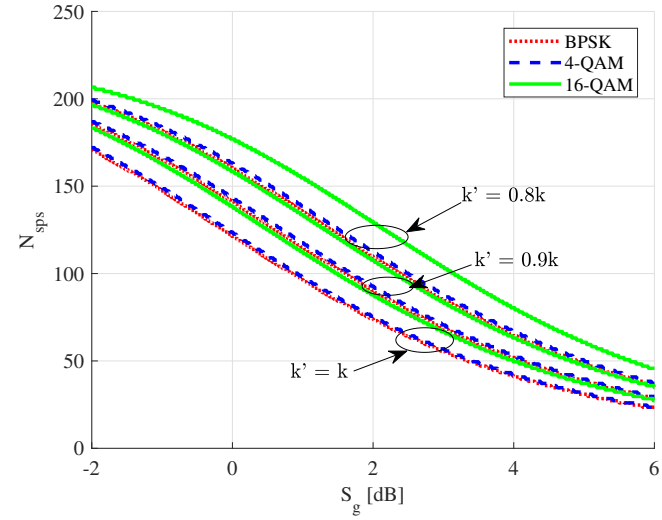


(a)

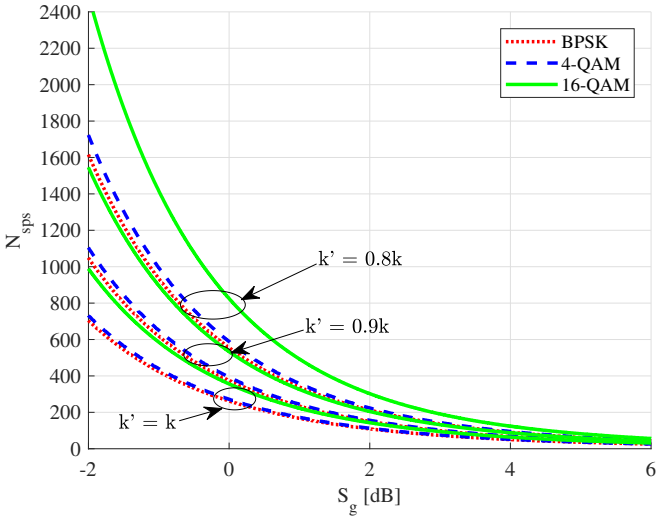


(b)

Figure 3.9: Number of single packet sessions needed to achieve 128-bit SS in terms of SNR gap with WiMax LDPC codes having length $n = 2304$, BPSK and different rates, for the cases of $k' = 0.8k$, $k' = 0.9k$ and $k' = k$, and shape factor $m = 3$, using: (**a**) the OOT-FP protocol and (**b**) the OOT protocol.



(a)



(b)

Figure 3.10: Number of single packet sessions needed to achieve 128-bit SS in terms of SNR gap with WiMax LDPC codes having length $n = 2304$, rate $R_c = 1/2$ and three different modulations, for the cases of $k' = 0.8k$, $k' = 0.9k$ and $k' = k$, and shape factor $m = 3$, using: (**a**) the OOT-FP protocol and (**b**) the OOT protocol.

3.8 Summary

We have defined a set of tools to assess the secrecy performance of practical coding and modulation schemes used for transmissions over fading wiretap channels. Fading has been modeled by means of the Nakagami-m distribution, this way representing a number of different scenarios. Such tools allow to find the requirements in terms of Bob's and Eve's channels SNR to achieve a fixed level of mutual information security with practical codes, as well as to compare their performance with that achievable with optimal codes.

Then we have proposed an OOT protocol based on coding and AONTs to achieve some desired level of SS over fading wiretap channels by using classical and practical transmission techniques. We have introduced the use of fake packets in the proposed protocol and assessed the resulting benefits in terms of performance.

We have provided some examples considering coding and modulation schemes compliant with the IEEE 802.16e (WiMax) standard, which also demonstrated the feasibility of the proposed approach. Our results show that such practical and widespread coding and modulation schemes can indeed achieve a secrecy performance close to that corresponding to optimal codes when high code rates and low order modulations are used. Instead, if one aims at reducing the SNR gap required between Bob's and Eve's channels in order to achieve some given level of mutual information security, our results show that it is beneficial to use low code rates and low order modulations. Moreover, it has been shown that applying our protocol with practical coding and modulation schemes it is possible to achieve satisfactory SS levels, even when the average quality of the eavesdropper channel is not worse than that of the main channel.

Chapter 4

Optimal resource allocation for secure Gaussian parallel relay channels with practical conditions

In this chapter we provide some results on resource allocation for confidential communications over the Gaussian parallel relay channels, expanding the work done in [5]. That work was limited to an ideal scenario, while we address a more realistic model by including the more practical constraints of finite-length coding and discrete constellations.

We focus on securing the transmission over a wireless channel between two users that are communicating through trusted relays, i.e., other users that are allowed to receive and transmit the secret message. We assume that relay nodes decode and forward (DF) the messages they receive, and all users and relays communicate over a set of parallel channel, e.g. obtained by orthogonal frequency division multiplexing (OFDM) modulation. While it is reasonable to assume that the channel state information (CSI) among Alice, Bob and the relays is known to all nodes, the CSI of channels to Eve may only be partially known. In some cases, it can be advantageous for Eve to perform some transmission, and we can assume to have full CSI (including channels to Eve). Otherwise, Eve is a passive attacker and legitimate nodes can only exploit some side information, e.g., on the minimum distance of Eve from transmitting nodes to get a partial CSI.

We first derive the achievable secrecy rate of this scheme under the assumption of full CSI by Alice and the relay nodes. Then in order to consider the impact of discrete constellations and finite-length coding we define an achievable secrecy rate under a constraint on the equivocation rate at Eve. Using an approximated formula of the achievable secrecy rate we derive the optimal power allocation for point-to-point confidential transmission. By exploiting the power and rate adaptation algorithm for the parallel relay channels of [5], we obtain a resource allocation algorithm coupling two Gale and Shapley algorithms to allocate resources over the parallel relay channels. We also consider the partial CSI scenario, in which Alice does not know the gains of her channels to Eve, while their statistics are known. In this case we only guarantee

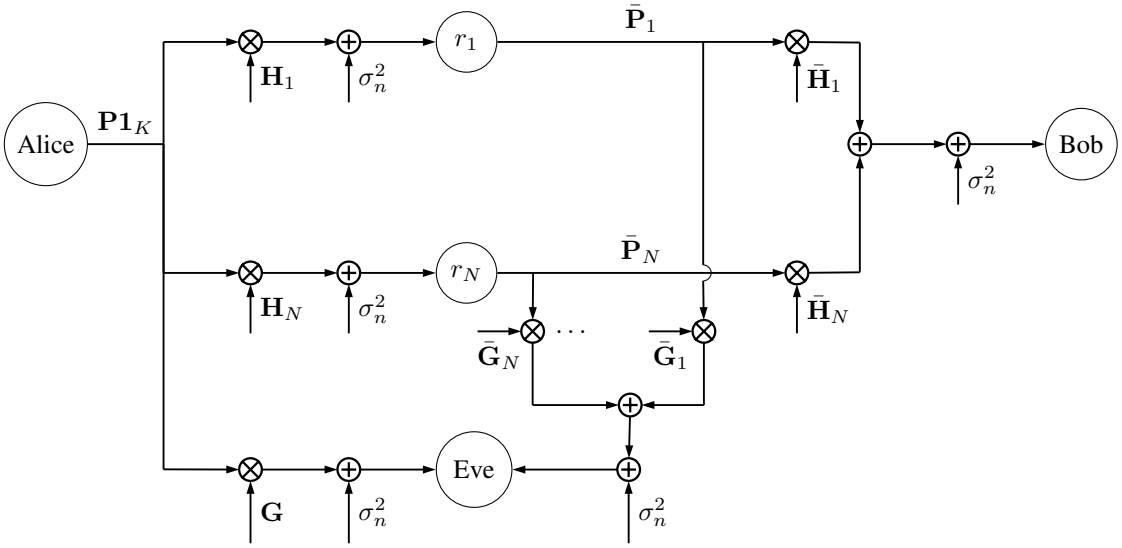


Figure 4.1: Power flow of the relay parallel channels with N relays, $r_1, \dots r_N$. Mixers $\otimes$ and adders $\uplus$ represent element-wise multiplication and addition of vectors, respectively.

an *outage* secrecy rate, i.e., the rate of a message that remains secret with a given probability. We show that the algorithm derived for full CSI can be easily adapted to the partial CSI scenario. Numerical results are provided, showing the merit of the proposed solution.

Notation: Vectors and matrices are written in bold letters. We denote the base-2 and natural-basis logarithm by $\log$ and $\ln$, respectively. We indicate the positive part of a real quantity x as $[x]^+ = \max\{x; 0\}$. $\mathbb{E}[X]$ denotes the expectation of the random variable X, $\mathbb{P}[\cdot]$ is the probability operator, and ^{T} denotes the matrix transpose operator. The entropy is denoted as $\mathbb{H}(\cdot)$, while the mutual information is denoted as $\mathbb{I}(\cdot, \cdot)$.

4.1 Related works

The physical layer security of messages transmitted over parallel channels with the assistance of trusted relays has already been addressed in the literature. Most works consider that relays can either forward the message or generate a noise signal to jam Eve. For links comprising a single channel, early works have addressed the relay selection problem [43–46], while various combinations of message forwarding and jamming are considered in [47–50] with multiple antennas nodes. In [51] multiple relays either jam or forward noise, i.e., they transmit random codewords from a globally

known codebook, that hurts more Eve than Bob.

We focus on links comprising parallel channels. For this scenario, in [52] rate-equivocation regions are derived by considering one relay only and assuming full CSI. In [53] OFDM is considered with a single relay, and Eve is equipped with multiple antennas under partial CSI: subcarriers, powers and rates are optimized to maximize the average secrecy outage capacity. In [54] the downlink of a cellular system is considered - where the multi-antenna base station performs both beamforming and jamming against a single multi-antenna eavesdropper, and an outage problem is formulated under partial CSI. The scenario is extended in [55], where multiple relays operate in DF mode and still an outage approach is considered. In [56] a single relay with parallel channels is considered, which performs cooperative jamming against Eve, under full CSI. When the single relay performs DF, resource optimization has been considered in [57]. More comprehensive results, considering also the direct transmission from Alice to Bob are obtained in [58]. Resource allocation for transmission over parallel channels assisted by DF relays without secrecy features has also been widely studied. Bit loading [59] and power and rate allocation [60] have been investigated, while the availability of multiple relays transmitting on a single sub-carrier is studied in [61], with efficient greedy algorithms provided in [62]. The resource allocation for parallel channels with secrecy outage constraint has been considered in [63] and [64], without taking into account the conditions imposed by the presence of relay nodes in the system.

Recently, optimal resource allocation for security purposes under different conditions has gained the attention of several authors. In [65], an optimization framework is proposed for two-hop communications that jointly optimizes source and relay powers, and transmission time in each hop, with the goal of maximizing the secrecy outage capacity in a massive multiple input multiple output (MIMO) scenario. In [66], optimal power allocation and pricing strategies are determined using a Stackelberg game model in order to maximize the players' utilities, under both of perfect and imperfect CSI assumptions and in the presence of multiple eavesdroppers. An optimal power strategy to maximize the achievable secrecy rates in wireless multi-hop DF relay networks with a power constraint is studied in [67], under the assumption of global CSI, and an iterative cooperative beamformer design is also proposed. The work is extended in [68] to the case of full-duplex relays, with cooperative beamforming to null out the signal at multiple eavesdroppers. In [69] a heuristic resource allocation iterative algorithm is presented, based on the proximal theory that maximizes the secure capacity of device-to-device communications in heterogeneous networks. Joint source-relay power optimization in a dual-hop communication using duality theory is performed in [70], with the aim of maximizing the overall secrecy rate, under individual power constraints and using an high SNR approximation. In [71], a robust resource allocation framework is proposed in the presence of an active eavesdropper, assuming that both the legitimate receiver and the eavesdropper are full-duplex:

the receiver sends jamming signals against the eavesdroppers, without the need for external helpers, and uncertain CSI on the links between the eavesdropper and the legitimate receivers is considered. Other works considers optimal power allocation for security purposes with the help of imperfect hardware analysis [72] and an external jammer [73]. Optimization algorithms for null space beamforming with full CSI have been proposed in [74], while in [75] authors propose a joint relay selection and optimal power allocation to maximize security in a cooperative network, considering the presence of untrusted relays and passive eavesdroppers, possibly colluding. These works however do not take into consideration the impact of practical limitations in the system.

The impact of finite constellation inputs on the achievable secrecy rate for Gaussian broadcast channel with confidential message (BCCM) is considered in [76]. However, the role of finite-length coding is not investigated, and the scenario is not extended to parallel relay channels nor to optimal power allocation. Also in [77] the authors take into account practical conditions, such as finite alphabets, in order to evaluate the power allocation that maximize the ergodic secrecy rate with a low computational complexity, but secrecy is helped with the use of artificial noise in order to degrade the eavesdropper's performances, and the power allocation scheme is based on a gradient search algorithm.

4.2 System Model

We consider a communication system to transmit a confidential message M from Alice to Bob through N trusted cooperating relays. Any link between a pair of devices is constituted by a set of K parallel AWGN channels. Eve is an eavesdropping device that overhears communications originated from both Alice and the relays. No direct link between Alice and Bob is available, and all devices operate in half-duplex mode. Therefore the message transmission comprises two phases:

1) Alice transmits to the relays, and

2) the relays transmit to Bob.

We also assume that in phase 2 at most one relay transmits on channel k and that the two phases have the same duration.

Fig. 4.1 shows the power flow of the considered scenario. We indicate with $P_{n,k}$ the transmit power of Alice on channel k to relay n in phase 1, while $\bar{P}_{n,k}$ is the transmit power of relay n on channel k in phase 2. The $N \times K$ matrix $\mathbf{P}$ ($\bar{\mathbf{P}}$) collects all transmit powers, having $P_{n,k}$ ($\bar{P}_{n,k}$) at entry n, k. In Fig. 4.1, $\mathbf{P}\mathbf{1}_K$ denotes the N-size column vector of transmit powers for each relay, with $\mathbf{1}_K$ being the K-size column vector of all ones. We consider power constraints for both Alice and the relays, i.e.,

$$\sum_{k=1}^{K} P_{n,k} \leq P_{tot,1}, \quad n = 1, \ldots, N\,.$$
$$\sum_{k=1}^{K} \bar{P}_{n,k} \leq P_{tot,2}\,, \quad n = 1, \ldots, N\,. \tag{4.1}$$

The power constraint per relay in phase (1) simplifies the power allocation in this phase and still provides an upper bound on the total transmit power from the source, that can not exceed $NP_{tot,1}$.

The link from Alice to relay n is represented by the K-size column vector $\mathbf{H}_n = [H_{n,1}, \ldots, H_{n,K}]^T$ containing the gains for each channel. The power of the data signal received by relay n on channel k is therefore $H_{n,k}P_{n,k}$. Similarly, the vector $\bar{\mathbf{H}}_n$ denotes the power gains of the link between relay n and Bob and $\bar{H}_{n,k}\bar{P}_{n,k}$ is the power of the data signal received by Bob from relay n on channel k. For links to Eve, $\mathbf{G}$ is the vector of power gains of the signal coming from Alice, while $\bar{\mathbf{G}}_n$ is the power gain vector of the signal coming from relay n.

The noise is assumed to be i.i.d., with zero mean and unitary ($\sigma_n^2 = 1$ in Fig. 4.1) variance for all channels. Therefore, the SNR at relay n for a transmission from Alice on channel k is $H_{n,k}P_{n,k}$, and similarly for a transmission from relay n on channel k the SNR at Bob in phase 2 is $\bar{H}_{n,k}\bar{P}_{n,k}$.

4.3 Achievable Secrecy Rate

We consider a per-channel encoding, i.e., Alice splits M into K messages M_k, $k = 1, \ldots, K$, each of which is separately encoded and transmitted on a channel. In [64] an in-depth analysis of this coding strategy is provided, showing that it performs similarly to the scheme with joint coding across channels, while being simpler to design. Therefore, each relay in general receives only a subset of the entire message bits. In the second phase again each relay splits the received secret bits into groups, which are separately encoded and transmitted on a different channel, among those assigned to the relay.

In both phases, secrecy is achieved through classical wiretap coding [2], based on adding random bits to the secret message and encoding the resulting block with capacity achieving codes. The *weak* secrecy rate of a point-to-point transmission is the rate of a message M that [2]: $i)$ is correctly decoded by Bob and $ii)$ has a rate of mutual information with the signal received by Eve Z that is vanishing for infinite codewords, i.e.,

$$\lim_{l \to \infty} \frac{1}{l} \mathbb{I}(Z, M) = 0\,, \tag{4.2}$$

where l is the message length in bits. Due to the per-channel encoding, the achievable

weak secrecy rate is the sum of the achievable secrecy rates on each used channel. Let $R_{n,k}$ be the secrecy rate on channel k, intended for relay n in phase 1, and $\bar{R}_{n,k}$ the secrecy rate on channel k transmitted by relay n in phase 2. The achievable secrecy rate between Alice and Bob is the minimum between the secrecy rates in both phases, i.e.,

$$R_{tot}(\mathbf{P}, \bar{\mathbf{P}}) = \frac{1}{2} \sum_{n=1}^{N} \min \left\{ \sum_{k=1}^{K} R_{n,k}(P_{n,k}), \sum_{k=1}^{K} \bar{R}_{n,k}(P_{n,k}) \right\}, \tag{4.3}$$

where the factor $1/2$ is due to the two phases of the same duration, and we have highlighted the dependence of the achievable rates on the transmit powers, the minimum reflects the fact that either of the two phases can be a bottleneck for transmission, and the sum over the subcarriers k takes into account the fact that we decode and re-encode the data signal at the relays, thus each relay demodulates all received signals and spit power and data among the subcarriers in its own way upon transmission in phase 2.

Since we assume that Alice is transmitting to a single relay per channel we also have

$$R_{n^*,k}(P_{n^*,k}) > 0 \rightarrow R_{n,k}(P_{n,k}) = 0\,,\ n \neq n^*\,, \tag{4.4}$$

and since we assume that at most one relay is transmitting in any channel in phase 2 we also have

$$\bar{R}_{n^*,k}(\bar{P}_{n^*,k}) > 0 \rightarrow \bar{R}_{n,k}(\bar{P}_{n,k}) = 0\,,\ n \neq n^*\,. \tag{4.5}$$

In the following we derive the achievable secrecy rates, when full CSI is available at Alice, taking into consideration infinite and finite-length coding, continuous and discrete modulation formats. Then with discuss the ϵ-outage achievable secrecy rates when Alice has only a partial CSI, i.e., she knows only the statistics of her channels to Eve.

4.3.1 Infinite-length coding with Gaussian Constellations

When infinite-length coding and Gaussian constellations are used, perfect secrecy, i.e., no information leakage to Eve, can be achieved [2]. In this case, the achievable secrecy rate can be written as

$$R_{n,k}(P_{n,k}) = C(P_{n,k}H_{n,k}) - C(P_{n,k}G_{n,k})\,, \tag{4.6}$$

where $C(x) = \log(1+x)$. Similar expressions are obtained for $\bar{R}_{n,k}(\bar{P}_{n,k})$ where $P_{n,k}$, $H_{n,k}$, and $G_{n,k}$ are replaced by $\bar{P}_{n,k}$, $\bar{H}_{n,k}$ and $\bar{G}_{n,k}$, respectively.

4.3.2 Finite-length Coding with Gaussian Constellations

A first limitation to the achievable secrecy rates introduced by practical systems is related to the use of codes working on finite-length blocks of symbols. In such a

setting, weak secrecy cannot be guaranteed and Eve can get some information on the secret message[1]. Moreover, the decodability condition at Bob cannot be guaranteed, and we must consider a non-zero codeword error rate (CER) κ.

Let $R_{n,k}$ and $\bar{R}_{n,k}$ be the message rates, for which we have a level of secrecy θ. In particular, in order to measure the information leakage to Eve we resort to the equivocation rate, i.e., Eve's uncertainty about the message after observing the transmitted codeword (through her channel). For relay n transmitting on channel k and using codewords of m symbols, the equivocation rate per symbol is

$$\rho_{n,k}(P_{n,k}) = \frac{1}{2m}\mathbb{H}(M_k|Z_{n,k})\,,$$

$$\bar{\rho}_{n,k}(\bar{P}_{n,k}) = \frac{1}{2m}\mathbb{H}(M_k|\bar{Z}_{n,k})\,,$$

where the factor 2 comes from the fact that we have two phases of the same duration. We have that

$$0 \leq \rho_{n,k}(P_{n,k}) \leq R_{n,k}(P_{n,k})\,,$$

$$0 \leq \bar{\rho}_{n,k}(\bar{P}_{n,k}) \leq \bar{R}_{n,k}(\bar{P}_{n,k}),$$

where the upper bound is achieved with infinitely-long codewords ($m \to \infty$). We consider that transmission is secure if

$$\frac{\rho_{n,k}(P_{n,k})}{R_{n,k}(P_{n,k})} \geq \theta\,, \quad \frac{\bar{\rho}_{n,k}(\bar{P}_{n,k})}{\bar{R}_{n,k}(\bar{P}_{n,k})} \geq \theta\,, \tag{4.7}$$

where $\theta \in (0, 1]$ is a suitably defined parameter that limits the gap with respect to weak secrecy conditions with infinite-length coding. Let us indicate with $\hat{M}_n$ the decoded version of message M_n. The *achievable secrecy rates* for finite-length coding are therefore the maximum rates satisfying condition (4.7), i.e.,

$$R_{n,k}(P_{n,k}) = \max_r r \tag{4.8}$$

s.t.

$$\frac{\rho_{n,k}(P_{n,k}, r)}{r} \geq \theta\,,$$

$$\mathbb{P}[M_n \neq \hat{M}_n] = \kappa\,.$$

A similar problem can be written for phase 2, for a given allocated power $\bar{P}_{n,k}$, i.e.,

$$\bar{R}_{n,k}(\bar{P}_{n,k}) = \max_r r \tag{4.9}$$

[1] Indeed, the definition of weak secrecy (4.2) entails a limit to infinity of the message length that can not be used in finite-codewords schemes.

s.t.

$$\frac{\bar{\rho}_{n,k}(\bar{P}_{n,k}, r)}{r} \geq \theta\,,$$
$$\mathbb{P}[M_n \neq \hat{M}_n] = \kappa\,.$$

For the computation of the equivocation rate we can resort to a lower bound. By the definition of entropy and mutual information, we have that Eve's equivocation rate can be rewritten as

$$\begin{aligned}\rho_{n,k}(P_{n,k}) =& \frac{1}{m}\left[\mathbb{H}(X_{n,k}) - \mathbb{I}(X_{n,k}; Z_{n,k}) + \right.\\ &\left. + \mathbb{H}(M_n|Z_{n,k}, X_{n,k}) - \mathbb{H}(X_{n,k}|M_n, Z_{n,k})\right]\,,\end{aligned} \tag{4.10}$$

where $X_{n,k}$ and $Z_{n,k}$ are the signals received by Bob and Eve in phase 1, respectively.

By the definition of spectral efficiency as upper bound to the mutual information, we have that

$$\mathbb{I}(X_{n,k}; Z_{n,k}) < mC(P_{n,k}G_{n,k}), \tag{4.11}$$

and

$$\mathbb{H}(M_n|Z_{n,k}, X_{n,k}) \leq \mathbb{H}(M_n|X_{n,k}) = 0.$$

On the other hand, the entropy of $X_{n,k}$ is the code rate, which in turns determines the (non null) CER at relay n, due to the use of finite-length coding. A bound on the code rate as a function of the CER for finite length coding is provided by [78], that in this scenario can be written as

$$\mathbb{H}(X_{n,k}) = m\gamma(P_{n,k}H_{n,k}) = m\left[C(P_{n,k}H_{n,k}) - \frac{\log e}{\sqrt{2m}}Q^{-1}(\kappa)\right]^+\,, \tag{4.12}$$

where κ is the target CER at relay n and $[x]^+ = x$ for $x \geq 0$ and 0 otherwise, and $Q(\cdot)$ is the complementary cumulative distribution function of the standard Gaussian variable.

Let $\eta(R_{n,k}, P_{n,k}G_{n,k})$ be the CER experienced by a fictitious receiver at the wire-tapper position trying to decode for $X_{n,k}$ from observing $Z_{n,k}$ and M. By the Fano inequality we have

$$\mathbb{H}(X_{n,k}|M_n, Z_{n,k}) \leq 1 + m\left(\gamma(P_{n,k}H_{n,k}) - R_{n,k}\right)\eta(R_{n,k}, P_{n,k}G_{n,k})\,. \tag{4.13}$$

Hence from (4.10) and (4.13) we have the following lower bound on $\rho_{n,k}(P_{n,k})$:

$$\begin{aligned}\rho_{n,k}(P_{n,k}) \geq \sigma_{n,k}(P_{n,k}) =& \\ &\gamma(P_{n,k}H_{n,k}) - C(P_{n,k}G_{n,k}) \\ &- (\gamma(P_{n,k}H_{n,k}) - R_{n,k})\eta(R_{n,k}, P_{n,k}G_{n,k}) - \frac{1}{m}\,.\end{aligned}$$

The above analysis can be simplified by considering that, for adherence to practical systems, deterministic coding instead of random coding can be used. In such a case, each l-bit block of data is univocally mapped into a codeword $\mathcal{C}_{n,k}$. This is opposed to either random or coset coding, which are often invoked in the literature for this kind of systems, but yield further issues (e.g., concerning the generation of randomness). In the case of deterministic coding, we no longer need to estimate the CER for the fictitious receiver and we can write a simpler lower bound on the equivocation rate, that is

$$\rho_{n,k}(P_{n,k}) = \frac{1}{m}\left[\mathbb{H}(\mathcal{C}_{n,k}) - \mathbb{I}(\mathcal{C}_{n,k}; Z_{n,k})\right] .$$

Resorting again to (4.11) and (4.12), we obtain the following approximation on Eve's equivocation rate

$$\rho_{n,k}(P_{n,k}) \simeq \xi_{n,k}(P_{n,k}) \triangleq \left[C(H_{n,k}P_{n,k}) - \frac{\log e}{\sqrt{2m}}Q^{-1}(\kappa) - C(G_{n,k}P_{n,k})\right]^{+}, \tag{4.14}$$

which does not depend on $R_{n,k}$.

By replacing $\rho_{n,k}(P_{n,k})$ with its approximated lower bound $\xi_{n,k}(P_{n,k})$ (and similarly for phase-2 equivocation rates) in problems (4.8) and (4.9) and removing the (already used) constraint on the error probability $\mathbb{P}[M_n \neq \hat{M}_n]$, we obtain the approximated achievable rates in the two phases; the solution can be easily obtained in a closed form as

$$R_{n,k}(P_{n,k}) = \frac{1}{\theta}\left[C(H_{n,k}P_{n,k}) - \frac{\log e}{\sqrt{2m}}Q^{-1}(\kappa) - C(G_{n,k}P_{n,k})\right]^{+}$$

Note that the obtained secrecy rate with finite-length coding is smaller than that obtained with infinite-length coding. In particular, $\frac{\log e}{\sqrt{2m}}Q^{-1}(\kappa)$ represents the secrecy rate loss due to finite-length coding, which decreases as either the code length m or the CER κ increase. Note that the choice of m is mostly dictated by implementation constraints as well as desired latency limitations, while κ is associated to the reliability of the transmission. In Fig. 4.2 we compare the results obtained for $R_{n,k}(P_{n,k})$ as a function of $G_{n,k}P_{n,k}$ for codes of different length, choosing $H_{n,k}/G_{n,k}$ of 20 dB and $\theta = 1$.

We consider a fitting of $R_{n,k}(P_{n,k})$ solution of (4.8) by the linear combination of logarithms of the powers, in order to ease resource allocation, i.e.,

$$\begin{aligned} R_{n,k}(P_{n,k}) \simeq\ & \alpha_1 + \alpha_2 \log(1 + \alpha_3 H_{n,k}P_{n,k}) - \\ & \alpha_4 \log(1 + \alpha_5 H_{n,k}P_{n,k}) - [\alpha_6 \log(1 + \alpha_7 G_{n,k}P_{n,k}) - \\ & \alpha_8 \log(1 + \alpha_9 G_{n,k}P_{n,k})] . \end{aligned} \tag{4.15}$$

Note that (4.15) directly models the achievable secrecy rate rather than the equivo-

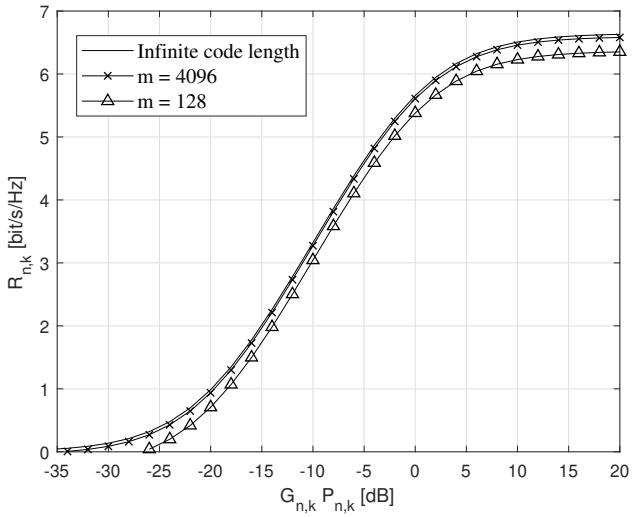


Figure 4.2: $R_{n,k}(P_{n,k})$ as a function of $G_{n,k}P_{n,k}$ for $H_{n,k}/G_{n,k}$ equal to 20 dB, for infinite length codes and codes of length 4096 and 128, with $\theta = 1$.

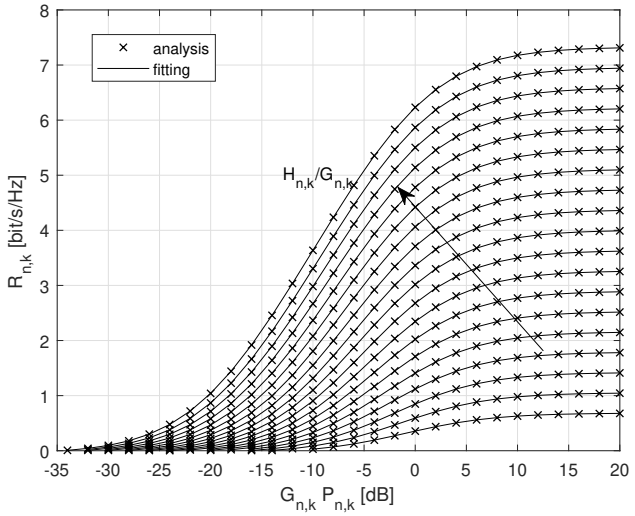


Figure 4.3: $R_{n,k}(P_{n,k})$ as a function of $G_{n,k}P_{n,k}$ for values of $H_{n,k}/G_{n,k}$ between 2 dB and 20 dB with a step of 1 dB, and results obtained with the fitting function (4.15), for codes of length 4096 and $\theta = 0.9$.

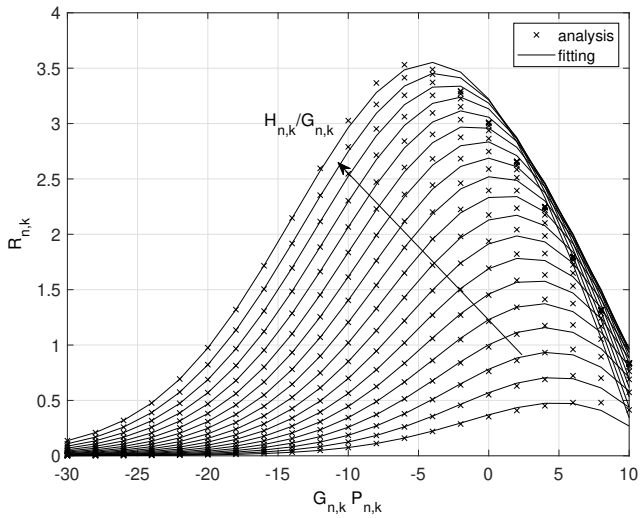


Figure 4.4: $R_{n,k}(P_{n,k})$ as a function of $G_{n,k}P_{n,k}$ for values of $H_{n,k}/G_{n,k}$ between 2 dB and 20 dB with a step of 1 dB, and results obtained with the fitting function (4.15), considering a 16-QAM constellation.

cation rate, and the parameters α_i are chosen at solution of problem (4.8). By this formulation the secrecy rates with ideal conditions can be seen as a sub-case of (4.15) with $\alpha_i = 1$ for $i = 2, 3, 6, 7$ and $\alpha_i = 0$ otherwise. The motivation behind the choice of this fitting will be better understood when using it in the resource optimization problem focus of this paper: indeed it will turn out (see Section 4.4) that with this choice the optimization problem boils down to finding the roots of suitable polynomials.

Fig. 4.3 shows $R_{n,k}(P_{n,k})$ as a function of $G_{n,k}P_{n,k}$ for values of $H_{n,k}/G_{n,k}$ between 2 dB and 20 dB with a step of 1 dB, and results obtained by the fitting function (4.15) with $\kappa = 10^{-3}$ and $m = 4,096$ and $\theta = 0.9$. We observe a good agreement of the fitting function with $R_{n,k}(P_{n,k})$, especially at low rates, and high values of $G_{n,k}P_{n,k}$, with a slight overestimation of the rate for intermediate values of $G_{n,k}P_{n,k}$ for high $H_{n,k}/G_{n,k}$ ratios.

4.3.3 Infinite-length Coding with Discrete Constellations

A second limitation of practical systems is the use of suboptimal constellations with discrete points taken from a finite alphabet. In this case, perfect secrecy can still be achieved, but we must consider the constellation-constrained spectral efficiency $\hat{C}(\cdot)$ instead of $C(\cdot)$, i.e., (4.6) becomes

$$R_{n,k}(P_{n,k}) = \hat{C}(P_{n,k}H_{n,k}) - \hat{C}(P_{n,k}G_{n,k})\,. \tag{4.16}$$

In order to obtain simple resource allocation algorithms we consider again (4.15) as a fitting of $R_{n,k}(P_{n,k})$. Fig. 4.4 shows the secrecy rate as a function of the SNR for a 16-QAM constellation, and its comparison with the exact function. We observe a good agreement between the approximated and the exact curves, with slight higher discrepancy for high values of $G_{n,k}P_{n,k}$. However note that in a power optimization process these high power values will not be used, since they provide a lower secrecy rate than lower power values. We still have a slight mismatch between the fitting and the analysis in correspondence of the maximum rate, which is however not so relevant (especially for increasing values of $H_{n,k}/G_{n,k}$).

4.3.4 Finite-length Coding with Discrete Constellations

Let us consider the limitations introduced in Sections 4.3.2 and 4.3.3 jointly, i.e., both finite-length coding and discrete constellations, which describe a practical scenario. Also in this case we resort to the equivocation rate for the definition of the achievable secrecy rate (see problems (4.8) and (4.9)), by replacing the spectral efficiency $C(P)$ with the constellation-constrained spectral efficiency $\hat{C}(P)$ in (4.14). On the other hand, since the approximation provided by [78] is valid for any input distribution, (4.12) still holds true.

As already done in the previous section, we propose to fit $R_{n,k}(P_{n,k})$ by the function (4.15). Fig. 4.5 shows $R_{n,k}(P_{n,k})$ for values of $H_{n,k}/G_{n,k}$ between 2 dB and 20 dB with a step of 1 dB, and results obtained by the fitting function (4.15) with $\kappa = 10^{-3}$, 16-QAM constellation and $m = 4,096$. In this case, we observe a good agreement between the approximated and the exact curves for low values of $G_{n,k}P_{n,k}$, while the curves shows a little difference for high values. As observed for the case of infinite-length coding, also in this case the high power values will not be used in the optimization.

4.3.5 ϵ-Outage Achievable Secrecy Rate

In many practical scenarios Alice and the relays have only a partial CSI of their channels to Eve. This is mainly due to the fact that Eve may not have an advantage in revealing its channels, e.g., by transmitting, unless this could be useful to increase the rate of other messages exchanged between her and the legitimate nodes. Indeed, in the absence of full CSI there is a non-zero probability (outage probability) that for any power allocation and choice of the secret message rate Eve may get some information on M.

In particular, we focus on the secrecy outage probability in each transmission phase and for each channel. Let $\pi_{n,k}$ and $\bar{\pi}_{n,k}$ be the secrecy outage probabilities on channel k with respect to relay n in the first and the second phase, when messages are transmitted at rates $R_{n,k}(P_{n,k})$ and $\bar{R}_{n,k}(\bar{P}_{n,k})$, respectively. We consider as design

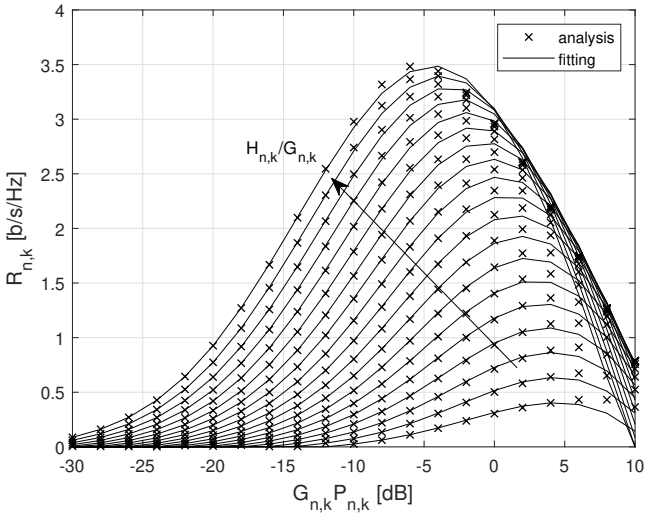


Figure 4.5: $R_{n,k}(P_{n,k})$ as a function of $G_{n,k}P_{n,k}$ for values of $H_{n,k}/G_{n,k}$ between 2 dB and 20 dB with a step of 1 dB, and results obtained with the fitting function (4.15), considering a 16-QAM constellation and $m = 4,096$.

criterion the limitation of the secrecy outage probability on each channel, i.e.,

$$\pi_{n,k} \leq \epsilon\,, \quad \bar{\pi}_{n,k} \leq \epsilon\,, \tag{4.17}$$

where ϵ is the target secrecy outage probability.

In the following we assume that the legitimate nodes know the statistics of both $G_{n,k}$ and $\bar{G}_{n,k}$, thus having a partial CSI. If $R_{n,k}(P_{n,k})$ is the achievable secrecy rate for Alice-Eve channel realization $G^*_{n,k}$, then the secrecy outage probability can be written as

$$\pi_{n,k} = \mathbb{P}[G_{n,k} > G^*_{n,k}]\,.$$

Similar expressions are obtained for the second phase. From (4.17) we define F_ϵ as the *outage gain*, i.e., the channel gain for which

$$\pi_{n,k} = \mathbb{P}[G_{n,k} > F_\epsilon] = \epsilon\,. \tag{4.18}$$

Then the ϵ-outage achievable secrecy rate can be obtained from the previous sections by considering $G_{n,k} = \bar{G}_{n,k} = F_\epsilon$.

Note that this approach works with any kind of fading (e.g., Rayleigh Rician, Nakagami fading). For example, for Rayleigh fading $\mathbb{P}[G_{n,k} \leq F_\epsilon] = \exp[-F_\epsilon/(d_{\mathrm{E}}^{-\zeta})]$, where ζ is the path-loss exponent, therefore

$$F_\epsilon = -d_{\mathrm{E}}^{-\zeta} \ln \epsilon\,.$$

For Nakagami fading $G_{n,k}$ is gamma distributed, i.e., $\mathbb{P}[G_{n,k} \leq F_\epsilon] = \frac{1}{\Gamma(\sigma)}\gamma(\sigma, \kappa F_\epsilon)$, where σ is the shape parameter, κ is the rate parameter, $\Gamma(\cdot)$ and $\gamma(\cdot)$ are the Gamma and lower-incomplete Gamma functions, respectively. Therefore we have

$$F_\epsilon = \frac{1}{\kappa}\gamma^{-1}(\sigma, \Gamma(\sigma)\epsilon)\,,$$

and $\gamma^{-1}(\sigma, x)$ is the inverse lower-incomplete Gamma function.

4.4 Single Link Power Optimization

We first consider the single-link power optimization, where we allocate powers that maximize the secrecy sum rate between two nodes, using parallel channels. This problem must be solved in both transmission phases and here we focus on the first phase, i.e., the optimization of the communication from Alice to a specific relay n, assuming that all power $P_{\text{tot},1}$ can be used on that link. In this situation we have $P_{n',k} = 0$ for $n' \neq n$, $n = 1, \ldots, N$, $k = 1, \ldots, K$ and we must solve

$$R_{\max} = \max_{\{P_{n,k}\}} \sum_{k=1}^{K} R_{n,k}(P_{n,k})\,, \quad \text{s.t. (4.1).} \tag{4.19}$$

The four cases of previous section are considered, i.e., $a)$ infinite-length coding with Gaussian constellation, $b)$ finite-length coding with Gaussian constellation, $c)$ infinite-length coding with discrete constellations, and $d)$ finite-length coding with discrete constellations. Moreover, we consider here the case of ϵ-outage rates discussed in Section 4.3.5, thus considering gain F_ϵ for all channels to Eve.

4.4.1 Infinite-length Coding with Gaussian Constellations

For infinite-length coding with Gaussian constellation, the optimization problem (4.19) has been proven to be convex and solved in [79, Theorem 1]. In particular, we immediately see that all channels for which $H_{n,k} < F_\epsilon$ must be switched off ($P_{n,k} = 0$) since they do not provide any secrecy rate. Let the set of used channels be

$$\mathcal{F} = \{k : H_{n,k} > F_\epsilon\}\,.$$

Then we have

$$\begin{aligned} P_{n,k} = & \left[-\frac{\lambda(H_{n,k}+F_\epsilon)}{2\lambda F_\epsilon H_{n,k}} + \right. \\ & \left. \frac{\sqrt{[\lambda(H_{n,k}+F_\epsilon)]^2 - 4\lambda F_\epsilon H_{n,k}(\lambda - H_{n,k} + F_\epsilon)}}{2\lambda F_\epsilon H_{n,k}} \right]^+ \\ = & \left[-\frac{H_{n,k}+F_\epsilon}{2F_\epsilon H_k} + \right. \\ & \left. \left(. \frac{\sqrt{(H_{n,k}+F_\epsilon)^2 - 4FH_{n,k}(1-(H_k-F_\epsilon)/\lambda)}}{2F_\epsilon H_k} \right] \right]^+ \\ = & \left[-\frac{H_{n,k}+F_\epsilon}{2F_\epsilon H_k} + \right. \\ & \left. \frac{\sqrt{(H_{n,k}-F_\epsilon)^2 + 4FH_{n,k}(H_{n,k}-F_\epsilon)/\lambda)}}{2F_\epsilon H_{n,k}} \right]^+ , \end{aligned} \tag{4.20}$$

where $\lambda \geq 0$ is the Lagrange multiplier to be optimized in order to satisfy the power constraint, which can be computed by a dichotomic search to find the unique optimal solution.

4.4.2 Finite-length Coding with Gaussian Constellations

For finite-length coding with Gaussian constellations we exploit the fitting (4.15) and the optimization problem (4.19) becomes

$$\begin{aligned} R_{max} = \max_{\{P_{n,k}\}} & \sum_{k=1}^{K} \alpha_1 + \alpha_2 \log(1+\alpha_3 H_{n,k} P_{n,k}) - \\ & \alpha_4 \log(1+\alpha_5 H_{n,k} P_{n,k}) - \alpha_6 \log(1+\alpha_7 F_\epsilon P_{n,k}) + \\ & \alpha_8 \log(1+\alpha_9 F_\epsilon P_{n,k}) , \end{aligned} \tag{4.21}$$

subject to (4.1).

We observe that (4.21) is a maximization problem of continuously differentiable objective functions with inequality constraints of continuously differentiable functions, thus satisfying the necessary conditions for the application of the Lagrange multipliers method, which provides the constrained maxima as one or more solutions of

$$\begin{aligned} & \frac{\alpha_2\alpha_3 H_{n,k}}{1+\alpha_3 H_{n,k} P_{n,k}} + \frac{\alpha_4\alpha_5 H_{n,k}}{1+\alpha_5 H_{n,k} P_{n,k}} + \\ & + \frac{\alpha_6\alpha_7 F_\epsilon}{1+\alpha_7 F_\epsilon P_{n,k}} + \frac{\alpha_8\alpha_9 F_\epsilon}{1+\alpha_9 F_\epsilon P_{n,k}} - \lambda = 0 , \end{aligned} \tag{4.22}$$

where $\lambda \geq 0$ is the Lagrange multiplier to be chosen in order to satisfy the power

constraint.

By using the common denominator of the four fractions in (4.22), by simple algebraic steps we separate the terms of different order and we define

$$A_{n,k} = \lambda \ln(2)\alpha_3\alpha_5\alpha_7\alpha_9 H_{n,k}^2 F_\epsilon^2,$$

$$\begin{aligned} B_{n,k} = {} & \alpha_3\alpha_5\alpha_6\alpha_7\alpha_9 H_{n,k}^2 F_\epsilon^2 - \alpha_3\alpha_5\alpha_7\alpha_8\alpha_9 H_{n,k}^2 F_\epsilon^2 - \\ & \alpha_2\alpha_3\alpha_5\alpha_7\alpha_9 H_{n,k}^2 F_\epsilon^2 + \alpha_3\alpha_4\alpha_5\alpha_7\alpha_9 H_{n,k}^2 F_\epsilon^2 + \\ & \lambda \ln(2)\alpha_5\alpha_7\alpha_9 H_{n,k} F_\epsilon^2 + \lambda \ln(2)\alpha_3\alpha_7\alpha_9 H_{n,k} F_\epsilon^2 + \\ & \lambda \ln(2)\alpha_3\alpha_5\alpha_9 H_{n,k}^2 F_\epsilon + \lambda \ln(2)\alpha_3\alpha_5\alpha_7 H_{n,k}^2 F_\epsilon, \end{aligned}$$

$$\begin{aligned} C_{n,k} = {} & \lambda \ln(2)\alpha_3\alpha_5 H_{n,k}^2 - \alpha_2\alpha_3\alpha_5\alpha_9 H_{n,k}^2 F_\epsilon + \\ & \lambda \ln(2)\alpha_7\alpha_9 F_\epsilon^2 + \lambda \ln(2)\alpha_5\alpha_7 H_{n,k} F_\epsilon - \\ & \alpha_2\alpha_3\alpha_5\alpha_7 H_{n,k}^2 F_\epsilon + \alpha_4\alpha_5\alpha_7\alpha_9 H_{n,k} F_\epsilon^2 + \\ & \alpha_3\alpha_4\alpha_5\alpha_9 H_{n,k}^2 F_\epsilon + \lambda \ln(2)\alpha_3\alpha_9 H_{n,k} F_\epsilon + \\ & \lambda \ln(2)\alpha_3\alpha_7 H_{n,k} F_\epsilon + \alpha_3\alpha_6\alpha_7\alpha_9 H_{n,k} F_\epsilon^2 + \\ & \alpha_3\alpha_5\alpha_6\alpha_7 H_{n,k}^2 F_\epsilon - \alpha_5\alpha_7\alpha_8\alpha_9 H_{n,k} F_\epsilon^2 - \\ & \alpha_3\alpha_7\alpha_8\alpha_9 H_{n,k} F_\epsilon^2 - \alpha_3\alpha_5\alpha_8\alpha_9 H_{n,k}^2 F_\epsilon + \\ & \lambda \ln(2)\alpha_5\alpha_9 H_{n,k} F_\epsilon - \alpha_2\alpha_3\alpha_7\alpha_9 H_{n,k} F_\epsilon^2 + \\ & \alpha_3\alpha_4\alpha_5\alpha_7 H_{n,k}^2 F_\epsilon + \alpha_5\alpha_6\alpha_7\alpha_9 H_{n,k} F_\epsilon^2, \end{aligned}$$

$$\begin{aligned} D_{n,k} = {} & \alpha_4\alpha_5\alpha_7 H_{n,k} F_\epsilon - \alpha_2\alpha_3\alpha_9 H_{n,k} F_\epsilon - \\ & \alpha_5\alpha_8\alpha_9 H_{n,k} F_\epsilon - \alpha_3\alpha_8\alpha_9 H_{n,k} F_\epsilon + \\ & \alpha_5\alpha_6\alpha_7 H_{n,k} F_\epsilon - \alpha_2\alpha_3\alpha_7 H_{n,k} F_\epsilon + \\ & \alpha_4\alpha_5\alpha_9 H_{n,k} F_\epsilon - \alpha_2\alpha_3\alpha_5 H_{n,k}^2 + \\ & \alpha_3\alpha_6\alpha_7 H_{n,k} F_\epsilon + \alpha_6\alpha_7\alpha_9 F_\epsilon^2 - \alpha_7\alpha_8\alpha_9 F_\epsilon^2 + \\ & \lambda \ln(2)\alpha_9 F_\epsilon + \lambda \ln(2)\alpha_7 F_\epsilon + \lambda \ln(2)\alpha_5 H_{n,k} + \\ & \alpha_3\alpha_4\alpha_5 H_{n,k}^2 + \lambda \ln(2)\alpha_3 H_{n,k}, \end{aligned}$$

$$\begin{aligned} E_{n,k} = {} & \lambda \ln(2) - \alpha_2\alpha_3 H_{n,k} - \alpha_8\alpha_9 F_\epsilon + \\ & \alpha_6\alpha_7 F_\epsilon + \alpha_4\alpha_5 H_{n,k} \end{aligned}$$

the Lagrangian (4.22) becomes

$$\begin{aligned} & A_{n,k} P_{n,k}^4 + B_{n,k} P_{n,k}^3 + C_{n,k} P_{n,k}^2 \\ & \qquad + D_{n,k} P_{n,k} + E_{n,k} = 0\,. \end{aligned} \tag{4.24}$$

For a given $\lambda \geq 0$, for all real positive roots of the polynomial, we compute (4.16) and select the root yielding the highest secrecy rate. When no real roots are found, it

means that the secrecy rate is strictly decreasing for $P_{n,k} > 0$, thus $P_{n,k} = 0$ and a null secrecy rate is achieved.

We can now appreciate the value of the fitting (4.15), which provides the simple polynomial (4.24) whose roots can be obtained using well-established algorithms. Note that the algorithm must include a dichotomic search over $\lambda \geq 0$ in order to satisfy the power constraints. Again, note that the solution to problem (4.21) is a generalization of the solution (4.20) for ideal transmission conditions.

4.4.3 Discrete Constellations

For infinite-length coding with discrete constellations, the optimization problem (4.19) using the fitting (4.15) becomes (4.21), hence by applying also in this case the Lagrange multiplier method we obtain again (4.24).

For finite-length coding with discrete constellations, the optimization problem (4.19) using the fitting (4.15) becomes (4.21) and the Lagrange multiplier methods leads to (4.24).

4.5 Maximum Rate Power Allocation

We now consider the power allocation problem at Alice and Bob with the aim of maximizing the secrecy rate, i.e.,

$$R_{max} = \max_{\mathbf{P},\bar{\mathbf{P}}} R_{tot}(\mathbf{P}, \bar{\mathbf{P}}) \,, \tag{4.25}$$

$$\text{subject to power constraints (4.1)}, \tag{4.26}$$

$$\text{and rate constraints (4.4) and (4.5)}.$$

Since the total rate is obtained by summing the minimum between the phase 1 and phase 2 rates [see (4.3)], the objective function (4.25) can be replaced by the maximization of the sum rate in phase 2, under the constraint that the sum incoming and outgoing rates in the two phases are the same, i.e.,

$$R^*_{tot}(\boldsymbol{P}^*, \bar{\boldsymbol{P}}^*) = \max_{\boldsymbol{P},\bar{\boldsymbol{P}}} \sum_{n=1}^{N} \sum_{k^1}^{K} \bar{R}_{n,k}(\bar{P}_{n,k}) \,,$$

$$\sum_{k=1}^{K} \bar{R}_{n,k}(P_{n,k}) \leq \sum_{k=1}^{K} R_{n,k}(P_{n,k}) \quad n = 1, \ldots, N \,. \tag{4.27}$$

As observed in [5], this is a mixed-integer programming problem, and for its solution we resort to the iterative approach of [5] based on the game-theoretic Gale and

Shapley algorithm for the stable matching problem [80]. Next we summarize the algorithm, while referring the reader to [5] for its detailed description.

The stable matching problem aims at matching dames to cavaliers, without having a dame and a cavalier belonging to two different couples both preferring to be matched: cavaliers first propose themselves to dames and dames, instead of directly choosing the cavalier, discard the worst proposal. Then cavaliers re-do their proposal (without being allowed to re-present a discarded proposal) and dames again discard the worst. The process is iterated until each dame has at most one proposal.

In our context, we have to match channels and relays, and we have two of such matches, for the two phases, i.e., for each $k = 1, \ldots, K$ we must find a relay index $n(k)$ for phase 1 and relay index $\bar{n}(k)$ for phase 2. The proposals are the rates offered by the relays. The dames are the channels at Bob and Alice. We perform the two matches in cascade: we choose the matching for phase 2, but proposals will be computed by solving the rate matching problem for phase 1.

The general solution of the problem in reported in Algorithm 1. The algorithm works iteratively, where at each iteration all relays provide Bob a proposal for each channel. Then Bob discards the request providing the lowest rate for all channels where at least two proposals have been received. At next iteration a relay that has seen his proposal discarded will update the rates on all other channels and send the new proposal to Bob. The process stops when Bob receives at most one proposal for each channel.

The output of the algorithm is the allocated power matrices $\boldsymbol{P}$ and $\bar{\boldsymbol{P}}$, having as (n, k) entry $P_{n,k}$ and $\bar{P}_{n,k}$, respectively. Matrices $\boldsymbol{R}$ and $\bar{\boldsymbol{R}}$ collect the rates in the two phases and are defined analogously. The set $\bar{\mathcal{Q}}_k$, $k = 1, \ldots,$ is iteratively updated and at each iteration collects the indices of relays that are not allowed to transmit on channel k in phase 2. Set

$$\bar{\mathcal{S}}_n = \{k : n \notin \bar{\mathcal{Q}}_k\},$$

instead collects all channels available for transmission to relay n in phase 2.

Initially, all these sets are empty as all relays are potentially free to transmit on any channel in phase 2. Note that the constraint of having at most one relay transmitting in each channel is not taken into account initially. However, at each iteration a competing relay for one channel is prevented from transmitting and the process stops exactly when there is at most one relay transmitting on each channel. At each iteration the routine `Rate_Offer_Phase_2` is run, which provides the power allocation that maximizes the total rate under the power constraints, the channel availability in phase 2 (i.e., sets $\bar{\mathcal{Q}}_k$). Then, we remove the relay that provides the minimum rate on one channel, i.e., we select the relay/channel index

$$(n', k') = \text{ar}gmin_{(n,k):|\bar{\mathcal{Q}}_k|<(N-1) \text{ and } n\notin\bar{\mathcal{Q}}_k} \bar{R}_{n,k} \tag{4.28}$$

and insert its index in $\bar{\mathcal{Q}}_k$.

Algorithm 1: General Resource Allocation Algorithm

output: $\boldsymbol{P}$, $\bar{\boldsymbol{P}}$

Set $\bar{\mathcal{Q}}_k = \emptyset$;

while $(\exists k : |\bar{\mathcal{Q}}_k| < N - 1)$ **do**

  ($\boldsymbol{P}$, $\bar{\boldsymbol{P}}$, $\bar{\boldsymbol{R}}$) = `Rate_Offer_Phase_2` $(\{\bar{\mathcal{Q}}_k\})$;

  Find relay channel indices (n', k') from (4.28);

  $\bar{\mathcal{Q}}_{k'} = \bar{\mathcal{Q}}_{k'} \cup \{n'\}$

end

return

4.5.1 The Rate_Offer_Phase_2 Algorithm

We now detail the Rate_Offer_Phase_2 algorithm that provides the rate offers for phase 2, given channel availability $\bar{\mathcal{S}}_n$ for each relay $n = 1, \ldots, N$. In formulas, this is problem (4.25)-(4.26) subject to the additional constraint

$$\sum_{k=1}^{K} R_{n,k}(P_{n,k}) = \sum_{k \in \bar{\mathcal{S}}_n} \bar{R}_{n,k}(\bar{P}_{n,k}) \,, \quad n = 1, \ldots, N \,.$$

Solution of the problem is achieved by computing the rates achieved in phase 2 by calling function MACalPowRate, which takes into account channel availability. Let $\bar{R}(\bar{P}_{n,k})$ be obtained solution. Then we compute the maximum rates achievable in phase 1 that the rate matching constraint (4.27) by function Rate_Offer_Phase_1. Lastly, we consider the rates obtained in phase 1 as a constraint to re-compute the optimal power allocation in phase 2, under the channel availability constraint. This is achieved by calling MACalPowRate for each relay, with the additional constraint that the rate in phase 2 cannot exceed that in phase 1, i.e., $\sum_{k=1}^{K} R_{n,k}$.

The resulting algorithm is reported in Algorithm 2, where $\bar{\boldsymbol{R}}_{n,\cdot}$ $(\bar{\boldsymbol{P}}_{n,\cdot})$ denotes the nth row of $\bar{\boldsymbol{R}}$ $(\bar{\boldsymbol{P}}_{n,\cdot})$. Matrix $\bar{\boldsymbol{H}}_{n,\bar{\mathcal{S}}_n}$ collects the columns of matrix $\bar{\boldsymbol{H}}$ with indices in $\bar{\mathcal{S}}_n$.

4.5.2 The MACalPowRate Algorithm

We now detail the MACalPowRate algorithm, which aims at maximizing the total secret rate over channel set $\mathcal{S}_n$, in point to point transmission, under a) a power constraint and b) a total rate constraint. Here we focus on the solution of the point-to-point problem for relay n in phase 1, but it can be applied also in phase n. In formulas, the problem can be written as

$$\max_{P_{n,k}, k \in \mathcal{S}_n} R_{n,k}(P_{n,k}) \quad \text{s.t. (4.1) and}$$

Algorithm 2: Rate_Offer_Phase_2

Input : $\{\bar{\mathcal{Q}}_k\}$
Output: $\boldsymbol{P}$, $\bar{\boldsymbol{P}}$, $\bar{\boldsymbol{R}}$
Data: ϵ, $\bar{\boldsymbol{H}}$

2.1 $\mathcal{S}_n = \{k : n \notin \bar{\mathcal{Q}}_k\}$
2.2 **for** $n = 1$ **to** N **do**
2.3 $\quad (\bar{\boldsymbol{P}}_{n,\cdot}, \bar{\boldsymbol{R}}_{n,\cdot})$ = `MACalPowRate` $(\bar{\boldsymbol{H}}_{\bar{\mathcal{S}}_n,n}, P_{tot}, \infty)$;
2.4 **end**
2.5 $(\boldsymbol{P}, \boldsymbol{R})$ = `Rate_Offer_Phase_1` $(\bar{\boldsymbol{R}})$;
2.6 **for** $n = 1$ **to** N **do**
2.7 $\quad (\bar{\boldsymbol{P}}_{n,\cdot}, \bar{\boldsymbol{R}}_{n,\cdot})$ = `MACalPowRate` $(\bar{\boldsymbol{H}}_{\bar{\mathcal{S}}_n,n}, P_{tot}, \sum_{k=1}^{K} R_{n,k})$;
2.8 **end**
2.9 **return**

Algorithm 3: MACalPowRate

Input : $\boldsymbol{h}$, P_{tot}, ρ_{max}
Output: $\boldsymbol{\pi}$, $\boldsymbol{\rho}$
Data: ϵ

3.1 $w_{\min} = 0, w_{\max} = 1, w = 1.$
3.2 $(\boldsymbol{\pi}, \boldsymbol{\rho})$ = `MACAllocation` $(\boldsymbol{h}, P_{tot})$;
3.3 **if** $\sum_k \rho_k > \rho_{\max}$ **then**
3.4 $\quad$ **while** $|\sum_k \rho_k - \rho_{\max}| > \epsilon$ **do**
3.5 $\quad\quad$ **if** $\sum_k \rho_k > \rho_{\max}$ **then**
3.6 $\quad\quad\quad w_{\max} = w$
3.7 $\quad\quad$ **else**
3.8 $\quad\quad\quad w_{\min} = w$
3.9 $\quad\quad$ **end**
3.10 $\quad\quad w = (w_{\max} + w_{\min})/2;$
3.11 $\quad\quad (\boldsymbol{\pi}, \boldsymbol{\rho})$ = `MACAllocation` $(\boldsymbol{h}, wP_{tot})$;
3.12 $\quad$ **end**
3.13 **end**
3.14 **return**

$$\sum_{k \in \mathcal{S}_n} R_{n,k}(P_{n,k}) \leq \rho_{\max}\,, \tag{4.29}$$

where $\rho_{\max}$ is the rate constraint. Note that when constraint (4.29) is active, not all power will be used. Therefore, the solution is obtained by first considering the point to point optimization problem without the rate constraint, i.e., the problem addressed in Section 4.4. If this solution does not satisfy the rate constraint, then the power constraint is modified (i.e., the available power is reduced) until the rate constraint is satisfied.

The resulting algorithm is reported in Algorithm 3, where $\boldsymbol{h}$ denotes the channel matrix (possibly being a sub-matrix of $\boldsymbol{H}$ or $\bar{\boldsymbol{H}}$). Note that the point to point solution of Section 4.4 is indicated by function MACAllocation, and that the algorithm performs a dichotomic search between zero allocated power and P_{tot} in order to find the intermediate power constraint that yields a rate satisfying the rate constraint.

4.5.3 The Rate_Offer_Phase_1 Algorithm

We now detail the Rate_Offer_Phase_1 algorithm, which aims providing the offers for rates in phase 1, given the maximum rates that can be supported in phase 2, $\bar{\boldsymbol{R}}$.

Solution is achieved by applying again the Gale Shapley approach, where now at each iteration relays offer rates for phase 1 and Alice discards the worst proposal. In this process we must take into account the power constraint on Alice. However, as for the resource allocation the constraint on the unique use of a channel was progressively enforced, this occurs also for the power constraint. In particular, the rate proposal to relay n in phase 1 is obtained by power allocation that maximizes the secrecy rate to relay n, assuming that all power is available to transmit to relay n. This is achieved by calling N times MACAllocation, one time for each Alice-relay link, obtaining power allocations $P^*_{n,k}$. Among all proposals to all relay we select the worst one, corresponding to the channel-relay couple $(n', k') = argmin_{(n,k)} R_{n,k}(P^*_{n,k})$, and we discard it, by preventing relay n' from allocating power on channel k'. The process is iterated by updating the power allocation for user n'. However, to take into account the power constraint of phase 1 that operates across the relays, the maximum power available to relay n' is reduced, since channel k' is used by another relay. We however do not know which relay at the end will use the channel, thus we reduce the total power available for user n' of the minimum among all power allocations, i.e., $P_{tot} - \min_{n \neq n'} P_{n,k^*}$. The process is iterated until for each channel Alice transmits at most one relay. At a generic iteration, let $\mathcal{Q}_k$ be the set of relays to which Alice transmits on channel k, and $\mathcal{S}_n$ the set of channels that can be used for relay n. Then, the discarded proposal is that of relay/channel couple

$$(n', k') = \min_{k, n \in \mathcal{Q}_k} R_{n,k}(P^*_{n,k})\,, \tag{4.30}$$

and the maximum power to be used for transmission to relay n is

$$P_n^{(tot)} = P_{tot} - \sum_{k \neq \mathcal{S}_n} \min_{n \in \mathcal{Q}_k} P_{n,k} \,. \tag{4.31}$$

The resulting algorithm is reported in Algorithm 4.

Algorithm 4: Rate_Offer_Phase_1

input : $\bar{\boldsymbol{R}}$
output: $\boldsymbol{P}$, $\boldsymbol{R}$

Set $\mathcal{Q}_k = \emptyset$, $\mathcal{S}_n = \{1, \ldots, K\}$, $P_n^{(tot)} = P_{tot}$;
while $(\exists k : |\mathcal{Q}_k| < N - 1)$ **do**
  for $n = 1$ **to** N **do**
    $(\boldsymbol{P}_{\cdot,n}, \boldsymbol{R}_{\cdot,n})$ = `MACalPowRate(`$\boldsymbol{H}_{\mathcal{S}_n,n}$, $P_n^{(tot)}$, $\sum_{k=1}^{K} \bar{R}_{n,k}$`)`;
  end
  Compute (n', k^*) using (4.30);
  $\mathcal{Q}_{k^*} = \mathcal{Q}_k \cup \{n^*\}$;
  $\mathcal{S}_n = \{k : n \notin \mathcal{Q}_k\}$;
  for $n = 1$ **to** N **do**
    Compute $P_n^{(tot)}$ from (4.31);
  end
end
return

4.6 Numerical Results

Let us consider the scenario reported in Fig. 4.6, where the relay nodes are positioned along a line that is orthogonal to the segment between Alice to Bob, intersecting it at a distance d_{I} and d_{II} from Alice and Bob, respectively. Moreover, relays are equispaced with a distance Δ between any two adjacent relays. We further assume that the eavesdropper is at least at a distance d_{E} from any transmitting node, i.e., it is outside of the dashed circles surrounding Alice and the relays.

The $K = 16$ channels between any couple of nodes are assumed independent Rayleigh fading. We also consider $P_{\mathrm{tot},1} = P_{\mathrm{tot},2} = 1$. The average SNR at unitary distance is of 0 dB, and the path loss coefficient is 3.5, thus the average SNR at distance d is $d^{-3.5}$. About the eavesdropper, since it is assumed to be at a minimum distance d_{E} from any transmitting node, the outage gain is obtained from (4.18). For finite-length coding we assume a CER at Bob $\kappa = 10^{-3}$ and $m = 128$ and $4,096$.

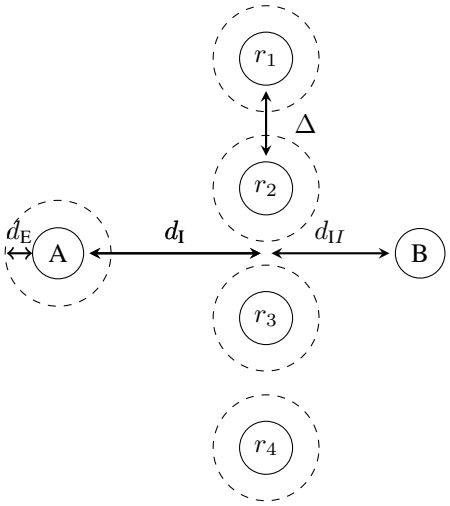


Figure 4.6: Node position diagram. A: Alice, B: Bob, r_n: relay n.

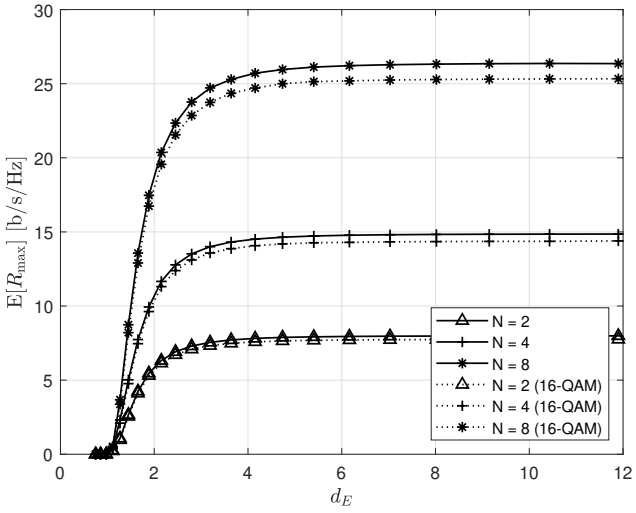


Figure 4.7: Average maximum secrecy rate as a function of d_{E} with infinite-length coding, both Gaussian and discrete (16-QAM) constellations and various values of N.

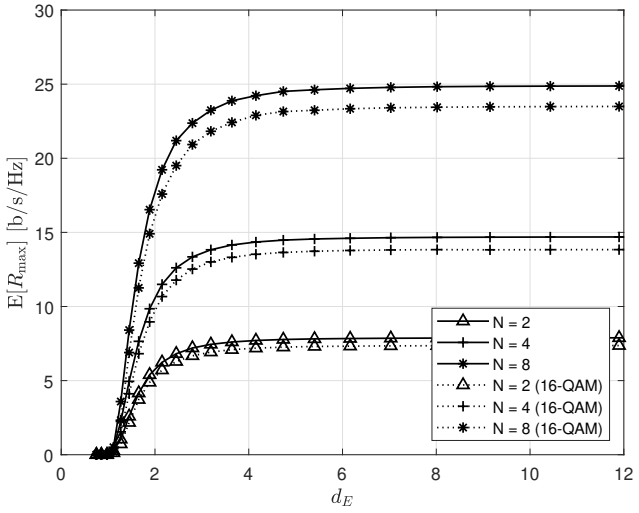


Figure 4.8: Average maximum outage secrecy rate as a function of d_{E} with finite-length coding ($m = 4,096$), both Gaussian and discrete (16-QAM) constellations and various values of N.

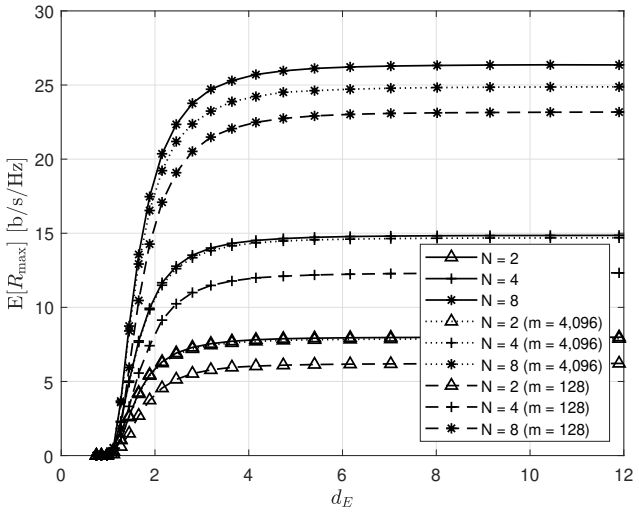


Figure 4.9: Average maximum outage secrecy rate as a function of d_{E} with Gaussian constellation, both infinite- and finite-length ($m = 128$ and $m = 4,096$) coding and various values of N.

4.6.1 Impact of Eve's distance

We first consider a scenario in which each relay has the same distance from Alice and Bob, i.e., $d_{\mathrm{I}} = d_{\mathrm{II}} = 0.8$, the separation between relays is $\Delta = 0.05$, the number of relays is $N = 2$, 4 or 8.

Figs. 4.7 and 4.8 show the average maximum outage secrecy rate $\mathbb{E}[R_{\max}]$, averaged over channel realizations, as a function of d_{E}, for a target secrecy outage probability $\epsilon = 10^{-4}$, and comparing different coding and constellations settings. Fig. 4.7 reports results for a transmission using infinite-length coding and both Gaussian and discrete constellations. We observe that the use of 16-QAM does not decreases significantly the performance in this case, since the maximum rate of 16 channels with 16-QAM is 256 b/s/Hz, well above the average secrecy rate of 25 b/s/Hz actually achieved in the considered setting even with the Gaussian constellation. Moreover, by increasing the number of relays, the average maximum outage secrecy rate increases, as a diversity gain is available on the links among legitimate nodes. Fig. 4.8 shows results for finite-length coding and both Gaussian and discrete constellations. For the Gaussian constellation, comparing Figs. 4.7 and 4.8 we note a negligible performance degradation for a codeword length $m = 4,096$ with respect to infinite-length coding, since $Q^{-1}(\kappa) = 3.1$ and from (4.14) the loss is of the order of $K \cdot 10^{-3} \approx 10^{-2}$. About Fig. 4.8 we further observe that finite-length coding further increases the gap with respect to Gaussian constellation: this is due to the fact that proper matching in the two phases of relaying must be found to achieve a end-to-end secrecy rate and adding constraints further limit this performance, in a non-linear fashion. Lastly, as $d_{\mathrm{E}} \to \infty$ we note that the rate curves flattens in correspondence of the insecure rate of the relay parallel channels, as in this case security conditions are always met and the performance is limited only by the legitimate channel conditions.

4.6.2 Impact of Codeword Length

In Fig. 4.9 the impact of the codeword length for finite-length coding with Gaussian constellation is investigated. Note that the performance of codes with long codewords ($m = 4,096$) is comparable to that of infinite-length coding. Considering a 16-QAM, differences between infinite-length coding and 4,096-length coding are negligible as the average maximum outage secrecy rates coincide for all numbers of relay nodes. Similar results are obtained with discrete constellations, not reported here for the sake of conciseness, where, as seen before, the impact of finite-length coding is stronger than for the Gaussian constellation.

4.6.3 Impact of Relative Node Distances

We then study the impact of the relative distances among legitimate nodes. In particular, we fix the Alice-Bob distance to $d_{\mathrm{I}} + d_{\mathrm{II}} = 2$, and we let the ratio between the

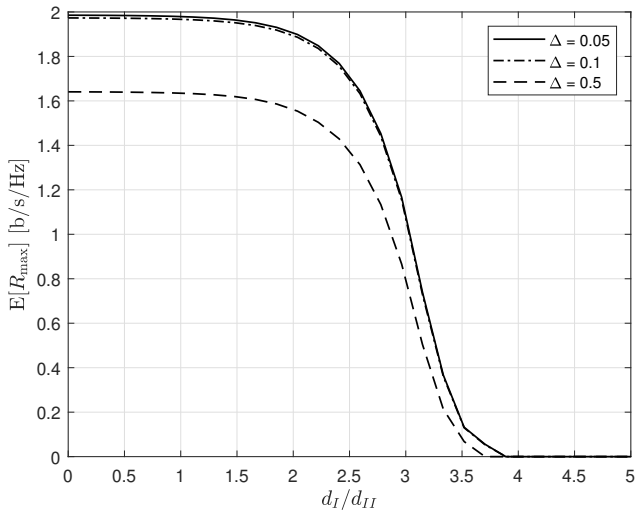


Figure 4.10: Average maximum outage secrecy rate as a function of $d_{\mathrm{I}}/d_{\mathrm{II}}$, various values of Δ, and finite-length coding ($m = 4,096$) with discrete (16-QAM) constellations, for $d_{\mathrm{E}} = 10$, $\epsilon = 10^{-4}$ and $N = 4$ relays.

two distances $d_{\mathrm{I}}/d_{\mathrm{II}}$ and the distance among the relays vary, i.e., $\Delta = \{0.05, 0.1, 0.5\}$, for $d_{\mathrm{E}} = 10$, $\epsilon = 10^{-4}$ and $N = 4$ relays.

Fig. 4.10 shows the average maximum outage secrecy rate as a function of $d_{\mathrm{I}}/d_{\mathrm{II}}$, and finite-length coding ($m = 4,096$) with Gaussian constellation. We observe that for decreasing values of Δ the curves tend asymptotically to a maximum average secrecy rate of 2 b/s/Hz. On the other hand, as $d_{\mathrm{I}}/d_{\mathrm{II}}$ tends to infinity, the average maximum outage secrecy rate tends to zero, as the Alice-relay links will provide vanishing data rates. When the distance Δ tends to zero all the relay nodes are squeezed in the same point between Alice and Bob, which represents the optimal relaying configuration.

4.6.4 Comparison With Other Solutions

Figs. 4.11 and 4.12 provide a comparison between our resource allocation (denoted as Gale-Shapley, or GS) strategy and two suboptimal solutions, respectively uniform power allocation over the K channels and water-filling allocation. Various scenarios are considered, i.e. infinite-length codes with Gaussian signaling and finite-length coding with discrete constellations. We also consider that each relay has the same distance from Alice and Bob, i.e., $d_{\mathrm{I}} = d_{\mathrm{II}} = 0.8$, the separation between relays is $\Delta = 0.05$, the number of relays is $N = 2$, 4 or 8. Water-filling provides the best possible power allocation in Eve's absence, since it assigns more power to the channels presenting better gains. However, this solution is not convenient from a

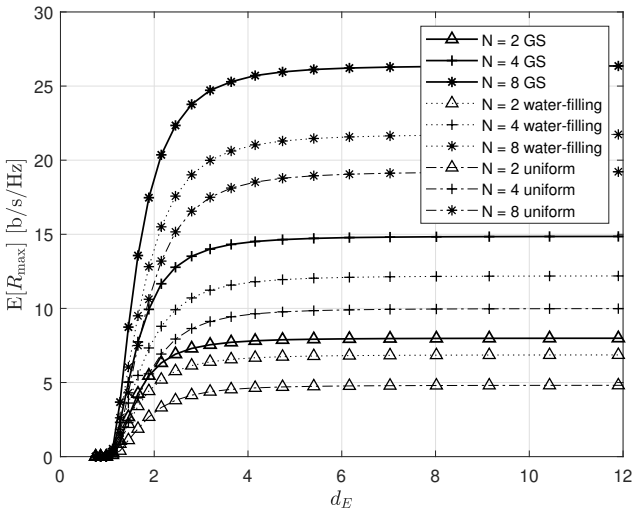


Figure 4.11: Average maximum outage secrecy rate as a function of d_E, with infinite-length coding ($m = 4,096$) and Gaussian constellations, obtained using: GS power allocation, water-filling and uniform power allocation.

security standpoint, since channels that are good for the legitimate receiver could also be good for the attacker, thus degrading the secrecy performance. As predictable, uniform allocation leads to the worst average secrecy rate for all the considered cases.

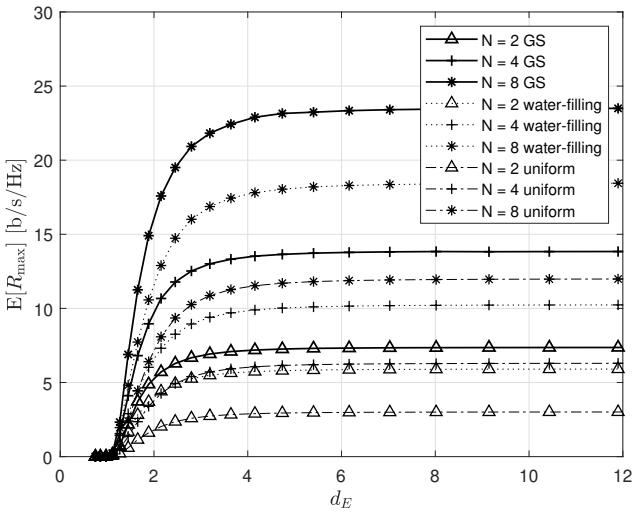


Figure 4.12: Average maximum outage secrecy rate as a function of d_E, with finite-length coding ($m = 4,096$) and discrete (16-QAM) constellations, obtained using: optimal power allocation, water-filling and uniform power allocation..

4.7 Summary

In this chapter we have derived the secrecy rate of the Gaussian relay parallel channel under finite-length coding and discrete constellation constraints, defined as the maximum rate for which a minimum equivocation rate is achieved at Eve. Moreover, we have applied a coupled version of the Gale and Shapley algorithm to allocate power within each channel in order to maximize the secrecy rate. Numerical results show the effectiveness of the resource allocation approach we consider, and show that moderate sizes of both the constellation alphabet and the codewords are sufficient to achieve close-to-optimal secrecy rates for typical wireless transmission scenarios.

Chapter 5

Security in heterogeneous distributed storage systems

Distributed storage systems and caching systems are becoming widespread, and this motivates the increasing interest on assessing their achievable performance in terms of reliability for legitimate users and security against malicious users. While the assessment of reliability takes benefit of the availability of well established metrics and tools, assessing security is more challenging. The classical cryptographic approach aims at estimating the computational effort for an attacker to break the system, and ensuring that it is far above any feasible amount. This has the limitation of depending on attack implementation and advances in computing power. The information-theoretic approach instead exploits capacity measures to ensure unconditional security against attackers, but often does not provide practical recipes to achieve such a condition. We propose a mixed cryptographic/information-theoretic approach to reach a twofold goal: estimating the levels of information-theoretic security and defining a practical scheme able to achieve them. Proper dimensioning of the scheme's parameters, in order to optimize the security metrics, is achieved through the application of an effective probabilistic model checking, thus removing most of the limitations related to more conventional methods.

The distributed storage systems (DSSs) analyzed in this work are well-known and our goal is to characterize them, in order to derive explicit and practical design rules, through the mixed cryptographic/information theoretic approach, by taking advantage of the powerful tool provided by probabilistic model checking (PMC). More precisely, we model the DSSs as Markov decision processs (MDPs). Then we use PRISM [81], a state of the art probabilistic model checker, to measure the likelihood of an attack. By exploiting a result of Massey [82], we later estimate the eavesdropper equivocation about the secret message, and define the conditions under which it may equal the message entropy, thus achieving perfect secrecy. This results in the definition of some sets of the system parameters that permit us to achieve perfect secrecy in practical conditions.

5.1 Related works

Formal verification of security requirements in communication protocols is a well-established practice. In particular, model-checking [83–86] enables to verify automatically all the possible interleaved runs of the protocols in the presence of an adversary that can intercept, remove, modify the original messages as well as inject new messages. In this respect, the Dolev-Yao intruder model [87] is considered the most general model (the worst case) as it assumes a non-deterministic attacker in full control of the communication channels.

Traditionally, model checking can verify whether a system can be attacked or not and are not suitable for verifying security protocols in systems characterized by uncertainty or using randomized algorithms. We *assume*, instead, that every component of a cloud system may be attacked with some probability, and we wish to measure the likelihood of such attacks. This motivated us to define custom probabilistic intruder models, in place of the Dolev-Yao intruder.

Our scenario requires the use of probabilistic model-checking that provides a quantitative measure of security in terms of the probability of reaching a bad state. Examples of probabilistic model-checking applications in security can be found in the recent literature [88–91]. Probabilistic (as also the traditional one) model-checking suffers in general the state-explosion problem, making the verification of real-world security protocols and systems sometimes unfeasible.

One way to address this problem is to trade memory for computation by statistically [92] measuring the probability to satisfy or to violate a property over a set of traces generated by randomly sampling the model. This method in general requires a large number of samples to measure the probability of a rare event such as a security breach.

Another direction, that we have also pursued in this paper, is to find a suitable *abstraction technique* [93] that reduces the description of the system to a feasible state-space, still preserving the properties of interest. For example, distributed systems consisting of several instances of identical communicating components can benefit by proving a *small model theorem* that guarantees the existence of a bound in the number of the identical components for which it is sufficient to solve the verification problem to prove the correctness also for any larger number of components. Although this approach, also referred to as *parameterized verification*, has gained a lot of interest to verify (non-)deterministic systems [94–97] to the best of our knowledge it is still scarcely explored in the probabilistic setting [98, 99].

Finally, let us remark that also the structure of the attacker may determine the feasibility of the verification of security properties. In our work we have employed a passive intruder model, and indeed several authors agree that this is enough when analyzing confidentiality requirements. For example, Li and Pang [100], and Shmatikov [101] used passive intruders to verify anonymity of protocols, a special case of confidential-

ity. The latter work also considers probabilistic attacks. As far as we know, the use of a probabilistic passive attacker model for the analysis of data confidentiality is original.

5.2 System Overview

The scheme we consider for reliable and secure storage of user data into a distributed system, depicted in Fig. 5.1, is based on three main steps: i) encryption, ii) slicing and iii) encoding. They are described next.

5.2.1 Encryption

Let us suppose that the user wishes to upload a file of x bits into the system. This block of x data bits is first subject to a randomized encryption step based on an AONT [40, 102].

The distinctive feature of AONTs is that when the whole amount of encrypted data is available, the encryption map can be painlessly inverted without the need of any pre-shared secret key, and the plaintext data easily recovered. However, in presence of one or more missing or erroneous bits in the ciphertext, inverting the AONT does not allow to recover any useful information concerning the plaintext data. Therefore, any user able to collect the whole ciphertext is able to get the plaintext, otherwise it cannot be even partly recovered.

The classical way of implementing an AONT relies on computational security concepts. In fact, as described in [40], an AONT can be obtained starting from a classical symmetric cipher like the advanced encryption standard (AES) [1]. In this case, a random encryption key is chosen and the message is divided in blocks of size compliant with the chosen symmetric cipher. Then, each block is encrypted with the symmetric cipher using the random key, and a hash digest of each encrypted block is computed. All hash digests are XORed together and with the random key itself, and the result is appended as a last block. Any user who is able to collect all encrypted blocks is also able to compute their hash digests and retrieve the random key, that is then used to decrypt all blocks and obtain the plaintext. Missing one or more blocks does not allow to recover the random key and, hence, the plaintext data. Other implementations of AONTs can be found in the literature. For example, it has been shown in [41] that

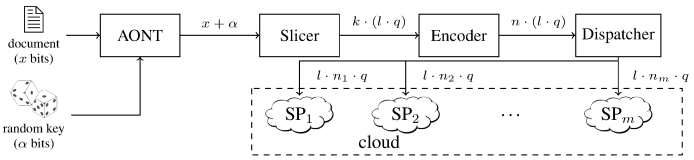


Figure 5.1: Block diagram of AONT-RS (with data length expressed in bits)

an AONT can also be obtained under the paradigm of unconditional security.

In the system we consider, the x-bit secret message is appended with y pseudo-random bits before entering an AONT. Such bits are generated through a pseudo-random numbers generator (PRNG) of sufficient length (the meaning of "sufficient" will be clarified afterwards). At the output of the AONT we get $z = x + y + \alpha$ encrypted bits, where α represents the ciphertext expansion due to the AONT (e.g., for adding the last block containing the random key).

5.2.2 Dispersal

After encryption, data are divided into slices that have to be dispersed among nodes (we consider that one slice is sent to each node, but the analysis can be extended to other scenarios). The slicer deals with the fragmentation of the bits into k slices, where k is the dimension of the linear block code subsequently used in the encoding step.

Each slice contains l blocks of q bits, where $l = \left\lceil \frac{z}{kq} \right\rceil$. The choice of k depends on the type of encoder that follows the slicer, as we will see in the next subsection. If necessary, some stuffing bits are included. However, for the sake of simplicity, we suppose that z is equal to klq, so that the slicer can exactly divide the z bits of the message in l blocks.

After encoding, the number of slices increases from k to $n > k$ due to redundancy added by the code. Slices are then ready to be distributed among nodes in the cloud system.

5.2.3 Encoding

The code used is an erasure code with length n and dimension k, defined over the Galois field of order 2^q. We consider two families of codes: full-length Reed-Solomon (RS) codes with $q > 1$ and $n = 2^q - 1$ and Luby transform (LT) codes with $q = 1$ and several values of n (that, for LT codes, does not depend on q). In this sense, n represents the number of encoded symbols in case of RS codes, while it represents the number of encoded bits for LT codes. The code rate $R = k/n$ represents the fraction of information symbols/bits over encoded symbols/bits.

In the case of RS codes, the code is able to recover a number of erased symbols smaller than or equal to $n-k$, independently of their positions, while it cannot recover a number of erasures greater than $n - k$. In the case of LT codes, instead, for any number of erased bits, there is some probability that the code is able to recover them. As done in [103], we can define two threshold values, k_1 and k_2, with $k_1 < k_2$, such that when the number of available bits is below k_1, we are reasonably sure that the code is unable to correct erasures, and when the number of available bits is above k_2, we are reasonably sure that the code is able to correct all erasures. For a number of available bits between k_1 and k_2, we estimate the erasure recovery probability through

the results of the analysis in [103], by extrapolating them to several code lengths and rates, with a linear approximation.

The encoder generates a block of encoded data having size w equal to the ratio between the overall information length z and the code rate R, i.e., $w = nlq$.

5.2.4 Dispatcher

It sends each of the n slices to one of the m independent storage nodes (the service providers (SPs) of Fig. 5.1) through the network. Each SP receives a number of slices equal to n_i, with $\sum_{i=1}^{m} n_i = n$.

5.3 Security Analysis

We assume that an attacker has the following options to recover the secret message:

1. Exhaustive searching within the message space;
2. Breaking the AONT:
3. Gathering a sufficient number of slices to recover the message.

We also assume that nodes are grouped in different classes, each of them having a different probability to be violated. For this reason, we can refer to the case of heterogeneous systems.

For the sake of simplicity, we suppose that the secret message M is ideally compressed, and hence it has maximum entropy, that is, $\mathbb{H}(M) = x$. This assumption can easily be removed in order to take into account some practical, suboptimal compression function. Under this hypothesis, exhaustive searching within the message space requires 2^x attempts by an attacker, and this obviously represents the ultimate upper bound on the attack complexity.

The second approach that an attacker could try for recovering the message is to break the AONT. In this case, part of the secret message could be decrypted even without having collected a sufficient number of slices. We suppose that the AONT is built by using strong cryptographic components. Under such hypothesis, a reasonable assumption is that the number of attempts required by an attacker to break the AONT is in the order of 2^α. We always fix $\alpha \geq x$, such that for an attacker to break the AONT is not more favorable than exhaustive searching within the message space.

The third kind of attacks we consider consists in recovering a sufficient number of slices in such a way that an attacker could reconstruct the secret message by inverting the AONT. However, we suppose that the attacker has no access to the storage system; therefore, the only way for him to collect slices is to intercept them during transmission. Each slice s_i, $i = 1, 2, 3, \ldots n$, is transmitted over a (wired or wireless) link from which an eavesdropper, noted by Eve in the following, is able to intercept

it with probability p_i, $i = 1, 2, 3, \ldots n$. The probability that Eve successfully recovers a sufficient number of slices and decodes the message from them is denoted as P. Since the system we consider is heterogeneous, the p_i values can be different one each other. Unlike homogeneous systems, for heterogeneous systems it is very difficult to compute P through combinatorial arguments, as also observed in [104, 105]. For this purpose, we propose an approach based on PMC, that is described in Section 5.5.

Let us suppose that the attacker is observing the transmission of a large number of slices corresponding to the storage of a large amount of one or more users' data. We also assume that the attacker has only one chance to intercept as many slices as possible within a set of n slices corresponding to each block. This is realistic since in large storage systems the reading and writing rate of each block of data is very small compared to the overall data transfer speed. Therefore, it is realistic to assume that each attack attempt is independent of the outcome of previous attacks.

Under these hypotheses, the probability that an attacker is able to successfully recover a sufficient number of slices and decodes the message after i attempts is

$$P_i = P \cdot (1 - P)^{i-1}. \tag{5.1}$$

Based on (5.1), the average number of attempts required by an eavesdropper to recover the message through this attack is

$$A = \sum_{i=1}^{\infty} i \cdot P_i = 1/P.$$

Because P_i in (5.1) follows a geometric distribution, from [82] we have

$$\mathbb{H}(M|Z) = E \geq \log_2 (A - 1) + \log_2 (e) = E^*, \tag{5.2}$$

where e is Euler's number, M is the secret message, Z is Eve's observation and E is Eve's equivocation about the secret message.

Based on these considerations, through (5.2) we obtain a lower bound E^* on Eve's equivocation. On the other hand, for an x-bit random message M, we have $E \leq x$. Therefore, when $E^* = x$, we necessarily have $E = x$, that means we achieve perfect secrecy. In the following we show that the perfect secrecy condition can indeed be achieved with practical choices of the system parameters.

5.4 Attack scenarios

According to Fig. 5.1, the m SPs are distributed over the Internet, therefore users must exploit network connectivity in order to dispatch the n slices over them. In this work we assume the user is connected to the Internet from its local area network (LAN) through a gateway or an edge router. Furthermore, the protocol used by the

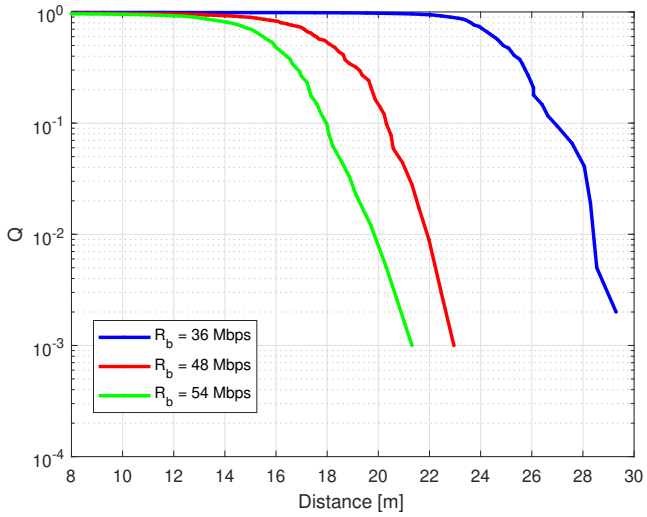


Figure 5.2: Probability of correct reception in a WLAN (L=200 bytes, S_T=12 dBm)

Dispatcher may rely on end-to-end security techniques, like Secure Sockets Layer/-Transport Layer Security (SSL/TLS). Nevertheless, let us investigate the worst case scenario where (1) the intruder can sniff the LAN, and (2) end-to-end encryption does not apply or does not work. These assumptions are reasonable because: (i) a LAN, especially if wireless, seems the most exposed one to eavesdropping; (ii) SSL/TLS requires a public key infrastructure, the related certificate management, and this may not be affordable in some cases (e.g., when SPs are small and cheap storage nodes); (iii) SSL/TLS can be affected by implementation bugs or other vulnerabilities [106, 107].

In this work we focus on wireless attacks. The wireless LAN (WLAN) may be open, or the attacker may be an insider of the network and possess all the credentials. This occurs any time the user is in an open wireless network or in a network protected via techniques based on pre-shared keys (e.g., WEP or WPA-PSK), so that all the other users can successfully acquire all her/his packets.Two scenarios will be analyzed: a user connected only to a WLAN, or a user connected to both a WLAN and a wired LAN.

In wireless connections, packets are subjected to channel conditions such as noise and signal degradation. Of course, a wireless attack depends on how far the eavesdropper is from the source, since the channel quality decreases inversely to the distance. Moreover, the legitimate receiver is authorized to ask for re-transmission of lost packets (e.g. in presence of bad channel conditions), while the eavesdropper cannot, otherwise it would be revealed.

Let us consider a setting where errors affecting one slice are independent of errors affecting another slice. This assumption is realistic in many real-world WLAN de-

ployments, e.g., in the presence of relatively fast fading (coherence time not longer than the duration of a slice). Therefore, the probability that a single slice attack is successful corresponds to the probability that a slice is received by the attacker without errors. This probability depends on different parameters which act on the wireless channel, and we are interested in evaluating it as a function of the distance of the eavesdropper from the user. Let us denote by P the channel packet error rate, and by $Q = (1 - P)$ the probability that a slice is received without errors. The value of Q depends on the transmitted power S_T and on the path loss model. Examples are reported in [108], for the case of indoor communications. Numerical values for P are given in [109], where data is collected through a network simulator and packets with $L = 200$ bytes of application data and $S_T = 12$ dBm are considered. By assuming these parameters, Fig. 5.2 reports the average values of the probability of correct reception for a distance between 8 m and 30 m, considering different data rates R_b, for a typical IEEE 802.11g wireless connection.

5.5 Probabilistic model checking

To model the considered systems and attack scenarios, we use MDPs (see e.g. [110, Ch. 10.5]).

Definition 1 (Markov Decision Processes) *Assume a finite set of atomic propositions AP. A MDP is a tuple $\mathcal{M} = (S, Act, Pr, \iota, L)$ where:*

- $S = \{s_1, s_2, \ldots\}$ *is a finite set of states,*
- $Act = \{\alpha_1, \alpha_2, \ldots\}$ *is a finite set of actions,*
- $Pr : S \times Act \times S \to [0, 1]$ *is a probabilistic transition function such that, for all states $s \in S$ and actions $\alpha \in Act$, $\sum_{s' \in S} Pr(s, \alpha, s') \in \{0, 1\}$;*
- $\iota : S \to [0, 1]$ *is the initial distribution probability of states, such that $\sum_{s \in S} \iota(s) = 1$;*
- $L : S \to 2^{AP}$ *is a labeling function.*

The probabilistic transition function can be extended to sets of states as follows: $\Pr(s, \alpha, T) = \sum_{t \in T} \Pr(s, \alpha, t)$, for all $s \in S$, $\alpha \in Act$, and $T \subseteq S$. Let us call *parametric MDP* any MDP $\mathcal{M}(p_1, \ldots, p_w)$ whose transitions refer to probabilities as parameters $p_1, \ldots, p_w$. We will say that a MDP $\mathcal{M}'$ *instantiates* $\mathcal{M}(p_1, \ldots, p_w)$ if there exist real values $0 \leq q_1, \ldots, q_w \leq 1$ such that $\mathcal{M}'$ is obtained from $\mathcal{M}$ by replacing p_i with q_i, for all i.

Definition 2 (Parallel composition) *Given MDPs $\mathcal{M}_1 = (S_1, Act_1, Pr_1, \iota_1, L_1)$ and $\mathcal{M}_2 = (S_2, Act_2, Pr_2, \iota_2, L_2)$, we write $\mathcal{M}_1 \parallel \mathcal{M}_2$ to denote the MDP $(S_1 \times S_2, Act_1 \cup Act_2, Pr, \iota, L)$ obtained as follows:*

- $Pr((s,t),\alpha,(s',t')) = \begin{cases} Pr_1(s,\alpha,s') & \text{if } \alpha \in Act_1 \setminus Act_2, t = t' \\ Pr_2(t,\alpha,t') & \text{if } \alpha \in Act_2 \setminus Act_1, s = s' \\ Pr_1(s,\alpha,s') \cdot Pr_2(t,\alpha,t') & \text{if } \alpha \in Act_1 \cap Act_2 \end{cases}$
- $\iota((s,s')) = \iota_1(s) \cdot \iota_2(s')$, *for all* $s \in S_1, s' \in S_2$
- $L((s,s')) = L_1(s) \cup L_2(s')$

We say that $\mathcal{M}_1 \parallel \mathcal{M}_2$ *is the* parallel composition *of* $\mathcal{M}_1$ *and* $\mathcal{M}_2$.

Let us write $\mathcal{M}[\alpha/\alpha']$ to denote the MDP where action α has been replaced by action α'. Formally: $(S, Act, \Pr, \iota, L)[\alpha/\alpha'] = (S, Act', \Pr', \iota, L)$ where:

- $Act' = (Act \setminus \{\alpha\}) \cup \{\alpha'\}$, and
- $\Pr'(s,\beta,t) = \begin{cases} \Pr(s,\alpha',t) & \text{if } \beta = \alpha, \\ \Pr(s,\beta,t) & else \end{cases}$

To express properties of probabilistic systems, the probabilistic temporal logic *PCTL** [110, Ch. 10] can be used. We report its grammar:

$$\Phi \quad ::= \quad true \mid p \mid \Phi \wedge \Phi \mid \neg\Phi \mid \mathbb{P}_J(\varphi)$$

$$\varphi \quad ::= \quad \Phi \mid \varphi \wedge \varphi \mid \neg\varphi \mid X\varphi \mid G\varphi \mid F\varphi$$

where $p \in AP$ and $J \subseteq [0,1]$ is a rational interval. Terms of Φ are *state formulae*, while terms of φ are *path formulae*. Intuitively, formula $G\varphi$ (resp. $F\varphi$) holds w.r.t. some path iff every (resp. some) state visited along the path satisfies sub-formula φ. Given an MDP $\mathcal{M}$, we write $\mathcal{M} \models \Phi$ expressing that *all the initial states* of $\mathcal{M}$ satisfy Φ. Given a *PCTL** *path* formula φ and an MDP $\mathcal{M}$, we write $\mathcal{P}_{max}(\varphi, \mathcal{M})$ (resp. $\mathcal{P}_{min}(\varphi, \mathcal{M})$) denoting the maximum (resp. minimum) probability with which the specification φ is satisfied. Such a value can be computed in polynomial time w.r.t. its input [110, Ch. 10.5].

Given two MDPs $\mathcal{M}_1$ and $\mathcal{M}_2$, one can show that they are indistinguishable if (i) every transition to equivalent states on one system is mimicked on the other system, and (ii) equivalent states are reached with the same probability on the two systems. This is captured by *probabilistic bisimulation* [110, Ch. 10.5].

Definition 3 (Probabilistic bisimulation) *Given MDPs* $(S_1, Act_1, Pr_1, \iota_1, L_1)$ *and* $(S_2, Act_2, Pr_2, \iota_2, L_2)$*, a probabilistic bisimulation is any relation* $R \subseteq S \times S$ *such that*

$R(s,s')$ *iff* $L(s) = L(s')$ *and* $Pr(s,\alpha,T) = Pr(s',\alpha,T)$*, for each action* $\alpha \in Act$*, equivalence class* $T \in S/R$*, and* $s,s' \in S$*, where* $(S, Act, Pr, \iota, L) = (S_1, Act_1, Pr_1, \iota_1, L_1) \parallel (S_2, Act_2, Pr_2, \iota_2, L_2)$.

The probabilistic bisimulation is an equivalence relation, thus given two states s,t, let us write $s \approx_R t$ if R is a probabilistic bisimulation and $R(s,t)$ holds. When R is clear from the context, we may omit it. It is known that bisimilar MDPs satisfy the same *PCTL** formulae [110, Ch. 10.5].

Given two MDPs $\mathcal{M}_1$ and $\mathcal{M}_2$ and a sequence of action pairs $\Gamma = \alpha_1/\alpha'_1, \ldots, \alpha_n/\alpha'_n$, let us write $\mathcal{M}_1 \approx_\Gamma \mathcal{M}_2$ to denote that $\mathcal{M}_1[\alpha_1/\alpha'_1]\ldots[\alpha_m/\alpha'_m] \approx \mathcal{M}_2[\alpha_1/\alpha'_1]\ldots[\alpha_m/\alpha'_m]$. We call the relation $\approx_\Gamma$ *probabilistic bisimulation up-to action replacement* and intuitively denotes the fact that $\mathcal{M}_1$ and $\mathcal{M}_2$ are bisimilar modulo a simple operation of renaming their actions.

Theorem 1 *Given two MDPs* $\mathcal{M}_1$*,* $\mathcal{M}_2$ *such that* $\mathcal{M}_1 \approx \mathcal{M}_2$*, then* $\mathcal{M}_1 \models \Phi$ *iff* $\mathcal{M}_2 \models \Phi$*, for any* $\Phi \in PCTL^\star$.

Corollary 1 *Given two MDPs* $\mathcal{M}_1$*,* $\mathcal{M}_2$ *s.t.* $\mathcal{M}_1 \approx \mathcal{M}_2$*, then* $\mathcal{P}_{max}(\varphi, \mathcal{M}_1) = \mathcal{P}_{max}(\varphi, \mathcal{M}_2)$ *and* $\mathcal{P}_{min}(\varphi, \mathcal{M}_1) = \mathcal{P}_{min}(\varphi, \mathcal{M}_2)$ *for any* $\varphi \in PCTL^\star$.

Property 1 (Associativity) *Given MDPs* $\mathcal{M}_1 = (S_1, Act_1, Pr_1, \iota_1, L_1)$*,* $\mathcal{M}_2 = (S_2, Act_2, Pr_2, \iota_2, L_2)$*, and* $\mathcal{M}_3 = (S_3, Act_3, Pr_3, \iota_3, L_3)$*, then:*

$$(\mathcal{M}_1 \parallel \mathcal{M}_2) \parallel \mathcal{M}_3 \approx \mathcal{M}_1 \parallel (\mathcal{M}_2 \parallel \mathcal{M}_3)$$

Proof. Assume $s_1, s'_1 \in S_1$, $s_2, s'_2 \in S_2$, and $s_3, s'_3 \in S_3$. The property is a consequence of the follogin facts: 1) $(S_1 \times S_2) \times S_3$ is isomorphic to $S_1 \times (S_2 \times S_3)$. 2) By product associativity $(\iota_1(s_1) \cdot \iota_2(s_2)) \cdot \iota_3(s_3) = \iota_1(s_1) \cdot (\iota_2(s_2) \cdot \iota_3(s_3))$. 3) By set union associativity $(L_1(s_1) \cup L_2(s_2)) \cup L_3(s_3) = (L_1(s_1) \cup L_2(s_2)) \cup L_3(s_3)$. 4) Finally, for any $\alpha \in Act$ and $p \in [0,1]$, one shows by cases on the definition of Pr that $\Pr(((s_1,s_2),s_3),\alpha,((s'_1,s'_2),s'_3)) = p$ iff $\Pr((s_1,(s_2,s_3)),\alpha,(s'_1,(s'_2,s'_3))) = p$. ■

Property 2 (Commutativity) *Given two MDPs* $\mathcal{M}_1 = (S_1, Act_1, Pr_1, \iota_1, L_1)$ *and* $\mathcal{M}_2 = (S_2, Act_2, Pr_2, \iota_2, L_2)$ *then:*

$$\mathcal{M}_1 \parallel \mathcal{M}_2 \approx \mathcal{M}_2 \parallel \mathcal{M}_1$$

Proof. Assume $s_1 \in S_1$ and $s_2 \in S_2$. The property is a consequence of the following facts: 1) $S_1 \times S_2$ is isomorphic to $S_2 \times S_1$. 2) By product commutativity $\iota_1(s_1) \cdot \iota_2(s_2) = \iota_2(s_2) \cdot \iota_1(s_1)$. 3) By set union $L_1(s_1) \cup L_2(s_2) = L_2(s_2) \cup L_1(s_1)$. 4) Finally, for any $\alpha \in Act$ and $p \in [0,1]$, one shows by cases on the definition of Pr that $\Pr((s_1,s_2),\alpha,(s'_1,s'_2)) = p$ iff $\Pr((s_2,s_1),\alpha,(s'_2,s'_1)) = p$. ■

5.6 Assessment methodology

Here we describe the MDPs modeling user, links to storage nodes and attacker. Later we show how the system can be verified for *any number* of links.

5.6.1 Modeling

At every modification, the AONT-RS schema encrypts the file with a fresh and random key. Then the data is encoded and dispersed. This means that a *write* operation invalidates the slices of previous versions of the same file. On the other side, a *read* operation gives a new opportunity to the attacker for collecting new slices up to the threshold k. We refer this model to as *write-once/read-many*.

At a first sight, the considered model might resemble the threshold public key encryption systems [111], where a private key is distributed among n decryption servers, so that at least k servers are needed for decryption. In reality, between the two systems there are important differences. In particular, in the AONT-based scheme we have only one user and the algorithm exploits symmetric ciphering. The only analogy is on the concept of threshold that, however, while in the case of threshold encryption is applied to the number of users that aim to decipher, in the present case is applied to the number of recovered slices.

In Figs. 5.3, 5.4 and 5.5 we represent the relevant MDPs using straight variable names for edge pre-conditions, while primed variable names are considered edge post-conditions. When the edge does not have a synchronization label (resp. a probability), we assume that the transition is asynchronous (resp. it has probability 1). When the boolean formula is omitted, we assume it is a tautology.

The user. Fig. 5.3 shows the MDP $\textsc{User}(d_1, \ldots, d_m)$ where m is the number of storage nodes. Its variables are: *nw* counts the number of written slices, c tracks the next node to write to, *nra* counts the number of remaining read attempts of the previously written message, $nr_1 \ldots nr_m$ count the slices to be read from the node i, in this attempt of reading the message. Note that the dispatch probabilities make a probability distribution, i.e. $\sum_{1 \leq i \leq m} d_i = 1$. USER also reads variables $ctr_1, \ldots, ctr_m$ from MDPs $\textsc{Link}_1, \ldots, \textsc{Link}_m$ to know the number of slices hosted by each node. USER starts by dispatching the n slices across the available m node links. When done, it goes to the reading stage, where it loops n_reads times reading back the previously written message.

The node links. MDP $\textsc{Link}_i^c(a_i)$ depicted in Fig. 5.4 models any link to some storage node that could host up to c slices. It stores in ctr_i the amount of hosted slices *not known* by the attacker so far, and uses a binary flag $leak_i$ to remember whether the last written or read slice was intercepted by the intruder. Note that only one of the m copies of $\textsc{Link}_i$ can have $leak_i = 1$ at any time. This forces the attacker to intercept the leaked slice before the user tries to write or read another one (possibly leaking it again). Probability a_i represents the likelihood of a slice being intercepted

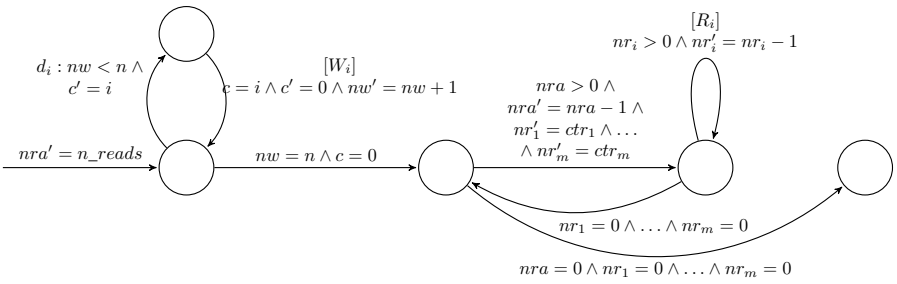


Figure 5.3: The USER$(d_1, \ldots, d_m)$

when traveling between the user and the storage node, while $1 - a_i$ is the probability of not being intercepted. The node link synchronizes with the user through actions R_i and W_i, and with the attacker using action L_i.

The attacker. ATTACKER is the MDP modeling the intruder depicted in Fig. 5.5. It has a single local variable ctr_a counting the number of collected slices so far. The attack proceeds by collecting the slices leaked by the m copies of LINK$_i$ in the system. An attack is successful when $ctr_a \geq k$.

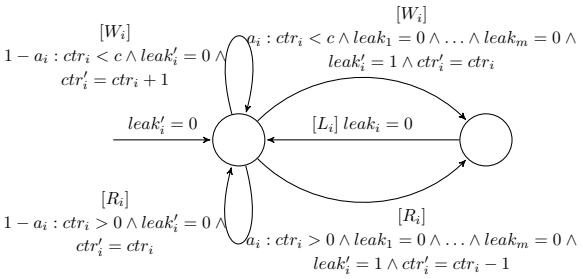


Figure 5.4: The LINK$_i^c(a_i)$

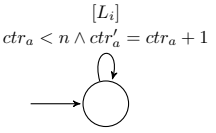


Figure 5.5: The ATTACKER

5.6.2 Security assessment

It is evident that assessing the security of dispersal cloud storage algorithms is inherently a parameterized problem. Indeed, by allowing an arbitrarily large number of read operations by the user, the attacker has probability 1 of intercepting more than k slices (every read the attacker has one more chance of intercepting the missing slices, until it intercepts all of them). Similarly, assuming the secret is split into an arbitrarily large number of slices gives the attacker a negligible probability of succeeding in his/her attack. Between these ends lie all the parameters values of the actual implementations of AONT-based algorithms. Very often such values are not bound to a

clearly stated security metric.

Our approach exploits bounded and probabilistic model checking to compute the likelihood of a successful attack, specified as a *PCTL** formula, for several parameter configurations. The collected data allow us to draw a multi-dimensional graph relating the probability of a successful attack with the parameter values.

For any $m \in \mathbb{N}$, the dispersal cloud storage algorithms is modeled by:

$$\mathcal{M}_m^{\text{AONT}} := \textsc{User}(d_1, \ldots, d_m) \parallel \textsc{Link}_1^n(a_1) \parallel \ldots \parallel \textsc{Link}_m^n(a_m) \parallel \textsc{Attacker}.$$

Finally, the probabilistic model checker is repeatedly invoked to solve the following problem varying parameter values:

$$\mathcal{P}_{max}(F(ctr_a \geq k), \mathcal{M}_m^{\text{AONT}})$$

5.6.3 Small model theorem for node links

A small model theorem allows to verify a class of infinite state systems by only checking a finite size system. The key observation is that, in a system where slices are intercepted when traveling between user and storage nodes, two or more node links with the same attack probability are indistinguishable from a single node link having the same attack probability, modulo some technicalities.

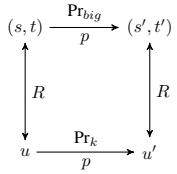
Lemma 1 (Reduction) *For any natural numbers $c, d, i, j, k > 0$ such that $i \neq j$, any probability a. Given the MDPs $\textsc{Link}_i^c(a)$, $\textsc{Link}_j^d(a)$, and $\textsc{Link}_k^{c+d}(a)$:*

$$\textsc{Link}_i^c(a) \parallel \textsc{Link}_j^d(a) \approx_\Gamma \textsc{Link}_k^{c+d}(a)$$

where $\Gamma = R_i/R_k, R_j/R_k, W_i/W_k, W_j/W_k, L_i/L_k, L_j/L_i$.

Sketched. Fix $\textsc{Link}_i^c(a) = (S_i, Act_i, \Pr_i, \iota_i, L_i), \textsc{Link}_j^d(a) = (S_j, Act_j, \Pr_j, \iota_j, L_j), \textsc{Link}_k^{c+d}(a) = (S_k, Act_k, \Pr_k, \iota_k, L_k)$. One shows that there exists a relation $R \subseteq (S_i \times S_j) \times S_k$ that is indeed a probabilistic bisimulation up-to action replacing Γ. Take $R := \{((s,t),u) \ : \ s.ctr_i + t.ctr_j = u.ctr_k, s.leak_i = 1 \iff t.leak_j = 1 \vee u.leak_k = 1\}$.

Call $\Pr_{big}$ the probabilistic transition function of the composed MDP $\textsc{Link}_i^c(a) \parallel \textsc{Link}_j^d(a)$. Now one proves that the following commutative diagram holds:



We first prove one direction. Fix any s, t, u such that $R((s,t),u)$ and then reason by cases on $\Pr_{big}((s,t),\alpha,(s',t')) = p$.

- Case $\alpha = R_i$ (i.e. read on i): if $p = a$ (i.e. a leaking happened) then $s.ctr_i > 0$, $s'.ctr_i = s.ctr_i - 1$, $s.leak_i = t.leak_j = t'.leak_j = 0$, $s'.leak_i = 1$. By $R((s,t),u)$ we know that $u.ctr_k = s.ctr_i + t.ctr_j$ and $u.leak_k = \min(s.leak_i + t.leak_j, 1) = 0$. Let us name u' the (unique) state satisfying the following: $u'.ctr_k = u.ctr_k - 1$ and $u'.leak_k = 1$. It is evident that $R((s',t'),u')$, concluding this branch of the case. If $p = 1 - a$ (i.e. no leaking happened) one similarly observes that $s.ctr_i > 0$, $s'.ctr_i = s.ctr_i$, and $s.leak_i = t.leak_j = s'.leak = t'.leak = 0$. Now, under our assumptions, let us define u' to be the (unique) state where $u'.ctr_k = u.ctr_k = s'.ctr_i + t'.ctr_j$ and $u'.leak_k = 0 = \min(s'.leak + t'.leak, 1)$. Observing that $R((s',t'),u')$ ends this case.
- Case $\alpha = R_j$ (i.e. read on j): it is symmetric to the previous one.
- Cases $\alpha \in \{W_i, W_j, L_i, L_j\}$ are straightforward to check following the reasoning for the case $\alpha = R_i$.

Finally, fix any s, t, u such that $R((s,t),u)$ and reason by cases on $\Pr_k(u,\alpha,u') = p$ to show that the opposite direction holds. ■ Given a sorted list of numbers $a_1, \ldots, a_m$ s.t. $a_1 \leq \ldots \leq a_m$, let us call its *distinction* the list of indices $i_1, \ldots, i_{q+1}$ satisfying the following:

- $i_1 = 1$, $i_{q+1} = m$, and $i_1 < \ldots < i_{q+1}$,
- $\forall j \in [1,q].\forall k \in [i_j, i_{j+1} - 1].\ a_{i_j} = a_k$, and
- $\forall j \in [1,q].\ a_{i_j} < a_{i_{j+1}}$.

Such constraints mean that the list $a_1, \ldots, a_m$ can be partitioned into q sub-lists, each containing identical values, and each pair of lists containing distinct values. For example, the distinction of the sorted list of probabilities $0.00, 0.00, 0.05, 0.10, 0.10, 0.10, 0.15$, is the list of indices $1, 3, 4, 7$.

The core theoretical contribution of this work shows that one can do parameterized probabilistic model checking of systems with any number of LINKs, by considering only a finite number of them. Such number is often called *cutoff*.

Theorem 2 (Small Model Theorem) *For any naturals $m, c_1, \ldots, c_m > 0$ and probabilities $a_1, \ldots, a_m$. Given the MDPs $\text{LINK}_1^{c_1}(a_1), \ldots, \text{LINK}_m^{c_m}(a_m)$. For any MDP $\mathcal{M}$ and formula $\Phi \in$ PCTL* the following holds:*

$$\mathcal{M} \parallel \text{LINK}_1^{c_1}(a_1) \parallel \ldots \parallel \text{LINK}_m^{c_m}(a_m) \models \Phi \Leftrightarrow$$

$$\mathcal{M} \parallel \text{LINK}_1^{c_{i_1}}(a_{i_1}) \parallel \ldots \parallel \text{LINK}_q^{c_{i_q}}(a_{i_q}) \models \Phi$$

where, for some $0 < q \leq m$*, the list of indices* $i_1, \ldots, i_q$ *is a distinction of the list* $a_1, \ldots, a_m$ *(assume w.l.o.g. that the latter is sorted), the dispatch probabilities are given by* $d_{i_j} = \sum_{k=i_j}^{i_{j+1}-1} d_k$ *while the capacities are defined as* $c_{i_j} = \sum_{k=i_j}^{i_{j+1}-1} c_k$*.*

Proof. Let us recursively apply Lemma 1. The latter reduces any pair of LINKs with identical attack probabilities to a single LINK. The procedure ends when all the LINKs have distinct attack probabilities. By Lemma 1, for every $j \in [1, q]$, the LINK having attack probability a_{i_j} has capacity c_{i_j} (resp. dispatch probability d_{i_j}) defined as the sum of the capacities (resp. of the dispatch probabilities) of all the original LINKs with identical attack probability. By Lemma 1 and Theorem 1 they satisfy the same *PCTL** formulae. ■

5.7 Numerical results

In this section, we analyze two case studies. In the first case study, we measure the efficiency of channel usage necessary to achieve perfect secrecy, in terms of the ratio between the size of the original message x and size of the message after being processed w . In the second one, we consider two typical scenarios of real-world implementations of dispersal cloud storage systems: (a) the user is connected to a wireless LAN, and (b) the user is connected to two LANs, one wired and one wireless (combined scenario).

5.7.1 First case study

The procedure COMPUTEMETRICS shown in 5.15 summarizes how the security metrics presented in Section 5.3 are computed by our framework. It takes as input a template process $\mathcal{M}^T$ describing the system, a specification $\mathcal{R}_{att}$ describing the attack states, the number of slices n, the threshold values k_1 and k_2, with $k_1 \leq k_2$ and q.

The PMC (in our case PRISM [81]) is repeatedly invoked, for increasing values of l, to compute the probability that an attack is successful. Then, according to the analysis in Section 5.3, it assumes that the source message has maximal entropy and perfect secrecy is reached. The ratio x/w between the message size and the total size is then plotted, which has the meaning of channel usage. To maximize this parameter is a premium, as it means that, for a given value of w, the amount of information bits transferred to the nodes with perfect secrecy is maximum. Incidentally, we observe that the procedure indirectly fixes the value of y, so clarifying the meaning of the claim in Section 5.2.1, where it was said that the output of the PRNG must be sufficiently long.

For the simple case of a homogeneous system, that can be easily modeled through combinatorial arguments, the framework has been validated replacing the call to PMC with a different description of the same system expressed via an independent analysis

Algorithm 5: ComputeMetrics($\mathcal{M}^T$,R_{att},n,k_1,k_2,q)

5.1 $l \leftarrow 1$
5.2 **while** true **do**
5.3 $\quad p_{guess} \leftarrow 1/2^{l \cdot q}$ $p_{attack} \leftarrow PMC(\mathcal{M}^T(n, k_1, k_2, p_{guess}), R_{att})$
5.4 $\quad E^\star \leftarrow \log_2(e \cdot (1/p_{attack} - 1))$ ▷ cfr. [82]
5.5 $\quad$ ▷ Assume key and message have maximal entropy
5.6 $\quad x \leftarrow \alpha \leftarrow \lfloor E^\star \rfloor$
5.7 $\quad z \leftarrow k \cdot l \cdot q$ ▷ Message length
5.8 $\quad y \leftarrow z - x - \alpha$ ▷ Random padding
5.9 $\quad w \leftarrow n \cdot l \cdot q$ ▷ Total size
5.10 $\quad$ **if** $y < 0$ **then**
5.11 $\quad\quad y \leftarrow 0$ $x \leftarrow \alpha \leftarrow \lfloor z/2 \rfloor$
5.12 $\quad$ **end**
5.13 $\quad$ GraphRatio(l, x/w)
5.14 $\quad l \leftarrow l + 1$
5.15 **end**

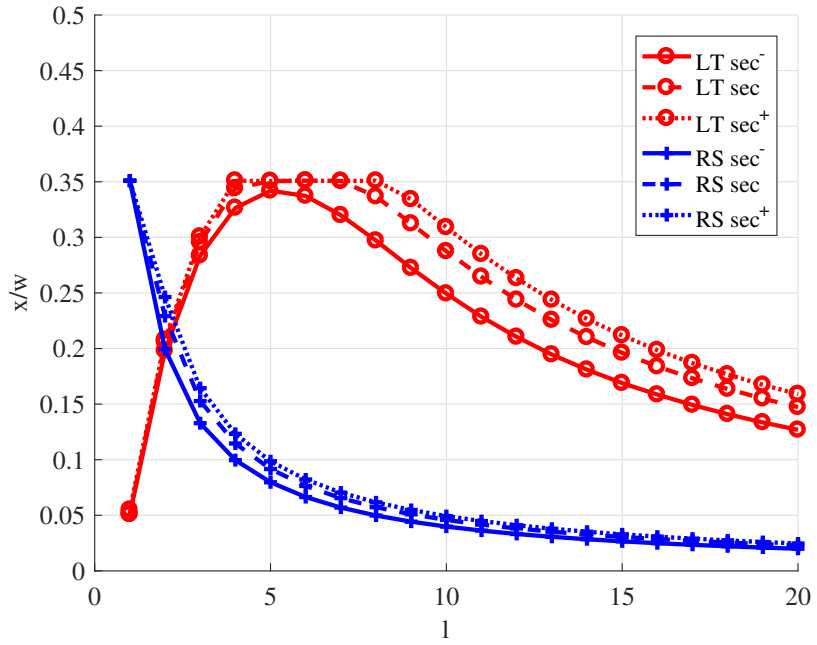
in Matlab. We verified on small instances of the problem (the only treatable through Matlab) that the obtained probabilities were equivalent.

It is well known that one critical issue of (probabilistic) model checking is the so-called *state-space explosion*. Namely, every added variable in the model causes an exponential increase in the number of states in the model itself, and consequently in the computation time. For such reason we use this section also to show the feasibility of the proposed framework on problem instances of realistic size. Each call of the model checker terminates within a time ranging from few seconds to 10 minutes, after which it timeouts, interrupting the experiment for that series.

Our first example compares AONT-RS and AONT-LT channels with $n = 255$ slices and different slice dispatching policies. Namely, for each channel we test three scenarios: in the first one (sec$^-$), it is more likely that slices are sent towards groups of nodes with higher intercept probabilities; the second (sec) tend to dispatch slices towards groups of nodes having smaller intercept probability; the third one (sec$^+$) sends even more slices than (sec) to more secure nodes.

In Fig. 5.6 we graph how the channel usage (and then, indirectly the security metric) varies w.r.t. the parameter l. For the AONT-LT case, the greatest channel usage is obtained by choosing l between 4 and 7, depending on the dispatching policy. For the AONT-RS case, instead, the best channel usage is reached with $l = 1$. Surprisingly, for the latter case dispatching slices through more secure links does not improve significantly the efficiency of the system.

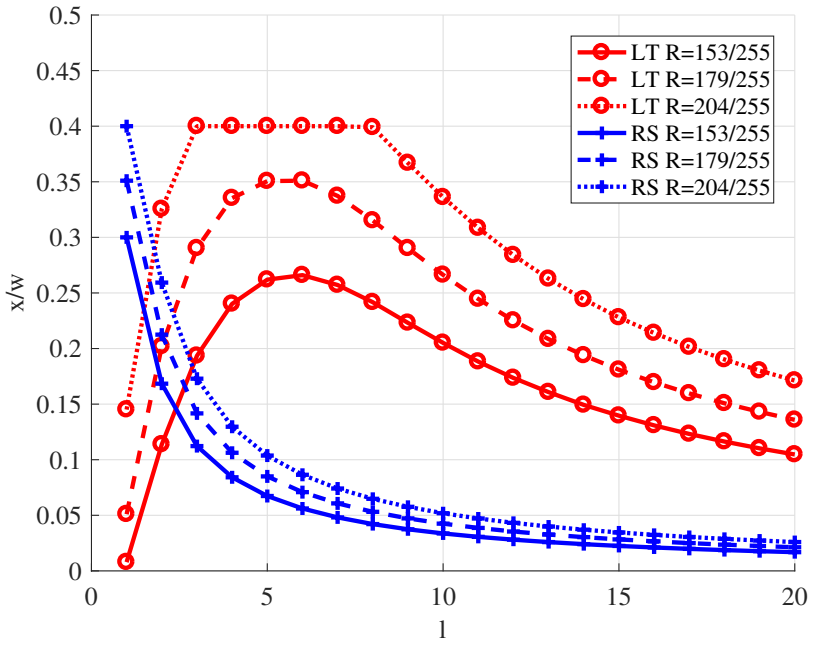
In a second example we compare channels again with $n = 255$ slices, choosing different code rates R. It is clear that, the greater the rate, the less the system is exposed to attacks. On the other side, the lower the rate, the more the system is



Parameters	LT	RS
n	255	
k_1	$0.6 \cdot n$	$0.7 \cdot n$
k_2	$0.8 \cdot n$	$0.7 \cdot n$
m	5	
d_i for $i \in [1, m]$	sec^- : $[0.1, 0.1, 0.2, 0.2, 0.4]$ sec : $[0.4, 0.2, 0.2, 0.1, 0.1]$ sec^+ : $[0.6, 0.2, 0.1, 0.05, 0.05]$	
a_i for $i \in [1, m]$	$5 \cdot i/1000$	

Figure 5.6: The efficiency of channel usage compared under different dispatch policies.

resilient to node failures. It is not immediate to see, though, which of the two ends may ensure a higher level of equivocation and consequently a more efficient usage of the channel itself, when ensuring perfect secrecy. The graph in Fig. 5.7 shows that, under the considered set of parameters, when the code rates are closer to 1, both AONT schemes achieve a higher efficiency, as the ratio between message size and total size x/w shows. In the case of AONT-RS the maximum efficiency is reached with $l = 1$, while for AONT-LT with $4 \leq l \leq 8$, depending on the actual rates. This is somehow expected, since higher coding rates yield a more efficient use of the channel. However, we are also taking into account the security performance; therefore, this conclusion is not trivial.



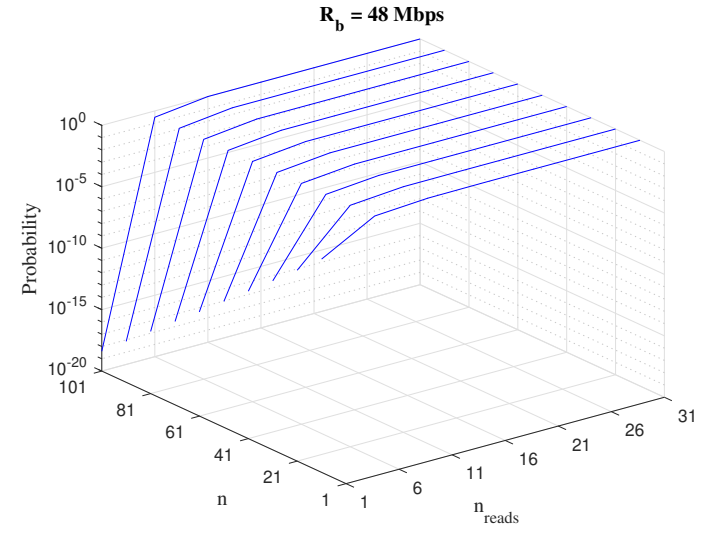
Parameters	LT	RS
n	255	
m	5	
d_i for $i \in [1, m]$	0.2	
a_i for $i \in [1, m]$	$5 \cdot i/1000$	

Figure 5.7: The efficiency of channel usage compared under different code rates $R = k/n$.

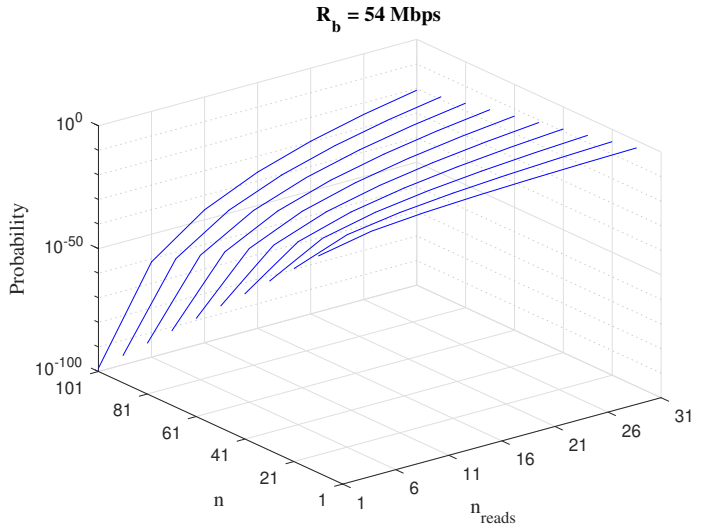
5.7.2 Second case study

Let now consider the second case study, where two different kinds of attack are implemented, one through a wireless LAN and the other in a combined scenario, in which the user is connected to two LANs, one wired and one wireless.

In our experiments we choose the number of slices n to range between 10 and 100, and the number of read events n_reads between 1 and 31. As explained in Sec. 5.4, the probability of an attack in the wireless network depends on the distance of the attacker from the user. Here we set such a distance to 20 m which, according to Fig. 5.2, corresponds to a probability of intercepting a slice of 0.009 (resp. 0.148) for a network operating at 54 Mbps (resp. 48 Mbps). The successful attack probability of the wired LAN is instead assumed to be zero. Consequently, we can apply the small



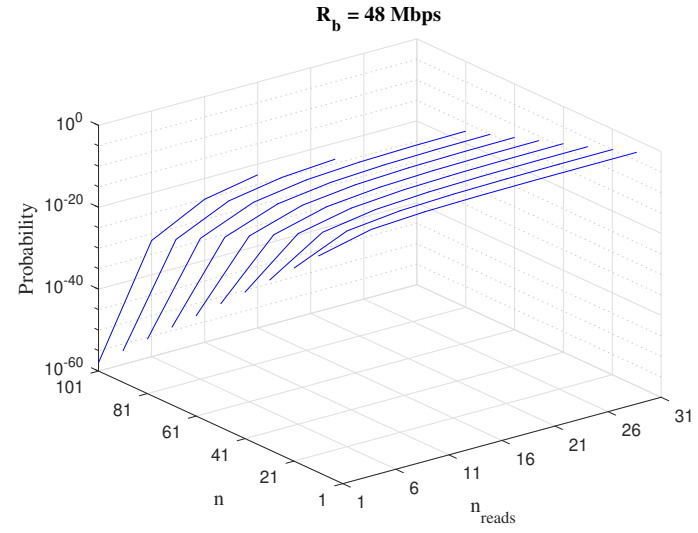
(a)



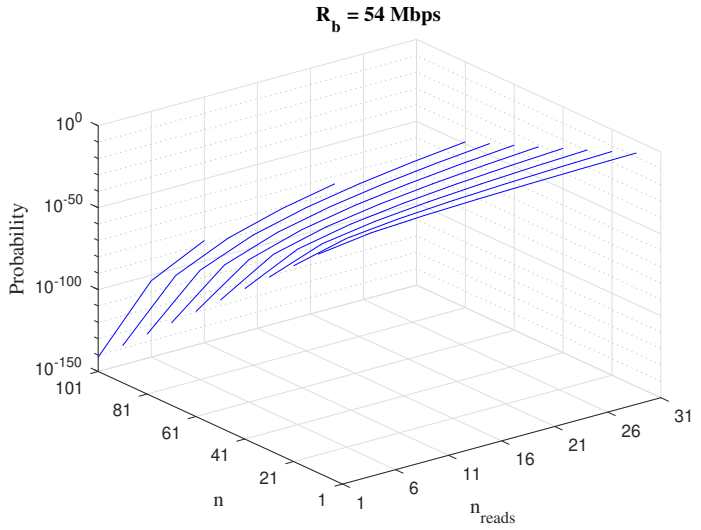
(b)

Figure 5.8: Probabilities of a successful attack to confidentiality in a wireless scenario.

model theorem discussed in Sec. 5.6.3, and reduce the actual number m of LINKs in the model checked system to a fixed value, i.e. the number of different intercept probabilities in the system. Thus, m equals the number of considered LANs. Finally, in the presented case studies we consider a threshold $k = 0.7 \cdot n$ and for the combined scenario that slices are routed more likely to the wired LAN (75%) than to the wireless



(a)



(b)

Figure 5.9: Probabilities of a successful attack to confidentiality in a combined scenario.

LAN (25%).

To assess confidentiality in both scenarios we used SecMC[1], an open-source modular tool allowing to define model checking workflows. The tool repeatedly invokes

[1] https://bitbucket.org/fcloseunivpm/secmc

the PRISM model checker [81]. Each invocation instantiates a parametric MDP and returns the probability of a successful attack (as defined by the security metric (5.6.2)). Figures 5.8 and 5.9 plot the obtained results. In them, each line corresponds to a given number of slices used to split the message, and every line relates the likelihood of a successful attack to the number of read attempts by the client.

Our analysis reveals that while the considered cloud dispersal protocols protect against untrusted storage nodes, they do not ensure a high level of confidentiality against an eavesdropper in the same wireless LAN of the user, in the case that some storage nodes do not use end-to-end cryptography, or the implementation of the latter is broken. Fixing the same parameters n and n_reads, the measure of confidentiality may be several orders of magnitude bigger in the case of 54 Mbps networks w.r.t. 48 Mbps. But, especially in the wireless scenario, the probability of an attack grows too fast with the number of file reads. Furthermore, in networks with 48 Mbps rate or lower, even using 100 or more slices, it is enough to force the user to read the file 6 or more times to reach a probability of reconstructing the file close to 100%.

The methodology can assess security metrics while designing cloud dispersal algorithms. Assume the designer fixes this reference scenario: 54 Mbps wireless LAN, an eavesdropper 20 m far from the user. Let assume that the probability of an attack should be bounded by 10^{-21}. This provided, the methodology suggests to limit the number of reads before overwriting and redistributing the file between 11 and 15 (resp. between 26 and 30) in case of 51 slices (resp. 91 slices).

5.8 Summary

We have presented a framework to design and assess DSSs based on data dispersal algorithms able to achieve a given level of security. The framework is based on a joint use of computational and information-theoretic security notions. In our model we have considered a client that can read and write in the storage nodes and a passive intruder that can steal individual slices from heterogeneous communication channels without compromising the storage nodes. The presented case studies show how our approach, based on a solid PMC formulation, can be exploited to find, when implementing the DSS, the parameter values that optimize the security metrics of interest.

From the developed analysis we argue that the considered architectures of AONT based DSSs are capable of expressing realistic settings appearing in practice, which are often characterized by heterogeneous distributions and network topologies.

Our methodology can be applied to (1) formally specify a custom security metric for the system under consideration, and (2) to certify at design time the parameter assignments ensuring a given level of security.

Chapter 6

Private information retrieval for caching at the edge

Until now we focused on security issues concerning the transmission of a message in a network, first from a physical layer point of view and then in a dispersal cloud storage. In other words, our aim was to prevent an illegal user to have access to a secret message. Strictly related to the concept of security is the one of privacy. In this case an attacker is not interested in knowing the content of the file, but rather which file has been downloaded from the legitimate user among a class of files. Designing a storage system in which a file can be downloaded without revealing any information of which file is actually downloaded to the servers storing it is usually referred to as *private information retrieval* (PIR).

In this chapter, we consider PIR of content from a cellular network. In particular, we consider the private retrieval of content from a library of files that have different popularities. We consider a similar scenario as in [112] where, to reduce the backhaul usage, content is cached in small-cell base stations (SBSs) using maximum distance separable (MDS) codes. We propose a PIR scheme for this scenario that achieves privacy against a number of spy SBSs that possibly collude. The proposed PIR scheme is an extension of Protocol 3 in [113] to the case of multiple code rates, suitable for the scenario where files have different popularities. We also propose an MDS-coded content placement slightly different than the one in [112] but that is more adapted to the PIR case. We show that, for the conventional content retrieval scenario with no privacy, the proposed content placement is equivalent to the one in [112], in the sense that it yields the same average backhaul rate. We then derive the backhaul rate for the PIR case as a function of the content placement. We prove that uniform content placement, i.e., all files that are cached are encoded with the same code rate, is optimal. This is a somewhat surprising result, in contrast to the case where no PIR is considered, where optimal content placement is far from uniform [112]. We further consider the minimization of a weighted sum of the backhaul rate and the communication rate from the SBSs, relevant for the case where limiting the communication from the SBSs is also important. We finally report numerical results for both the scenario where SBSs are placed regularly in a grid and for a Poisson point process (PPP) deployment model

where SBSs are distributed over the plane according to a PPP. We show numerically that popular content placement is optimal for some system parameters. To the best of our knowledge, PIR for the wireless caching scenario has not been considered before.

Notation: We use lower case bold letters to denote vectors, upper case bold letters to denote matrices, and calligraphic upper case letters to denote sets. For example, $\mathbf{x}$, $\mathbf{X}$, and $\mathcal{X}$ denote a vector, a matrix, and a set, respectively. We denote a submatrix of $\mathbf{X}$ that is restricted in columns by the set $\mathcal{I}$ by $\mathbf{X}|_{\mathcal{I}}$. $\mathcal{C}$ will denote a linear code over the finite field $\mathrm{GF}(q)$. The multiplicative subgroup of $\mathrm{GF}(q)$ (not containing the zero element) is denoted by $\mathrm{GF}(q)^{\times}$. We use the customary code parameters (n, k) to denote a code $\mathcal{C}$ of blocklength n and dimension k. A generator matrix for $\mathcal{C}$ will be denoted by $\mathbf{G}^{\mathcal{C}}$ and a parity-check matrix by $\mathbf{H}^{\mathcal{C}}$. A set of coordinates of $\mathcal{C}$, $\mathcal{I} \subseteq \{1, \ldots, n\}$, of size k is said to be an *information set* if and only if $\mathbf{G}^{\mathcal{C}}|_{\mathcal{I}}$ is invertible. The Hadamard product of two linear subspaces $\mathcal{C}$ and $\mathcal{C}'$, denoted by $\mathcal{C} \circ \mathcal{C}'$, is the space generated by the Hadamard products $\mathbf{c} \circ \mathbf{c}' \triangleq (c_1 c'_1, \ldots, c_n c'_n)$ for all pairs $\mathbf{c} \in \mathcal{C}$, $\mathbf{c}' \in \mathcal{C}'$. The inner product of two vectors $\mathbf{x}$ and $\mathbf{x}'$ is denoted by $\langle \mathbf{x}, \mathbf{x}' \rangle$, while $w_{\mathsf{H}}(\mathbf{x})$ denotes the Hamming weight of $\mathbf{x}$. $(\cdot)^{\top}$ represents the transpose of its argument, while $\mathbb{H}(\cdot)$ represents the entropy function. With some abuse of language, we sometimes interchangeably refer to binary vectors as erasure patterns under the implicit assumption that the ones represent erasures. An erasure pattern (or binary vector) $\mathbf{x}$ is said to be correctable by a code $\mathcal{C}$ if matrix $\mathbf{H}^{\mathcal{C}}|_{\chi(\mathbf{x})}$ has rank $|\chi(\mathbf{x})|$.

6.1 Related works

Recently, private information retrieval (PIR) has attracted a significant interest in the research community [113–123]. In PIR, a user would like to retrieve data from a distributed storage system (DSS) in the presence of spy nodes, without revealing any information about the piece of data she is interested in to the spy nodes. PIR was first studied by Chor *et al.* [124] for the case where a binary database is replicated among n servers (nodes) and the aim is to privately retrieve a single bit from the database in the presence of a single spy node (referred to as the noncolluding case), while minimizing the total communication cost. In the last few years, spurred by the rise of DSSs, research on PIR has been focusing on the more general case where data is stored using a storage code.

The PIR capacity, i.e., the maximum achievable PIR rate, was studied in [113, 119, 120, 122, 123]. In [120, 123], the PIR capacity was derived for the scenario where data is stored in a DSS using a repetition code. In [122], for the noncolluding case, the authors derived the PIR capacity for the scenario where data is stored using an (single) MDS code, referred to as the MDS-PIR capacity. For the case where several spy nodes collaborate with each other, referred to as the colluding case, the MDS-PIR capacity is in general still unknown, except for some special cases [119] (and for repetition codes [123]). PIR protocols for DSSs have been proposed in [113, 115, 117, 118, 121].

In [117], a PIR protocol for MDS-coded DSSs was proposed and shown to achieve the MDS-PIR capacity for the case of noncolluding nodes when the number of files stored in the DSS goes to infinity. PIR protocols for the case where data is stored using non-MDS codes were proposed in [113, 118, 121]. There are also several works on PIR that have gone beyond the classical distributed storage model. In [125], the authors considered PIR with side information. The works [126–132] further generalized the system model in [125] and presented appropriate PIR schemes. A common theme across these papers is that they consider multiple servers and replication. In [129,130], the authors presented PIR schemes for DSSs where the servers are constrained in storage capacity.

6.2 System Model

We consider a cellular network where a macro-cell is served by a macro-cell base station (MBS). Mobile users wish to download files from a library of F files that is always available at the MBS through a backhaul link. We assume all files of equal size.[1] In particular, each file consists of βL bits and is represented by a $\beta \times L$ matrix $\boldsymbol{X}^{(i)}$,

$$\boldsymbol{X}^{(i)} = \begin{pmatrix} \tilde{\boldsymbol{x}}_1^{(i)} \\ \vdots \\ \tilde{\boldsymbol{x}}_\beta^{(i)} \end{pmatrix}$$

where upperindex $i = 1, \ldots, F$ is the file index. Therefore, each file can be seen as divided into β stripes $\tilde{\boldsymbol{x}}_1^{(i)}, \ldots, \tilde{\boldsymbol{x}}_\beta^{(i)}$ of L bits each. The file library has popularity distribution $\boldsymbol{p} = (p_1, \ldots, p_F)$, where file $\boldsymbol{X}^{(i)}$ is requested with probability p_i. We also assume that N_{SBS} SBSs are deployed to serve requests and offload traffic from the MBS whenever possible. To this purpose, each SBS has a cache size equivalent to M files. The considered scenario is depicted in Fig. 6.1.

6.2.1 Content Placement

File $\boldsymbol{X}^{(i)}$ is partitioned into βk_i packets of size L/k_i bits and encoded before being cached in the SBSs. In particular, each packet is mapped onto a symbol of the field $\mathrm{GF}(q^{\delta_i})$, with $\delta_i \geq \frac{L}{k_i \log_2 q}$. For simplicity, we assume that $\frac{L}{k_i \log_2 q}$ is integer and set $\delta_i = \frac{L}{k_i \log_2 q}$. Thus, stripe $\tilde{\boldsymbol{x}}_a^{(i)}$ can be equivalently represented by a stripe

[1] Assuming files of equal size is without loss of generality, since content can always be divided into chunks of equal size.

$\boldsymbol{x}_a^{(i)}$, $a = 1, \ldots, \beta$, of symbols over $\mathrm{GF}(q^{\delta_i})$. Each stripe $\boldsymbol{x}_a^{(i)}$ is then encoded using an (N_{SBS}, k_i) MDS code $\mathcal{C}_i$ over $\mathrm{GF}(q)$ into a codeword $\boldsymbol{c}_a^{(i)} = (c_{a,1}^{(i)}, \ldots, c_{a,N_{\mathsf{SBS}}}^{(i)})$, where code symbols $c_{a,j}^{(i)}$, $j = 1, \ldots, N_{\mathsf{SBS}}$, are over $\mathrm{GF}(q^{\delta_i})$. For later use, we define $k_{\mathsf{min}} \triangleq \min\{k_i\}$, $k_{\mathsf{max}} \triangleq \max\{k_i\}$, and $\delta_{\mathsf{max}} \triangleq \frac{L}{k_{\mathsf{min}} \log_2 q}$.

The encoded file can be represented by a $\beta \times N_{\mathsf{SBS}}$ matrix $\boldsymbol{C}^{(i)} = (c_{a,j}^{(i)})$. Code symbols $c_{a,j}^{(i)}$ are then stored in the j-th SBS (the ordering is unimportant). Thus, for each file $\boldsymbol{X}^{(i)}$, each SBS caches one coded symbol of each stripe of the file, i.e., a fraction $\mu_i = 1/k_i$ of the i-th file. As $k_i \in \{1, \ldots, N_{\mathsf{SBS}} - 1\}$,

$$\mu_i \in \mathcal{M} \triangleq \{0, 1/(N_{\mathsf{SBS}} - 1), \ldots, 1/2, 1\},$$

where $\mu_i = 0$ implies that file $\boldsymbol{X}^{(i)}$ is not cached. Note that, to achieve privacy, $k_i < N_{\mathsf{SBS}}$, i.e., files need to be cached with redundancy. As a result, $\mu_i = 1/N_{\mathsf{SBS}}$ is not allowed. This is in contrast to the case of no PIR, where $k_i = N_{\mathsf{SBS}}$ (and hence $\mu_i = 1/N_{\mathsf{SBS}}$) is possible.

Since each SBS can cache the equivalent of M files, the μ_i's must satisfy

$$\sum_{i=1}^{F} \mu_i \leq M.$$

We define the vector $\mu = (\mu_{\mathbf{1}}, \ldots, \mu_{\mathbf{F}})$ and refer to it as the *content placement*. Also, we denote by $\mathcal{C}_{\mathsf{MDS}}^{\mu}$ the caching scheme that uses MDS codes $\{\mathcal{C}_i\}$ according to the content placement μ. For later use, we define $\mu_{\mathsf{min}} \triangleq \min\{\mu_i | \mu_i \neq 0\}$ and $\mu_{\mathsf{max}} \triangleq \max\{\mu_i\}$.

We remark that the content placement above is slightly different than the content placement proposed in [112]. In particular, we assume fixed code length (equal to the number of SBSs, N_{SBS}) and variable k_i, such that, for each file cached, each SBS caches a single symbol from each stripe of the file. In [112], the content placement is done by first dividing each file into k symbols and encoding them using an $(\tilde{n}_i, k)$ MDS code, where $\tilde{n}_i = k + (N_{\mathsf{SBS}} - 1)m_i$, $m_i \leq k$. Then, m_i (different) symbols of the i-th file are stored in each SBS and the MBS stores $k - m_i$ symbols.[2] Our formulation is perhaps a bit simpler and more natural from a coding perspective. Furthermore, we will show in Section 6.4 that the proposed content placement is equivalent to the one in [112], in the sense that it yields the same average backhaul rate.

6.2.2 File Request

Mobile devices request files according to the popularity distribution $\boldsymbol{p} = (p_1, \ldots, p_F)$. Without loss of generality, we assume $p_1 \geq p_2 \geq \ldots \geq p_F$. The user request is ini-

[2] This is because the model in [112] assumes that one SBS is always accessible to the user. If this is not the case, the MBS must store all k symbols of the file. Here, we consider the case where the MBS must store all k symbols because it is a bit more general.

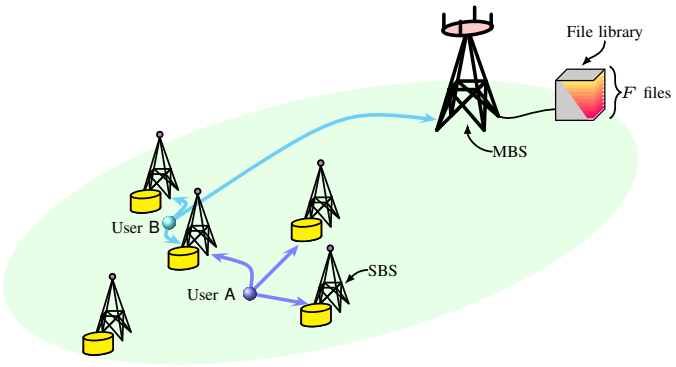


Figure 6.1: A wireless network for content delivery consisting of a MBS and five SBSs. Users download files from a library of F files. The MBS has access to the library through a backhaul link. Some files are also cached at SBSs using a $(5, 3)$ MDS code. User A retrieves a cached file from the three SBSs within range. User B retrieves a fraction $2/3$ of a cached file from the two SBSs within range and the remaining fraction from the MBS.

tially served by the SBSs within communication range. We denote by γ_b the probability that the user is served by b SBSs and define $\gamma = (\gamma_0, \ldots, \gamma_{N_{\mathsf{SBS}}})$. If the user is not able to completely retrieve $\boldsymbol{X}^{(i)}$ from the SBSs, the additional required symbols are fetched from the MBS. Using the terminology in [112], the average fraction of files that are downloaded from the MBS is referred to as the backhaul rate, denoted by R, and defined as

$$\mathsf{R} \triangleq \frac{\text{average no. of bits downloaded from the MBS}}{\beta L}.$$

Note that for the case of no caching $\mathsf{R} = 1$.

As in [112], we assume that the communication is error free.

6.2.3 Private Information Retrieval and Problem Formulation

We assume that some of the SBSs are spy nodes that (potentially) collaborate with each other. On the other hand, we assume that the MBS can be trusted. The users wish to retrieve files from the cellular network, but do not want the spy nodes to learn any information about which file is requested by the user. The goal is to retrieve data from the network privately while minimizing the use of the backhaul link, i.e., while minimizing R. Thus, the goal is to optimize the content placement μ to minimize R.

6.3 Private Information Retrieval Protocol

In this section, we present a PIR protocol for the caching scenario. The PIR protocol proposed here is an extension of Protocol 3 in [113] to the case of multiple code rates.[3]

Assume without loss of generality that the user wants to download file $\boldsymbol{X}^{(i)}$. To retrieve the file, the user generates $n \leq N_{\mathsf{SBS}}$ query matrices, $\boldsymbol{Q}^{(l)}$, $l = 1, \ldots, n$, where $\boldsymbol{Q}^{(1)}, \ldots, \boldsymbol{Q}^{(b)}$ are the queries sent to the b SBSs within visibility and the remaining $n - b$ queries $\boldsymbol{Q}^{(b+1)}, \ldots, \boldsymbol{Q}^{(n)}$ are sent to the MBS. Note that n is a parameter that needs to be optimized. Each query matrix is of size $d \times \beta F$ symbols (from $\mathrm{GF}(q)$) and has the following structure,

$$\boldsymbol{Q}^{(l)} = \begin{pmatrix} \boldsymbol{q}_1^{(l)} \\ \boldsymbol{q}_2^{(l)} \\ \vdots \\ \boldsymbol{q}_d^{(l)} \end{pmatrix} = \begin{pmatrix} q_{1,1}^{(l)} & q_{1,2}^{(l)} & \cdots & q_{1,\beta F}^{(l)} \\ q_{2,1}^{(l)} & q_{2,2}^{(l)} & \cdots & q_{2,\beta F}^{(l)} \\ \vdots & \vdots & \cdots & \vdots \\ q_{d,1}^{(l)} & q_{d,2}^{(l)} & \cdots & q_{d,\beta F}^{(l)} \end{pmatrix}.$$

The query matrix $\boldsymbol{Q}^{(l)}$ consists of d subqueries $\boldsymbol{q}_j^{(l)}$, $j = 1, \ldots, d$, of length βF symbols each. In response to query matrix $\boldsymbol{Q}^{(l)}$, a SBS (or the MBS) sends back to the user a response vector $\boldsymbol{r}^{(l)} = (r_1^{(l)}, \ldots, r_d^{(l)})^\top$ of length d, computed as

$$\boldsymbol{r}^{(l)} = (r_1^{(l)}, \ldots, r_d^{(l)})^\top = \boldsymbol{Q}^{(l)} \big(c_{1,l}^{(1)}, \ldots, c_{\beta,l}^{(1)}, \ldots, c_{\beta,l}^{(F)}\big)^\top. \tag{6.1}$$

We will denote the j-th entry of the response vector $\boldsymbol{r}^{(l)}$, i.e., $r_j^{(l)}$, as the j-th subresponse of $\boldsymbol{r}^{(l)}$. Each response vector consists of d subresponses, each being a linear combination of βF symbols. Note that the operations are performed over the largest extension field, i.e., $\mathrm{GF}(q^{\delta_{\mathsf{max}}})$, and the subresponses are also over this field, i.e., each subresponse is of size $L/k_{\mathsf{min}} = L\mu_{\mathsf{max}}$ bits and hence each response is of size $dL\mu_{\mathsf{max}}$ bits.

The queries and the responses must be such that privacy is ensured and the user is able to recover the requested file. More precisely, information-theoretic PIR in the context of wireless caching with spy SBSs is defined as follows.

Definition 4 *Consider a wireless caching scenario with N_{SBS} SBSs that cache parts of a library of F files and in which a set $\mathcal{T}$ of T SBSs act as colluding spies. A user wishes to retrieve the i-th file and generates queries $\boldsymbol{Q}^{(l)}$, $l = 1, \ldots, n$. In response to the queries the SBSs and (potentially) the MBS send back the responses $\boldsymbol{r}^{(l)}$. This*

[3] Protocol 3 in [113] is based on and improves the protocol in [121], in the sense that it achieves higher PIR rates.

scheme achieves perfect information-theoretic PIR if and only if

$$\textit{Privacy:} \qquad \mathbb{H}\big(i|\boldsymbol{Q}^{(l)}, l \in \mathcal{T}\big) = \mathbb{H}(i); \tag{6.2a}$$

$$\textit{Recovery:} \qquad \mathbb{H}\big(\boldsymbol{X}^{(i)}|\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(n)}\big) = 0. \tag{6.2b}$$

Condition (6.2a) means that the spy SBSs gain no additional information about which file is requested from the queries (i.e., the uncertainty about the file requested after observing the queries is identical to the a priori uncertainty determined by the popularity distribution), while Condition (6.2b) guarantees that the user is able to recover the file from the n response vectors.

We define the (n, k_i) code $\mathcal{C}'_i$, $i = 1, \ldots, F$, as the code obtained by puncturing the underlying (N_{SBS}, k_i) storage code $\mathcal{C}_i$, and by $\mathcal{C}'_{\mathsf{max}}$ the code with parameters (n, k_{max}).[4] For the protocol to work, we require that k_{min} divides k_i for all i, i.e., $k_{\mathsf{min}} \mid k_i$. This ensures that $\mathrm{GF}(q^{\delta_i}) \subseteq \mathrm{GF}(q^{\delta_{\mathsf{max}}})$. Furthermore, we require the codes $\mathcal{C}'_i$ to be such that $\mathcal{C}'_i \subseteq \mathcal{C}'_{\mathsf{max}}$. The protocol is characterized by the codes $\{\mathcal{C}'_i\}$ and by two other codes, $\bar{\mathcal{C}}$ and $\tilde{\mathcal{C}}$. Code $\bar{\mathcal{C}}$ (over $\mathrm{GF}(q)$) has parameters $(n, \bar{k})$ and characterizes the queries sent to the SBSs and the MBS, while code $\tilde{\mathcal{C}}$ (defined below) defines the responses sent back to the user from the SBSs and the MBS. The designed protocol achieves PIR against a number of colluding SBSs $T \leq d_{\mathsf{min}}^{\bar{\mathcal{C}}^\perp} - 1$, where $d_{\mathsf{min}}^{\bar{\mathcal{C}}^\perp}$ is the minimum Hamming distance of the dual code of $\bar{\mathcal{C}}$.

6.3.1 Query Construction

The queries must be constructed such that privacy is preserved and the user can retrieve the requested file from the n response vectors $\boldsymbol{r}^{(l)}$, $l = 1, \ldots, n$. In particular, the protocol is designed such that the subresponses $r_j^{(l)}$, $l = 1, \ldots, n$, corresponding to the n subqueries $\boldsymbol{q}_j^{(1)}, \ldots, \boldsymbol{q}_j^{(n)}$ recover Γ unique code symbols of the file $\boldsymbol{X}^{(i)}$.

The queries are constructed as follows. The user chooses βF codewords $\bar{\boldsymbol{c}}_m^{(i)} = (c_{m,1}^{(i)}, \ldots, c_{m,n}^{(i)}) \in \bar{\mathcal{C}}$, $m = 1, \ldots, \beta$, $i = 1, \ldots, F$, independently and uniformly at random. Then, the user constructs n vectors,

$$\mathring{\boldsymbol{c}}_l = (\mathring{\boldsymbol{c}}_l^{(1)}, \ldots, \mathring{\boldsymbol{c}}_l^{(F)}), \quad l = 1, \ldots, n, \tag{6.3}$$

where $\mathring{\boldsymbol{c}}_l^{(i)}$ collects the l-th coordinates of the β codewords $\bar{\boldsymbol{c}}_m^{(i)}$, $m = 1, \ldots, \beta$, i.e., $\mathring{\boldsymbol{c}}_l^{(i)} = (\bar{c}_{1,l}^{(i)}, \ldots, \bar{c}_{\beta,l}^{(i)})$.

Assume that the user wants to retrieve file $\boldsymbol{X}^{(i)}$. Then, subquery $\boldsymbol{q}_j^{(l)}$ is constructed as

$$\boldsymbol{q}_j^{(l)} = \mathring{\boldsymbol{c}}_l + \boldsymbol{\delta}_j^{(l)}, \tag{6.4}$$

[4] Without loss of generality, to simplify notation we assume that the last coordinates of the code are punctured.

where

$$\boldsymbol{\delta}_j^{(l)} = \begin{cases} \boldsymbol{\omega}_{\beta(i-1)+s_j^{(l)}} & \text{if } l \in \mathcal{J}_j, \\ \boldsymbol{\omega}_0 & \text{otherwise,} \end{cases} \tag{6.5}$$

for some set $\mathcal{J}_j$ that will be defined below. Vector $\boldsymbol{\omega}_t$, $t = 1, \ldots, \beta F$, denotes the t-th (βF)-dimensional unit vector, i.e., the length-βF vector with a one in the t-th coordinate and zeroes in all other coordinates, and $\boldsymbol{\omega}_0$ the all-zero vector. The meaning of index $s_j^{(l)}$ will become apparent later.

According to (6.4), each subquery vector is the sum of two vectors, $\mathring{\boldsymbol{c}}_l$ and $\boldsymbol{\delta}_j^{(l)}$. The purpose of $\mathring{\boldsymbol{c}}_l$ is to make the subquery appear random and thus ensure privacy (i.e., Condition (6.2a)). On the other hand, the vectors $\boldsymbol{\delta}_j^{(l)}$ are deterministic vectors which must be properly constructed such that the user is able to retrieve the requested file from the response vectors (i.e., Condition (6.2b)). Similar to Protocol 3 in [113], the vectors $\boldsymbol{\delta}_j^{(l)}$ are constructed from a $d \times n$ binary matrix $\hat{\boldsymbol{E}}$ where each row represents a weight-Γ erasure pattern that is correctable by $\tilde{\mathcal{C}}$ and where the weights of its columns are determined from β information sets $\mathcal{I}_m$, $m = 1, \ldots, \beta$, of $\mathcal{C}'_{\mathsf{max}}$.

The construction of $\hat{\boldsymbol{E}}$ is addressed below. We define the set $\mathcal{F}_l$ as the index set of information sets $\mathcal{I}_m$ that contain the l-th coordinate of $\mathcal{C}'_{\mathsf{max}}$, i.e., $\mathcal{F}_l = \{m : l \in \mathcal{I}_m\}$. To allow the user to recover the requested file from the response vectors, $\hat{\boldsymbol{E}}$ is constructed such that it satisfies the following conditions.

C1. The user should be able to recover Γ unique code symbols of the requested file $\boldsymbol{X}^{(i)}$ from the responses to each set of n subqueries $\boldsymbol{q}_j^{(l)}$, $l = 1, \ldots, n$. This is to say that each row of $\hat{\boldsymbol{E}}$ should have exactly Γ ones. We denote by $\mathcal{J}_j$ the support of the j-th row of $\hat{\boldsymbol{E}}$.

C2. The user should be able to recover $\Gamma d \geq \beta k_i$ unique code symbols of the requested file $\boldsymbol{X}^{(i)}$, at least k_i symbols from each stripe. This means that each row $\hat{\boldsymbol{e}}_j = (\hat{e}_{j,1}, \ldots, \hat{e}_{j,n})$, $j = 1, \ldots, d$, of $\hat{\boldsymbol{E}}$ should correspond to an erasure pattern that is correctable by $\tilde{\mathcal{C}}$.

C3. Let $\boldsymbol{t}_l$, $l = 1, \ldots, n$, be the l-th column vector of $\hat{\boldsymbol{E}}$. The protocol should be able to recover $w_{\mathsf{H}}(\boldsymbol{t}_l)$ unique code symbols from the l-th response vector, which means that it is required that $w_{\mathsf{H}}(\boldsymbol{t}_l) = |\mathcal{F}_l|$. We call the vector $(w_{\mathsf{H}}(\boldsymbol{t}_1), \ldots, w_{\mathsf{H}}(\boldsymbol{t}_n))$ the *column weight profile of* $\hat{\boldsymbol{E}}$.

Finally, from $\hat{\boldsymbol{E}}$ we construct the vectors $\boldsymbol{\delta}_j^{(l)}$ in (6.5). In particular, index $s_j^{(l)}$ in (6.5) is such that $s_j^{(l)} \in \mathcal{F}_l$ and $s_j^{(l)} \neq s_{j'}^{(l)}$ for $j \neq j'$, $j, j' = 1, \ldots, d$.

6.3.2 Response Vectors

The j-th subresponse corresponding to subquery $\boldsymbol{q}_j^{(l)}$, $j = 1, \ldots, d$, is (see (6.1))

$$r_j^{(l)} = \langle \boldsymbol{q}_j^{(l)}, (c_{1,l}^{(1)}, \ldots, c_{\beta,l}^{(F)}) \rangle.$$

The user collects the n subresponses $r_j^{(l)}$, $l = 1, \ldots n$, in the vector $\boldsymbol{\rho}_j$,

$$\boldsymbol{\rho}_j = \begin{pmatrix} r_j^{(1)} \\ r_j^{(2)} \\ \vdots \\ r_j^{(n)} \end{pmatrix} = \sum_{m=1}^{\beta} \underbrace{\begin{pmatrix} \bar{c}_{m,1}^{(1)} c_{m,1}^{(1)} \\ \bar{c}_{m,2}^{(1)} c_{m,2}^{(1)} \\ \vdots \\ \bar{c}_{m,n}^{(1)} c_{m,n}^{(1)} \end{pmatrix}}_{\in \left\{ \boldsymbol{x} \in (\mathrm{GF}(q^{\delta_{\max}}))^n : \boldsymbol{H}^{\mathcal{C}'_1 \circ \bar{\mathcal{C}}} \boldsymbol{x} = \boldsymbol{0} \right\}} + \underbrace{\begin{pmatrix} \bar{c}_{m,1}^{(2)} c_{m,1}^{(2)} \\ \bar{c}_{m,2}^{(2)} c_{m,2}^{(2)} \\ \vdots \\ \bar{c}_{m,n}^{(2)} c_{m,n}^{(2)} \end{pmatrix}}_{\in \left\{ \boldsymbol{x} \in (\mathrm{GF}(q^{\delta_{\max}}))^n : \boldsymbol{H}^{\mathcal{C}'_2 \circ \bar{\mathcal{C}}} \boldsymbol{x} = \boldsymbol{0} \right\}}$$

$$+ \cdots + \underbrace{\begin{pmatrix} \bar{c}_{m,1}^{(F)} c_{m,1}^{(F)} \\ \bar{c}_{m,2}^{(F)} c_{m,2}^{(F)} \\ \vdots \\ \bar{c}_{m,n}^{(F)} c_{m,n}^{(F)} \end{pmatrix}}_{\in \left\{ \boldsymbol{x} \in (\mathrm{GF}(q^{\delta_{\max}}))^n : \boldsymbol{H}^{\mathcal{C}'_{\max} \circ \bar{\mathcal{C}}} \boldsymbol{x} = \boldsymbol{0} \right\}} + \begin{pmatrix} o_j^{(1)} \\ o_j^{(2)} \\ \vdots \\ o_j^{(n)} \end{pmatrix}, \tag{6.6}$$

where symbol $o_j^{(l)}$ represents the code symbol from file $\boldsymbol{X}^{(i)}$ downloaded in the j-th subresponse from the l-th response vector. Due to the structure of the queries obtained from $\hat{\boldsymbol{E}}$, the user retrieves Γ code symbols from the set of n subresponses to the j-th subqueries. Consider a retrieval code $\tilde{\mathcal{C}}$ of the form

$$\tilde{\mathcal{C}} = \sum_{i=1}^{F} \mathcal{C}'_i \circ \bar{\mathcal{C}} \overset{(a)}{=} \left(\sum_{i=1}^{F} \mathcal{C}'_i \right) \circ \bar{\mathcal{C}}, \tag{6.7}$$

where $\mathcal{C}'_i + \mathcal{C}'_j$ denotes the sum of subspaces $\mathcal{C}'_i$ and $\mathcal{C}'_j$, resulting in the set consisting of all elements $\boldsymbol{c} + \boldsymbol{c}'$ for any $\boldsymbol{c} \in \mathcal{C}'_i$ and $\boldsymbol{c}' \in \mathcal{C}'_j$, and where (a) follows due to the fact that the Hadamard product is distributive over addition.

The symbols requested by the user are then obtained solving the system of linear

equations defined by

$$\boldsymbol{H}^{\tilde{\mathcal{C}}}\boldsymbol{\rho}_j = \boldsymbol{H}^{\tilde{\mathcal{C}}} \begin{pmatrix} o_j^{(1)} \\ o_j^{(2)} \\ \vdots \\ o_j^{(n)} \end{pmatrix}.$$

6.3.3 Privacy

For the retrieval, we require $\tilde{\mathcal{C}}$ to be a valid code, i.e., it must have a code rate strictly less than 1. For a given number of colluding SBSs T, the combination of conditions on $\bar{\mathcal{C}}$ and $\tilde{\mathcal{C}}$ restricts the choice for the underlying storage codes $\{\mathcal{C}_i\}$. In the following theorem, we present a family of MDS codes, namely generalized Reed-Solomon (GRS) codes [133, Ch. 5], that work with the protocol.

Lemma 2 *Given an (n, k_{max}) GRS code $\mathcal{C}_{\mathsf{max}}$, for all $k < k_{\mathsf{max}}$, there exists an (n, k) GRS code that is a subcode of $\mathcal{C}_{\mathsf{max}}$.*

Proof. GRS codes of length n and dimension k_{max} over a finite field $\mathrm{GF}(q)$ are weighted evaluation codes with a weighting vector $\boldsymbol{v} = (v_1, \ldots, v_n) \in (\mathrm{GF}(q)^{\times})^n$ [133, Ch. 5]. Let $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_n) \in (\mathrm{GF}(q)^{\times})^n$ satisfy $\kappa_i \neq \kappa_j$ for all $i \neq j$. The canonical generator matrix for an (n, k_{max}) GRS code $\mathcal{C}_{\mathsf{max}}$ with weighting vector $\boldsymbol{v} \in (\mathrm{GF}(q)^{\times})^n$ is given by

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \kappa_1 & \kappa_2 & \ldots & \kappa_n \\ \vdots & \vdots & \ldots & \vdots \\ \kappa_1^{k_{\mathsf{max}}-1} & \kappa_2^{k_{\mathsf{max}}-1} & \ldots & \kappa_n^{k_{\mathsf{max}}-1} \end{pmatrix} \begin{pmatrix} v_1 & 0 & \ldots & 0 \\ 0 & v_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & v_n \end{pmatrix}. \tag{6.8}$$

Clearly, taking the first k rows of the leftmost matrix of (6.8) and multiplying it with the rightmost diagonal matrix generates an (n, k) subcode of $\mathcal{C}_{\mathsf{max}}$ which by itself is a GRS code with the same weighting vector $\boldsymbol{v}$. Thus, GRS codes are naturally nested, and the result follows. ■

Theorem 3 *Let $\mathcal{C}^{\mu}_{\mathsf{MDS}}$ be a caching scheme with GRS codes $\{\mathcal{C}_i\}$ of parameters (N_{SBS}, k_i) and let $\mathcal{C}'_i$ be the (n, k_i) code obtained by puncturing $\mathcal{C}_i$. Also, let $\bar{\mathcal{C}}$ be an (n, T) GRS code. Then, for $\beta = \Gamma = n - (k_{\mathsf{max}} + T - 1)$ and $d = k_{\mathsf{max}}$, the protocol achieves PIR against T colluding SBSs.*

Proof. The proof is given in the Appendix 8. ■

We remark that, with some slight modifications, the proposed protocol can be adapted to work with non-MDS codes.

6.3.4 Example

As an example, consider the case of $F = 2$ files, $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$, both of size βL bits. The first file $\boldsymbol{X}^{(1)}$ is stored in the SBSs according to 6.2 using an $(N_{\mathsf{SBS}} = 6, k_1 = 1)$ binary repetition code $\mathcal{C}_1$. Similarly, the second file $\boldsymbol{X}^{(2)}$ is stored (again according to 6.2) using an $(N_{\mathsf{SBS}} = 6, k_2 = 5)$ binary single parity-check code $\mathcal{C}_2$. Assume $n = N_{\mathsf{SBS}} = 6$ (i.e., no puncturing) and that none of the SBSs collude, i.e., $T = 1$. Furthermore, we assume that the user wants to retrieve $\boldsymbol{X}^{(1)}$ and is able to contact $b = n = 6$ SBSs (i.e., we consider the extreme case where the user is not contacting the MBS). According to 3, we can choose $\beta = \Gamma = n - (k_{\mathsf{max}} + T - 1) = 6 - (5 + 1 - 1) = 1$ and $d = k_{\mathsf{max}} = 5$. Finally, we choose $\bar{\mathcal{C}}$ as an $(n = 6, T = 1)$ binary repetition code.

According to (6.7), the retrieval code $\tilde{\mathcal{C}} = (\mathcal{C}_1 + \mathcal{C}_2) \circ \bar{\mathcal{C}} = \mathcal{C}_1 + \mathcal{C}_2 = \mathcal{C}_2$ and can be generated by

$$\boldsymbol{G}^{\tilde{\mathcal{C}}} = \boldsymbol{G}^{\mathcal{C}_2} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Moreover, let

$$\hat{\boldsymbol{E}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{I}_1 = \{1, 2, 3, 4, 5\},$$

where $\mathcal{I}_1$ is an information set of $\mathcal{C}_{\mathsf{max}} = \mathcal{C}_2$ (the submatrix $\boldsymbol{G}_{\mathcal{I}_1}^{\mathcal{C}_2}$ has rank $k_2 = 5$). Note that $\hat{\boldsymbol{E}}$ satisfies all three conditions C1–C3 and has column weight profile $(1, 1, 1, 1, 1, 0) = (|\mathcal{F}_1|, \ldots, |\mathcal{F}_6|)$.

Query Construction. The user generates $\beta F = 2$ codewords $\bar{\boldsymbol{c}}_1^{(1)}$ and $\bar{\boldsymbol{c}}_1^{(2)}$ independently and uniformly at random from $\bar{\mathcal{C}}$. Without loss of generality, let $\bar{\boldsymbol{c}}_1^{(1)} = \bar{\boldsymbol{c}}_1^{(2)} = (1, \ldots, 1)$. Next, the $n = 6$ subqueries $q_1^{(l)}$, $l = 1, \ldots, 6$, are constructed according to (6.4), (6.5) as

$$\boldsymbol{q}_1^{(l)} = \begin{cases} \mathring{\boldsymbol{c}}_l + (1, 0) & \text{if } l = 1, \\ \mathring{\boldsymbol{c}}_l + (0, 0) & \text{otherwise,} \end{cases}$$

where $\mathring{\boldsymbol{c}}_l$ is defined in (6.3).

File Retrieval. Consider the $n = 6$ subresponses $r_1^{(l)}$, $l = 1, \ldots, 6$. Then, according

to (6.6),

$$\boldsymbol{\rho}_1 = \begin{pmatrix} r_1^{(1)} \\ r_1^{(2)} \\ r_1^{(3)} \\ r_1^{(4)} \\ r_1^{(5)} \\ r_1^{(6)} \end{pmatrix} = \underbrace{\begin{pmatrix} \bar{c}_{1,1}^{(1)} c_{1,1}^{(1)} \\ \bar{c}_{1,2}^{(1)} c_{1,2}^{(1)} \\ \vdots \\ \bar{c}_{1,6}^{(1)} c_{1,6}^{(1)} \end{pmatrix}}_{\in \left\{ \boldsymbol{x} \in (\mathrm{GF}(2^5))^n : \boldsymbol{H}^{\mathcal{C}_1' \circ \bar{\mathcal{C}}} \boldsymbol{x} = \boldsymbol{0} \right\}} + \underbrace{\begin{pmatrix} \bar{c}_{1,1}^{(2)} c_{1,1}^{(2)} \\ \bar{c}_{1,2}^{(2)} c_{1,2}^{(2)} \\ \vdots \\ \bar{c}_{1,6}^{(2)} c_{1,6}^{(2)} \end{pmatrix}}_{\in \left\{ \boldsymbol{x} \in (\mathrm{GF}(2^5))^n : \boldsymbol{H}^{\mathcal{C}_2' \circ \bar{\mathcal{C}}} \boldsymbol{x} = \boldsymbol{0} \right\}} + \begin{pmatrix} o_1^{(1)} \\ o_1^{(2)} \\ \vdots \\ o_1^{(6)} \end{pmatrix}$$

$$= \begin{pmatrix} x_{1,1}^{(1)} \\ x_{1,1}^{(1)} \\ x_{1,1}^{(1)} \\ x_{1,1}^{(1)} \\ x_{1,1}^{(1)} \\ x_{1,1}^{(1)} \end{pmatrix} + \begin{pmatrix} x_{1,1}^{(2)} \\ x_{1,2}^{(2)} \\ x_{1,3}^{(2)} \\ x_{1,4}^{(2)} \\ x_{1,5}^{(2)} \\ \sum_{l=1}^{5} x_{1,l}^{(2)} \end{pmatrix} + \begin{pmatrix} x_{1,1}^{(1)} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

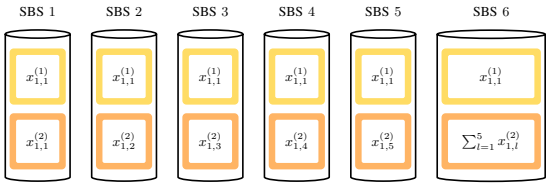


Figure 6.2: Wireless caching scenario in which there are $N_{\mathsf{SBS}} = 6$ SBSs. The SBSs store $F = 2$ files, $\boldsymbol{X}^{(1)} = (x_{1,1}^{(1)}) \in \mathrm{GF}(2^5)^{1\times 1}$ and $\boldsymbol{X}^{(2)} = (x_{1,1}^{(2)}, x_{1,2}^{(2)}, x_{1,3}^{(2)}, x_{1,4}^{(2)}, x_{1,5}^{(2)}) \in \mathrm{GF}(2)^{1\times 5}$, of $\beta L = 5$ bits each. The first file $\boldsymbol{X}^{(1)}$ is encoded using an $(N_{\mathsf{SBS}} = 6, k_1 = 1)$ binary repetition code $\mathcal{C}_1$, while the second file $\boldsymbol{X}^{(2)}$ is encoded using an $(N_{\mathsf{SBS}} = 6, k_2 = 5)$ binary single parity-check code $\mathcal{C}_2$.

and the code symbol $x_{1,1}^{(1)}$ of the file $\boldsymbol{X}^{(1)}$ is recovered from

$$\boldsymbol{H}^{\tilde{\mathcal{C}}}\boldsymbol{\rho}_1 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_{1,1}^{(1)} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = x_{1,1}^{(1)}.$$

Note that in order to retain privacy across the two files of the library, we need to send $d = k_{\mathsf{max}} = 5$ subqueries to each SBS, thus generating 5 subresponses from each SBS (even if the first file can be recovered from the $n = 6$ subresponses $r_1^{(l)}, l = 1, \ldots, 6$).

6.4 Backhaul Rate Analysis: No PIR Case

In this section, we derive the backhaul rate for the proposed caching scheme for the case of no PIR, i.e., the conventional caching scenario where PIR is not required.

Proposition 1 *The average backhaul rate for the caching scheme $\mathcal{C}_{\mathsf{MDS}}^{\mu}$ in Section 6.2*

for the case of no PIR is

$$\mathsf{R}_{\mathsf{noPIR}} = \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \max\left(0, 1/\mu_i - b\right) \mu_i + \sum_{i=1}^{F} p_i \lfloor 1 - \mu_i \rfloor. \tag{6.9}$$

Proof. To download file $\boldsymbol{X}^{(i)}$, if the user is in communication range of a number of SBSs, b, larger than or equal to $1/\mu_i$, the user can retrieve the file from the SBSs and there is no contribution to the backhaul rate. Otherwise, if $b < 1/\mu_i$, the user retrieves a fraction $L/k_i = L\mu_i$ of the file from each of the b SBSs, i.e., a total of $b\beta L\mu_i$ bits, and downloads the remaining $(1/\mu_i - b)\beta L\mu_i$ bits from the MBS. Averaging over $\boldsymbol{\gamma}$ and $\boldsymbol{p}$ (for the files cached) and normalizing by the file size βL, the contribution to the backhaul rate of the retrieval of files that are cached in the SBSs is

$$\sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \max\left(0, 1/\mu_i - b\right) \mu_i. \tag{6.10}$$

On the other hand, the files that are not cached are retrieved completely from the MBS, and their contribution to the backhaul rate is

$$\sum_{i=1}^{F} p_i \lfloor 1 - \mu_i \rfloor. \tag{6.11}$$

Combining (6.10) and (6.11) completes the proof. ■

We denote by $\mathsf{R}^*_{\mathsf{noPIR}}$ the maximum PIR rate resulting from the optimization of the content placement. $\mathsf{R}^*_{\mathsf{noPIR}}$ can be obtained solving the following optimization problem,

$$\begin{aligned} \mathsf{R}^*_{\mathsf{noPIR}} = \min_{\mu_i \in \mathcal{M}'} \; & \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \max\left(0, 1/\mu_i - b\right) \mu_i \\ & + \sum_{i=1}^{F} p_i \lfloor 1 - \mu_i \rfloor \\ \text{s.t.} \; & \sum_{i=1}^{F} \mu_i \leq M, \end{aligned}$$

where $\mathcal{M}' = \mathcal{M} \cup \{1/N_{\mathsf{SBS}}\}$, as $\mu_i = 1/N_{\mathsf{SBS}}$ is a valid value for the case where PIR is not required.

In the following lemma, we show that the proposed content placement is equivalent to the one in [112], in the sense that it yields the same average backhaul rate.

Lemma 3 *The average backhaul rate given by* (6.9) *for the caching scheme* $\mathcal{C}^{\mu}_{\mathsf{MDS}}$ *in*

Section 6.2 is equal to the one given by the caching scheme in [112], i.e., the two content placements are equivalent.

Proof. We can rewrite (6.9) using simple math as

$$\begin{aligned}
&\mathsf{R}_{\mathsf{noPIR}} \\
&= \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \max\big(0, 1/\mu_i - b\big)\mu_i + \sum_{i=1}^{F} p_i \lfloor 1-\mu_i \rfloor \\
&= \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \max\big(0, 1 - b\mu_i\big) + \sum_{i=1}^{F} p_i \lfloor 1-\mu_i \rfloor \\
&= \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \big(1 - \min\big(1, b\mu_i\big)\big) + \sum_{i=1}^{F} p_i \lfloor 1-\mu_i \rfloor \\
&\overset{(a)}{=} \sum_{i=1}^{F} p_i (\lceil \mu_i \rceil + \lfloor 1-\mu_i \rfloor) \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \big(1 - \min\big(1, b\mu_i\big)\big) \\
&= \sum_{i=1}^{F} p_i \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \big(1 - \min\big(1, b\mu_i\big)\big),
\end{aligned}$$

which is the expression in [112, eq. (1)]. (a) follows from the fact that we can write $p_i \lfloor 1-\mu_i \rfloor$ as $p_i \lfloor 1-\mu_i \rfloor \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \big(1-\min\big(1, b\mu_i\big)\big)$. For $0 < \mu_i \leq 1$ both expressions are zero, while for $\mu_i = 0$ both expressions boil down to p_i as $p_i \lfloor 1-\mu_i \rfloor \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b \big(1-\min\big(1, b\mu_i\big)\big) = p_i \sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b$ and $\sum_{b=0}^{N_{\mathsf{SBS}}} \gamma_b = 1$. ■

For popular content placement, i.e., the case where the M most popular files are cached in all SBSs (this corresponds to caching the M most popular files using an $(N_{\mathsf{SBS}}, 1)$ repetition code, i.e., $\mu_i = 1$ for $i \leq M$ and $\mu_i = 0$ for $i > M$), the backhaul rate is given by

$$\mathsf{R}_{\mathsf{noPIR}}^{\mathsf{pop}} = \gamma_0 \sum_{i=1}^{M} p_i + \sum_{i=M+1}^{F} p_i. \tag{6.12}$$

6.5 Backhaul Rate Analysis: PIR Case

In this section, we derive the backhaul rate for the case of PIR (i.e., when the user wishes to download content privately) and we prove that uniform content placement (under the PIR protocol in Section 6.3 with GRS codes) is optimal. The average backhaul rate is given in the following proposition.

Proposition 2 *The average backhaul rate for the caching scheme $\mathcal{C}_{\mathsf{MDS}}^{\mu}$ in Section 6.2*

(with GRS codes) for the PIR case is

$$\mathsf{R}_{\mathsf{PIR}} = \frac{\mu_{\mathsf{max}}}{\mu_{\min}(n-T+1)-1} \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{n} \gamma_b (n-b) + \sum_{i=1}^{F} p_i \lfloor 1-\mu_i \rfloor. \tag{6.13}$$

Proof. To download file $\boldsymbol{X}^{(i)}$, the user generates n query matrices. If the user is in communication range of b SBSs, it receives b responses (one from each SBS). The responses to the remaining $n-b$ query matrices need to be downloaded from the MBS. Since each response consists of d subresponses of size $L\mu_{\mathsf{max}}$ bits, the user downloads $(n-b)dL\mu_{\mathsf{max}}$ bits from the MBS. Averaging over $\boldsymbol{\gamma}$ and $\boldsymbol{p}$ (for the files cached) and normalizing by the file size βL, the contribution to the backhaul rate of the retrieval of files that are cached in the SBSs is

$$\frac{1}{\beta} \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{n} \gamma_b (n-b) d\mu_{\mathsf{max}}. \tag{6.14}$$

Now, using the fact that $\beta = \Gamma = n - (k_{\mathsf{max}} + T - 1) = \frac{\mu_{\mathsf{min}}(n-T+1)-1}{\mu_{\mathsf{min}}}$ and $d = k_{\mathsf{max}} = 1/\mu_{\mathsf{min}}$ (see Theorem 3), we can rewrite (6.14) as

$$\frac{\mu_{\mathsf{max}}}{\mu_{\min}(n-T+1)-1} \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{n} \gamma_b (n-b). \tag{6.15}$$

On the other hand, the files that are not cached are retrieved completely from the MBS, and their contribution to the backhaul rate is (as for the no PIR case)

$$\sum_{i=1}^{F} p_i \lfloor 1-\mu_i \rfloor. \tag{6.16}$$

Combining (6.15) and (6.16) completes the proof. ■

6.5.1 Optimal Content Placement

Let $\mathsf{R}^*_{\mathsf{PIR}}$ be the maximum PIR rate resulting from the optimization of the content placement. $\mathsf{R}^*_{\mathsf{PIR}}$ can be obtained solving the following optimization problem,

$$\begin{aligned}\mathsf{R}^*_{\mathsf{PIR}} = \min_{\substack{\mu_i \in \mathcal{M} \\ n \in \mathcal{A}}} & \frac{\mu_{\mathsf{max}}}{\mu_{\min}(n-T+1)-1} \sum_{i=1}^{F} p_i \lceil \mu_i \rceil \sum_{b=0}^{n} \gamma_b (n-b) \\ & + \sum_{i=1}^{F} p_i \lfloor 1-\mu_i \rfloor \\ \text{s.t.} & \sum_{i=1}^{F} \mu_i \leq M \text{ and } k_{\mathsf{min}} \mid k_i,\end{aligned} \tag{6.17}$$

where $\mathcal{A} = \{1/\mu_{\mathsf{min}} + T, \ldots, N_{\mathsf{SBS}}\}$ and the minimum value that n can take on, i.e., $1/\mu_{\mathsf{min}} + T$, comes from the fact that $\mu_{\min}(n-T+1)-1$ has to be positive.

Lemma 4 *Uniform content allocation, i.e., $\mu_i = \mu$ for all files that are cached, is optimal. Furthermore, the optimal number of files to cache is the maximum possible, i.e., $\mu_i = \mu$ for $i \leq \min(M/\mu, F)$.*

Proof. We first prove the first part of the lemma. We need to show that either the optimal solution to the optimization problem in (6.17) is the all-zero vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_F) = (0, \ldots, 0)$, or there exists a nonzero optimal solution $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_F)$ for which $\mu_{\mathsf{max}} = \mu_{\mathsf{min}}$. Consider the second case, and let $\boldsymbol{\mu}$ denote any nonzero feasible solution to (6.17), i.e., a nonzero solution that satisfies the cache size constraint. Furthermore, let $\boldsymbol{\mu}' = (\mu'_1, \ldots, \mu'_F)$ denote the length-F vector obtained from $\boldsymbol{\mu}$ as $\mu'_i = \mu_{\mathsf{min}}$ for $\mu_i \neq 0$ and $\mu'_i = 0$ otherwise. Clearly, $\boldsymbol{\mu}'$ satisfies the cache size constraint as well. Note that $\mu'_{\mathsf{max}} = \mu'_{\mathsf{min}} = \mu_{\mathsf{min}}$. Thus,

$$\begin{aligned}\frac{\mu'_{\mathsf{max}}}{\mu'_{\mathsf{min}}(n-T+1)-1} &= \frac{\mu_{\mathsf{min}}}{\mu_{\mathsf{min}}(n-T+1)-1} \\ &\leq \frac{\mu_{\mathsf{max}}}{\mu_{\mathsf{min}}(n-T+1)-1}.\end{aligned}$$

Furthermore, since both the double summation in the first term of the objective function in (6.17) and the second term in (6.17) only depend on the support of $\boldsymbol{\mu}$, it follows that the value of the objective function for $\boldsymbol{\mu}'$ is smaller than or equal to the value of the objective function for $\boldsymbol{\mu}$. Thus, for any nonzero feasible solution $\boldsymbol{\mu}$ there exists another at least as good nonzero feasible solution $\boldsymbol{\mu}'$ for which all nonzero entries are the same (i.e., $\mu'_{\mathsf{min}} = \mu'_{\mathsf{max}} = \mu$), and the result follows by applying the above procedure to a (nonzero) optimal solution to (6.17).

We now prove the second part of the lemma. Caching a file helps in reducing the

backhaul rate if

$$\frac{\mu}{\mu(n-T+1)-1}\sum_{b=0}^{n}\gamma_b(n-b)<1, \tag{6.18}$$

for some $n \in \mathcal{A}$ and $\mu \in \mathcal{M}$. This is independent of the file index i. Thus, if the optimal solution is to cache at least one file ($\boldsymbol{\mu} \neq \mathbf{0}$), (6.18) is met for some $n \in \mathcal{A}$ and caching other files (as many files as permitted up to the cache size constraint, with decreasing order of popularity) is optimal as it further reduces the backhaul rate. ■

Following Lemma 4, the optimization problem in (6.17) can be rewritten as

$$\mathsf{R}^{*}_{\mathsf{PIR}} = \min_{\substack{\mu\in\mathcal{M}\\ n\in\mathcal{A}}} \frac{\mu}{\mu(n-T+1)-1} \sum_{i=1}^{\min(M/\mu,F)} p_i \sum_{b=0}^{n}\gamma_b(n-b) + \sum_{i=M/\mu+1}^{F} p_i. \tag{6.19}$$

6.5.2 Popular Content Placement

For popular content placement, the backhaul rate is given by

$$\mathsf{R}^{\mathsf{pop}}_{\mathsf{PIR}} = \min_{n\in\mathcal{A}} \frac{1}{n-T}\sum_{i=1}^{M} p_i \sum_{b=0}^{n}\gamma_b(n-b) + \sum_{i=M+1}^{F} p_i. \tag{6.20}$$

Note that the optimization over n is still required.

6.6 Weighted Communication Rate

So far, we have considered only the backhaul rate. However, it might also be desirable to limit the communication rate from SBSs to the user. We thus consider the weighted communication rate, $\mathsf{C}_{\mathsf{PIR}}$, defined as[5]

$$\mathsf{C}_{\mathsf{PIR}} = \mathsf{R}_{\mathsf{PIR}} + \theta \mathsf{D}_{\mathsf{PIR}},$$

where $\mathsf{D}_{\mathsf{PIR}}$ is the average communication rate (normalized by the file size βL) from the SBSs, and θ is a weighting parameter. We consider $\theta \leq 1$, stemming from the fact that the bottleneck is the backhaul. Note that minimizing the average backhaul rate corresponds to $\theta = 0$.

Proposition 3 *The average communication rate from the SBSs for the caching scheme*

[5]For the case of no PIR, a linear scalarization of the MBS and SBS download delays was considered in [134]. The communication rate is directly related to the download delay.

$\mathcal{C}^{\mu}_{\mathsf{MDS}}$ in Section 6.2 (with GRS codes) for the PIR case is

$$\mathsf{D}_{\mathsf{PIR}} = \frac{\mu_{\mathsf{max}}}{\mu_{\min}(n-T+1)-1} \sum_{b=0}^{n} \tilde{\gamma}_b b, \tag{6.21}$$

where $\tilde{\gamma}_b = \gamma_b$ for $b < n$ and $\tilde{\gamma}_n = \sum_{b=n}^{N_{\mathsf{SBS}}} \gamma_b$.

Proof. To ensure privacy, the user needs to download data from the SBSs within visibility regardless whether the requested file is cached or not. This is in contrast to the case of no PIR. Note that, if the user queries the SBSs only in the case the requested file is cached, then the spy SBSs would infer that the user is interested in one of the files cached, thus gaining some information about the file requested. In other words, the user sends dummy queries and downloads data that is useless for the retrieval of the file but is necessary to achieve privacy. The user receives b responses from the b SBSs within communication range, each of size $dL\mu_{\mathsf{max}}$ bits. Let $\tilde{\gamma}_b$ denote the probability to receive responses from b SBSs. For $b < n$, $\tilde{\gamma}_b$ is equal to the probability that b SBSs are within communication range, i.e., $\tilde{\gamma}_b = \gamma_b$. On the other hand, the probability to receive responses from n SBSs, $\tilde{\gamma}_n$, is the probability that at least n SBSs are within communication range, i.e., $\tilde{\gamma}_n = \sum_{b=n}^{N_{\mathsf{SBS}}} \gamma_b$. Averaging over $\tilde{\boldsymbol{\gamma}}$ and $\boldsymbol{p}$ (for all files, cached and not cached) and normalizing by the file size βL, the contribution to the communication rate of the retrieval of a file from the SBSs is

$$\frac{1}{\beta} \sum_{i=1}^{F} p_i \sum_{b=0}^{n} \tilde{\gamma}_b b d \mu_{\mathsf{max}}. \tag{6.22}$$

Now, using the fact that $\beta = \Gamma = n - (k_{\mathsf{max}} + T - 1) = \frac{\mu_{\mathsf{min}}(n-T+1)-1}{\mu_{\mathsf{min}}}$ and $d = k_{\mathsf{max}} = 1/\mu_{\mathsf{min}}$ (see Theorem 3), we can rewrite (6.22) as (6.21). ■

The corresponding optimization problem is

$$\begin{aligned} \mathsf{C}^{*}_{\mathsf{PIR}} = \min_{\substack{\mu_i \in \mathcal{M} \\ n \in \mathcal{A}}} \; & \mathsf{R}_{\mathsf{PIR}} + \theta \mathsf{D}_{\mathsf{PIR}} \\ \text{s.t.} \sum_{i=1}^{F} \mu_i \leq M \text{ and } & k_{\mathsf{min}} \mid k_i, \end{aligned} \tag{6.23}$$

where $\mathsf{R}_{\mathsf{PIR}}$ is given in (6.13).

Lemma 5 *Uniform content allocation, i.e., $\mu_i = \mu$ for all files that are cached, is optimal. Furthermore, the optimal number of files to cache is the maximum possible, i.e., $\mu_i = \mu$ for $i \leq \min(M/\mu, F)$.*

Proof. The proof of Lemma 4 applies to both terms in (6.23) and the result follows. ■

Following Lemma 5, the optimization problem in (6.23) can be rewritten as

$$\mathsf{C}^*_{\mathsf{PIR}} = \min_{\substack{\mu\in\mathcal{M}\\ n\in\mathcal{A}}} \frac{\mu}{\mu(n-T+1)-1} \sum_{i=1}^{\min(M/\mu,F)} p_i \sum_{b=0}^{n} \gamma_b (n-b) + \sum_{i=M/\mu+1}^{F} p_i + \theta \frac{\mu}{\mu(n-T+1)-1} \sum_{b=0}^{n} \tilde{\gamma}_b b. \tag{6.24}$$

6.7 Numerical Results

For the numerical results in this section, we assume that the files popularity distribution $\boldsymbol{p}$ follows the Zipf law [135], i.e., the popularity of file $\boldsymbol{X}^{(i)}$ is

$$p_i = \frac{1/i^{\alpha}}{\sum_{\ell} 1/\ell^{\alpha}},$$

where $\alpha \in [0.5, 1.5]$ is the skewness factor [112] and by definition $p_1 \geq p_2 \geq \ldots \geq p_F$. In Figs. 6.3 and 6.4, we consider a network topology where SBSs are deployed over a macro-cell of radius D meters according to a regular grid with distance d meters between them [112, 134]. Each SBS has a communication radius of r meters. Let $\mathcal{R}_b$ be the area where a user can be served by b SBSs. Then, assuming that the users are uniformly distributed over the macro-cell area with density ϕ users per square meter, the probability that a user is in communication range of b SBSs can be calculated as in [112]

$$\gamma_b = \frac{\phi \mathcal{R}_b}{\phi \sum_{a=1}^{N_{\mathsf{max}}} \mathcal{R}_a},$$

where the areas $\mathcal{R}_b$ can be easily obtained by simple geometrical evaluations, and N_{max} is the maximum number of SBSs within communication range of a user.

For the results in Figs. 6.3 and 6.4, the system parameters (taken from [112]) are $D = 500$ meters, which results in $N_{\mathsf{SBS}} = 316$ over the macro-cell area, $F = 200$ files, $\alpha = 0.7$, and $r = 60$ meters. This results in $\boldsymbol{\gamma} = (0, 0, 0.1736, 0.5113, 0.3151, 0, \ldots, 0)$, i.e., the maximum number of SBSs in visibility of a user is $N_{\mathsf{max}} = 4$.

In Fig. 6.3, we plot the optimized backhaul rate $\mathsf{R}^*_{\mathsf{PIR}}$ (red, solid lines) according to (6.19) as a function of the cache size constraint M for the noncolluding case ($T = 1$) and $T = 2$ and $T = 3$ colluding SBSs. The curves in Fig. 6.3 should be interpreted as the minimum backhaul rate that is necessary in order to achieve privacy against T spy SBSs out of the n SBSs that are contacted by the user. For the particular system parameters considered, the optimal value of n is 3 for $T = 1$ and $T = 2$, and all values of M, i.e., the scheme yields privacy against T spy SBSs out of the $n = 3$ SBSs contacted. For $T = 3$ the optimal value of n is 4 for all values of M, and thus the scheme yields privacy against 3 spy SBSs out of $n = 4$ SBSs. We also plot the

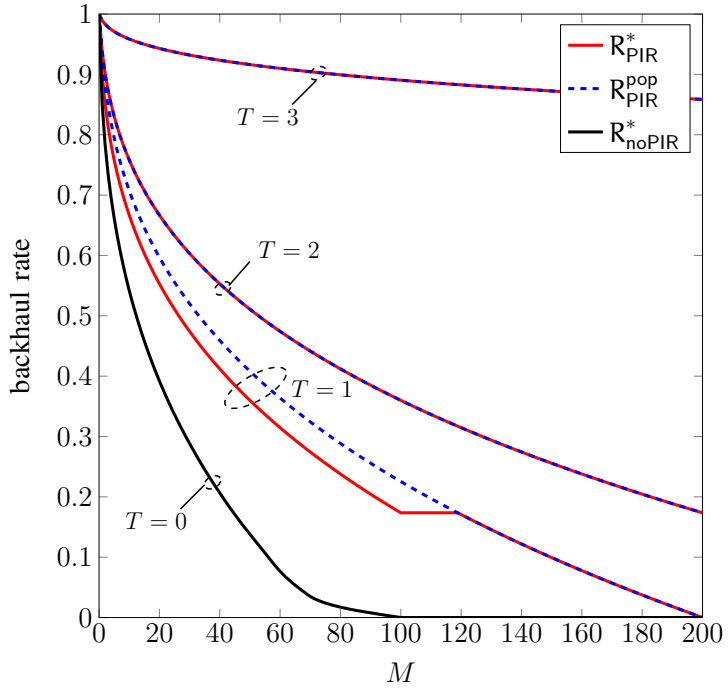


Figure 6.3: Backhaul rate as a function of the cache size constraint M for a system with $F = 200$ files, $N_{\mathsf{SBS}} = 316$, and $\alpha = 0.7$.

optimized backhaul rate $\mathsf{R}^*_{\mathsf{noPIR}}$ for the case of no PIR.[6] As can be seen in the figure, caching helps in significantly reducing the backhaul rate for $T = 1$ and $T = 2$. For $T = 3$ caching also helps in reducing the backhaul rate, but the reduction is smaller. Also, as expected, compared to the case of no PIR ($\mathsf{R}^*_{\mathsf{noPIR}}$, black, solid line) achieving privacy requires a higher backhaul rate. The required backhaul rate increases with the number of colluding SBSs T.

For $M \geq 100$ and no PIR, the backhaul rate is zero, as all files can be downloaded from the SBSs. Indeed, for $M = 100$, we can select $k_i = 2\ \forall i$ and cache one coded symbol from each stripe of each file in each SBS (thus satisfying the constraint $\sum_{i=1}^{F} \mu_i \leq M$ as $\sum_{i=1}^{200} \mu_i = \sum_{i=1}^{200} 1/k_i = \sum_{i=1}^{200} 0.5 = 100$). Since for no PIR to retrieve each stripe of a file it is enough to download 2 symbols from each stripe of the file (due to the MDS property) and according to $\boldsymbol{\gamma}$ at least 2 SBSs are within range, for $M = 100$ (and hence for $M > 100$ as well) the user can always retrieve the file from the SBSs and the backhaul rate is zero. For the case of PIR and $T = 1$, on the other hand, the required backhaul rate is positive unless all complete files can be cached in all SBSs, i.e., $M = F$. For $T = 2$ and $T = 3$, even for $M = F$ the

[6] The curve $\mathsf{R}^*_{\mathsf{noPIR}}$ in the figure is identical to that in [112, Fig. 4]. As proved in Lemma 3, while the proposed content placement is different from the one in [112], they are equivalent in terms of average backhaul rate.

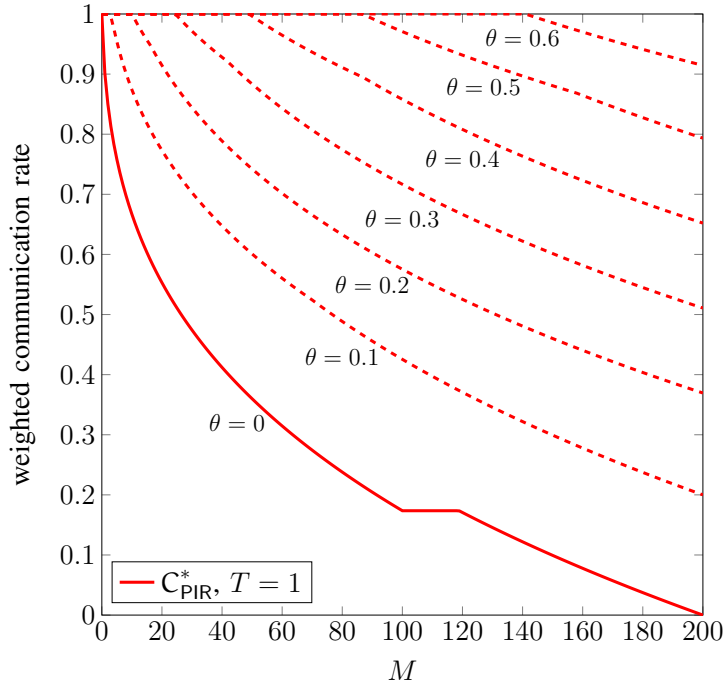


Figure 6.4: Optimized weighted communication rate as a function of the cache size constraint M for a system with $T = 1$ spy SBS, $F = 200$ files, $N_{\mathsf{SBS}} = 316$, $\alpha = 0.7$, and several values of θ.

backhaul rate is not zero. This is because in this case the user needs to receive $n = 3$ and $n = 4$ responses $\boldsymbol{r}^{(l)}$, $l = 1, \ldots, n$, respectively (from the SBSs or the MBS). However, for the considered system parameters the probability that the user has $b \geq 3$ SBSs within range is not one, thus the user always needs to download data from the MBS to recover the file and the backhaul rate is positive.

For comparison purposes, in the figure we also plot the backhaul rate for the case of popular content placement $\mathsf{R}^{\mathsf{pop}}_{\mathsf{PIR}}$ in (6.20) (blue, dashed lines). In this case, the optimal value of n is 2, 3, and 4 for $T = 1$, $T = 2$, and $T = 3$, respectively. We remark that the curve $\mathsf{R}^{\mathsf{pop}}_{\mathsf{PIR}}$ for $T = 1$ overlaps with the curve $\mathsf{R}^{\mathsf{pop}}_{\mathsf{noPIR}}$. This is due to the fact that for $T = 1$, $n = 2$, and $\gamma_0 = \gamma_1 = 0$, $\mathsf{R}^{\mathsf{pop}}_{\mathsf{PIR}}$ in (6.20) boils down to $\sum_{M+1}^{F} p_i$, which is $\mathsf{R}^{\mathsf{pop}}_{\mathsf{noPIR}}$ in (6.12). However, for the general case, i.e., other $\boldsymbol{\gamma}$, $\mathsf{R}^{\mathsf{pop}}_{\mathsf{PIR}}$ and $\mathsf{R}^{\mathsf{pop}}_{\mathsf{noPIR}}$ may differ. As already shown in [112], for no PIR the optimized content placement yields significantly lower backhaul rate than popular content placement. For the PIR case and $T = 1$, up to $M = 118$ the optimized content placement also yields some performance gains with respect to popular content placement, albeit not as significant as for the case of no PIR. Interestingly, as shown in the figure, for $M \geq 119$, PIR popular content placement is optimal. Furthermore, as shown in the figure, for $T = 2$

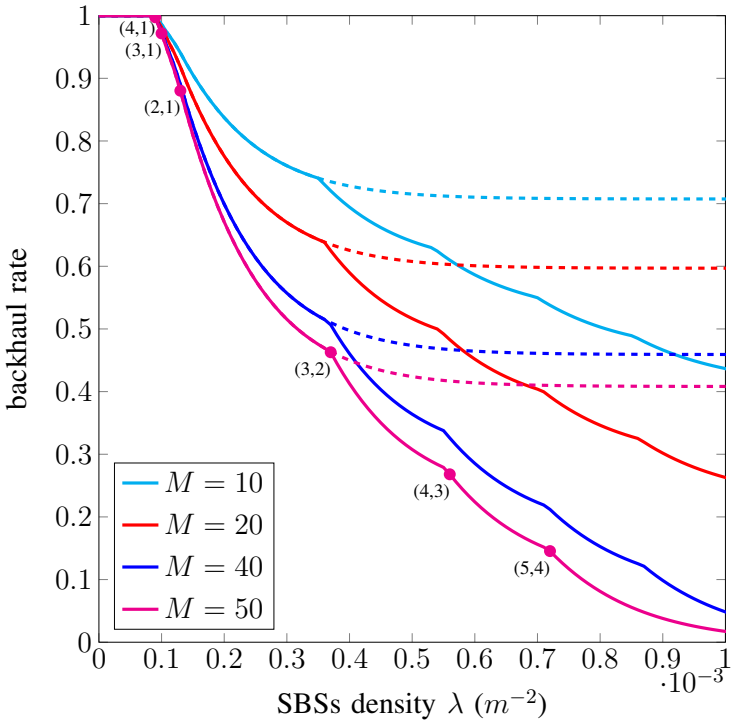


Figure 6.5: Backhaul rate as a function of the density of SBSs λ and several values M for the scenario where SBSs are distributed according to a PPP and $T = 1$. $F = 200$ files and $\alpha = 0.7$. Solid lines correspond to optimal content placement ($\mathsf{R}^{*}_{\mathsf{PIR}}$ in (6.19)) and dashed lines to popular content placement ($\mathsf{R}^{\mathsf{pop}}_{\mathsf{PIR}}$ in (6.20)).

and $T = 3$ popular content placement is optimal for all M.

In Fig. 6.4, we plot the optimized weighted communication rate $\mathsf{C}^{*}_{\mathsf{PIR}}$ in (6.24) for the noncolluding case ($T = 1$) as a function of the cache size constraint M and several values of θ. For the considered system parameters, caching is still useful for small values of θ if the cache size is big enough. For example, for $\theta = 0.5$ caching helps in reducing the weighted communication rate with respect to no caching for $M \geq 87$. For $\theta \geq 0.7$, caching does not bring any reduction of the weighted communication rate.

In Figs. 6.5 and 6.6, we plot the backhaul rate for a PPP deployment model where SBSs are distributed over the plane according to a PPP and a user at an arbitrary location in the plane can connect to all SBSs that are within radius r_{u}. Let λ be the density of SBSs per square meter. For this scenario, the probability that a user is in

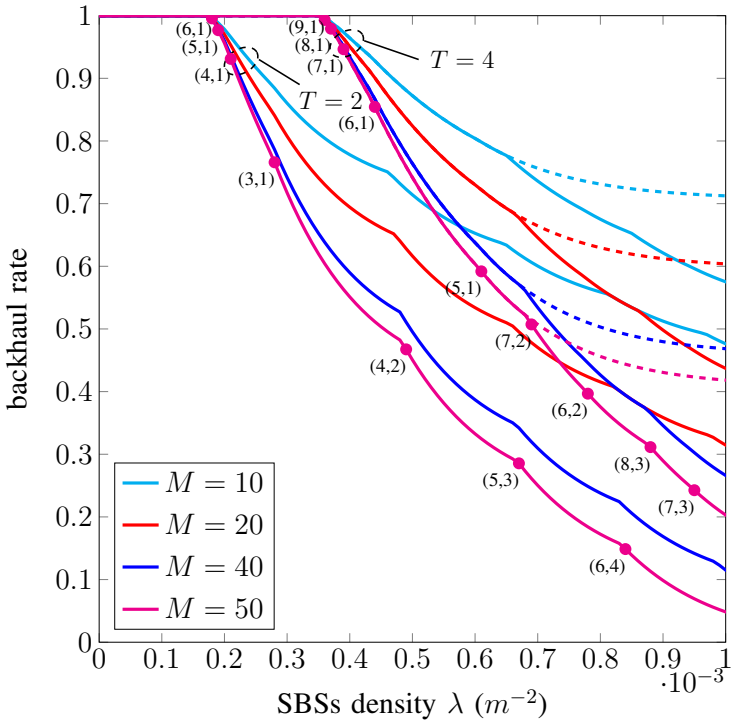


Figure 6.6: Backhaul rate as a function of the density of SBSs λ and several values of M for the scenario where SBSs are distributed according to a PPP and $T = 2$ and $T = 4$. $F = 200$ files and $\alpha = 0.7$. Solid lines correspond to optimal content placement ($\mathsf{R}^{*}_{\mathsf{PIR}}$ in (6.19)) and dashed lines to popular content placement ($\mathsf{R}^{\mathsf{pop}}_{\mathsf{PIR}}$ in (6.20)).

communication range of b SBSs is given by [136]

$$\gamma_b = \mathrm{e}^{-\psi} \frac{\psi^b}{b!},$$

where $\psi = \lambda \pi r_{\mathsf{u}}^2$. In Fig. 6.5, we plot the optimized backhaul rate ($\mathsf{R}^{*}_{\mathsf{PIR}}$ in (6.19), solid lines) as a function of the density λ for $F = 200$ files, $\alpha = 0.7$, $r_{\mathsf{u}} = 60$ meters, different cache size constraint M, and a single spy SBS, i.e., $T = 1$. For small densities, caching does not help in reducing the backhaul rate. However, as expected, the required backhaul rate diminishes by increasing the density of SBSs. For comparison purposes, we also plot the backhaul rate for popular content placement ($\mathsf{R}^{\mathsf{pop}}_{\mathsf{PIR}}$ in (6.20), dashed lines). Interestingly, popular content placement is optimal up to a given density of SBSs, after which optimizing the content placement brings a significant reduction of the required backhaul rate. Similar results are observed for $T = 2$ and $T = 4$ colluding SBSs in Fig. 6.6 with the same system parameters as in Fig. 6.5. In Figs. 6.5 and 6.6, for each M the optimal value of n and μ depends on

the density of SBSs. Typically, a pair (n, μ) is optimal for a range of densities. In the figures, we give the optimal values of n and k for $M = 50$ (in particular we give the pair (n, k), with $k = 1/\nu$, which is also the code parameters of the punctured code $\mathcal{C}'$). For convenience, in the figures we only give the parameters for the densities where the optimal pair (n, k) changes. The values should be read as follows: In Fig. 6.5, walking the curve from top-left to bottom-right, no caching is optimal for densities up to $\lambda = 8 \cdot 10^{-5}$. For $\lambda = 9 \cdot 10^{-5}$, $(4, 1)$ is optimal. Then, $(3, 1)$ is optimal for densities $\lambda = 10^{-4}$ to $\lambda = 1.2 \cdot 10^{-4}$. From $\lambda = 1.3 \cdot 10^{-4}$ to $\lambda = 3.2 \cdot 10^{-4}$ the optimal value is $(2, 1)$, and so on (the curves are plotted with steps of 10^{-5}).

6.8 Summary

In this chapter, we have proposed a private information retrieval scheme that allows to download files of different popularities from a cellular network, where to reduce the backhaul usage content is cached at the wireless edge in SBSs, while achieving privacy against a number of spy SBSs. We derived the backhaul rate for this scheme and formulated the content placement optimization. We showed that, as for the no PIR case, up to a number of spy SBSs caching helps in reducing the backhaul rate. Interestingly, contrary to the no PIR case, uniform content placement is optimal. Furthermore, popular content placement is optimal for some scenarios. Although uniform content placement is optimal, the proposed PIR scheme for multiple code rates may be useful in other scenarios, e.g., for distributed storage where data is stored using codes of different rates.

Chapter 7

Conclusions

In this thesis, we have studied the secrecy and privacy performances of practical systems, using an information-theoretic approach. Such an approach may have relevance especially in the Internet of Things era, where wireless devices have to fulfil strict resource constraints difficult to satisfy using classic cryptographic techniques.

We first have defined a set of tools to assess the secrecy performance of practical coding and modulation schemes used for transmissions over fading wiretap channels. We found the requirements in terms of Bob's and Eve's channels SNR to achieve a fixed level of mutual information security with practical codes, as well as to compare their performance with that achievable with optimal codes. We have shown that practical coding and modulation schemes (as the ones compliant with the IEEE WiMax standard) can achieve a good secrecy performance, comparable to optimal codes when high code rates and low order modulations are used, while low code rates and low order modulations help to reduce the SNR gap required between the legitimate receiver's and the eavesdropper's channels in order to achieve some given level of mutual information security. Moreover we have proposed an on-off transmission protocol based on coding and AONT to achieve some desired level of semantic security, which also exploits the possibility to send a fake packet when channel conditions are below a prefixed threshold. The application of our protocol with practical coding and modulation schemes permits to achieve satisfactory semantic security levels, even when the average quality of the eavesdropper channel is not worse than that of the main channel. Specifically, we observed that using high rate codes permits us to achieve the target security level with a smaller number of time slots with respect to low rate codes, while using high order modulation schemes seem not to be beneficial, although high order modulation schemes may be needed in order to ensure that the channel can be considered stationary during each single packet session.

Then we have analyzed what level of security is obtainable in a Gaussian relay parallel channel under finite-length coding and discrete constellation constraints. We have derived the secrecy rate, defining it as the maximum rate for which a minimum equivocation rate is achieved by the eavesdropper. Moreover, we have applied a coupled version of the Gale and Shapley algorithm to allocate power within each channel in order to maximize the secrecy rate. We have shown that moderate sizes of both the

constellation alphabet and the codewords are sufficient to achieve close-to-optimal secrecy rates for typical wireless transmission scenarios. By comparing our resource allocation strategy with uniform power allocation and water-filling allocation we have demonstrated that our solution is the most convenient from a security standpoint, while water-filling provides the best possible power allocation in the absence of the attacker.

Using an adapted version of the AONT-based protocol shown in the first chapters, we have presented a framework that uses joint computational and information-theoretic security notions to design and assess heterogeneous distributed storage systems based on data dispersal algorithms able to achieve a given level of security. Such system architectures represent a realistic setting which often appears in practice, and has been implemented in multiple real cloud storage systems. In our model we have considered a user that can read and write files in the storage nodes and a passive attacker that can steal individual slices from heterogeneous communication channels without compromising the storage nodes. By exploiting a probabilistic model checker we found the parameter values that optimize the security metrics of interest.

Concerning privacy, we have proposed a private information retrieval scheme that allows to download files of different popularities from a cellular network, at the presence of a number of spy nodes. The content is cached at the wireless edge in small-cell base stations in order to reduce the communication costs. We derived the backhaul rate for this scheme and formulated the content placement optimization. We showed that, as for the no PIR case, up to a number of spy small-cell base stations caching helps in reducing the backhaul rate. Interestingly, contrary to the no PIR case, uniform content placement is optimal. Furthermore, popular content placement is optimal for some scenarios. Although uniform content placement is optimal, the proposed PIR scheme for multiple code rates may be useful in other scenarios, e.g., for distributed storage where data is stored using codes of different rates. We also have considered the minimization of a weighted sum of the backhaul rate and the communication rate from the small-cell base stations, relevant for the case where limiting the communication costs from the small-cell base stations is also important.

Chapter 8

Appendix

8.1 Proof of Theorem3 of Chapter 6

To prove that the protocol described in 6.3 achieves PIR against T colluding SBSs, we need to prove that both the privacy condition in (6.2a) and the recovery condition in (6.2b) are satisfied. We first prove that the recovery condition in (6.2b) is satisfied.

According to Lemma 2 of Chapter 6, GRS codes are naturally nested. Furthermore, puncturing a GRS code results in another GRS code, since GRS codes are weighted evaluation codes [133, Ch. 5]. Thus, $\mathcal{C}'_i \subseteq \mathcal{C}'_{\mathsf{max}}$ for all i, and it follows from 6.7 that

$$\tilde{\mathcal{C}} = \left(\sum_{i=1}^{F} \mathcal{C}'_i \right) \circ \bar{\mathcal{C}} = \mathcal{C}'_{\mathsf{max}} \circ \bar{\mathcal{C}}.$$

Furthermore, it can easily be shown that the Hadamard product of two GRS codes is also a GRS code with dimension equal to the sum of the dimensions minus 1. Thus, $\tilde{\mathcal{C}}$ is a GRS code of dimension $k_{\mathsf{max}} + T - 1$. As $\tilde{\mathcal{C}}$ is an $(n, k_{\mathsf{max}} + T - 1)$ MDS code (GRS codes are MDS codes), it can correct arbitrary erasure patterns of up to $\Gamma = n - (k_{\mathsf{max}} + T - 1)$ erasures. This implies that one can construct a valid $k_{\mathsf{max}} \times n$ ($d = k_{\mathsf{max}}$) matrix $\hat{\boldsymbol{E}}$ (satisfying conditions C1–C3) from $\beta = \Gamma$ information sets $\{\mathcal{I}_m\}$ of $\mathcal{C}'_{\mathsf{max}}$ as shown below.

Let $\mathcal{J}_j = \{j, \ldots, (j + \Gamma - 1) \bmod n\}$, $j = 1, \ldots, k_{\mathsf{max}}$. Construct $\hat{\boldsymbol{E}}$ in such a way that $\mathcal{J}_j$ is the support of the j-th row of $\hat{\boldsymbol{E}}$. Hence, C1 is satisfied. Furthermore, since $\tilde{\mathcal{C}}$ is an $(n, k_{\mathsf{max}} + T - 1)$ MDS code and $\Gamma = n - (k_{\mathsf{max}} + T - 1)$, all rows of $\hat{\boldsymbol{E}}$ are correctable by $\tilde{\mathcal{C}}$, and thus C2 is satisfied. Finally, run 6, which constructs $\beta = \Gamma$ information sets $\{\mathcal{I}_m\}$ of $\mathcal{C}'_{\mathsf{max}}$ (and the corresponding sets $\{\mathcal{F}_l\}$) such that C3 is satisfied. Note that since $\mathcal{C}'_{\mathsf{max}}$ is an MDS code, all coordinate sets of size k_{max} are information sets of $\mathcal{C}'_{\mathsf{max}}$, and hence 6 will always succeed in constructing a valid set of information sets of $\mathcal{C}'_{\mathsf{max}}$ (the inequalities in Lines 6 and 7 together with the fact that the overall weight of $\hat{\boldsymbol{E}}$ is Γk_{max} ensure that $\beta = \Gamma$ valid information sets for $\mathcal{C}'_{\mathsf{max}}$ are constructed). In particular, the while-loop in Line 6 will always terminate.

From the constructed matrix $\hat{\boldsymbol{E}}$, the user is able to recover $\Gamma d \geq \beta k_i$ unique code symbols of the requested file $\boldsymbol{X}^{(i)}$, at least k_i symbols from each stripe. Furthermore,

Algorithm 6: Construction of $\{\mathcal{I}_m\}$ for Theorem 3

```
   Input  : Ê, β, n, k_max
   Output: {I_m}, {F_l}
5.1 for m ∈ {1, ..., β} do
5.2 |  I_m ← ∅
5.3 end
5.4 for l ∈ {1, ..., n} do
5.5 |  F_l ← ∅, m ← 1
5.6 |  while |F_l| ≤ w_H(t_l) do
5.7 |  |  if |I_m] < k_max then
5.8 |  |  |  F_l ← F_l ∪ {m}
5.9 |  |  |  I_m ← I_m ∪ {l}
5.10|  |  end
5.11|  |  m ← m + 1
5.12|  end
5.13 end
```

a set of k_i recovered code symbols from each stripe corresponds to an information set of $\mathcal{C}'_i$ (any subset of size k_i of any information set of size k_{max} of $\mathcal{C}'_{\mathsf{max}}$ is an information set of $\mathcal{C}'_i$), and the requested file $\boldsymbol{X}^{(i)}$ can be recovered. This can be seen following a similar argument as in the proof of [113, Th. 6], and it follows that the recovery condition in (6.2b) is satisfied.

Secondly, we consider the privacy condition in (6.2a). A reasoning similar to the proof of [113, Lem. 6] shows that it is satisfied, and we refer the interested reader to this proof for further details. The fundamental reason is that addition of a deterministic vector in (6.5) does not change the joint probability distribution of $\{\boldsymbol{Q}^{(l)}, l \in \mathcal{T}\}$ for any set $\mathcal{T}$ size T, and the proof follows the same lines as the proof of [121, Th. 8]. However, note that there is a subtle difference in the sense that independent instances of the protocol may query different sets of SBSs. However, since the set of SBSs that are queried is independent of the requested file and depends only on which SBSs that are within communication range, this fact does not leak any additional information on which file is requested by the user.

Bibliography

[1] "Announcing the advanced encryption standard", Federal Information Processing Standards Publication 197, United States National Institute of Standards and Technology (NIST), November 2001.

[2] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*, Cambridge University Press, first edition, 2011.

[3] A. D. Wyner, "The wire-tap channel", *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, October 1975.

[4] M. Bloch, M. Hayashi, and A. Thangaraj, "Error-control coding for physical-layer secrecy", *Proc. IEEE*, vol. 103, no. 10, pp. 1725–1746, October 2015.

[5] S. Tomasin, "A Gale-Shapley algorithm for allocation of relayed parallel wire-tap coding channels", in *Proc. IEEE Conference on Communications and Network Security (CNS)*, September 2015, pp. 119–124.

[6] C. E. Shannon, "A mathematical theory of communication", *The Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

[7] C. E. Shannon, "Communication theory of secrecy systems", *The Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, October 1949.

[8] Y. Liang, H. V. Poor, and S. Shamai, "Information theoretic security", *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4–5, pp. 355–580, 2008.

[9] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free", in *Advances in Cryptology - EUROCRYPT 2000*, B. Preenel, Ed. 2000, vol. 1807 of *Lecture Notes in Computer Science*, pp. 351–368, Springer.

[10] I. Csiszár, "Almost independence and secrecy capacity", *Probl. of Inform. Transmission*, vol. 32, no. 1, pp. 40–47, January–March 1996.

[11] S. Goldwasser and S. Micali, "Probabilistic encryption", *Journal of Computer and System Sciences*, vol. 28, no. 2, pp. 270 – 299, 1984.

[12] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel", in *Advances in Cryptology - CRYPTO 2012*, R. Savafi-Naini and R Canetti, Eds. 2012, vol. 7417 of *LNCS*, pp. 294–311, Springer Berlin Heidelberg.

[13] A. Thangaraj, S. Dihidar, A.R. Calderbank, S.W. McLaughlin, and J-M. Merolla, "Applications of LDPC codes to the wiretap channel", *IEEE Trans. Inform. Theory*, vol. 53, no. 8, pp. 2933–2945, August 2007.

[14] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes", *IEEE Trans. Inform. Theory*, vol. 57, no. 10, pp. 6428–6443, October 2011.

[15] C. Ling and J.-C. Belfiore, "Achieving AWGN channel capacity with lattice Gaussian coding", *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 5918–5929, October 2014.

[16] D. Klinc, Jeongseok Ha, S.W. McLaughlin, J. Barros, and Byung-Jae Kwak, "LDPC codes for the Gaussian wiretap channel", *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 532–540, September 2011.

[17] M. Baldi, M. Bianchi, and F. Chiaraluce, "Coding with scrambling, concatenation, and HARQ for the AWGN wire-tap channel: A security gap analysis", *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 883–894, June 2012.

[18] M. Baldi, F. Chiaraluce, N. Laurenti, S. Tomasin, and F. Renna, "Secrecy transmission on parallel channels: Theoretical limits and performance of practical codes", *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 11, pp. 1765–1779, November 2014.

[19] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channel", *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4687–4698, October 2008.

[20] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communications over fading channels", *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2470–2492, June 2008.

[21] F. Renna, N. Laurenti, and H. V. Poor, "Physical layer secrecy for OFDM transmissions over fading channels", *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 4, pp. 1354–1367, August 2012.

[22] H. V. Poor and R. Schaefer, "Wireless physical layer security", in *Proceedings of the National Academy of Science, USA*, 2017, vol. 114, pp. 19–26.

[23] J. P. Vilela, M. Gomes, W. K. Harrison, D. Sarmento, and F. Dias, "Interleaved concatenated coding for secrecy in the finite blocklength regime", *IEEE Signal Processing Letters*, vol. 23, no. 3, pp. 356–360, March 2016.

[24] Willie K. Harrison, Dinis Sarmento, João P. Vilela, and Marco A. C. Gomes, "Analysis of short blocklength codes for secrecy", *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, pp. 255, Oct 2018.

[25] B. He and X. Zhou, "Secure on-off transmission design with channel estimation errors", *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 12, pp. 1923–1936, Dec 2013.

[26] J. Hu, W. Yang, N. Yang, X. Zhou, and Y. Cai, "On off-based secure transmission design with outdated channel state information", *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6075–6088, Aug 2016.

[27] P. Mu, Z. Li, and B. Wang, "Secure on off transmission in slow fading wiretap channel with imperfect csi", *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 9582–9586, Oct 2017.

[28] J. Choi, "On channel-aware secure harq-ir", *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 2, pp. 351–362, Feb 2017.

[29] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "On the throughput of secure hybrid-ARQ protocols for Gaussian block-fading channels", *IEEE Trans. Inform. Theory*, vol. 55, no. 4, pp. 1575–1591, April 2009.

[30] S. Tomasin and N. Laurenti, "Secure harq with multiple encoding over block fading channels: Channel set characterization and outage analysis", *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 10, pp. 1708–1719, Oct 2014.

[31] P. Cuff, "A framework for partial secrecy", in *Proc. IEEE Global Telecommunications Conference (GLOBECOM 2010)*, Miami, FL, December 2010.

[32] M. Nakagami, "The m-distribution: A general formula of intensity distribution of rapid fading", in *Statistical Methods in Radio Wave Propagation*, W.C. Hoffman, Ed., pp. 3–36. Pergamon Press, New York, 1960.

[33] D. Klinc, Jeongseok Ha, S.W. McLaughlin, J. Barros, and Byung-Jae Kwak, "LDPC codes for the Gaussian wiretap channel", in *Proc. IEEE Information Theory Workshop (ITW 2009)*, Taormina, Italy, October 2009, pp. 95–99.

[34] C. W. Wong, Tan F. Wong, and John M. Shea, "Secret-sharing LDPC codes for the BPSK-constrained Gaussian wiretap channel", *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 551–564, September 2011.

[35] M. Baldi, G. Ricciutelli, N. Maturo, and F. Chiaraluce, "Performance assessment and design of finite length LDPC codes for the Gaussian wiretap channel", in *Proc. IEEE ICC 2015 - Workshop on Wireless Physical Layer Security*, London, UK, June 2015, pp. 446–451.

[36] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime", *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[37] T. Erseghe, "On the evaluation of the Polyanskiy-Poor-Verdù converse bound for finite block-length coding in AWGN", *IEEE Trans. Inform. Theory*, vol. 61, no. 12, pp. 6578–6590, December 2015.

[38] 802.16e 2005, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands", December 2005.

[39] R. Yang, "LDPC-coded modulation for transmission over AWGN and flat Rayleigh fading channels", Master's thesis, Université Laval, 2010.

[40] R. L. Rivest, "All-or-nothing encryption and the package transform", in *Fast Software Encryption*. 1997, vol. 1267 of *LNCS*, pp. 210–218, Springer Berlin Heidelberg.

[41] D. R. Stinson, "Something about all or nothing (transforms)", *Designs, Codes and Cryptography*, vol. 22, no. 2, pp. 133–138, 2001.

[42] J. L. Massey, "Shift-register synthesis and BCH decoding", *IEEE Trans. Inform. Theory*, vol. 15, no. 1, pp. 122–127, January 1969.

[43] Chin-Liang Wang, Ting-Nan Cho, and Kai-Jie Yang, "A new cooperative transmission strategy for physical-layer security with multiple eavesdroppers", in *Proc. 75th IEEE Vehicular Technology Conference (VTC Spring)*, 2012, pp. 1–5.

[44] Yulong Shen, Xiaohong Jiang, Jianfeng Ma, and Weisong Shi, "Secure and reliable transmission with cooperative relays in two-hop wireless networks", in *Proc. Information Technology Convergence*, James J. (Jong Hyuk) Park, Leonard Barolli, Fatos Xhafa, and Hwa-Young Jeong, Eds., vol. 253 of *Lecture Notes in Electrical Engineering*, pp. 397–406. Springer Netherlands, 2013.

[45] S. Luo, H. Godrich, A. Petropulu, and H.V. Poor, "A knapsack problem formulation for relay selection in secure cooperative wireless communication", in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 2512–2515.

[46] Yulong Shen, Xiaohong Jiang, Jianfeng Ma, and Weisong Shi, "Exploring relay cooperation for secure and reliable transmission in two-hop wireless networks", *CoRR*, vol. abs/1212.0287, 2012.

[47] Zhiguo Ding, Mai Xu, Jianhua Lu, and Fei Liu, "Improving wireless security for bidirectional communication scenarios", *IEEE Trans. on Vehicular Technology*, vol. 61, no. 6, pp. 2842–2848, 2012.

[48] Lun Dong, Zhu Han, A.P. Petropulu, and H.V. Poor, "Improving wireless physical layer security via cooperating relays", *IEEE Trans. on Signal Processing*, vol. 58, no. 3, pp. 1875–1888, 2010.

[49] Jiangyuan Li, A.P. Petropulu, and S. Weber, "On cooperative relaying schemes for wireless physical layer security", *IEEE Trans. on Signal Processing*, vol. 59, no. 10, pp. 4985–4997, 2011.

[50] X. Chen, D. W. K. Ng, W. H. Gerstacker, and H. Chen, "A survey on multiple-antenna techniques for physical layer security", *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 1027–1053, Secondquarter 2017.

[51] R. Bassily and S. Ulukus, "Deaf cooperation and relay selection strategies for secure communication in multiple relay networks", *IEEE Trans. on Signal Processing*, vol. 61, no. 6, pp. 1544–1554, 2013.

[52] Z. H. Awan, A. Zaidi, and L. Vandendorpe, "Secure communication over parallel relay channel", *IEEE Trans. on Information Forensics and Security*, vol. 7, no. 2, pp. 359–371, April 2012.

[53] D.W.K. Ng, E.S. Lo, and R. Schober, "Secure resource allocation and scheduling for ofdma decode-and-forward relay networks", *IEEE Trans. on Wireless Communications*, vol. 10, no. 10, pp. 3528–3540, 2011.

[54] D.W.K. Ng, E.S. Lo, and R. Schober, "Resource allocation for secure OFDMA networks with imperfect csit", in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, 2011, pp. 1–6.

[55] D.W.K. Ng and R. Schober, "Resource allocation for secure ofdma decode-and-forward relay networks", in *Proc. 12th Canadian Workshop on Information Theory (CWIT)*, 2011, pp. 202–205.

[56] Zhangju Yu, Yayan Ma, Baoyun Wang, and Junxi Zhao, "Optimal resource allocation for ofdm wiretap channel with cooperative jammer", in *Proc. Int. Conf. on Wireless Communications Signal Processing (WCSP)*, 2012, pp. 1–4.

[57] Cheol Jeong and Il-Min Kim, "Optimal power allocation for secure multi-carrier relay systems", in *Proc. 8th Int. Workshop on Multi-Carrier Systems Solutions (MC-SS)*, 2011, pp. 1–4.

[58] Cheol Jeong and Il-Min Kim, "Optimal power allocation for secure multicarrier relay systems", *IEEE Trans. on Signal Processing*, vol. 59, no. 11, pp. 5428–5442, 2011.

[59] Bo Gui and Leonard Jr. Cimini, "Bit loading algorithms for cooperative OFDM systems", *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, no. 1, pp. 476797, 2008.

[60] Luc Vandendorpe, Jerome Louveaux, Onur Oguz, and Abdellatif Zaidi, "Rate-optimized power allocation for DF-relayed OFDM transmission under sum and individual power constraints", *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, pp. 814278, 2009.

[61] Tao Wang and L. Vandendorpe, "Sum rate maximized resource allocation in multiple DF relays aided OFDM transmission", *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1559–1571, September 2011.

[62] K. Bakanoglu, S. Tomasin, and E. Erkip, "Resource allocation for the parallel relay channel with multiple relays", *IEEE Trans. on Wireless Communications*, vol. 10, no. 3, pp. 792–802, March 2011.

[63] N. Laurenti, S. Tomasin, and F. Renna, "Resource allocation for secret transmissions on parallel Rayleigh channels", in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2014.

[64] M. Baldi, F. Chiaraluce, N. Laurenti, S. Tomasin, and F. Renna, "Secrecy transmission on parallel channels: Theoretical limits and performance of practical codes", *IEEE Trans. on Information Forensics and Security*, vol. 9, no. 11, pp. 1765–1779, November 2014.

[65] J. Chen, X. Chen, W. H. Gerstacker, and D. W. K. Ng, "Resource allocation for a massive MIMO relay aided secure communication", *IEEE Trans. on Information Forensics and Security*, vol. 11, no. 8, pp. 1700–1711, August 2016.

[66] He Fang, Li Xu, and Kim-Kwang Raymond Choo, "Stackelberg game based relay selection for physical layer security and energy efficiency enhancement in cognitive radio networks", *Applied Mathematics and Computation*, vol. 296, pp. 153 – 167, 2017.

[67] J. H. Lee, "Optimal power allocation for physical layer security in multi-hop df relay networks", *IEEE Trans. on Wireless Communications*, vol. 15, no. 1, pp. 28–38, January 2016.

[68] J.-H. Lee, I. Sohn, and Y.-H. Kim, "Transmit power allocation for physical layer security in cooperative multi-hop full-duplex relay networks", *Sensors*, vol. 16, no. 10, pp. 1726, 2016.

[69] K. Zhang, M. Peng, P. Zhang, and X. Li, "Secrecy-optimized resource allocation for device-to-device communication underlaying heterogeneous networks", *IEEE Trans. on Vehicular Technology*, vol. 66, no. 2, pp. 1822–1834, February 2017.

[70] Waqas Aman, Guftaar Ahmad Sardar Sidhu, Haji M. Furqan, and Zain Ali, "Enhancing physical layer security in af relay-assisted multicarrier wireless transmission", *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 6, pp. 1–14, June 2018.

[71] M. R. Abedi, N. Mokari, H. Saeedi, and H. Yanikomeroglu, "Robust resource allocation to enhance physical layer security in systems with full-duplex receivers: Active adversary", *IEEE Trans. on Wireless Communications*, vol. 16, no. 2, pp. 885–899, February 2017.

[72] A. Kuhestani, A. Mohammadi, K. Wong, P. L. Yeoh, M. Moradikia, and M. R. A. Khandaker, "Optimal power allocation by imperfect hardware analysis in untrusted relaying networks", *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4302–4314, July 2018.

[73] A. Kuhestani, A. Mohammadi, and P. L. Yeoh, "Optimal power allocation and secrecy sum rate in two-way untrusted relaying networks with an external jammer", *IEEE Transactions on Communications*, vol. 66, no. 6, pp. 2671–2684, June 2018.

[74] Mohanad Obeed and Wessam Mesbah, "Efficient algorithms for physical layer security in two-way relay systems", *Physical Communication*, vol. 28, pp. 78–88, 2018.

[75] A. Kuhestani, A. Mohammadi, and M. Mohammadi, "Joint relay selection and power allocation in large-scale mimo systems with untrusted relays and passive eavesdroppers", *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 2, pp. 341–355, Feb 2018.

[76] Z. Mheich, F. Alberge, and P. Duhamel, "Achievable secrecy rates for the broadcast channel with confidential message and finite constellation inputs", *IEEE Transactions on Communications*, vol. 63, no. 1, pp. 195–205, Jan 2015.

[77] X. Liu, D. Ma, J. Xiong, W. Li, and L. Cheng, "Power allocation for an-aided beamforming design in miso wiretap channels with finite-alphabet signaling", in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, Sept 2016, pp. 1–6.

[78] Y. Polyanskiy, "Saddle point in the minimax converse for channel coding", *IEEE Trans. on Information Theory*, vol. 59, no. 5, pp. 2576–2595, 2013.

[79] E. A. Jorswieck and A. Wolf, "Resource allocation for the wire-tap multi-carrier broadcast channel", in *Proc. of Int. Workshop on Multiple Access Communications (MACOM)*, Saint Petersburg, Russia, June 2008.

[80] D. Gale and L. S. Shapley, "College admissions and the stability of marriage", *American Mathematical Monthly*, vol. 69, pp. 9–15, 1962.

[81] M. Kwiatkowska, G. Norman, and D. Parker, *PRISM 4.0: Verification of Probabilistic Real-Time Systems*, Springer, 2011.

[82] J. L. Massey, "Guessing and entropy", in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Trondheim, Norway, 27 Jun.–1 Jul. 1994, p. 204.

[83] David A. Basin, Cas Cremers, and Catherine A. Meadows, "Model checking security protocols", in *Handbook of Model Checking*. Springer, 2017.

[84] Santiago Escobar, Catherine A. Meadows, and José Meseguer, "A rewriting-based inference system for the NRL protocol analyzer and its meta-logical properties", *Theor. Comput. Sci.*, vol. 367, no. 1-2, pp. 162–202, 2006.

[85] Francesco Pagliarecci, Luca Spalazzi, and Francesco Spegni, "Model checking grid security", *Future Generation Comp. Syst.*, vol. 29, no. 3, pp. 811–827, 2013.

[86] Maurizio Panti, Luca Spalazzi, Simone Tacconi, and Salvatore Valenti, "Automatic verification of security in payment protocols for electronic commerce", in *ICEIS 2002, Proc. of the 4th Int. Conf. on Enterprise Information Systems*, 2002, pp. 968–974.

[87] Danny Dolev and Andrew Chi-Chih Yao, "On the security of public key protocols", *IEEE Trans. Inf. Theory*, vol. 29, no. 2, pp. 198–207, 1983.

[88] Ezio Bartocci, Radu Grosu, Panagiotis Katsaros, C. R. Ramakrishnan, and Scott A. Smolka, "Model repair for probabilistic systems", in *Proc. of TACAS 2011: the 17th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. 2011, vol. 6605 of *LNCS*, pp. 326–340, Springer.

[89] Gabriele Lenzini, Sjouke Mauw, and Samir Ouchani, "Security analysis of socio-technical physical systems", *Computers & Electrical Engineering*, vol. 47, pp. 258–274, 2015.

[90] Samir Ouchani and Mourad Debbabi, "Specification, verification, and quantification of security in model-based systems", *Computing*, vol. 97, no. 7, pp. 691–711, 2015.

[91] Fan Yang, Guowu Yang, and Yujie Hao, "The modeling library of eavesdropping methods in quantum cryptography protocols by model checking", *International Journal of Theoretical Physics*, vol. 55, no. 7, pp. 3414–3427, 2016.

[92] Axel Legay, Benoît Delahaye, and Saddek Bensalem, "Statistical model checking: An overview", in *Proc. of RV 2010: the First International Conference on Runtime Verification*. 2010, vol. 6418 of *LNCS*, pp. 122–135, Springer Berlin Heidelberg.

[93] Edmund M. Clarke, Orna Grumberg, and David E. Long, "Model checking and abstraction", *ACM Trans. Program. Lang. Syst.*, vol. 16, no. 5, pp. 1512–1542, 1994.

[94] Benjamin Aminof, Tomer Kotek, Sasha Rubin, Francesco Spegni, and Helmut Veith, "Parameterized model checking of rendezvous systems", in *Proc. of CONCUR 2014: the 25th International Conference on Concurrency Theory*. 2014, vol. 8704 of *LNCS*, pp. 109–124, Springer.

[95] Roderick Bloem, Swen Jacobs, Ayrat Khalimov, Igor Konnov, Sasha Rubin, Helmut Veith, and Josef Widder, "Decidability in parameterized verification", *SIGACT News*, vol. 47, no. 2, pp. 53–64, 2016.

[96] Taylor T. Johnson and Sayan Mitra, "A small model theorem for rectangular hybrid automata networks", in *Proc. of FORTE 2012: Inter. Conf. on Formal Techniques for Distributed Objects, Components, and Systems*. 2012, vol. 7273 of *LNCS*, pp. 18–34, Springer.

[97] Luca Spalazzi and Francesco Spegni, "Parameterized model-checking of timed systems with conjunctive guards", in *Proc. of VSTTE 2014: the 6th Int. Conf. on Verified Software: Theories, Tools and Experiments*. 2014, vol. 8471 of *LNCS*, pp. 235–251, Springer.

[98] Nathalie Bertrand and Paulin Fournier, "Parameterized verification of many identical probabilistic timed processes", in *Proc. of FSTTCS 2013: the IARCS Annual Conf. on Found. of Soft. Tech. and Theor. Computer Science*, 2013, vol. 24 of *LIPIcs*, pp. 501–513.

[99] Ondrej Lengál, Anthony Widjaja Lin, Rupak Majumdar, and Philipp Rümmer, "Fair termination for parameterized probabilistic concurrent systems", in *Proc. of TACAS 2017*, 2017, vol. 10205 of *LNCS*, pp. 499–517.

[100] Yongjian Li and Jun Pang, "Formalizing provable anonymity in Isabelle/HOL", *Formal Aspects of Computing*, vol. 27, no. 2, pp. 255–282, 2015.

[101] Vitaly Shmatikov, "Probabilistic analysis of an anonymity system", *Journal of Computer Security*, vol. 12, no. 3-4, pp. 355–377, 2004.

[102] V. Boyko, "On the security properties of OAEP as an all-or-nothing transform", in *Advances in Cryptology – CRYPTO' 99*, vol. 1666 of *Lecture Notes in Computer Science*, pp. 503–518. Springer, 1999.

[103] M. Baldi, N. Maturo, E. Montali, and F. Chiaraluce, "AONT-LT: A data protection scheme for cloud and cooperative storage systems", in *Proc. Int. Conf. on High Performance Computing & Simulation (HPCS 2014)*, Bologna, Italy, July 2014, pp. 566–571.

[104] T. Ernvall, S. El Rouayheb, C. Hollanti, and H. V. Poor, "Capacity and security of heterogeneous distributed storage systems", *IEEE J. Select. Areas Commun.*, vol. 31, no. 12, pp. 2701–2709, December 2013.

[105] M. L. Merani, C. Barcellona, and I. Tinnirello, "Multi-cloud privacy preserving schemes for linear data mining", in *Proc. IEEE International Conference on Communications (ICC 2015)*, London, UK, June 2015.

[106] Z. Durumeric, J. Kasten, D. Adrian, J. A. Halderman, M. Bailey, F. Li, N. Weaver, J. Amann, J. Beekman, M. Payer, and V. Paxson, "The matter of heartbleed", in *Proc. of the 2014 Internet Measurement Conference*. 2014, pp. 475–488, ACM.

[107] M. Georgiev, S. Iyengar, S. Jana, R. Anubhai, D. Boneh, and V. Shmatikov, "The most dangerous code in the world: validating SSL certificates in non-browser software", in *Proc. of the ACM Conf. on Comp. and Commun. Sec.*, 2012, pp. 38–49.

[108] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings", *IEEE Trans. Microwave Theory Tech.*, vol. 40, no. 2, pp. 202–217, February 1992.

[109] G. Pei and T.R. Henderson, "Validation of OFDM error rate model in ns-3", `www.nsnam.org/~pei/80211ofdm.pdf`, 2010.

[110] C. Baier and J.-P. Katoen, *Principles Of Model Checking*, Springer, 2008.

[111] Y. Desmedt, "Threshold cryptosystems", in *Advances in Cryptology — AUSCRYPT '92, Seberry J., Zheng Y. (eds)*. 1993, vol. 718 of *LNCS*, pp. 1–14, Springer Berlin Heidelberg.

[112] V. Bioglio, F. Gabry, and I. Land, "Optimizing MDS codes for caching at the edge", in *Proc. IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, December 2015.

[113] S. Kumar, H.-Y. Lin, E. Rosnes, and A. Graell i Amat, "Achieving maximum distance separable private information retrieval capacity with linear codes", arXiv:1712.03898v3 [cs.IT], December 2017.

[114] Y. Ishai, E. Kushilevitz, R. Ostrovsky, and A. Sahai, "Batch codes and their applications", in *Proc. 36th Annual ACM Symp. Theory Comput. (STOC)*, Chicago, IL, June 2004, pp. 262–271.

[115] N. B. Shah, K. V. Rashmi, and K. Ramchandran, "One extra bit of download ensures perfectly private information retrieval", in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Honolulu, HI, June/July 2014, pp. 856–860.

[116] Terence H. Chan, Siu-Wai Ho, and Hirosuke Yamamoto, "Private information retrieval for coded storage", in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Hong Kong, China, June 2015, pp. 2842–2846.

[117] Razan Tajeddine and Salim El Rouayheb, "Private information retrieval from MDS coded data in distributed storage systems", in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, July 2016, pp. 1411–1415.

[118] S. Kumar, E. Rosnes, and A. Graell i Amat, "Private information retrieval in distributed storage systems using an arbitrary linear code", in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 1421–1425.

[119] Hua Sun and Syed A. Jafar, "Private information retrieval from MDS coded data with colluding servers: Settling a conjecture by Freij-Hollanti et al.", in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, June 2017, pp. 1893–1897.

[120] Hua Sun and Syed Ali Jafar, "The capacity of private information retrieval", *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4075–4088, July 2017.

[121] R. Freij-Hollanti, O. W. Gnilke, C. Hollanti, and D. A. Karpuk, "Private information retrieval from coded databases with colluding servers", *SIAM J. Appl. Algebra Geom.*, vol. 1, no. 1, pp. 647–664, November 2017.

[122] K. Banawan and S. Ulukus, "The capacity of private information retrieval from coded databases", *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1945–1956, March 2018.

[123] H. Sun and S. A. Jafar, "The capacity of robust private information retrieval with colluding databases", *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 2361–2370, April 2018.

[124] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval", in *Proc. 36th IEEE Symp. Found. Comp. Sci. (FOCS)*, Milwaukee, WI, October 1995, pp. 41–50.

[125] S. Kadhe, B. Garcia, A. Heidarzadeh, S. El Rouayheb, and A. Sprintson, "Private information retrieval with side information", arXiv:1709.00112v1 [cs.IT], September 2017.

[126] Y.-P. Wei, K. Banawan, and S. Ulukus, "Fundamental limits of cache-aided private information retrieval with unknown and uncoded prefetching", arXiv:1709.01056v1 [cs.IT], September 2017.

[127] Z. Chen, Z. Wang, and S. Jafar, "The capacity of private information retrieval with private side information", arXiv:1709.03022v1 [cs.IT], September 2017.

[128] Y.-P. Wei, K. Banawan, and S. Ulukus, "The capacity of private information retrieval with partially known private side information", arXiv:1710.00809v2 [cs.IT], October 2017.

[129] M. Abdul-Wahid, F. Almoualem, D. Kumar, and R. Tandon, "Private information retrieval from storage constrained databases – coded caching meets PIR", arXiv:1711.05244v1 [cs.IT], November 2017.

[130] M. A. Attia, D. Kumar, and R. Tandon, "The capacity of private information retrieval from uncoded storage constrained databases", arXiv:1805.04104v2 [cs.IT], May 2018.

[131] Y.-P. Wei, K. Banawan, and S. Ulukus, "Cache-aided private information retrieval with partially known uncoded prefetching: Fundamental limits", *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1126–1139, Jun. 2018.

[132] S. Li and M. Gastpar, "Converse for multi-server single-message PIR with side information", arXiv:1809.09861v1 [cs.IT], September 2018.

[133] W. Cary Huffman and Vera Pless, Eds., *Fundamentals of Error-Correcting Codes*, Cambridge University Press, Cambridge, UK, 2010.

[134] K. Shanmugam, N. Golrezaei, A. G. Dimakis, A. F. Molisch, and G. Caire, "Femtocaching: Wireless content delivery through distributed caching helpers", *IEEE Trans. on Information Theory*, vol. 59, no. 12, pp. 8402–8413, December 2013.

[135] L. Breslau, Pei Cao, Li Fan, G. Phillips, and S. Shenker, "Web caching and Zipf-like distributions: Evidence and implications", in *Proc. IEEE Joint Conf. Comput. Commun. Soc. (INFOCOM)*, New York, NY, March 1999, pp. 126–134.

[136] B. Serbetci and J. Goseling, "On optimal geographical caching in heterogeneous cellular networks", in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, San Francisco, CA, March 2017.