ABSTRACT

Title of dissertation: RANDOM CODES AND GRAPHS FOR SECURE COMMUNICATION

Nagaraj Prasanth Anthapadmanabhan
Doctor of Philosophy, 2009

Dissertation directed by: Professor Alexander Barg
Department of Electrical and Computer Engineering

This dissertation considers two groups of problems related to secure communication. The *first* line of research is devoted to theoretical problems of copyright protection of digital content. Embedding identification data in the content is a well-developed technique of content protection known under the name of fingerprinting. Schemes that provide such protection are known as *fingerprinting codes* in the literature. We study limits of the number of users of a fingerprinting system as well as constructions of low-complexity fingerprinting codes that support a large number of users. The *second* problem that is addressed in the dissertation relates to connectivity analysis of ad hoc wireless networks. One of the basic requirements in such environments is to ensure that none of the nodes are completely isolated from the network. We address the problem of characterizing threshold parameters for node isolation that enable the system designer to choose the power needed for network operation based on the outage probability of links in the network.

The methods of this research draw from coding theory, information theory and random graphs. An idea that permeates most results in this dissertation is the application of randomization both in the analysis of fingerprinting and node isolation.

The main contributions of this dissertation belong in the area of fingerprinting and are described as follows. We derive new lower and upper bounds on the optimal trade-off between the number of users and the length of the fingerprints required to ensure reliability of the system, which we call fingerprinting capacity. Information-theoretic techniques employed in our proofs of bounds on capacity originate in coding theorems for channels with multiple inputs. Constructions of fingerprinting codes draw on methods of coding theory related to list decoding and code concatenation.

We also analyze random graph models for ad hoc networks with link failures and secure sensor networks that employ randomized key distribution. We establish a precise zero-one law for node isolation in the model with link failures for nodes placed on the circle. We further generalize this result to obtain a one-law for secure sensor networks on some surfaces.

RANDOM CODES AND GRAPHS FOR SECURE COMMUNICATION

by

Nagaraj Prasanth Anthapadmanabhan

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee:
Professor Alexander Barg, Chair/Advisor
Professor Armand Makowski
Professor Prakash Narayan
Professor Aravind Srinivasan
Professor Min Wu

Dedication

To my mother

Acknowledgments

It is with utmost gratitude that I express my thanks to my research advisor Prof. Alexander Barg. I have learned several skills, both related to research and life in general, through our many interactions. I appreciate the trust he placed in me allowing me to explore topics of my choosing. I'm thankful for the many meetings which would extend into the late hours and some even on weekends, which resulted in several key chapters of this thesis. For all these things and more, I'm forever grateful to Prof. Barg.

I am fortunate to have collaborated with Professors Ilya Dumer, Armand Makowski and Prakash Narayan which led to several results obtained in this thesis. I'm indebted to Prof. Ilya Dumer for his insight regarding code constructions for fingerprinting. I'm grateful to Prof. Armand Makowski for introducing me to random graphs, and for the technical guidance and support he provided during our joint work. The results on fingerprinting capacity would not have been possible without the direction provided by Prof. Prakash Narayan and his enthusiastic participation in numerous discussions. I'm thankful to each of them as they have all contributed in some way to my development as a researcher.

I thank Professors Armand Makowski, Prakash Narayan, Aravind Srinivasan and Min Wu for serving on my advisory committee, taking time to read several parts of the thesis, and for their helpful suggestions and advice. A special mention is appropriate for Prof. Adrian Papamarcou for the guidance he provided during my initial years in graduate school and for directing me to the Information and Coding theory seminars which initiated my interest in this field. I wish to thank Professors Pierre Moulin and Damianos Karakos for some useful discussions on topics related to the thesis. I'm also very grateful to Professors Alexander Barg, Armand Makowski and Prakash Narayan for their patience in writing several letters, and their assistance with my job search.

A friendly chat with fellow students can perk up one's day and also direct one to avenues that were not explored before. In this respect, I specially thank Arya, Punarbasu, Sirin and Ravi for eagerly listening to my research problems, while also introducing me to theirs. I also thank my friends and colleagues Anna, Brooke, Chunxuan, Ersen, Himanshu, Kaushik, Nan, Osman and Wei for the good times we spent in the department and at various research meetings.

I would like to thank the staff at ECE department, in particular Dr. Tracy Chung, Maria Hoo, Melanie Miller, Vivian Lu and Ronald Jean, who ensured a smooth process in completing all the academic and administrative formalities.

A tight circle of friends was especially important as an international student with a rare opportunity to visit family. I thank my good friends Narayanan and Som for the cheerful experiences we shared as housemates for over three years. I'm thankful to all my friends that used to gather at 8003 University Square who were very helpful in times of need and rejuvenated the spirits on dull days. My thanks also to Aravind and Rajeev who helped me get accustomed to an unfamiliar land when I first arrived in the United States.

This dissertation would not have been possible without the love and support from my family. My parents and my brother have been supportive and understanding of all my decisions. I'm forever grateful for everything they have done. My wife Buvana provided me the inspiration and motivation as we experienced the thick and thin of graduate student life together. Even in times of crisis, I could always count on her to bring a smile to my face.

Grants: I gratefully acknowledge the funding provided by the National Science Foundation under grants CCR 0310961, CCF 0515124 and CCF 0635271 that supported my research at University of Maryland.

College Park, April 2009 *Prasanth Anthapadmanabhan*

Table of Contents

List of Tables viii

List of Figures ix

List of Abbreviations x

1 Introduction 1
- 1.1 Secure Content Distribution . . . 2
  - 1.1.1 Motivation . . . 2
  - 1.1.2 Previous Work . . . 5
  - 1.1.3 Objectives . . . 7
  - 1.1.4 Main Contributions . . . 7
- 1.2 Security in Sensor Networks . . . 9
  - 1.2.1 Motivation . . . 9
  - 1.2.2 Previous Work . . . 11
  - 1.2.3 Main contributions . . . 12
- 1.3 Basic Notation and Conventions . . . 13
- 1.4 Useful Identities . . . 14

2 Preliminaries on Fingerprinting 15
- 2.1 The Fingerprinting Problem . . . 15
  - 2.1.1 Deterministic Codes . . . 16
  - 2.1.2 Randomized Codes . . . 19
- 2.2 Fingerprinting Capacity . . . 21
- 2.3 Properties of Fingerprinting Capacity . . . 22
- 2.4 Related Combinatorial Properties . . . 24
- 2.5 Concatenated Codes . . . 25
- 2.6 Appendix . . . 26
  - 2.6.1 A lemma on the size of coalitions . . . 26
  - 2.6.2 Proof of Proposition 2.12 . . . 28

3 Lower Bounds on Fingerprinting Capacity 30
- 3.1 Lower Bound I . . . 31
  - 3.1.1 Code Generation . . . 31
  - 3.1.2 Minimal Configurations . . . 31
  - 3.1.3 Analyzing the Error Probability . . . 32
  - 3.1.4 Coalitions of Size 3 . . . 34
  - 3.1.5 Coalitions of Arbitrary Size . . . 38
- 3.2 Lower Bound II . . . 39
  - 3.2.1 Code Generation . . . 40
  - 3.2.2 Analyzing the Error Probability . . . 40
  - 3.2.3 Coalitions of Size 2 . . . 41
  - 3.2.4 Coalitions of Size 3 . . . 43

3.3 Summary and Comparisons 46

4 Interlude: The "Decode-One" Multiple Access Channel 48
4.1 Single Known Channel 49
4.1.1 Problem Statement 49
4.1.2 Main Results 50
4.2 Arbitrarily Varying Channel 62
4.2.1 Problem Statement 62
4.2.2 Main Results 64
4.3 Concluding Remarks 70
4.4 Appendix 71
4.4.1 Proof of Lemma 4.10 71

5 Upper Bounds on Fingerprinting Capacity 73
5.1 Upper Bound I 74
5.1.1 The general case 75
5.1.2 The binary case 79
5.2 Upper Bound II 80
5.2.1 The general case 80
5.2.2 The binary case 81
5.3 Summary and Recent Results 82
5.4 Appendix 83
5.4.1 Proof of Theorem 5.3 83

6 Randomized Frameproof Codes 87
6.1 Problem Definition 88
6.2 Lower Bounds for Binary Frameproof Codes 89
6.3 Linear Frameproof Codes 91
6.3.1 Linear Construction for $t = 2$ 92
6.3.2 Connection to Minimal Vectors 93
6.3.3 Linear Codes for Larger t 95
6.4 Polynomial-time Validation for Larger t 95
6.5 Concluding Remarks 99

7 Two-level Fingerprinting 100
7.1 Introduction 100
7.2 Problem Statement 101
7.2.1 Deterministic Codes 102
7.2.2 Randomized Codes 104
7.3 Traceability and Frameproof Codes: Simple Facts 106
7.4 Fingerprinting Codes 107
7.4.1 Code Generation 107
7.4.2 Useful Facts 108
7.4.3 $(t, 1)$-fingerprinting 109
7.4.4 $(t, 2)$-fingerprinting 113

7.5 Concluding Remarks . . . 116

8 Fingerprinting Codes with Polynomial-time Tracing 117
8.1 One-level Codes . . . 118
8.1.1 Code Construction . . . 118
8.1.2 Main Result . . . 120
8.2 Two-level Codes . . . 123
8.2.1 Code Construction . . . 124
8.2.2 Main Result . . . 127

9 Beyond the Disk Model of Wireless Networks 132
9.1 Model and Assumptions . . . 133
9.1.1 WSNs with random link failures . . . 134
9.1.2 Secure WSNs . . . 135
9.1.3 Objectives . . . 137
9.2 Previous Work . . . 137
9.3 Main Results . . . 139
9.4 Method of First and Second Moments . . . 141
9.5 Calculation of First Moments . . . 143
9.6 Proof of the One Laws . . . 145
9.7 Calculation of Second Moments . . . 147
9.8 Proof of the Zero Laws . . . 152
9.9 Simulation Results . . . 160
9.10 Concluding Remarks . . . 161

Bibliography 163

List of Tables

3.1 Bad assignments for degree-3 minimal configurations of Type 3 36

3.2 Bad assignments for degree-3 minimal configurations of Type 4 37

6.1 Achievable rates of randomized frameproof codes 91

List of Figures

1.1 Fingerprinting system model 3

3.1 Types of minimal configurations of degree 3 35

5.1 The uniform channel with 2 pirates 79

5.2 To the proof that $a(x_0) > a(0)$. 85

7.1 Achievable rate region for binary $(2, 1)$-fingerprinting 110

9.1 Simulation results for WSNs with random link failures 160

List of Abbreviations

AG code	Algebraic-geometry code
i.i.d.	independent and identically distributed
indep.	independent
IPP	Identifiable Parent Property
MAC	Multiple-Access Channel
MD	Minimum Distance
resp.	respectively
r.-h.s.	right-hand side
RS code	Reed-Solomon code
r.v.	random variable
TA	Traceability
W.l.o.g.	Without loss of generality
WSN	Wireless Sensor Network

Chapter 1

Introduction

A central theme in this dissertation is applications of *randomization* in problems of secure content distribution and communication. For many complex problems in system design, deterministic solutions either do not exist, or exhibit poor performance and/or implementation complexity. Randomization allows us to introduce probabilistic schemes which can yield simple solutions with improved performance and complexity. Sometimes randomization is the only alternative for certain problems where deterministic schemes cannot work. In adversarial scenarios, it is especially important to randomize in order to confound the attackers.

Nevertheless, in order to obtain practical schemes, we also require some structure balancing the randomization. Coding theory and combinatorial designs, in essence, provide matrices or *codes* with a special structure. On the other hand, *graphs* are a useful tool to model and visualize this structure. In this dissertation, we attempt to combine randomization together with codes and graphs to tackle some problems of interest in secure communication. We focus on the two following classes of problems:

(a) Random codes with application to distributing copyrighted content in a secure manner so as to prevent piracy,

(b) Random graph models for connectivity analysis of secure networks.

The *main accomplishments* of this research concern the copyright protection problem and its variations. We characterize fundamental performance limits for schemes used to fight piracy of digital content, present explicit constructions of such

schemes with performance close to the limits, and extend these results to a number of related problems in information-theoretic cryptography.

Another aspect of secure communication addressed in this thesis is concerned with random graph models for secure sensor networks, and networks with random link failures. In this, somewhat separate, part of our research, we characterize the threshold parameter scalings that guarantee the absence of isolated nodes in the network.

In the next two sections we introduce these two topics and discuss previous work and our results in more detail. We will adhere to an informal discussion style throughout this introduction, giving formal definitions and statements of results in later chapters.

1.1 Secure Content Distribution

1.1.1 Motivation

The distribution of licensed digital content (e.g., software, movies, music etc.) has become increasingly popular in recent times. With this comes the need to protect the copyright of the distributor against unauthorized redistribution of the content, commonly known as piracy. To introduce the problem, we begin with an informal description.

Suppose the distributor has some content which he would like to distribute among a set of licensed users. One can think of a simple scheme where each licensed copy is identified by a unique mark (*fingerprint*) which is embedded in the content and is imperceptible to the users of the system. Note that the distributed copies are identical except for the fingerprints. If a naive user distributes a copy of his fingerprinted content, then the pirated copy can easily be traced back to the guilty user and hence he will be exposed. Tracing the guilty user becomes more difficult

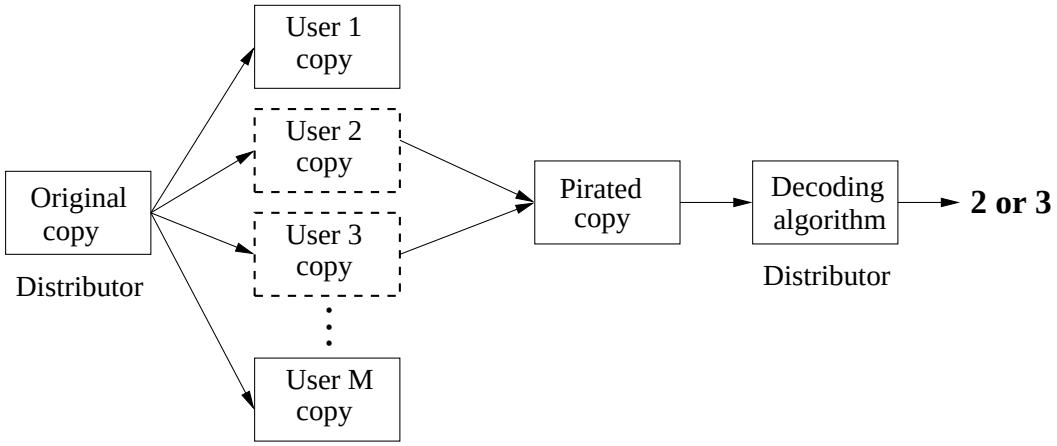


Figure 1.1: Model of the fingerprinting system. If users 2 and 3 collaborate to create the pirated copy, we require the decoding algorithm to output either 2 or 3.

when a collection of users form a *coalition* to detect the fingerprints and modify/erase them before illegally distributing the data. The members of the coalition will be referred to as *pirates*.

Digital fingerprinting is a technique that assigns to each user a mark in a way that enables the distributor to identify at least one of the pirates as long as the coalition size does not exceed a certain threshold t, which is a parameter of the problem. The model of a general fingerprinting system is shown in Figure 1.1. As an example, if users 2 and 3 collaborate to create the pirated copy, the objective of the distributor will be to output either 2 or 3 using a decoding algorithm[1]. The distributor commits a decoding error if it is unable to identify any user as a member of the pirate coalition or if a user outside the coalition is identified as a pirate.

There are two main setups considered for the fingerprinting problem in the literature. The *distortion* setting is commonly used in applications relating to multimedia fingerprinting. See [62, 82, 85, 52] among others for work on multimedia fingerprinting. In this model, the fingerprint is usually a "covert signal" which is superimposed on the original "host" data in such a way that the difference, or distortion, between the original and the fingerprinted copies is smaller than some

[1]The terms decoding, tracing, and pirate identification will be used interchangeably throughout.

threshold. The coalitions are restricted to creating a forgery which has distortion less than some threshold from at least one of the colluders' fingerprinted copies.

The line of research into the construction of fingerprinting schemes pursued in this dissertation applies to systems designed to protect digital content and relies on the so-called ***marking assumption*** setting introduced by Boneh and Shaw [26]. In this case, the fingerprint is a set of redundant digits which are distributed in some random positions (unknown to the users) across the information digits of the original content. The fingerprint positions remain the same for all users. It is assumed that these redundant digits do not affect the functionality of the content, while tampering with an information digit damages the content permanently. The motivation for this assumption arises in applications to protecting software where modifying arbitrary digits can damage its functionality. Note that the problem of embedding such fingerprints in the software is a separate non-trivial problem which we will not consider in this thesis.

The pirate coalition attempts to uncover some of the fingerprint positions by comparing their copies of the data for differences. Because the users' copies are identical except for the fingerprints, once such a difference is located in some position, it is guaranteed to be a redundant fingerprint digit. The comparison procedure does not reveal any information regarding the bits that are identical in all the data copies of the coalition members, which can be either information or fingerprint digits. The *marking assumption* states that to preserve functionality of the data, the coalitions may modify only those positions where they find a difference in their fingerprinted copies. For this reason, in analyzing this model, it becomes sufficient to look *only at the fingerprint positions* and ignore the information digits.

The collection of fingerprint assignments for all the users of the system together with the decoding algorithm is called a *code* below. The design objective of a fingerprinting system is to provide a code construction that will guarantee a low

probability of identification error. It is known [26] that any fixed assignment of fingerprints (a fixed code) cannot satisfy this requirement: namely, there exist attack strategies of the coalition that will result in the error probability bounded away from zero irrespective of the decoding employed. For this reason it becomes necessary for the distributor to use some form of ***randomization*** in constructing such codes, where the random key is known only to the distributor (see Section 2.1 for details).

Informally speaking, a randomized code family is said to be t***-fingerprinting*** if given a pirated copy produced by any coalition of size at most t, the probability of decoding error approaches 0 as the length of the fingerprints increases. Such codes are also termed *collusion-secure* against t pirates in the literature.

Fingerprinting codes also find applications in the context of broadcast encryption (e.g., pay-per-view TV) [27, 28]. In this problem, all the users receive the same broadcast data. To prevent unauthorized users from accessing the content, the data is encrypted before broadcasting. The broadcast is divided into a number, say n, of encrypted segments, each of which can be decrypted using one of q keys. Each licensed user is given a particular collection of n keys, one to decrypt each segment. A coalition of licensed users may attempt to create a new set of keys (designed for redistribution) in which for each segment one of the colluders contributes his assigned key. It is easy to see that the marking assumption is satisfied in producing such an unlicensed set of keys. Therefore, if the keys are assigned according to a fingerprinting code, then at least one of the guilty users can be traced from this key set with high probability.

1.1.2 Previous Work

Though the first works on fingerprinting date back to the 1980s [83, 22], it was the works of Chor et al. [27, 28] on traitor tracing for broadcast encryption and of Boneh and Shaw [26] on collusion-secure fingerprinting codes that brought it to

the attention of the research community. The paper [26] also was the first to give a construction of code families of increasing length with vanishing error probability. Further general constructions were proposed by Barg et al. [15] and Tardos [79], followed by [21, 68, 70, 72] and many others.

The case of deterministic codes with (exactly) zero error probability was considered independently by Hollmann et al. [53] who called them codes with the *identifiable parent property*, or IPP codes. As was pointed above, such code constructions are possible only under some assumptions regarding the problem which will be discussed in detail later in this thesis. IPP codes were further studied in [16, 17, 3, 19] among others.

The *rate* of a fingerprinting code quantifies the tradeoff between the number of users that can be supported by the system and the fingerprint length required to make it workable. The rate is given by the ratio of the logarithm of the number of licensed users to the length of the fingerprints. A ***fundamental question*** in understanding the fingerprinting problem is as follows: Given the maximum number of colluders, what is the largest rate attainable by fingerprinting codes such that there is a tracing algorithm that makes an error with an arbitrarily small probability? This maximum attainable code rate will be called ***capacity*** in this thesis. At the time the dissertation research was carried out, the only lower bounds on the capacity available were implied by the constructions of [26, 15, 79] and the existence results of [21]; no upper bounds were known.

At the same time, in the distortion setting, some results on the information-theoretic limits were known (see [63, 75, 76, 77]). In particular, Somekh-Baruch and Merhav [75, 77] obtained upper and lower bounds for the optimal rate which differ by a factor corresponding to the maximum coalition size. However, the methods used to establish these bounds are insufficient to establish tight bounds on the capacity of the digital fingerprinting problem or to construct code families and decoding algorithms

for it. For this reason the two problem formulations complement each other rather than compete in the context of copyright protection systems. Ahlswede and Cai [2] considered the marking assumption setting, but addressed a simpler problem whose results do not directly apply to fingerprinting.

1.1.3 Objectives

The main objectives of this dissertation are to investigate the fundamental information-theoretic limits of fingerprinting codes and to develop code constructions exhibiting performance close to these limits. Specifically, our goal is to establish new tighter bounds on the capacity of fingerprinting codes for a given coalition size. In addition, we also aim to construct codes having high rates, low error probability of identification, and efficient decoding algorithms.

1.1.4 Main Contributions

The fingerprinting problem. We contribute new bounds on the fingerprinting capacity, analyzing this problem both for coalitions of small size and of an arbitrary given size t. This is done because the more precise methods proposed for the former case encounter substantial technical obstacles and are replaced by other methods for the general case. We also study several problems related to fingerprinting, proposing new code constructions and identification methods in each of them.

- We introduce a new derivation method for lower bounds on the fingerprinting capacity which takes account only of the collusion events that occur with high probability, discarding "atypical" coalitions. This is a novel idea in fingerprinting which enabled us to establish capacity bounds that are substantially better than previously known results. This is the contents of Chapter 3.

- Motivated by fingerprinting, in Chapter 4 we study a new problem in multi-user information theory. Specifically, we analyze the capacity of a multiple-access channel (MAC) in which (a) all channel inputs use the same code, and (b) the decoder recovers *only one* of the transmitted messages. These differences render the problem much harder to solve than the usual MAC capacity. We derive upper and lower bounds for the capacity of the above channel expressed in a form that involves only single-letter mutual information quantities and thus facilitates numerical computation of the capacity bounds.

- Relying on the insights developed for the MAC problem described, in Chapter 5 we employ information-theoretic methods to obtain the first known upper bounds on fingerprinting capacity. Evaluating the bounds is a difficult optimization problem which we address separately in the general case and for coalitions of small size. This leads to new capacity bounds in each of these cases.

- In order to obtain practical codes with high rates and efficient algorithms, we study a variation of the fingerprinting model known under the name of "frameproof codes" [26, 78]. In the modified system, whenever a user tries to access his copy, the fingerprint is submitted to a validation algorithm to verify that it is indeed permissible before the copy can be executed. In this setup, we are faced with the simplified objective of designing codes such that the pirates cannot forge the fingerprint of an innocent user. In Chapter 6, we construct such randomized codes with high rates and with validation complexity polynomial in the fingerprint length.

- In Chapter 7, we introduce *two-level* fingerprinting codes which provide partial information about the pirates even when the coalition size exceeds the designed limit. In this setting, the users are organized in a hierarchical manner by

classifying them into various groups. We begin with formalizing the two-level fingerprinting problem and then move to the construction problem for two-level codes. The codes that we construct have the following property: As in traditional fingerprinting codes, if the coalition size is at most t, the tracing algorithm determines one of the guilty users correctly with high probability. In addition, even when a larger number s ($> t$) of pirates participate, the algorithm provides partial information by tracing one of the *groups* containing a guilty user.

- Finally, in Chapter 8 we construct polynomial-complexity one- and two-level fingerprinting codes. Combining randomized codes with code concatenation and the list decoding approach of [15], we obtain binary fingerprinting codes with polynomial-time tracing algorithms having the best known rate of $\Omega(1/t^4)$ with coalition size t. We also construct a family of two-level fingerprinting codes of asymptotically positive rate which performs two-level tracing in time polynomial in the length of the fingerprints.

The ideas regarding the fingerprinting capacity introduced in the thesis were very recently adapted and developed by Amiri and Tardos [4], Dumer [36], and Moulin (with Huang) [61, 54], leading in particular, to tight results for fingerprinting capacity. We elaborate on these developments in Chapters 3 and 5.

1.2 Security in Sensor Networks

1.2.1 Motivation

A wireless sensor network (WSN) consists of a large collection of sensors distributed over some terrain and communicating in an ad hoc manner. In order to establish secure communication between a pair of sensors they need to be equipped with a shared secret key. A *probabilistic* key sharing scheme that addresses this

requirement was proposed by Eschenauer and Gligor [38]. An informal description of the scheme follows.

The system constructor has a large pool of secret keys. Before deployment, each sensor is provided with a certain number of keys selected at random from this key pool. Assuming that the sensors have only a finite communication range (as is the case in practice) and no interference between transmissions, any two nodes can now establish a secure link if they are located within the range and share a secret key. A typical objective is to construct, with high probability, a network which ensures that any two nodes can communicate securely, perhaps using multiple hops over intermediate links.

We model the connectivity problem described as a graph whose vertices correspond to the sensors and two vertices are connected by an edge if the pair of sensors can directly communicate with each other. The objective of enabling any two nodes in the network to reach each other, is accomplished by making sure that the corresponding graph model is *connected*, i.e., that there exists a path linking any two nodes. Clearly, the above condition requires that the network contains no isolated nodes (a node is *isolated* if it has no edges to any of the other nodes).

In a number of contexts, including the application just mentioned, *random* graph models have been found to be more appropriate in order to account for the inherent randomness. A common objective in random graph theory is to identify the critical scalings (or thresholds) of the graph parameters at which various monotone graph properties (e.g., disappearance of isolated nodes, graph connectedness, containment of a given subgraph etc.) emerge. Such thresholds are identified by what are known as "zero-one laws" in the literature.

The classical random graphs were introduced by Erdős and Rényi in their groundbreaking paper [37]. In wireless networking, *random geometric graphs* (also known as disk models) [46, 50, 66] have been proposed to model the effects of

geometry and limited communication range. In effect, if the security constraint is removed in our problem, then connectivity between two sensors depends only on whether they are located within communication range, and this is captured by random geometric graphs. In the complementary case when there is "full visibility", i.e., all nodes are within range of each other, the presence of a secure link between two nodes depends only on the existence of a shared secret key. Such networks can be modeled as *random key graphs* [32, 33, 86].

For secure WSNs, however, the random graph model should take into account both the communication range and the shared-key constraints. This can be viewed as taking the intersection of the edge sets of the corresponding random geometric and key graphs.

To proceed with the above model, we examine a related simpler problem. Consider a WSN where the nodes have a finite communication range. Suppose that the link between each pair of nodes can fail independently with a certain probability (which can scale with the number of nodes). In this case, an edge is present between two nodes if they are located within range and the pairwise link is indeed active. The advantage of this model, compared to the latter problem of secure WSNs, is that it eases analysis by eliminating the dependencies between the edges of the key graph. Therefore, this model will serve as a precursor to, and hopefully provide some insight for, the more complicated situation. As an added motivation to study this graph model, note that it can also be viewed as a simple method to include fading in the disk model of networks by thinking of fading as link outage.

1.2.2 Previous Work

Connectivity properties of random key graphs, which apply for secure WSNs under full visibility, have been analyzed by Di Pietro et al. [32, 33] and Yağan and

Makowski [86]. Although the case of partial visibility is mentioned in [32], their analysis assumes the range to be *constant* and does not allow an arbitrary scaling.

In [87], Yi et al. consider a WSN in which each *node* independently becomes inactive with a certain *fixed* probability which does not vary with the number of nodes. For this case, their results go beyond identifying the critical scalings for node isolation; they also provide the asymptotic distribution of the number of isolated nodes for given parameters. Their techniques are further extended in [88] to allow independent *link* failures where the failure probability may scale with the number of nodes. The same flavor of results also apply for secure WSNs. However, the results are derived under additional (non-trivial) technical assumptions on the scalings, and therefore, do not provide a complete characterization of the zero-one laws for this setting.

1.2.3 Main contributions

WSNs with random link failures. In this part of the research, covered in Chapter 9, we study WSNs with random link failures where the sensors are located on a circle. We provide a complete zero-one law establishing the exact threshold scaling which guarantees that with high probability none of the sensors are isolated. In contrast to earlier works cited above, our result uses no additional assumptions on the parameter scalings. We also obtain more general results. For both secure WSNs and networks with random link failures, if the nodes are located on a sphere or a torus in $\mathbb{R}^d$, we prove a one law which establishes sufficient conditions for the scalings so that with high probability, the WSN does not contain isolated nodes.

1.3 Basic Notation and Conventions

Random variables (r.v.'s) will be denoted by capital letters and their realizations by lower-case letters. The probability distribution of an r.v. X will be denoted by P_X. If X and Y are independent r.v.'s, then their joint distribution is written as $P_X \times P_Y$. For positive integers l, m, X_l^{l+m} denotes the collection of r.v.'s $\{X_l, X_{l+1}, \ldots, X_{l+m}\}$, and the shorthand $[l]$ is used to denote the set $\{1, \ldots, l\}$. We use the notation $=_{st}$ to indicate distributional equality. The indicator function of an event E is denoted by $\mathbf{1}[E]$.

Vectors are denoted by boldface letters. For example, $\mathbf{X}$ denotes a random vector, while $\mathbf{x}$ denotes a fixed vector. The Hamming *distance* between vectors $\mathbf{x}, \mathbf{y}$ is defined as $d_H(\mathbf{x}, \mathbf{y}) := \sum_i \mathbf{1}[x_i \neq y_i]$. If S is a set of vectors, we abbreviate $\min_{\mathbf{y} \in S} d_H(\mathbf{x}, \mathbf{y})$ as $d_H(\mathbf{x}, S)$.

The entropy of a random variable with finitely supported distribution will be denoted by $H(X)$. The mutual information of r.v.'s X and Y is written as $I(X;Y)$. For the definition and properties of these and other information-theoretic quantities used below we refer to the books of Csiszar and Körner [31] or Cover and Thomas [30]. All such quantities are defined with logarithms to the base q, which denotes the common support size of the r.v.'s in question. The entropy of a random variable X with p.m.f. $P_X(0) = 1 - x, P_X(1) = \cdots = P_X(q-1) = x/(q-1)$ will be called the q-ary entropy function, denoted $h_q(x)$. Explicitly, $h_q(x) := -x \log_q x/(q-1) - (1-x)\log_q(1-x)$. The information divergence of two such random variables equals $D_q(x\|y) := x \log_q(x/y) + (1-x)\log_q((1-x)/(1-y))$. For $q = 2$, we write simply $h(x)$ and $D(x\|y)$.

For two functions $f(n), g(n)$, we write $f(n) \sim g(n)$ if $\lim_{n\to\infty} f(n)/g(n) = 1$, and $f(n) \doteq g(n)$ if $\lim_{n\to\infty} n^{-1} \log(f(n)/g(n)) = 0$.

1.4 Useful Identities

The following are some standard identities which will be useful in various proofs throughout the dissertation.

1) Let Z be a binomial r.v. with parameters (n, p). Then

$$\mathbf{P}\left[Z \geq n\sigma\right] \leq q^{-nD_q(\sigma\|p)}, \qquad \text{if } \sigma > p \tag{1.1}$$

$$\mathbf{P}\left[Z \leq n\sigma\right] \leq q^{-nD_q(\sigma\|p)}, \qquad \text{if } \sigma < p. \tag{1.2}$$

2)

$$\sum_{k=0}^{\sigma n} \binom{n}{k} (q-1)^k \doteq q^{nh_q(\sigma)}, \qquad \text{if } \sigma \in \left(0, {}^{(q-1)}\!/_{q}\right). \tag{1.3}$$

Chapter 2

Preliminaries on Fingerprinting

This section introduces a formal statement of the fingerprinting problem. It then proceeds to give a definition of the fingerprinting capacity and to derive some of its simple properties that were obtained in the course of the present research.

2.1 The Fingerprinting Problem

Suppose that we need to distribute identical copies of some copyrighted material to M users. The distributor embeds an imperceptible fingerprint in each legal copy of the data. The fingerprints are assumed to be located inside the host message so that their precise position is unknown to the system users. The location of the fingerprints, however, remains the same for all users.

Let n denote the length of the fingerprints. Let $\mathcal{Q}$ denote an alphabet of (finite) size q, usually taken to be $\{0, \ldots, q-1\}$ with modulo q addition. The case $q = 2$ is called the binary alphabet. Assume that there is some ordering of the users and denote their set by $[M] = \{1, \ldots, M\}$.

Definition 2.1. An $(n, M)_q$ *code* (C, D) is a pair of encoding and decoding mappings

$$C : [M] \to \mathcal{Q}^n, \quad D : \mathcal{Q}^n \to [M] \cup \{0\}, \tag{2.1}$$

where the decoder output 0 signifies a decoding failure.

The image of $U \subseteq [M]$ under C is written as $C(U)$. Also, for convenience, we sometimes abuse terminology by calling the range of C a code (or codebook), and use the same notation C for it. Hopefully this ambiguity can be resolved

by the context. Following standard coding-theoretic terminology, the vectors in the codebook are called codewords, and the quantities n and M will be referred to as the code *length* and *size* respectively. The *rate* of this code is given by $R = n^{-1} \log_q M$. The *minimum distance* of C is defined as the smallest (Hamming) distance between any two codewords. If $\mathcal{Q}$ is a finite field, then an $[n, k, d]_q$ *linear code* is defined to be a vector subspace of $\mathcal{Q}^n$ with dimension k and minimum distance d. Please see MacWilliams and Sloane's book [59] for formal definitions and properties of linear codes.

The distributor's strategy of assigning fingerprints to the users may be either deterministic or randomized as explained in the following subsections.

2.1.1 Deterministic Codes

A deterministic assignment of fingerprints is given by an $(n, M)_q$ code (C, D) as defined above in (2.1). A *coalition* of t users is an arbitrary t-subset of $[M]$. Following accepted usage, we will refer to the members of the coalition as "pirates". We assume that the code (C, D) is public and can be used by the pirates in designing their attack. Suppose that the collection of fingerprints assigned to a coalition U, namely $C(U)$, is $\{\mathbf{x}_1, \ldots, \mathbf{x}_t\}$. The coalition attempts to create a pirated copy with a forged fingerprint $\mathbf{y} \in \mathcal{Q}^n$ so as to conceal their identities from the distributor.

Note that although the fingerprint locations are not available to the pirates, they may detect some of these locations by comparing their copies for differences and modify the detected positions. Thus, coordinate i of the fingerprints is called *undetectable* for the coalition U if

$$x_{1i} = x_{2i} = \cdots = x_{ti}$$

and is called *detectable* otherwise.

Definition 2.2 (Marking assumption)**.** The *marking assumption* states that for any fingerprint $\mathbf{y}$ created by the coalition U, $y_i = x_{1i} = x_{2i} = \cdots = x_{ti}$ in every coordinate i that is undetectable.

In other words, in creating $\mathbf{y}$, the pirates can modify only detectable positions. For a given set of observed fingerprints $\{\mathbf{x}_1, \ldots, \mathbf{x}_t\}$, the set of forgeries that can be created by the coalition is called the *envelope*. Its definition depends on the exact rule the coalition should follow to modify the detectable positions:

- If the coalition is restricted to use only a symbol from their assigned fingerprints in the detectable positions, we obtain the *narrow-sense envelope*:

$$\mathcal{E}_N(\mathbf{x}_1, \ldots, \mathbf{x}_t) = \{\mathbf{y} \in \mathcal{Q}^n | y_i \in \{x_{1i}, \ldots, x_{ti}\}, \forall i\}; \tag{2.2}$$

- If the coalition can use any symbol from the alphabet in the detectable positions, we obtain the *wide-sense envelope*:

$$\mathcal{E}_W(\mathbf{x}_1, \ldots, \mathbf{x}_t) = \{\mathbf{y} \in \mathcal{Q}^n | y_i = x_{1i}, \forall i \text{ undetectable}\}. \tag{2.3}$$

We remark that there are further generalizations of the rules above where coalitions are also allowed to erase the symbols in detectable positions. This generalization is not considered below; we refer the interested reader to [15]. In the following, we will use $\mathcal{E}(\cdot)$ to denote the envelope from any of the rules or their generalizations mentioned above.

Remark 2.3. Note that different problems can arise for each definition of the envelope. The binary alphabet is of special interest because of its wide use in practical digital applications. For this special case, it is easy to see that the narrow-sense and wide-sense envelopes are exactly the same.

Given a pirated copy with a forged fingerprint, the distributor's goal is to identify *one of the pirates* using a tracing algorithm. Note that it is impossible to trace all members of the coalition as some members may essentially contribute nothing to the forgery. Naturally, we require the error probability in tracing to be arbitrarily small. However, it was shown in [26] that this objective is unattainable with deterministic codes under the wide-sense formulation (2.3). Therefore, we restrict ourselves to the narrow-sense rule (2.2) in the case of deterministic codes.

Consider the following simple decoding strategy. Given a forgery $\mathbf{y}$, the distributor performs an exhaustive search for all coalitions $T \subseteq [M]$ of size at most t such that $\mathbf{y} \in \mathcal{E}_N(C(T))$ and outputs a user that is common to all these coalitions (an error is declared if there is no common user). Obviously, if there exists such a common user, it is definitely a pirate. This motivates the following definition.

Definition 2.4 (t-IPP)**.** A code C has *t-identifiable parent property* (or is t-IPP) if for any $U \subseteq C$ of size at most t and any $\mathbf{y} \in \mathcal{E}_N(U)$, the following holds:

$$\bigcap_{\substack{T \subseteq C:\ |T| \leq t,\\ \mathbf{y} \in \mathcal{E}_N(T)}} T \neq \emptyset.$$

Since any user in the above intersection is guaranteed to have participated in creating $\mathbf{y}$, this property allows us to identify a guilty user with zero probability of decoding error.

Remark 2.5. For the alphabet $\mathcal{Q} = \{0, \ldots, q-1\}$, the code $C = \{(0, \ldots, 0), \ldots, (q-1, \ldots, q-1)\}$ is a trivial q-IPP code (for any code length). We say a q-ary t-IPP code is *non-trivial* if its size is at least $\max(t, q) + 1$. It was shown in [78] that non-trivial IPP codes do not exist if $t \geq q$.

2.1.2 Randomized Codes

In this section we define randomized fingerprinting codes. Randomization is a powerful method of constructing fingerprinting codes both for the narrow-sense and wide-sense rules (2.2)-(2.3). Moreover, as remarked above, for the latter case any nontrivial fingerprinting code construction with a small error probability must rely on randomization. Even for the narrow-sense rule, randomization can increase the attainable code rates compared to deterministic codes at the cost of a small error probability.

Let $\mathcal{K}$ be a finite set whose size may depend on n. We will refer to the elements of $\mathcal{K}$ as *keys*. A randomized strategy to assign fingerprints is defined as the following random experiment. The distributor has at its disposal a family of codes $\{(C_k, D_k), k \in \mathcal{K}\}$, where each (C_k, D_k) is an $(n, M)_q$ code as defined in (2.1). The distributor chooses a key k according to a probability distribution function $(\pi(k), k \in \mathcal{K})$ and assigns the fingerprints according to C_k. On receiving a forged fingerprint, the distributor uses the tracing strategy D_k, corresponding to the selected key, to determine one of the guilty users. The code resulting from this random experiment is called a *randomized code* and is denoted by $(\mathcal{C}, \mathcal{D})$.

Following the standard convention in cryptography of the system design being publicly available, we allow the users to have knowledge of the family of codes $\{(C_k, D_k)\}$ and the distribution $\pi(\cdot)$, while the exact key choice is kept secret by the distributor.

Consider a coalition U of size t. Suppose that U relies on a randomized strategy $V(\cdot|\cdot, \ldots, \cdot)$ to create a new fingerprint, where $V(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_t)$ is the probability that the coalition creates $\mathbf{y}$ given that it observes the fingerprints $\mathbf{x}_1, \ldots, \mathbf{x}_t$. Our interest is in a special class of strategies which satisfy one of the restrictions (2.2), (2.3), depending on the application, in creating a forgery.

A strategy V is called *admissible* if

$$V(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_t) > 0 \text{ only if } \mathbf{y} \in \mathcal{E}(\mathbf{x}_1, \ldots, \mathbf{x}_t). \tag{2.4}$$

Let $\mathcal{V}_t$ denote the corresponding class of admissible strategies. Since the sequences $\mathbf{y}, \mathbf{x}_i$ are formed of entries chosen from a finite alphabet, such randomized strategies model any general attack the coalition is capable of, and also facilitate mathematical analysis.

Let the random vector $\mathbf{Y}_{\mathcal{C},U,V}$ represent the forgery generated by U using the strategy V in this manner. Assuming that the key k was chosen, the distributor employs the decoder D_k to trace one of the pirates from the observed forged fingerprint. The probability of error for a given coalition U and strategy V averaged over the family of codes is defined as follows:

$$e(\mathcal{C}, \mathcal{D}, U, V) = \mathbf{P}\left[\mathcal{D}(\mathbf{Y}_{\mathcal{C},U,V}) \notin U\right] = \mathbf{E}_K \sum_{\mathbf{y}: D_K(\mathbf{y}) \notin U} V(\mathbf{y}|C_K(U)),$$

where the expectation is taken with respect to the distribution $\pi(k)$.

Definition 2.6 (t-fingerprinting)**.** A randomized code $(\mathcal{C}, \mathcal{D})$ is said to be *t-fingerprinting with ε-error* if

$$\max_{U:\, |U|=\tau} \max_{V \in \mathcal{V}_\tau} e(\mathcal{C}, \mathcal{D}, U, V) \leq \varepsilon, \quad \forall \tau \leq t. \tag{2.5}$$

Remark 2.7. The fingerprinting problems arising from the above definition are of different nature for each of the two envelope restrictions described. However, for the binary alphabet, either choice leads to the same definition since the narrow-sense and wide-sense envelopes are exactly the same.

2.2 Fingerprinting Capacity

We now formulate the fingerprinting problem as a communication problem. The set of messages for the communication problem are identified with the set of users of the fingerprinting system. Each message is mapped to a codeword that corresponds to the fingerprint of the user. Any set of t messages (a coalition) can be chosen, and they are transmitted over an *unknown* communication channel with t inputs and a single output that is defined by the strategy of the coalition. The class of possible channels will be defined by the marking assumption. The output of the channel (that represents the strategy) gives the forged fingerprint. The task of the decoder is to recover at least *one* of the transmitted messages that have produced the channel output.

For readers familiar with the information-theoretic model called a Multiple-Access Channel (MAC), we observe that the above model differs from the traditional t-user MAC because: (a) the decoder makes an error only when its output does not match *any of the transmitted messages*, and (b) all channel inputs are required to use the same codebook.

For a given t-user strategy V, the *maximum* probability of error is given by

$$e_{\max}(\mathcal{C}, \mathcal{D}, V) = \max_{u_1,\ldots,u_t \in [M]} e(\mathcal{C}, \mathcal{D}, \{u_1, \ldots, u_t\}, V). \tag{2.6}$$

It is straightforward to see that the t-fingerprinting condition (2.5) can now be expressed as

$$e_{\max}(\mathcal{C}, \mathcal{D}, V) \leq \varepsilon \text{ for every } V \in \mathcal{V}_t. \tag{2.7}$$

Note that in maximization above the users $u_1, \ldots, u_t$ are not necessarily distinct.

Definition 2.8 (Fingerprinting capacity)**.** A number $R \geq 0$ is an *achievable rate for q-ary t-fingerprinting* if there exists a sequence of $(n, q^{nR_n})_q$ randomized codes

$(\mathcal{C}_n, \mathcal{D}_n)$ such that

$$\liminf_{n\to\infty} R_n = R, \quad \lim_{n\to\infty} \max_{V\in\mathcal{V}_{t,n}} e_{\max}(\mathcal{C}_n, \mathcal{D}_n, V) = 0.$$

The supremum of all such achievable rates is called the *capacity of q-ary t-fingerprinting*, and is denoted by $\mathsf{C}_{t,q}$.

Remark 2.9. A more accurate way to write the above definition involves a sequence of codes $(\mathcal{C}_i, \mathcal{D}_i), i = 1, 2, \ldots$ of (growing) length n_i and rate R_i such that $\liminf_{i\to\infty} R_i = R$. Similar qualifiers apply to the other quantities whose limits are considered above. Following established practice, we use the streamlined notation in relation to capacity and other similar notions throughout the thesis.

2.3 Properties of Fingerprinting Capacity

For later analysis let us establish some simple properties of the fingerprinting capacity $\mathsf{C}_{t,q}$.

Coalitions of size t. Below it will be convenient to rely on coalitions of size exactly t as opposed to $\leq t$. We argue that this restriction does not change the value of $\mathsf{C}_{t,q}$.

Given any t-user strategy V, define the maximum probability of error corresponding to *coalitions of size t alone* as

$$\tilde{e}_{\max}(\mathcal{C}, \mathcal{D}, V) = \max_{U:\, |U|=t} e(\mathcal{C}, \mathcal{D}, U, V). \tag{2.8}$$

The capacity value $\tilde{\mathsf{C}}_{t,q}$ corresponding to the above criterion is defined by substituting $\tilde{e}_{\max}(\mathcal{C}_n, \mathcal{D}_n, V)$ in place of $e_{\max}(\mathcal{C}_n, \mathcal{D}_n, V)$ in Definition 2.8.

Proposition 2.10.

$$\mathsf{C}_{t,q} = \tilde{\mathsf{C}}_{t,q}.$$

Clearly, $\mathsf{C}_{t,q} \leq \tilde{\mathsf{C}}_{t,q}$. The proof of the opposite inequality is also almost obvious because any coalition of t pirates can simply ignore any subset of $t-\tau$ pirates when devising a forged fingerprint $\mathbf{y}$. A formal version of this argument is provided by Lemma 2.17 in the Appendix.

Average vs. maximum error probability. Let us consider the *average* error probability

$$e_{\text{avg}}(\mathcal{C}, \mathcal{D}, V) = \frac{1}{M^t} \sum_{u_1,\ldots,u_t \in [M]} e(\mathcal{C}, \mathcal{D}, \{u_1, \ldots, u_t\}, V). \tag{2.9}$$

and the probability

$$\tilde{e}_{\text{avg}}(\mathcal{C}, \mathcal{D}, V) = \frac{1}{\binom{M}{t}} \sum_{U:\ |U|=t} e(\mathcal{C}, \mathcal{D}, U, V). \tag{2.10}$$

for coalitions of size exactly t. Define the capacities $\mathsf{C}^a_{t,q}$ and $\tilde{\mathsf{C}}^a_{t,q}$ respectively in accordance with the above average error probabilities.

Clearly, the average error criterion is weaker compared to the maximum one and so, we have

Fact 2.11.

$$\mathsf{C}_{t,q} \leq \mathsf{C}^a_{t,q}.$$

For a fixed, known single-user channel, allowing randomized codes makes the capacity under the maximum error probability criterion equal to that under the average error probability criterion (see, e.g., [31, p.223, Prob. 5]). We now extend this argument to the current context of multi-user channels and fingerprinting to show that both (2.8) and (2.10) lead to the same capacity value.

Proposition 2.12.

$$\tilde{\mathsf{C}}_{t,q} = \tilde{\mathsf{C}}^{a}_{t,q}.$$

A formal proof is available in the Appendix. The intuitive argument behind it relies on the fact that here we simply use a randomized code $(\mathcal{C}, \mathcal{D})$, which also includes all $M!$ permutations of any specific realization of $(\mathcal{C}, \mathcal{D})$. Because of the symmetry introduced by this, the error probability $e(\mathcal{C}, \mathcal{D}, U, V)$ is the same for all coalitions for a given V, and hence the average and the maximum probability are the same.

2.4 Related Combinatorial Properties

In this section, we recall a few other combinatorial properties of deterministic codes that are related to t-IPP and t-fingerprinting, and have been previously studied in the literature.

Definition 2.13 (t-TA)**.** [27, 78] A code C has *t-traceability property* (or is t-TA) if for any $U \subseteq C$ of size at most t and any $\mathbf{y} \in \mathcal{E}_N(U)$, the following holds:

$$d_H(U, \mathbf{y}) < d_H(C \backslash U, \mathbf{y}).$$

In essence, the above definition implies that for a t-TA code, we can trace one of the pirates (with zero error) by simply finding the user whose fingerprint is the closest to the forgery.

Definition 2.14 (t-frameproof)**.** [78] A code C is *t-frameproof* if for any $U \subseteq C$ of size at most t, it holds that:

$$\mathcal{E}_N(U) \cap (C \backslash U) = \emptyset.$$

In other words, any t pirates cannot produce the fingerprint of an innocent user not part of the actual coalition. Observe that, in relation to t-IPP, the t-TA property is stricter, while the latter t-frameproof property is weaker. This is summarized below.

Fact 2.15. [78] *t-TA $\Rightarrow$ t-IPP $\Rightarrow$ t-frameproof.*

The following sufficient conditions are well-known for t-TA and t-frameproof codes.

Theorem 2.16. [27] *For a code C of length n if the minimum distance d satisfies*

(a) $d > n(1 - 1/t^2)$, then C is t-TA.

(b) $d > n(1 - 1/t)$, then C is t-frameproof.

2.5 Concatenated Codes

A commonly used technique in the construction of codes with efficient decoding algorithms is Forney's *code concatenation* [41]. Concatenated codes are obtained as follows. We start with two codes, namely, an "outer" code $(C_{\text{out}}, D_{\text{out}})$ over an alphabet of size q and an "inner" code $(C_{\text{in}}, D_{\text{in}})$ over a smaller (say binary) alphabet with q codewords. To encode a given message index, we first use C_{out} to obtain a q-ary codeword. Next, each of the q-ary symbols in the resulting codeword is encoded using C_{in} to finally produce a longer binary word. Decoding follows the opposite sequence where D_{in} first retrieves the q-ary symbol corresponding to the binary word at every outer level coordinate. The resulting q-ary vector is then decoded using D_{out} to obtain the message index. The effective rate of the concatenated code is the product of the inner and outer code rates.

To give a formal definition, let $M = q^K$ be the set of messages to be encoded with the concatenated code. The encoding mapping C is defined as a composition

of two maps $C_{\text{out}} : [M] \to \mathcal{Q}^N$ and $C_{\text{in}} : \mathcal{Q} \to \{0,1\}^m$ acting as follows: for $i \in M$

$$C(i) = C_{\text{in}}^{\times N}(C_{\text{out}}(i)), \tag{2.11}$$

where $C_{\text{in}}^{\times N}$ is the N-fold extension of the inner encoding mapping applied in each coordinate of the codevector $C_{\text{out}}(i)$. The decoding mapping $D : \{0,1\}^{Nm} \to [M]$ is the corresponding composition of the maps D_{in} and D_{out} applied to $\mathbf{x} \in \{0,1\}^{Nm}$ as follows:

$$D(\mathbf{x}) = D_{\text{out}}(D_{\text{in}}^{\times N}(\mathbf{x})). \tag{2.12}$$

Typically, the inner code is chosen randomly from some ensemble of codes with high rate, while the outer level is a structured code with polynomial-time decoding. Concatenated codes combine structure and randomness to realize the benefits of both high rates and efficient algorithms.

In the case of deterministic codes, it is a simple and well-known fact that if both the inner and outer codes are t-frameproof (resp., t-IPP) then the concatenated code is also t-frameproof (resp., t-IPP). We employ concatenation in Chapters 6 and 8 for constructing efficient randomized codes for copyright protection.

2.6 Appendix

2.6.1 A lemma on the size of coalitions

Lemma 2.17. *Let $(\mathcal{C}, \mathcal{D})$ be a randomized code of size at least $2t-1$. Assume that*

$$\tilde{e}_{\max}(\mathcal{C}, \mathcal{D}, V) \leq \varepsilon \text{ for every } V \in \mathcal{V}_t. \tag{2.13}$$

Then for any $\tau \leq t$,

$$\tilde{e}_{\max}(\mathcal{C}, \mathcal{D}, V) \leq 2\varepsilon \text{ for every } V \in \mathcal{V}_\tau.$$

Proof. For simplicity of presentation we take $\tau = t - 1$. The general case of $1 \leq \tau < t$ can be established with only minor changes to the proof below. For any $V \in \mathcal{V}_{t-1}$, let us define a $V' \in \mathcal{V}_t$ where

$$V'(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}, \mathbf{x}_t) = V(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}), \quad \forall \mathbf{x}_1, \ldots, \mathbf{x}_t, \mathbf{y} \in \mathcal{Q}^n.$$

Then, for any coalition U of size $t-1$, and any user $u \notin U$,

$$\begin{aligned} e(\mathcal{C}, \mathcal{D}, U, V) &= \mathbf{E}_K \sum_{\substack{\mathbf{y}: \\ D_K(\mathbf{y}) \notin U}} V(\mathbf{y}|C_K(U)) \\ &= \mathbf{E}_K \sum_{\substack{\mathbf{y}: \\ D_K(\mathbf{y}) \notin U}} V'(\mathbf{y}|C_K(U), C_K(u)) \\ &= \mathbf{E}_K \left[\sum_{\substack{\mathbf{y}: \\ D_K(\mathbf{y}) \notin U'}} V'(\mathbf{y}|C_K(U')) + \sum_{\substack{\mathbf{y}: \\ D_K(\mathbf{y}) = u}} V'(\mathbf{y}|C_K(U')) \right], \end{aligned} \tag{2.14}$$

where $U' = U \cup \{u\}$. The first term in the last equation satisfies

$$e(\mathcal{C}, \mathcal{D}, U', V') \leq \varepsilon \tag{2.15}$$

by the assumption of the lemma. We will next show that the second term in (2.14) is also at most ε. Suppose for the sake of contradiction that

$$\mathbf{E}_K \sum_{\substack{\mathbf{y}: \\ D_K(\mathbf{y}) = u}} V'(\mathbf{y}|C_K(U')) > \varepsilon.$$

Let $u' \notin U'$ and $U'' = U \cup \{u'\}$ (we assume that the size of the code is at least $t+2$, or at least $2t-1$ in the general case). Then

$$\begin{aligned} e(\mathcal{C}, \mathcal{D}, U'', V') &= \mathbf{E}_K \sum_{\substack{\mathbf{y}: \\ D_K(\mathbf{y}) \notin U''}} V'(\mathbf{y}|C_K(U'')) \\ &\geq \mathbf{E}_K \sum_{\mathbf{y}: D_K(\mathbf{y})=u} V'(\mathbf{y}|C_K(U')) > \varepsilon. \end{aligned}$$

But this contradicts our initial assumption (2.13). ■

2.6.2 Proof of Proposition 2.12

It is clear that $\tilde{\mathsf{C}}_{t,q} \leq \tilde{\mathsf{C}}^a_{t,q}$. Therefore, it is enough to show that for every randomized code $(\mathcal{C}, \mathcal{D})$, there exists another randomized code $(\mathcal{C}^*, \mathcal{D}^*)$ of the same rate such that $\tilde{e}_{\max}(\mathcal{C}^*, \mathcal{D}^*, V) = \tilde{e}_{\text{avg}}(\mathcal{C}, \mathcal{D}, V)$ for every channel V.

We are given $\{(C_k, D_k), k \in \mathcal{K}\}$. Let $\sigma \in \Sigma$ identify a particular permutation from the set of all permutations on the message set $[M]$. Choose σ uniformly at random from Σ and construct a new key $\kappa \triangleq (k, \sigma)$. Define

$$C^*_\kappa(\cdot) \triangleq C_k(\sigma(\cdot)), \quad D^*_\kappa(\cdot) \triangleq \sigma^{-1}(D_k(\cdot)).$$

Let $(\mathcal{C}^*, \mathcal{D}^*)$ be the randomized code corresponding to the family $\{(C^*_\kappa, D^*_\kappa), \kappa \in \mathcal{K} \times \Sigma\}$. Then, for every channel V, $\tilde{e}_{\text{avg}}(\mathcal{C}^*, \mathcal{D}^*, V) = \tilde{e}_{\text{avg}}(\mathcal{C}, \mathcal{D}, V)$. Furthermore, for any $U \subseteq [M]$, $|U| = t$,

$$\begin{aligned} &e(\mathcal{C}^*, \mathcal{D}^*, U, V) \\ &= \frac{1}{M!} \sum_{\sigma \in \Sigma} \sum_{k \in \mathcal{K}} \pi(k) \sum_{\substack{\mathbf{y}: \\ D_k(\mathbf{y}) \notin \sigma(U)}} V(\mathbf{y}|C_k(\sigma(U))) \end{aligned}$$

which does not depend on the subset U because of the averaging over all permutations. This implies $\tilde{e}_{\max}(\mathcal{C}^*, \mathcal{D}^*, V) = \tilde{e}_{\text{avg}}(\mathcal{C}^*, \mathcal{D}^*, V)$. ■

Chapter 3

Lower Bounds on Fingerprinting Capacity

Lower bounds on fingerprinting capacity known in the literature are implied by various constructions of fingerprinting codes. Binary codes have been of particular interest due to their prevalence in practical applications, and thus will be our main focus in this chapter.

In [26, 15], the authors introduced randomness at some stages of code construction, relying on deterministic codes otherwise. It was later realized [79, 21] that better code rates are obtained by considering randomization over the entire family of possible binary codes. In particular, [79] constructed (binary) t-fingerprinting codes which established that $\mathsf{C}_{t,2} \geq 1/(100t^2)$. There were subsequent improvements of the constant 100, see for instance [23, 64, 65, 73, 74], all obtained by optimizing the parameters in [79]. For coalitions of size 2, [21] showed that one can obtain much higher code rates than that, proving the estimate $\mathsf{C}_{2,2} \geq 0.2075$.

In this chapter, we present ***new lower bounds*** for fingerprinting capacity obtained via two different techniques. The first method ties the ideas from [21] with the notion of minimal configurations to obtain a better lower bound for size-3 coalitions, which is also generalized further to an arbitrary number of pirates. Secondly, we improve upon the above results through a new idea which takes into account only typical allocations of fingerprints to coalitions, i.e., allocations that occur with high probability, discarding all the other possibilities. As it turns out, this idea generally leads to more powerful results than the method of minimal configurations.

3.1 Lower Bound I

3.1.1 Code Generation

Encoding: Fix a rate $R \in (0, 1]$ and define $M_n = \lfloor 2^{nR} \rfloor$. Consider the family of all possible binary (n, M_n) codes (encoding mappings). We choose a code at random from this family using the uniform distribution that assigns the probability 2^{-nM_n} to each of them (equivalently, we choose M_n fingerprints independently with $p(0) = p(1) = 1/2$ in each coordinate). For $R < 1/2$, the fingerprints will be distinct with high probability.

Decoding: We assume the following strategy to identify pirates. Suppose the assignment C_k is chosen. The distributor performs an exhaustive search for all coalitions $T \subseteq [M_n]$ of size at most t such that the suspect fingerprint $\mathbf{y} \in \mathcal{E}(C_k(T))$ and outputs users that are common to all these coalitions (and declares an error if there are no such users). The $(n, M_n)_2$ randomized code thus obtained is denoted by $(\mathcal{C}_n, \mathcal{D}_n)$.

At this point, the reader may notice a connection to t-IPP codes as the same decoding strategy was mentioned to motivate their definition in Chapter 2. Indeed, the choice of the above strategy enables us to use some elements of the analysis of IPP codes from [16] in the current problem.

3.1.2 Minimal Configurations

Definition 3.1. A set of coalitions $\mathcal{U} = \{U_i\}$ is called a *configuration* if they do not have a user common to all of them, i.e., $\bigcap_i U_i = \emptyset$. When every coalition in the configuration has at most t users, we call it a *configuration of degree* t.

The concept of configurations together with the identification strategy chosen enable us to estimate the probability of decoding error. We will use the following

Definition 3.2 (Minimal configuration)**.** A configuration $\mathcal{U} = \{U_1, \ldots, U_r\}$ is called *minimal* if the subset $\{U_1, \ldots, U_{i-1}, U_{i+1}, \ldots, U_r\}$ is not a configuration for all $i \in [r]$.

Remark 3.3. Any configuration contains a minimal configuration.

Configurations can be conveniently represented by hypergraphs. Each node in the hypergraph corresponds to a user and a hyperedge connects the nodes (users) in a coalition. We say that two minimal configurations are of the same *type* if their hypergraph representations are isomorphic.

Example 3.4. The minimal configurations of degree 2 can be of the two types, *Separation* or *Triangle*. Separation represents two disjoint subsets of size 2, while Triangle represents three subsets with the following structure: $\{a, b\}, \{b, c\}, \{c, a\}$.

The following result from [16] will be used below.

Theorem 3.5. *Let $\mathcal{U} = \{U_1, \ldots, U_r\}$ be a minimal configuration of degree t. Then $r \leq t + 1$ and the only minimal configuration with $r = t + 1$ is a complete t-uniform hypergraph with $t + 1$ edges (t-simplex). Moreover,*

$$\Big|\bigcup_{i=1}^{r} U_i\Big| \leq \Big\lfloor \Big(\frac{t}{2} + 1\Big)^2 \Big\rfloor.$$

3.1.3 Analyzing the Error Probability

Our aim is to establish that the probability of decoding error for the randomized codes $(\mathcal{C}_n, \mathcal{D}_n), n = 1, 2, \ldots,$ constructed above is decaying to zero (so long as R is less than a certain value). Let U_1 be a coalition of size t. As above, let $e(\mathcal{C}_n, \mathcal{D}_n, U_1, V)$ be the error probability of identification using some strategy V. For a vector $\mathbf{y} \in \{0, 1\}^n$ and coalitions $U_1, \ldots, U_r$ define the event

$$\tau_{\mathbf{y}}(U_1, \ldots, U_r) = \Big[\mathbf{y} \in \bigcap_{i=1}^{r} \mathcal{E}(\mathcal{C}_n(U_i))\Big]$$

where the randomness is in the selection of the code $\mathcal{C}_n$. The next proposition justifies the usefulness of minimal configurations.

Proposition 3.6. *Let $S = \{U_1, \ldots, U_r\}$ be a degree-t minimal configuration that contains U_1. For all $\mathbf{y} \in \{0,1\}^n$ if*

$$\mathbf{P}\left[\exists S:\ \tau_{\mathbf{y}}(S \backslash U_1) | \tau_{\mathbf{y}}(U_1)\right] \stackrel{n\to\infty}{\longrightarrow} 0 \tag{3.1}$$

then the error probability $e(\mathcal{C}_n, \mathcal{D}_n, U_1, V)$ approaches zero.

Proof. By Lemma 2.17, it suffices to consider only coalitions of size exactly t. Let U_1 be such a coalition and $V \in \mathcal{V}_t$ be an admissible strategy. The following calculation holds for every n, so for simplicity we drop n from the notation. We have

$$\begin{aligned} e(\mathcal{C}, \mathcal{D}, U_1, V) &= \mathbf{P}\left[\mathcal{D}(\mathbf{Y}_{\mathcal{C},U_1,V}) \notin U_1\right] \\ &= \sum_{\mathbf{y}\in\{0,1\}^n} \mathbf{P}\left[\mathbf{Y}_{\mathcal{C},U_1,V} = \mathbf{y}\right] \mathbf{P}\left[\mathcal{D}(\mathbf{y}) \notin U_1 \middle| \mathbf{Y}_{\mathcal{C},U_1,V} = \mathbf{y}\right]. \end{aligned} \tag{3.2}$$

Next, consider the inner conditional probability. Recall that $\mathcal{D}$ considers all the t-coalitions whose envelopes contain $\mathbf{y}$. Clearly U_1 is such a coalition. An error will occur if there exists a configuration each member of which can generate $\mathbf{y}$. (Note that if all the coalitions whose envelopes contain $\mathbf{y}$ intersect, then any member of the intersection is also a member of U_1, so no error is made). It suffices to consider only degree-t minimal configurations. We obtain

$$\begin{aligned} \mathbf{P}\left[\mathcal{D}(\mathbf{y}) \notin U_1 \middle| \mathbf{Y}_{\mathcal{C},U_1,V} = \mathbf{y}\right] &= \mathbf{P}\left[\exists S:\ \tau_{\mathbf{y}}(S\backslash U_1) \middle| \mathbf{Y}_{\mathcal{C},U_1,V} = \mathbf{y}\right]. \\ &= \mathbf{P}\left[\exists S:\ \tau_{\mathbf{y}}(S\backslash U_1) \middle| \tau_{\mathbf{y}}(U_1)\right]. \end{aligned}$$

To explain the second equality, observe that the condition $\mathbf{Y}_{\mathcal{C},U_1,V} = \mathbf{y}$ is formed of two events, one being $\tau_{\mathbf{y}}(U_1)$ and the other that $\mathbf{y}$ is selected by the members of the coalition U_1 relying on the strategy V. The second event can be dropped from the conditioning without changing the resulting probability. Substituting the result obtained in (3.2) and using the assumption of the proposition concludes the proof.

■

Now observe that

$$\mathbf{P}\left[\exists S:\ \tau_{\mathbf{y}}(S\backslash U_1)\middle|\tau_{\mathbf{y}}(U_1)\right] = \frac{\mathbf{P}\left[\exists S:\ \tau_{\mathbf{y}}(S)\right]}{\mathbf{P}\left[\tau_{\mathbf{y}}(U_1)\right]}. \tag{3.3}$$

Here the denominator equals

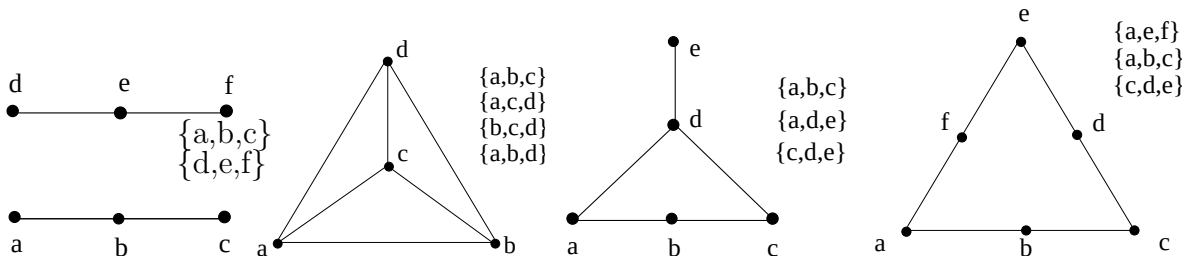
$$\mathbf{P}\left[\tau_{\mathbf{y}}(U_1)\right] = (1-2^{-t})^n,$$

since for each coordinate i only one set of values for the coalition's fingerprints $\{y_i+1,\ldots,y_i+1\}$ fails the marking assumption constraint. On the other hand, the numerator concerns the existence of a degree-t configuration containing the actual coalition U_1 such that all coalitions in the configuration are capable of creating the forgery $\mathbf{y}$.

To verify that the error probability (3.3) indeed approaches 0, it is sufficient to show that, for a given t, this holds true for degree-t *minimal* configurations of each possible type as long as the total number of types does not grow with n. We will follow this approach in the next sections.

3.1.4 Coalitions of Size 3

It was proved in [21] that $\mathsf{C}_{2,2} \geq 0.2075$. The proof goes by showing that for any rate $R < 1 - (1/2)\log_2 3 = 0.2075$, the sequence $(\mathcal{C}_n, \mathcal{D}_n)$ is 2-fingerprinting



(a) Separation (b) Simplex (c) Modified Triangle (d) Triangle

Figure 3.1: Types of minimal configurations of degree 3

with the error probability of identification falling exponentially with n. Here we analyze the resilience of this code sequence against coalitions of size 3 and establish the following result.

Theorem 3.7. *The capacity of 3-fingerprinting over the binary alphabet satisfies*

$$\mathsf{C}_{3,2} \geq 0.064.$$

Proof. The result is established by proving that the randomized code $(\mathcal{C}_n, \mathcal{D}_n)$ is 3-fingerprinting with error probability decaying exponentially in n for any rate

$$R < 1 - \frac{1}{3}\log_2 7. \tag{3.4}$$

The number of vertices (users) in a minimal configuration of degree 3 is at least 4 and at most 6 (by Theorem 3.5). By inspection, there are four different types of minimal configurations as shown in Figure 3.1 with coalitions listed in the figure.

For a given vector $\mathbf{y}$, let $p_e^{(i)}, i = 1, \ldots, 4$ be the error probability corresponding to (3.3) computed for minimal configurations of type i. By symmetry, $p_e^{(i)}$ does not depend on $\mathbf{y}$. Consider two disjoint size-3 coalitions, say $\{a, b, c\}$ and $\{d, e, f\}$, with fingerprints $\{\mathbf{x}(a), \mathbf{x}(b), \mathbf{x}(c)\}$, and $\{\mathbf{x}(d), \mathbf{x}(e), \mathbf{x}(f)\}$. In any coordinate i, the vector $\{x_i(a), x_i(b), \ldots, x_i(f)\}$ can take any of 2^6 equally probable values each of

Table 3.1: Bad assignments for degree-3 minimal configurations of Type 3. Boldface 0's indicate the coalition that cannot produce $y_i = 1$.

$x_i(a)$	$x_i(b)$	$x_i(c)$	$x_i(d)$	$x_i(e)$
0	**0**	**0**	0	0
0	**0**	**0**	0	1
0	**0**	**0**	1	0
0	**0**	**0**	1	1
0	0	1	**0**	**0**
0	1	0	**0**	**0**
0	1	1	**0**	**0**
1	0	**0**	**0**	**0**
1	1	**0**	**0**	**0**

which will be called an *assignment*. An assignment for some coordinate i for a given minimal configuration is "bad" if at least one coalition of the configuration cannot create the target y_i, i.e., the fingerprint vectors of the users in this coalition all have the value $y_i + 1$ in this coordinate. For Type 1, there are 15 such bad assignments. Hence, by the union bound, we have:

$$p_e^{(1)} \leq M_n^3 \frac{(1 - 15/64)^n}{(1 - 1/8)^n} \leq 2^{n3R} \left(\frac{7}{8}\right)^n.$$

For the simplex, whenever three or more of the four users have the value $y_i + 1$ in some coordinate i, the assignment is bad. Hence, we have 5 out of 16 equally probable values that are bad implying

$$p_e^{(2)} \leq M_n \frac{(1 - 5/16)^n}{(1 - 1/8)^n} \leq 2^{nR} \left(\frac{11}{14}\right)^n.$$

To estimate $p_e^{(i)}, i = 3, 4$ let us assume w.l.o.g. that $\mathbf{y} = 1^n$. The bad assignments are shown in Table 3.1 and Table 3.2 for $i = 3$ and $i = 4$, respectively (in each row, the boldface 0's indicate the coalition that cannot produce y_i).

Table 3.2: Bad assignments for degree-3 minimal configurations of Type 4. Boldface 0's indicate the coalition that cannot produce $y_i = 1$.

$x_i(a)$	$x_i(b)$	$x_i(c)$	$x_i(d)$	$x_i(e)$	$x_i(f)$
0	**0**	**0**	0	0	0
0	**0**	**0**	0	0	1
0	**0**	**0**	0	1	0
⋮	⋮	⋮	⋮	⋮	⋮
0	**0**	**0**	1	1	1
0	1	**0**	**0**	**0**	0
0	1	**0**	**0**	**0**	1
1	0	**0**	**0**	**0**	0
1	0	**0**	**0**	**0**	1
1	1	**0**	**0**	**0**	0
1	1	**0**	**0**	**0**	1
0	1	0	1	**0**	**0**
0	0	1	0	**0**	**0**
0	0	1	1	**0**	**0**
0	1	1	0	**0**	**0**
0	1	1	1	**0**	**0**

We conclude that

$$p_e^{(3)} \leq M_n^2 \frac{(1-9/32)^n}{(1-1/8)^n} \leq 2^{n2R} \left(\frac{23}{28}\right)^n,$$

$$p_e^{(4)} \leq M_n^3 \frac{(1-19/64)^n}{(1-1/8)^n} \leq 2^{n4R} \left(\frac{45}{56}\right)^n.$$

Taking R to satisfy (3.4), it is easy to see that $p_e^{(i)}$ tends to zero (exponentially fast) for all types i. Hence, the total probability of identification error goes to zero (exponentially fast). ■

3.1.5 Coalitions of Arbitrary Size

The case study undertaken for $t = 3$ does not extend readily to larger t because it is difficult to list out all possible types of minimal configurations. We will rely on the following lower bound on the number of bad assignments to work around this.

Lemma 3.8. *For any minimal configuration of degree t with N users, the number of bad assignments in any given coordinate is at least $(3/2)2^{N-t}$.*

Proof. W.l.o.g. let the target vector be $\mathbf{y} = 1^n$. Any minimal configuration (type) contains at least two coalitions, say U_1 and U_2. An assignment is bad if it assigns 0 to all members of one of the coalitions. Let b be the number of bad assignments. We have

$$b \geq 2^{N-|U_1|} + 2^{N-|U_1 \cup U_2|}(2^{|U_1 \setminus U_2|} - 1)$$

where the first term takes into account all assignments where U_1 is all-zero and the second term adds all assignments where U_2 is all-zero but U_1 is not all-zero. We obtain

$$b \geq 2^{N-t} + 2^{N-t}(1 - 2^{-|U_1 \setminus U_2|}) \geq \frac{3}{2}2^{N-t}.$$

■

Theorem 3.9. *The capacity of t-fingerprinting over the binary alphabet satisfies*

$$\mathsf{C}_{t,2} \geq -\frac{4}{t^2+4} \log_2 \left(1 - \frac{2^{-(t+1)}}{1 - 2^{-t}}\right).$$

Remark 3.10. This bound implies that $\mathsf{C}_{t,2} \approx \Omega\left(\frac{1}{t^2 2^t}\right)$.

Proof. We show that the randomized code $(\mathcal{C}_n, \mathcal{D}_n)$ is t-fingerprinting with error probability decaying exponentially in n for any rate

$$R < -\frac{4}{t^2+4} \log_2 \left(1 - \frac{2^{-(t+1)}}{1 - 2^{-t}}\right). \tag{3.5}$$

For a minimal configuration of a given type i denote by $N(i)$ the number of users in it. Let $b(i)$ be the number of bad assignments for such a configuration. The error probability can be bounded above as follows:

$$\begin{aligned} p_e^{(i)} &\leq M_n^{N(i)-t} \frac{(1-b(i)2^{-N(i)})^n}{(1-2^{-t})^n} \\ &\leq M_n^{N(i)-t} \left(\frac{1-\frac{3}{2}2^{-t}}{1-2^{-t}} \right)^n \\ &\leq 2^{nR(t^2+4)/4} \left(1 - \frac{2^{-(t+1)}}{1-2^{-t}} \right)^n \end{aligned}$$

where the last two inequalities were obtained using Lemma 3.8 and Theorem 3.5, respectively. Therefore, with R chosen as stated in (3.5), the error probability $p_e^{(i)}$ tends to 0 as n increases.

In order to prove that the total error probability tends to zero as well, it suffices to show that the number of types of degree-t minimal configurations (nonisomorphic t-uniform hypergraphs) does not grow with n. This is obvious because this number does not exceed the number of t-uniform hypergraphs on $s = (t/2+1)^2$ vertices with $i \leq t+1$ edges which can be crudely bounded above by

$$\sum_{i=2}^{t+1} \prod_{j=0}^{i-1} \left[\binom{s}{t} - j \right] \leq t \binom{s}{t}^{t+1}.$$

This completes the proof of the theorem. ■

3.2 Lower Bound II

In this section, we improve upon the previous lower bounds and obtain better capacity estimates for coalitions of size 2 and 3. The improvement is obtained by tailoring the decoder for the *typical* allocations of codewords to coalitions, i.e., the allocations of codewords that occur with high probability. (We say that an event

occurs with high probability if the probability that it fails is at most $\exp(-cn)$, where c is a positive constant.) The fingerprinting codes we construct have error probability decaying exponentially in the code length.

3.2.1 Code Generation

Encoding: The encoding method is the same as in Section 3.1.1. For a fixed rate $R \in (0, 1]$, the encoding mapping $\mathcal{C}_n$ is obtained by choosing $M_n = \lfloor 2^{nR} \rfloor$ fingerprints uniformly and independently (with replacement) from all the 2^n different vectors.

Decoding: The decoding strategy takes advantage of the coalitions that occur with high probability. Suppose that for every n there exists a set $\mathcal{T}_{t,n} \subseteq (\mathcal{Q}^n)^t$ such that for any coalition U of size t, the observed fingerprints $\mathcal{C}_n(U)$ belong to $\mathcal{T}_{t,n}$ with high probability. We will call a set with this property a *typical* set. An explicit characterization of typical sets for the cases $t = 2, 3$ will be given later.

Given a forged fingerprint $\mathbf{y}$, the decoder considers only coalitions corresponding to typical fingerprint assignments $(\mathbf{x}_1, \ldots, \mathbf{x}_t) \in \mathcal{T}_{t,n}$ from the codebook. Each such t-tuple is then discarded if the fingerprints in it simultaneously disagree with $\mathbf{y}$ in any position s, i.e., $x_{1s} = \cdots = x_{ts} \neq y_s$. Note that the remaining t-tuples contain $\mathbf{y}$ in their envelope. Finally, we pick any t-tuple $(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ from those remaining and complete the decoding by choosing a user whose fingerprint $\mathbf{x}_i$ has the smallest Hamming distance $d_i = d_H(\mathbf{y}, \mathbf{x}_i)$. The $(n, M_n)_2$ randomized code obtained by the above procedure is denoted by $(\mathcal{C}_n, \mathcal{D}_n)$ below.

3.2.2 Analyzing the Error Probability

As before, we wish to show that the probability of decoding error for the sequence of randomized codes $(\mathcal{C}_n, \mathcal{D}_n), n = 1, 2, \ldots,$ vanishes if the rate R is suffi-

ciently small. For any coalition U of size t and strategy $V \in \mathcal{V}_t$, we observe

$$\begin{aligned} e(\mathcal{C}_n, \mathcal{D}_n, U, V) &= \mathbf{P}\left[\mathcal{D}_n(\mathbf{Y}_{\mathcal{C}_n,U,V}) \notin U\right] \\ &\sim \sum_{(\mathbf{x}_1,\ldots,\mathbf{x}_t)\in\mathcal{T}_{t,n}} \mathbf{P}\left[\mathcal{C}_n(U) = (\mathbf{x}_1,\ldots,\mathbf{x}_t)\right] \sum_{\mathbf{y}\in\mathcal{E}(\mathbf{x}_1,\ldots,\mathbf{x}_t)} V(\mathbf{y}|\mathbf{x}_1,\ldots,\mathbf{x}_t) \\ &\qquad \times \mathbf{P}\left[\mathcal{D}_n(\mathbf{y}) \notin U \middle| \mathcal{C}_n(U) = (\mathbf{x}_1,\ldots,\mathbf{x}_t), \mathbf{Y}_{\mathcal{C}_n,U,V} = \mathbf{y}\right]. \end{aligned}$$

Thus, it suffices to study the conditions that allow us to obtain the probability

$$\mathbf{P}\left[\mathcal{D}_n(\mathbf{y}) \notin U \middle| \mathcal{C}_n(U) = (\mathbf{x}_1,\ldots,\mathbf{x}_t), \mathbf{Y}_{\mathcal{C}_n,U,V} = \mathbf{y}\right] \overset{n\to\infty}{\longrightarrow} 0$$

for any coalition U of size t, any typical t-tuple $(\mathbf{x}_1,\ldots,\mathbf{x}_t)$ of observed fingerprints, and any forgery $\mathbf{y} \in \mathcal{E}(\mathbf{x}_1,\ldots,\mathbf{x}_t)$.

3.2.3 Coalitions of Size 2

Theorem 3.11. *The capacity of 2-fingerprinting over the binary alphabet satisfies*

$$\mathsf{C}_{2,2} \geq 1/4.$$

Proof. We prove below that $(\mathcal{C}_n, \mathcal{D}_n)$ is 2-fingerprinting with exponentially falling error probability if the rate R satisfies

$$R < 1/4. \tag{3.6}$$

Given a small $\varepsilon > 0$, we define the typical set as the set of vector pairs which agree in l positions, where

$$l \in I_\varepsilon := \left[n\left(1/2 - \varepsilon\right), n\left(1/2 + \varepsilon\right)\right].$$

Consider any two pirates u_1 and u_2. Notice that their observed fingerprints $(\mathbf{x}_1, \mathbf{x}_2)$ belong to the typical set with high probability. Suppose $(\mathbf{x}_1, \mathbf{x}_2)$ is a typical pair with $l \in I_\varepsilon$ agreements. To create a forged fingerprint $\mathbf{y}$, the pirates must fill the remaining $n - l$ positions. Let $d_1 = d_H(\mathbf{y}, \mathbf{x}_1)$ and $d_2 = d_H(\mathbf{y}, \mathbf{x}_2)$. Then $n - l \in I_\varepsilon$ and therefore

$$d_1 + d_2 \in I_\varepsilon. \tag{3.7}$$

We now analyze the probability of decoding error. Obviously, the fingerprints $\mathbf{x}_1$ and $\mathbf{x}_2$ that belong to the factual pirates will not be discarded by the decoding algorithm. The following probabilistic analysis shows that for two innocent users, the decoder discards their observed fingerprints $(\mathbf{z}_1, \mathbf{z}_2)$ with high probability if the code rate satisfies (3.6).

Indeed, for $(\mathbf{z}_1, \mathbf{z}_2)$ to be typical, they should agree in $l \in I_\varepsilon$ positions. In all these positions, $\mathbf{z}_1, \mathbf{z}_2$ should also agree with $\mathbf{y}$ to fulfill the marking assumption. In each of the remaining $n - l$ positions, the vectors $\mathbf{z}_1, \mathbf{z}_2$ are represented by only two combinations, $(0, 1)$ or $(1, 0)$. The probability of choosing such a pair $(\mathbf{z}_1, \mathbf{z}_2)$ in our random code equals

$$P_l = \binom{n}{l} 2^{n-l}/2^{2n}$$

and has exponential order of $2^{-n/2}$ for any $l \in I_\varepsilon$. Furthermore, by the union bound, the total probability of choosing such a pair in a random code of size $M_n = \lfloor 2^{nR} \rfloor$ is at most

$$\binom{M_n}{2} \sum_{l \in I_\varepsilon} \binom{n}{l} 2^{n-l}/2^{2n}.$$

This probability tends to 0 exponentially fast for any rate $R < 0.25$.

Similarly, consider a coalition $(\mathbf{x}_1, \mathbf{z}_2)$ that includes the fingerprint $\mathbf{x}_1$ of an actual pirate and the fingerprint $\mathbf{z}_2$ of an innocent user. Recall that $\mathbf{x}_1$ disagrees with $\mathbf{y}$ in d_1 positions. Then to be output instead of $\mathbf{x}_1$, the fingerprint $\mathbf{z}_2$ must agree with $\mathbf{y}$ in these positions and disagree with it in another set of $d_2 \leq d_1$ positions.

The total probability of choosing such a fingerprint $\mathbf{z}_2$ is at most

$$M_n 2^{n-d_1}/2^n.$$

Since $d_1 + d_2 \in I_\varepsilon$ and $d_2 \leq d_1$, we have the restriction $d_2 \leq n/2\,(1/2 + \varepsilon)$. In this case, the above probability tends to 0 exponentially fast for any rate $R < 0.25$. Thus, with high probability, at least one pirate will be output by the decoder, and no remaining (innocent) users will be chosen as pirates. ■

3.2.4 Coalitions of Size 3

Theorem 3.12. *The capacity of 3-fingerprinting over the binary alphabet satisfies*

$$C_{3,2} \geq 1/12.$$

Proof. We will show that the error probability for $(\mathcal{C}_n, \mathcal{D}_n)$ with 3 pirates approaches 0 if the rate

$$R < 1/12. \tag{3.8}$$

For a triple $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, let

$$\mathcal{L} = \{s \in [n] : x_{1s} = x_{2s} = x_{3s}\},$$
$$\mathcal{L}_{ij} = \{s \in [n] : x_{is} = x_{js}\}, \;\; i, j = 1, 2, 3, \;\; i \neq j,$$

and let $l = |\mathcal{L}|$, $l_{ij} = |\mathcal{L}_{ij}|$. Given a small $\varepsilon > 0$, we say that $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ form a typical triple if

$$l \in J_\varepsilon := [n(1/4 - \varepsilon), n(1/4 + \varepsilon)], \tag{3.9}$$

$$l_{12}, l_{13}, l_{23} \in I_\varepsilon := [n(1/2 - \varepsilon), n(1/2 + \varepsilon)]. \tag{3.10}$$

For any three users u_1, u_2, u_3, note that the observed fingerprints form a typical triple with high probability.

Using the same idea as before, we now take the observed fingerprints $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ to be a typical triple. A forged fingerprint $\mathbf{y}$ agrees with all the three fingerprints on $\mathcal{L}$ and takes arbitrarily values $\{0, 1\}$ on the remaining subset $[n]\backslash\mathcal{L}$ positions. Let $d_i = d_H(\mathbf{y}, \mathbf{x}_i)$ for $i = 1, 2, 3$. Note that every position in $[n]\backslash\mathcal{L}$ contributes 1 or 2 to the sum $d_1 + d_2 + d_3$ implying

$$n(3/4 - \varepsilon) \leq d_1 + d_2 + d_3 \leq n(3/2 + 2\varepsilon). \tag{3.11}$$

Obviously, the fingerprints $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ corresponding to the factual pirates will not be discarded by the decoder. The following probabilistic analysis shows that a randomly chosen code with rate satisfying (3.8) enables the decoder to discard with high probability all typical triples $(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)$ of fingerprints formed by three innocent users.

Indeed, a typical triple can be identified only if the fingerprints in it simultaneously agree with $\mathbf{y}$ in some subset of $l \in J_\varepsilon$ positions. To simplify our analysis in this case, we can even ignore the extra conditions (3.10) in any of the remaining $n - l$ positions. Thus, we allow the vectors $(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)$ to take on any combination of binary symbols $\{0, 1\}$ different from all zeros or all ones. Given 6 such combinations, any typical triple is chosen with probability at most

$$P_l \leq \binom{n}{l} 6^{n-l}/2^{3n}.$$

We further observe that the total probability of choosing such a triple in a random code of size $M_n = \lfloor 2^{nR} \rfloor$ equals

$$\binom{M_n}{3} \sum_{l \in J_\varepsilon} \binom{n}{l} 6^{n-l}/2^{3n}$$

and tends to 0 exponentially fast if (3.8) holds.

Now consider a slightly more involved case when the decoder locates the pirate coalition $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ along with another coalition $(\mathbf{x}_1, \mathbf{z}_2, \mathbf{z}_3)$ that includes the fingerprint $\mathbf{x}_1$ of an actual pirate and the fingerprints $\mathbf{z}_2, \mathbf{z}_3$ of two innocent users. In what follows, we prove that a random code of rate (3.8) satisfies at least one of the following two conditions:

(i) The decoder chooses $\mathbf{x}_1$ in the coalition $(\mathbf{x}_1, \mathbf{z}_2, \mathbf{z}_3)$ with high probability.

(ii) The coalition $(\mathbf{x}_1, \mathbf{z}_2, \mathbf{z}_3)$ has vanishing probability.

Recall that $d_1 = d_H(\mathbf{y}, \mathbf{x}_1)$. Then an innocent user, $\mathbf{z}_2$ say, can be output by the decoder if $d_H(\mathbf{y}, \mathbf{z}_2) \leq d_1$. The probability that any such $\mathbf{z}_2$ is chosen among M_n random codewords is obviously at most

$$M_n 2^{-n} \sum_{i=0}^{d_1} \binom{n}{i}.$$

Given a code of rate (3.8), this probability vanishes if $d_1/n \leq 0.33$. Therefore, condition (i) above fails if

$$d_1/n > 0.33. \tag{3.12}$$

Now let us consider condition (ii) given this restriction. Consider a typical coalition $(\mathbf{x}_1, \mathbf{z}_2, \mathbf{z}_3)$. We have

$$l = |\{s \in [n] : x_{1s} = z_{2s} = z_{3s} = y_s\}|,$$
$$l' = |\{s \in [n] : z_{2s} = z_{3s} \neq x_{1s}\}|.$$

Thus, the vectors $\mathbf{z}_2, \mathbf{z}_3$ have fixed values on the one subset of size l, where these vectors are equal to $\mathbf{x}_1$, and on the other non-overlapping subset of size l', where the vectors $\mathbf{z}_2, \mathbf{z}_3$ are equal to the binary complement of $\mathbf{x}_1$. According to conditions

(3.9) and (3.10), $l \in J_\varepsilon$ and

$$l' = l_{23} - l \in J_{2\varepsilon}.$$

In the remaining $n - l - l'$ positions we have

$$(z_{2s}, z_{3s}) \in \{(1,0),(0,1)\}.$$

Summarizing the above arguments, we conclude that the total probability of choosing such vectors $\mathbf{z}_2, \mathbf{z}_3$ in the random code is bounded above as

$$\binom{M_n}{2} 2^{-2n} \sum_{l \in J_\varepsilon} \sum_{l' \in J_{2\varepsilon}} \binom{n-d_1}{l} \binom{n-l}{l'} 2^{n-l-l'}.$$

Straightforward verification shows that this quantity vanishes given conditions (3.9), (3.10), and (3.12) for a code of rate $R < 0.086$. Thus a random code of smaller rate (3.8) discards all mixed coalitions of the form $(\mathbf{x}_1, \mathbf{z}_2, \mathbf{z}_3)$ with high probability.

The last remaining case, of a mixed coalition $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{z}_3)$, is analyzed in a similar fashion (the analysis is simpler than the one above and will be omitted). ■

3.3 Summary and Comparisons

In this chapter, we studied lower bounds for fingerprinting capacity using random codes taking values over the entire space of possible binary codes. The main accomplishment of this chapter is proposing the idea of studying typical coalitions as opposed to all coalitions. This enabled us to improve the previous estimates of fingerprinting capacity for $t = 2, 3$ to the values $\mathsf{C}_{2,2} \geq 0.25$ and $\mathsf{C}_{3,2} \geq 0.083$, respectively. We have also examined the notion of minimal configurations (which had enabled us to improve the bound on $\mathsf{C}_{3,2}$ from what was known earlier). Of these two approaches the former one is more powerful and accounts for the strongest results of this research.

For $t = 2$, the previous best lower bound $\mathsf{C}_{2,2} \geq 0.2075$ was given by Kabatiansky [21], improved to 0.25 in Theorem 3.11 above. It has been shown recently [61, 54] that this rate matches the exact capacity value, and thus the proposed codes indeed achieve capacity. We elaborate further on these recent results in Chapter 5.

The case $t = 3$ was studied in [15, 57, 70, 68, 69, 79, 4, 54] of which the last two works followed the appearance of our paper [9, 10]. At the time when our bounds on $\mathsf{C}_{3,2}$ were obtained, they were the best known in the literature and improved the known results by an order of magnitude. Subsequent to our work (and relying partly on its ideas) the bound on $\mathsf{C}_{3,2}$ was improved from our 0.083 to 0.098 [4, 54]. The best presently known bound for general t is $\mathsf{C}_{t,2} \geq 1/(t^2 2 \ln 2)$, and is due to Amiri and Tardos, [4], Dumer [36], and Moulin and Huang [61, 54].

Publications: The results of this chapter have appeared in [5, 9, 10].

Chapter 4

Interlude: The "Decode-One" Multiple Access Channel

In the previous chapter, our objective was to compute lower bounds on the maximum attainable rates (capacity) of fingerprinting codes. In order to complement the previous results by finding upper bounds, we rely on information-theoretic techniques. Before studying this problem, we examine some related questions in multi-user information theory which are motivated by our interest in fingerprinting capacity. We later utilize the results and insight obtained here to derive upper bounds on fingerprinting capacity in Chapter 5.

In this chapter we consider a multiple-access channel (MAC) model that distinguishes itself from standard models of information theory in the following important aspects:

(a) the decoding task is to recover *one* of the transmitted messages to have produced the channel output,

(b) all channel inputs are required to use the same codebook (encoding).

We call this channel model *decode-one MAC* below. For simplicity of presentation we focus primarily on the case of *two* channel inputs. Most of the results are easily generalized to an arbitrary finite number of channel inputs.

In this chapter we derive ***new bounds for the capacity*** of the decode-one MAC. The chapter is organized into two parts, the first studying capacity of the decode-one MAC whose channel matrix is fixed and available to both communicating parties, and the second one addressing the case of the channel varying arbitrarily within a given family of transition matrices.

Throughout the chapter logarithms are taken to the base q.

4.1 Single Known Channel

4.1.1 Problem Statement

Let $\mathcal{Q}$ be an alphabet of finite size q and let $[M]$ denote the set of messages available for transmission. Each of the two transmitting terminals selects one message, say i and j respectively, from $[M]$. The corresponding length-n codewords are computed according to the encoding mapping of an $(n, M)_q$ code (C, D) as defined in (2.1). Both codewords are then transmitted symbol-by-symbol over a *channel* $W : \mathcal{Q} \times \mathcal{Q} \to \mathcal{Q}$, which is defined as a conditional probability distribution so that for all $x_1, x_2, y \in \mathcal{Q}$, $W(y|x_1, x_2)$ gives the probability that the symbol y is output given that x_1, x_2 are sent. The *memoryless* n-letter extension of W is written as $W^n : \mathcal{Q}^n \times \mathcal{Q}^n \to \mathcal{Q}^n$, where

$$W^n(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2) = \prod_{l=1}^{n} W(y_l|x_{1l}, x_{2l}). \tag{4.1}$$

Once the n output symbols are received, the decoder D attempts to recover *one* of the two transmitted messages. Let $\mathbf{Y}_{C,(i,j),W}$ denote the random vector output of the channel W when the input messages are i and j. Then the probability of error for (i, j) is

$$e(C, D, (i, j), W^n) = \mathbf{P}\left[D(\mathbf{Y}_{C,(i,j),W}) \notin \{i, j\}\right] = \sum_{\mathbf{y}:\ D(\mathbf{y}) \notin \{i,j\}} W^n(\mathbf{y}|C(i), C(j)).$$

The *average* probability of error is given by

$$e_{\text{avg}}(C, D, W^n) = \frac{1}{M^2} \sum_{i,j \in [M]} e(C, D, (i, j), W^n).$$

Definition 4.1 (Capacity of decode-one channel)**.** Given a q-ary decode-one channel W, we call $R \geq 0$ an *achievable rate* if there exists a sequence of $(n, q^{nR_n})_q$

(deterministic) codes (C_n, D_n) such that

$$\liminf_{n\to\infty} R_n = R, \quad \lim_{n\to\infty} e_{\text{avg}}(C_n, D_n, W^n) = 0.$$

The supremum of all such achievable rates is called the *capacity of the decode-one channel* W and is denoted by $\mathsf{C}_q(W)$.

Remark 4.2. In the above problem statement, both input terminals as well as the output use the same alphabet. Clearly, since both channel inputs have the same codebook, by definition the input alphabets are required to be identical. Here, we further assume the output alphabet to be the same as the input alphabet due to our underlying interest in the fingerprinting problem. The results are essentially of the same form even if the output alphabet is different.

4.1.2 Main Results

Theorem 4.3 (Single known channel: Lower bound)**.** *The capacity of the 2-input q-ary decode-one channel* W *satisfies*

$$\mathsf{C}_q(W) \geq \max_{\substack{P_{X_1X_2}:\\ X_1,X_2 \text{ i.i.d.}}} \max\left(I(X_1;Y), I(X_2;Y), \frac{1}{2}I(X_1,X_2;Y)\right), \tag{4.2}$$

where X_1, X_2, Y *are q-ary r.v.'s and* $P_{Y|X_1X_2} = W$*.*

Proof. *Encoding:* Fix a probability distribution P on $\mathcal{Q}$. Codewords $\mathbf{X}(i), i = 1, \ldots, M_n = \lfloor q^{nR} \rfloor$ of the code used for transmission are generated independently according to the distribution $\mathbf{P}[\mathbf{X}(i) = \mathbf{x}] = \prod_{l=1}^{n} P(x_l)$. Let $C_n : [M_n] \to \mathcal{Q}^n$ be the (random) encoding mapping defined by $C_n(i) = \mathbf{X}(i)$ for every $i \in [M_n]$.

Decoding: Given the channel W the decoder is defined as follows. Based on W, the decoder either attempts to decode jointly both transmitted messages or to

recover one of them as detailed below. The decoder finds which of the three mutual information quantities in (4.2) yields the largest rate. Suppose that this is $I(X_1;Y)$, then the decoder attempts at recovering only the message from Terminal 1. It pursues the recovery of the message from Terminal 2 or both messages if $I(X_2;Y)$ or $1/2 I(X_1,X_2;Y)$ is the largest of the values in (4.2), respectively.

First, we consider a method to jointly decode both transmitted messages. The proof technique is similar to the classical MAC with a few adjustments that account for the decode-one channel.

For q-ary r.v.'s X_1, X_2, X_3 with joint distribution $P_{X_1X_2X_3}$, the set $\mathcal{T}_n^{\varepsilon}(P_{X_1X_2X_3})$ of ε-*typical* sequences of length n is defined by

$$\begin{aligned}&\mathcal{T}_n^{\varepsilon}(P_{X_1X_2X_3})\\ &\quad=\left\{(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3):\left|-\frac{1}{n}\log P_{X_S}^n(\mathbf{x}_S)-H(X_S)\right|<\varepsilon,\ \ \forall S\subseteq\{1,2,3\}\right\},\end{aligned}\tag{4.3}$$

where $X_S=(X_i, i\in S)$, $\mathbf{x}_S=(\mathbf{x}_i, i\in S)$ and $P_{X_S}^n(\mathbf{x}_S)=\prod_{l=1}^n P_{X_S}(x_{Sl})$ (For background on jointly typical sequences and typicality based decoding, please refer to [30, §14.2, §8.7]).

Denote by $\mathcal{T}_n^{\varepsilon}:=\mathcal{T}_n^{\varepsilon}(P_{X_1X_2Y})$, where $P_{X_1X_2Y}=P\times P\times W$, the set of typical input-output triples $(\mathbf{x}_1,\mathbf{x}_2,\mathbf{y})$. On receiving the channel output, the decoder attempts to estimate *both* transmitted messages correctly (which is in fact more than what is required by the decode-one problem). The caveat, though, is that the decoder need not identify which specific input terminal each estimate corresponds to, i.e., the ordering of the receiver estimates is irrelevant! For any realization $\{\mathbf{x}(1),\ldots,\mathbf{x}(M_n)\}$ of the random codebook, and a given received vector $\mathbf{y}$, the decoder outputs the pair $\{i,j\}$ such that

$$i\neq j\quad\text{and}\quad(\mathbf{x}(i),\mathbf{x}(j),\mathbf{y})\in\mathcal{T}_n^{\varepsilon}$$

if such a pair exists and it is unique. An error is declared otherwise. The random code generated using the encoding and decoding described above is denoted by (C_n, D_n).

Error probability analysis: Observe that the expected average error probability for the random code is

$$\begin{aligned}
&\mathbf{E}\left[e_{\text{avg}}(C_n, D_n, W^n)\right] \\
&= \frac{1}{M_n^2}\left(\sum_{\substack{i,j\in[M_n]:\\ i\neq j}} \mathbf{E}\left[e(C_n, D_n, (i,j), W^n)\right] + \sum_{\substack{i,j\in[M_n]:\\ i=j}} \mathbf{E}\left[e(C_n, D_n, (i,j), W^n)\right]\right) \\
&\leq \frac{M_n - 1}{M_n}\mathbf{E}\left[e(C_n, D_n, (1,2), W^n)\right] + \frac{1}{M_n} \\
&\sim \mathbf{E}\left[e(C_n, D_n, (1,2), W^n)\right],
\end{aligned}$$

where the last two equations are due to the symmetry in the random code, and because $M_n \to \infty$ as $n \to \infty$.

Let $\mathbf{Y}$ denote the random received vector when messages (1,2) are sent. Define $E_{ij} = [(\mathbf{X}(i), \mathbf{X}(j), \mathbf{Y}) \in \mathcal{T}_n^{\varepsilon}]$. Then

$$\begin{aligned}
&\mathbf{E}\left[e(C_n, D_n, (1,2), W^n)\right] \\
&\leq \mathbf{P}\left[E_{12}^c\right] + \sum_{\substack{i\in\{1,2\}\\ j\notin\{1,2\}}} \mathbf{P}\left[E_{ij}\right] + \sum_{\substack{j\in\{1,2\}\\ i\notin\{1,2\}}} \mathbf{P}\left[E_{ij}\right] + \sum_{\substack{i,j\notin\{1,2\}\\ i\neq j}} \mathbf{P}\left[E_{ij}\right]. \qquad (4.4)
\end{aligned}$$

The first term above $\mathbf{P}\left[E_{12}^c\right] \to 0$ from the properties of typical sequences. Next, we examine the error event E_{ij} with $i \in \{1,2\}$, $j \notin \{1,2\}$.

$$
\begin{aligned}
\mathbf{P}\left[E_{ij}\right] &= \sum_{(\mathbf{x}_1,\mathbf{x}_2,\mathbf{y})\in\mathcal{T}_n^\varepsilon} \mathbf{P}\left[\mathbf{X}(j)=\mathbf{x}_2\right]\mathbf{P}\left[\mathbf{X}(i)=\mathbf{x}_1,\mathbf{Y}=\mathbf{y}\right] \\
&\le |\mathcal{T}_n^\varepsilon| q^{-n(H(X_2)-\varepsilon)} q^{-n(H(X_1,Y)-\varepsilon)} \\
&\le q^{-n(H(X_2)+H(X_1,Y)-H(X_1,X_2,Y)-3\varepsilon)} \\
&= q^{-n(I(X_2;X_1,Y)-3\varepsilon)} \\
&= q^{-n(I(X_2;Y|X_1)-3\varepsilon)},
\end{aligned}
$$

where X_1 and X_2 are i.i.d. with distribution P and $P_{Y|X_1,X_2} = W$. The inequalities above follow from the definition (4.3) of the typical set, and the fact that $|\mathcal{T}_n^\varepsilon| \le q^{H(X_1,X_2,Y)+\varepsilon}$.

The remaining two types of error events are analyzed similarly, yielding

$$
\begin{aligned}
&\mathbf{P}\left[E_{ij}\right] \le q^{-n(I(X_1;Y|X_2)-3\varepsilon)}, \quad j\in\{1,2\}, i\notin\{1,2\}, \\
&\mathbf{P}\left[E_{ij}\right] \le q^{-n(I(X_1,X_2;Y)-4\varepsilon)}, \quad i,j\notin\{1,2\}, i\ne j.
\end{aligned}
$$

Substituting the above in (4.4) and using the union bound, we get

$$
\begin{aligned}
&\mathbf{E}\left[e_{\text{avg}}(C_n,D_n,W^n)\right] \\
&\sim \mathbf{E}\left[e(C_n,D_n,(1,2),W^n)\right] \\
&\le \mathbf{P}\left[E_{12}^c\right] + q^{nR}q^{-n(I(X_2;Y|X_1)-3\varepsilon)} + q^{nR}q^{-n(I(X_1;Y|X_2)-3\varepsilon)} + q^{n2R}q^{-n(I(X_1,X_2;Y)-4\varepsilon)}.
\end{aligned}
$$

Since $\varepsilon > 0$ can be chosen arbitrarily small, the expected average error probability for the random code tends to 0 if

$$
R < \min\left(I(X_1;Y|X_2), I(X_2;Y|X_1), \frac{1}{2}I(X_1,X_2;Y)\right).
$$

This also establishes that the ensemble of random codes contains deterministic codes $(C_n^\star, D_n^\star), n = 1, 2, \ldots$ of rate R_n approaching the r.-h.s. of the above inequality and such that $e_{\text{avg}}(C_n^\star, D_n^\star, W^n) \to 0$.

On the other hand, we can obviously use a decoder which attempts to decode just the message from the first (resp., second) input all the time, treating the transmission of the other input as noise in the channel. Thus we effectively have a single-user point-to-point channel with respect to the selected input. It is therefore straightforward to obtain codes with rate $I(X_1; Y)$ (resp., $I(X_2; Y)$) by using a joint-typicality decoder for the chosen input.

Since the decoder is designed to achieve the maximum of the three rates, it is easy to verify that

$$\begin{aligned}
&\max\left(I(X_1;Y), I(X_2;Y), \min\left(I(X_1;Y|X_2), I(X_2;Y|X_1), \frac{1}{2}I(X_1,X_2;Y)\right)\right) \\
&\quad = \max\left(I(X_1;Y), I(X_2;Y), \frac{1}{2}I(X_1,X_2;Y)\right).
\end{aligned}$$

Finally, we optimize the achievable rate on the choice of the probability distribution P, thereby obtaining the claimed lower bound. ■

Remark 4.4. In the special case when W is *symmetric*, i.e., $W(y|x_1, x_2) = W(y|x_2, x_1)$ for all $x_1, x_2, y \in \mathcal{Q}$, the lower bound (4.2) above reduces to

$$\mathsf{C}_q(W) \geq \max_{\substack{P_{X_1X_2}: \\ X_1, X_2 \text{ i.i.d.}}} \frac{1}{2}I(X_1, X_2; Y).$$

Theorem 4.5 (Single known channel: Upper bounds)**.** *The capacity of the 2-input q-ary decode-one channel W satisfies*

$$\text{(a)} \quad \mathsf{C}_q(W) \leq \max_{\substack{P_{X_1X_2}: \\ X_1,X_2 \text{ i.i.d.}}} I(X_1, X_2; Y), \tag{4.5}$$

$$\text{(b)} \quad \mathsf{C}_q(W) \leq \max_{\substack{P_{X_1X_2}: \\ X_1,X_2 \text{ indep.}}} \max\left(I(X_1; Y|X_2), I(X_2; Y|X_1)\right), \tag{4.6}$$

where X_1, X_2, Y are q-ary r.v.'s and $P_{Y|X_1X_2} = W$.

Proof of Theorem 4.5: Upper bound (a).

Let $(C_n, D_n), n = 1, 2, \ldots$ be a sequence of n-length codes of rate R_n satisfying

$$\liminf_{n\to\infty} R_n = R, \quad e_{\text{avg}}(C_n, D_n, W^n) \leq \varepsilon_n, \tag{4.7}$$

where ε_n approaches 0 as n increases. Let U_1, U_2 be independent r.v.'s uniformly distributed over the message set $[M_n] = \{1, \ldots, q^{nR_n}\}$ and let

$$\mathbf{X}_1 := C_n(U_1), \quad \mathbf{X}_2 := C_n(U_2). \tag{4.8}$$

Also, let $\mathbf{Y}$ be such that $P_{\mathbf{Y}|\mathbf{X}_1,\mathbf{X}_2} = W^n$. Then, we have

$$\mathbf{P}\left[D_n(\mathbf{Y}) \notin \{U_1, U_2\}\right] \leq \varepsilon_n, \tag{4.9}$$

which follows from (4.7). We also have the following Markov chain

$$U_1, U_2 \leftrightarrow \mathbf{X}_1, \mathbf{X}_2 \overset{W^n}{\longleftrightarrow} \mathbf{Y}. \tag{4.10}$$

Now,

$$I(U_1, U_2; \mathbf{Y}) = 2nR_n - H(U_1, U_2|\mathbf{Y}), \tag{4.11}$$

because U_1 and U_2 are independent and uniformly distributed over $[M_n]$. The second term in (4.11) can be bounded above as follows. Define $E_i = \mathbf{1}\left[D_n(\mathbf{Y}) \neq U_i\right]$, $i = 1, 2$. Let $p_1 = \mathbf{P}\left[E_1 = 0, E_2 = 1\right]$ and $p_2 = \mathbf{P}\left[E_2 = 0, E_1 = 1\right]$. Since $D_n(\mathbf{Y}), E_1, E_2$ are known given $\mathbf{Y}, U_1, U_2$,

$$\begin{aligned} H(U_1, U_2|\mathbf{Y}) &= H(U_1, U_2, E_1, E_2|\mathbf{Y}, D_n(\mathbf{Y})) \\ &\overset{(a)}{\leq} 2\log 2 + H(U_1, U_2|\mathbf{Y}, D_n(\mathbf{Y}), E_1, E_2) \\ &\overset{(b)}{\leq} 2\log 2 + p_1 H(U_2|U_1, \mathbf{Y}, E_1 = 0, E_2 = 1) \\ &\qquad + p_2 H(U_1|U_2, \mathbf{Y}, E_2 = 0, E_1 = 1) + \varepsilon_n 2nR_n \\ &\leq 2\log 2 + nR_n + \varepsilon_n 2nR_n. \end{aligned}$$

The inequality (a) follows because E_1, E_2 are binary r.v.'s, and in (b), we have used the fact that U_i is known under the event considered in p_i. Using this in (4.11), we obtain

$$nR_n(1 - 2\varepsilon_n) \leq I(U_1, U_2; \mathbf{Y}) + 2\log 2. \tag{4.12}$$

Next, we exploit (4.10) and the memoryless property of the channel which gives

$$\begin{aligned} R_n &\leq \frac{1}{1 - 2\varepsilon_n}\left(\frac{1}{n} I(U_1, U_2; \mathbf{Y}) + \xi_n\right) \\ &\leq \frac{1}{1 - 2\varepsilon_n}\left(\frac{1}{n} I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}) + \xi_n\right) \\ &\leq \frac{1}{1 - 2\varepsilon_n}\left(\frac{1}{n}\sum_{l=1}^{n} I(X_{1l}, X_{2l}; Y_l) + \xi_n\right), \end{aligned}$$

where $\xi_n = 2\log 2/n$ approaches 0 as $n \to \infty$. Moreover, $\mathbf{X}_1$ and $\mathbf{X}_2$ are i.i.d. by (4.8). Therefore, for every $l \in [n]$, X_{1l}, X_{2l} are i.i.d. Hence,

$$R_n \leq \frac{1}{1 - 2\varepsilon_n}\left(\max_{\substack{P_{X_1X_2}: \\ X_1, X_2 \text{ i.i.d.}}} I(X_1, X_2; Y) + \xi_n\right).$$

Taking $n \to \infty$ and using (4.7), we obtain the stated upper bound (4.5). ■

Proof of Theorem 4.5: Upper bound (b).

We will actually prove a stronger statement than bound (b) establishing that this upper bound is true even for a fixed average error probability $\varepsilon \in (0,1)$ (not necessarily decaying to 0 as the code length increases). Results of this kind are generally known as *strong converse theorems*. We borrow techniques from [1] in this proof.

Consider a sequence $(C_n, D_n), n = 1, 2, \ldots$ of length-n codes of rate R_n satisfying

$$\liminf_{n\to\infty} R_n = R, \quad e_{\text{avg}}(C_n, D_n, W^n) \le \varepsilon, \tag{4.13}$$

where $0 < \varepsilon < 1$. Let $\mathbf{x}_i := C_n(i)$ be the fingerprints and $\mathbb{D}_i = \{\mathbf{y} : D_n(\mathbf{y}) = i\}$ denote the decoding regions for $i = 1, \ldots, M = q^{nR_n}$. (In this notation we have suppressed the dependence on n for simplicity.) Then the above error criterion can be written as follows:

$$\frac{1}{M^2} \sum_{i,j=1}^{M} W^n(\mathbb{D}_i \cup \mathbb{D}_j | \mathbf{x}_i, \mathbf{x}_j) \ge 1 - \varepsilon.$$

Consequently, either

$$\frac{1}{M^2} \sum_{i,j=1}^{M} W^n(\mathbb{D}_i | \mathbf{x}_i, \mathbf{x}_j) \ge \frac{1-\varepsilon}{2} \tag{4.14}$$

$$\text{or} \quad \frac{1}{M^2} \sum_{i,j=1}^{M} W^n(\mathbb{D}_j | \mathbf{x}_i, \mathbf{x}_j) \ge \frac{1-\varepsilon}{2} \tag{4.15}$$

must be true. Let us assume (4.14) is true. We first find a subset $\mathcal{A}$ of "good" pairs of messages for W. Define

$$\mathcal{A} := \{(i,j) : W^n(\mathbb{D}_i | \mathbf{x}_i, \mathbf{x}_j) \ge 1 - \bar{\varepsilon}, 1 \le i, j \le M\}, \tag{4.16}$$

where $\bar{\varepsilon}$ is such that $0 < 1 - \bar{\varepsilon} < (1-\varepsilon)/2$. Then

$$|\mathcal{A}| \geq (1-\varepsilon^*)M^2, \text{ where } \varepsilon^* := \frac{1+\varepsilon}{2\bar{\varepsilon}}. \tag{4.17}$$

Next, we derive a subset $\bar{\mathcal{A}}$ of the "good" pairs where approximate independence holds between the codewords corresponding to a pair of messages uniformly distributed over this subset. This is needed to restrict the maximization in the final result (4.6) to joint distributions where the r.v.'s are independent.

Lemma 4.6. *[1] Let $C = \{\mathbf{x}_1, \ldots, \mathbf{x}_M\} \subseteq \mathcal{Q}^n$, $\mathcal{A} \subset \{1, \ldots, M\} \times \{1, \ldots, M\}$ with $|\mathcal{A}| \geq (1-\varepsilon^*)M^2$, $0 < \varepsilon^* < 1$. Then for any $0 < \gamma < \varepsilon^*/(1-\varepsilon^*)$, $0 \leq \lambda < 1$, there exist $l_1, \ldots, l_r \in [n]$, where $r \leq \varepsilon^*/(\gamma(1-\varepsilon^*))$, and some $(\bar{x}_1, \bar{x}'_1), \ldots, (\bar{x}_r, \bar{x}'_r)$, such that for $\bar{\mathcal{A}} := \{(i,j) \in \mathcal{A} : x_{il_m} = \bar{x}_m, x_{jl_m} = \bar{x}'_m, \forall m \in [r]\}$*

(a) $\left|\bar{\mathcal{A}}\right| \geq \lambda^r |\mathcal{A}|$, and

(b) For all $x_1, x_2 \in \mathcal{Q}$, $l \in [n]$,

$$\begin{aligned}
&(1+\gamma)\mathbf{P}\left[\bar{X}_{1l} = x_1\right]\mathbf{P}\left[\bar{X}_{2l} = x_2\right] - \gamma - |\mathcal{Q}|^2\lambda \\
&\leq \mathbf{P}\left[\bar{X}_{1l} = x_1, \bar{X}_{2l} = x_2\right] \\
&\leq \max\left\{(1+\gamma)\mathbf{P}\left[\bar{X}_{1l} = x_1\right]\mathbf{P}\left[\bar{X}_{2l} = x_2\right], \lambda\right\},
\end{aligned}$$

where $(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2)$ are r.v.'s with uniform distribution on $\{(\mathbf{x}_i, \mathbf{x}_j) : (i,j) \in \bar{\mathcal{A}}\}$.

Applying Lemma 4.6 to $\mathcal{A}$ as in (4.16) with parameters $\gamma = n^{-1/2}$, $\lambda = n^{-1}$, we obtain

$$\left|\bar{\mathcal{A}}\right| \geq \lambda^r |\mathcal{A}|, \text{ for some } r \leq n^{1/2}\varepsilon^*/(1-\varepsilon^*). \tag{4.18}$$

For $j = 1, \ldots, M$, define $\mathcal{B}(j) = \{i : (i,j) \in \bar{\mathcal{A}}, 1 \leq i \leq M\}$. Observe that the subcode corresponding to $\mathcal{B}(j)$ is a "good" code for the single-user point-to-point

channel obtained by fixing the second input to j. Thus, the strong converse for the single-user channel given below holds for this subcode.

Lemma 4.7. *[14] If (C, D) is an $(n, M)_q$ code with codewords $\{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$ and decoding regions $\mathbb{D}_i, i = 1, \ldots, M$, for the (non-stationary) single-user memoryless channel $\{W_l\}_{l=1}^{\infty}$, such that for every $i = 1, \ldots, M$, $\mathbf{P}[\mathbb{D}_i|\mathbf{x}_i] \geq 1 - \bar{\varepsilon}$, $0 < \bar{\varepsilon} < 1$, then*

$$\log M \leq \sum_{l=1}^{n} I(X_l; Y_l) + O(n^{1/2}),$$

where $\mathbf{X}$ is distributed uniformly on the set of codewords, and $P_{Y_l|X_l} = W_l, l = 1, \ldots, n$.

Using Lemma 4.7 on the subcode $\mathcal{B}(j)$,

$$\log |\mathcal{B}(j)| \leq \sum_{l=1}^{n} I(\bar{X}_{1l}; \bar{Y}_l | \bar{X}_{2l} = x_{jl}) + O(n^{1/2}), \tag{4.19}$$

where $(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2)$ are distributed as in Lemma 4.6 and $P_{\bar{\mathbf{Y}}|\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2} = W^n$. Furthermore, using (4.19), we obtain

$$\begin{aligned}
&|\bar{\mathcal{A}}|^{-1} \sum_{(i,j)\in\bar{\mathcal{A}}} \log |\mathcal{B}(j)| \\
&\leq |\bar{\mathcal{A}}|^{-1} \sum_{(i,j)\in\bar{\mathcal{A}}} \sum_{l=1}^{n} I(\bar{X}_{1l}; \bar{Y}_l | \bar{X}_{2l} = x_{jl}) \sum_{x\in\mathcal{Q}} \mathbf{1}[x_{jl} = x] + O(n^{1/2}) \\
&= \sum_{l=1}^{n} \sum_{x\in\mathcal{Q}} |\bar{\mathcal{A}}|^{-1} \sum_{(i,j)\in\bar{\mathcal{A}}} \mathbf{1}[x_{jl} = x] I(\bar{X}_{1l}; \bar{Y}_l | \bar{X}_{2l} = x_{jl}) + O(n^{1/2}) \\
&= \sum_{l=1}^{n} I(\bar{X}_{1l}; \bar{Y}_l | \bar{X}_{2l}) + O(n^{1/2}),
\end{aligned} \tag{4.20}$$

since $\mathbf{P}\left[\bar{X}_{2l} = x\right] = |\bar{\mathcal{A}}|^{-1} \sum_{(i,j)\in\bar{\mathcal{A}}} \mathbf{1}[x_{jl} = x]$ for $l \in [n]$.

We next establish a lower bound on the left-side term in order to obtain an inequality for M. By the definition of $\mathcal{B}(j)$,

$$\begin{aligned}
|\bar{\mathcal{A}}|^{-1} \sum_{(i,j)\in\bar{\mathcal{A}}} \log|\mathcal{B}(j)| &= |\bar{\mathcal{A}}|^{-1} \sum_{j=1}^{M} |\mathcal{B}(j)| \log|\mathcal{B}(j)| \\
&\geq |\bar{\mathcal{A}}|^{-1} \sum_{j:|\mathcal{B}(j)|\geq \frac{1-\varepsilon^*}{n} M\lambda^r} |\mathcal{B}(j)| \log|\mathcal{B}(j)| \\
&\geq |\bar{\mathcal{A}}|^{-1} \log\left(\frac{1-\varepsilon^*}{n} M\lambda^r\right) \sum_{j:|\mathcal{B}(j)|\geq \frac{1-\varepsilon^*}{n} M\lambda^r} |\mathcal{B}(j)|. \qquad (4.21)
\end{aligned}$$

Now,

$$\begin{aligned}
\sum_{j:|\mathcal{B}(j)|\geq \frac{1-\varepsilon^*}{n} M\lambda^r} |\mathcal{B}(j)| &= \sum_{j=1}^{M} |\mathcal{B}(j)| - \sum_{j:|\mathcal{B}(j)|< \frac{1-\varepsilon^*}{n} M\lambda^r} |\mathcal{B}(j)| \\
&\geq |\bar{\mathcal{A}}| - \frac{1-\varepsilon^*}{n} M^2\lambda^r \\
&\geq |\bar{\mathcal{A}}| - \frac{1}{n}|\bar{\mathcal{A}}|
\end{aligned}$$

by using (4.17) and (4.18). Using this inequality in (4.21), we get

$$|\bar{\mathcal{A}}|^{-1} \sum_{(i,j)\in\bar{\mathcal{A}}} \log|\mathcal{B}(j)| \geq \left(1-\frac{1}{n}\right) \log\left(\frac{1-\varepsilon^*}{n} M\lambda^r\right). \qquad (4.22)$$

Combining (4.20), (4.22) and (4.18),

$$\begin{aligned}
\log M &\leq \left(1+\frac{1}{n-1}\right)\left(\sum_{l=1}^{n} I(\bar{X}_{1l};\bar{Y}_l|\bar{X}_{2l}) + O(n^{1/2})\right) \\
&\qquad - \log(1-\varepsilon^*) + \log n + \frac{\varepsilon^*}{1-\varepsilon^*} n^{1/2}\log n \\
&\leq \sum_{l=1}^{n} I(\bar{X}_{1l};\bar{Y}_l|\bar{X}_{2l}) + O(n^{1/2}\log n). \qquad (4.23)
\end{aligned}$$

Although the above inequality resembles what is needed in the theorem, note that $\bar{X}_{1l}$ and $\bar{X}_{2l}$ are not necessarily independent.

For $l \in [n]$, let (X_{1l}, X_{2l}, Y_l) be r.v.'s with distribution

$$\mathbf{P}\left[X_{1l} = x_1, X_{2l} = x_2, Y_l = y\right] = \mathbf{P}\left[\bar{X}_{1l} = x_1\right] \mathbf{P}\left[\bar{X}_{2l} = x_2\right] W(y|x_1, x_2)$$

for all $x_1, x_2, y \in \mathcal{Q}$. From Lemma 4.6(b), for $n^{-1/2} \geq |\mathcal{Q}|^2 n^{-1}$ and every $l \in [n]$

$$\begin{aligned}
&(1 + n^{-1/2})\mathbf{P}\left[\bar{X}_{1l} = x_1\right] \mathbf{P}\left[\bar{X}_{2l} = x_2\right] - 2n^{-1/2} \\
&\leq \mathbf{P}\left[\bar{X}_{1l} = x_1, \bar{X}_{2l} = x_2\right] \\
&\leq (1 + n^{-1/2})\mathbf{P}\left[\bar{X}_{1l} = x_1\right] \mathbf{P}\left[\bar{X}_{2l} = x_2\right] + n^{-1},
\end{aligned}$$

$$\text{i.e., } \left|\mathbf{P}\left[X_{1l} = x_1, X_{2l} = x_2\right] - \mathbf{P}\left[\bar{X}_{1l} = x_1, \bar{X}_{2l} = x_2\right]\right| \leq 2n^{-1/2}.$$

Thus, by the uniform continuity of mutual information, for all $l \in [n]$,

$$|I(X_{1l}; Y_l|X_{2l}) - I(\bar{X}_{1l}; \bar{Y}_l|\bar{X}_{2l})| \leq \alpha_n,$$

where $\alpha_n \to 0$ as $n \to \infty$. Together with (4.23) and dividing by n,

$$R_n \leq \max_{\substack{P_{X_1X_2}: \\ X_1, X_2 \text{ indep.}}} I(X_1; Y|X_2) + \beta_n, \tag{4.24}$$

where $\beta_n = \alpha_n + O(n^{-1/2} \log n) \to 0$ as $n \to \infty$. Similarly, assuming (4.15) is true, one can prove

$$R_n \leq \max_{\substack{P_{X_1X_2}: \\ X_1, X_2 \text{ indep.}}} I(X_2; Y|X_1) + \beta'_n. \tag{4.25}$$

Since either (4.24) or (4.25) always holds, taking $n \to \infty$ and using (4.13) completes the proof. ■

4.2 Arbitrarily Varying Channel

In the previous section, we considered a single channel which was known to both senders and the receiver. Therefore, it was possible to design the code to suit the given channel. We now consider a more complex situation where the channel in fact varies arbitrarily over a family of memoryless channels. On the other hand, we permit the use of a *randomized code* with the code selection available to both the input and output terminals. This problem is more in line with the fingerprinting scenario where the coalitions may choose an arbitrary admissible strategy, and the distributor uses a randomized code.

4.2.1 Problem Statement

As before, let $\mathcal{Q}$ be an alphabet of size q ($< \infty$) and let $[M]$ be the set of messages. We assume the use of a *randomized code* $(\mathcal{C}, \mathcal{D})$ similar to the description in Section 2.1.2. Recall that $(\mathcal{C}, \mathcal{D})$ is an r.v. over a family $\{(C_k, D_k)\}$ of $(n, M)_q$ codes. One of the codes is chosen at random and the selection is known to both senders and the receiver. Each input terminal picks one message and encodes it into a codeword according to the selected code C_k. Both codewords are then transmitted over a channel. The communication model is given by an arbitrarily varying channel (AVC) as defined below.

Let $\mathcal{W}$ denote a family of (memoryless) channels

$$\mathcal{W} = \{W_s : \mathcal{Q} \times \mathcal{Q} \to \mathcal{Q},\ \ s \in \mathcal{S}\}$$

indexed by $s \in \mathcal{S}$ which is called the "state" and is used to identify the particular channel in $\mathcal{W}$. The AVC model allows the channel state to vary from symbol to

symbol. For a given state sequence $\mathbf{s} \in \mathcal{S}^n$, the channel is given by

$$W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2) = \prod_{l=1}^{n} W_{s_l}(y_l|x_{1l}, x_{2l}). \tag{4.26}$$

The family of such channels $W_{\mathbf{s}}^n : \mathcal{Q}^n \times \mathcal{Q}^n \to \mathcal{Q}^n, \mathbf{s} \in \mathcal{S}^n$ is denoted by $\mathcal{W}^n$.

On receiving the channel output, the decoder D_k corresponding to the selected code is employed to recover *one* of the two transmitted messages. For a given state sequence $\mathbf{s}$, let $\mathbf{Y}_{\mathcal{C},(i,j),\mathbf{s}}$ denote the random channel output when the messages (i, j) are sent. The probability of error for messages (i, j) and state sequence $\mathbf{s}$ is

$$e(\mathcal{C}, \mathcal{D}, (i,j), \mathbf{s}) = \mathbf{P}\left[\mathcal{D}(\mathbf{Y}_{\mathcal{C},(i,j),\mathbf{s}}) \notin \{i,j\}\right] = \mathbf{E}_K \sum_{\mathbf{y}:\ D_K(\mathbf{y}) \notin \{i,j\}} W_{\mathbf{s}}^n(\mathbf{y}|C_K(i), C_K(j)).$$

The *average* probability of error for the state sequence $\mathbf{s}$ is given by

$$e_{\text{avg}}(\mathcal{C}, \mathcal{D}, \mathbf{s}) = \frac{1}{M^2} \sum_{i,j \in [M]} e(\mathcal{C}, \mathcal{D}, (i,j), \mathbf{s}).$$

Definition 4.8 (Capacity of decode-one AVC)**.** Given a class of q-ary decode-one channels $\mathcal{W}$, we call $R \geq 0$ an *achievable rate* for the decode-one AVC defined on $\mathcal{W}$ if there exists a sequence of $(n, q^{nR_n})_q$ randomized codes $(\mathcal{C}_n, \mathcal{D}_n)$ such that

$$\liminf_{n\to\infty} R_n = R, \quad \lim_{n\to\infty} \max_{\mathbf{s}\in\mathcal{S}^n} e_{\text{avg}}(\mathcal{C}_n, \mathcal{D}_n, \mathbf{s}) = 0.$$

The supremum of all such achievable rates is called the *capacity of the decode-one AVC* over $\mathcal{W}$ with randomized codes and is denoted by $\mathsf{C}_q(\mathcal{W})$.

4.2.2 Main Results

In the following, the *convex closure* $\overline{\mathcal{W}}$ of $\mathcal{W}$ is defined as the closure of probability distributions V of the form

$$V(y|x_1,x_2)=\sum_{s\in S}P(s)W_s(y|x_1,x_2),\quad x_1,x_2,y\in\mathcal{Q}$$

where S is any finite subset of $\mathcal{S}$, and P is any probability distribution over S. (A set of probability distributions is called convex/closed if it is convex/closed as a set of vectors in the real space of the corresponding dimension.)

Theorem 4.9 (Arbitrarily varying channel: Lower bound)**.** *The capacity of the 2-input q-ary decode-one AVC over $\mathcal{W}$ satisfies*

$$\mathsf{C}_q(\mathcal{W})\geq\max_{\substack{P_{X_1X_2}:\\ X_1,X_2\text{ i.i.d.}}}\max\left(\min_{W\in\overline{\mathcal{W}}}I(X_1;Y),\min_{W\in\overline{\mathcal{W}}}I(X_2;Y),\min_{W\in\overline{\mathcal{W}}}\frac{1}{2}I(X_1,X_2;Y)\right),$$

where X_1,X_2,Y are q-ary r.v.'s and $P_{Y|X_1X_2}=W$.

Proof. The proof is based on techniques from the single-user AVC proof given in [31, Lemma 6.10] and is in a number of aspects parallel to the proof of Theorem 4.3. In particular, we will consider three decoding strategies accounting for the recovery either one of the two or both transmitted messages.

Encoding: Fix a probability distribution P over $\mathcal{Q}$. We generate length-n codewords $\mathbf{X}_i, i=1,\ldots,M=\lfloor q^{nR}\rfloor$ independently, where the symbol in each coordinate is picked in an i.i.d. manner according to P. The message i is then encoded

as $\mathbf{X}_i$. We will use the following notation. For a distribution P and a channel W,

$$\begin{aligned}
PPW(y) &= \sum_{x_1,x_2\in\mathcal{Q}} P(x_1)P(x_2)W(y|x_1,x_2),\\
(PW)_1(y|x_1) &= \sum_{x_2} P(x_2)W(y|x_1,x_2),\\
(PW)_2(y|x_2) &= \sum_{x_1} P(x_1)W(y|x_1,x_2),\\
I(P\times P,W) &= \sum_{x_1,x_2,y\in\mathcal{Q}} P(x_1)P(x_2)W(y|x_1,x_2)\log\frac{W(y|x_1,x_2)}{PPW(y)}.
\end{aligned}$$

Decoding: We start by considering a joint decoder, i.e., a decoder which makes a decision by computing the likelihood of each possible message pair. Let $W^\star\in\overline{\mathcal{W}}$ be a channel attaining the minimum of $I(P\times P,W)$ for $W\in\overline{\mathcal{W}}$. For every $W\in\overline{\mathcal{W}}$ and $0\le\alpha\le 1$, we have

$$I(P\times P,\alpha W+(1-\alpha)W^\star)\ge I(P\times P,W^\star),$$

i.e.,

$$\lim_{\alpha\to 0}\frac{\partial}{\partial\alpha}I(P\times P,\alpha W+(1-\alpha)W^\star)\ge 0.$$

Since

$$\begin{aligned}
&\frac{\partial}{\partial\alpha}I(P\times P,\alpha W+(1-\alpha)W^\star)=\\
&\sum_{x_1,x_2,y} P(x_1)P(x_2)(W(y|x_1,x_2)-W^\star(y|x_1,x_2))\log\frac{\alpha W(y|x_1,x_2)+(1-\alpha)W^\star(y|x_1,x_2)}{\alpha PPW(y)+(1-\alpha)PPW^\star(y)},
\end{aligned}$$

it follows that

$$\sum_{x_1,x_2,y} P(x_1)P(x_2)W(y|x_1,x_2)\log\frac{W^\star(y|x_1,x_2)}{PPW^\star(y)}\ge I(P\times P,W^\star). \tag{4.27}$$

Consider a two-step maximum likelihood decoder tuned to the worst channel $W^\star \in \overline{\mathcal{W}}$ and operating as follows: For any realization of the random codebook $\{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$,

1. Find
$$(u_1, u_2) = \underset{i,j \in [M]:\, i \neq j}{\arg\max} \; W^{\star n}(\mathbf{y}|\mathbf{x}_i, \mathbf{x}_j)$$

2. Decode $\mathbf{y}$ as follows:
$$D(\mathbf{y}) = \begin{cases} u_1 & \text{if } (PW^\star)_1^n(\mathbf{y}|\mathbf{x}_{u_1}) \geq (PW^\star)_2^n(\mathbf{y}|\mathbf{x}_{u_2}) \\ u_2 & \text{otherwise.} \end{cases}$$

The randomized code obtained by the above procedure is denoted as $(\mathcal{C}, \mathcal{D})$.

Error probability analysis: Fix $\mathbf{s} \in \mathcal{S}^n$. As in Theorem 4.3, it is clear that the average probability of error

$$e_{\text{avg}}(\mathcal{C}, \mathcal{D}, \mathbf{s}) \sim e(\mathcal{C}, \mathcal{D}, (1,2), \mathbf{s})$$

as the code length increases due to the symmetry of the random code. Now,

$$\begin{aligned} & e(\mathcal{C}, \mathcal{D}, (1,2), \mathbf{s}) \\ & = \mathbf{E}_{\mathcal{C}} \left[\sum_{\mathbf{y}} W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{X}_1, \mathbf{X}_2) \mathbf{1}\left[\mathcal{D}(\mathbf{y}) \notin \{1,2\}\right] \right] \\ & = \sum_{\mathbf{x}_1, \mathbf{x}_2} P^n(\mathbf{x}_1) P^n(\mathbf{x}_2) \mathbf{E}_{\mathcal{C}} \left[\sum_{\mathbf{y}} W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{X}_1, \mathbf{X}_2) \mathbf{1}\left[\mathcal{D}(\mathbf{y}) \notin \{1,2\}\right] \Big| \mathbf{X}_1 = \mathbf{x}_1, \mathbf{X}_2 = \mathbf{x}_2 \right] \\ & = \sum_{\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}} P^n(\mathbf{x}_1) P^n(\mathbf{x}_2) W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2) \mathbf{P}\left[\mathcal{D}(\mathbf{y}) \notin \{1,2\} \big| \mathbf{X}_1 = \mathbf{x}_1, \mathbf{X}_2 = \mathbf{x}_2\right]. \end{aligned} \quad (4.28)$$

The last term in (4.28) can be bounded above as follows:

$$\begin{aligned}&\mathbf{P}\left[\mathcal{D}(\mathbf{y}) \notin \{1,2\} \middle| \mathbf{X}_1 = \mathbf{x}_1, \mathbf{X}_2 = \mathbf{x}_2\right] \\ &\leq \mathbf{P}\left[E_a\right] + \mathbf{P}\left[E_{b,1}\right] + \mathbf{P}\left[E_{b,2}\right] + \mathbf{P}\left[E_{c,1}\right] + \mathbf{P}\left[E_{c,2}\right] \end{aligned} \tag{4.29}$$

where

$$\begin{aligned}E_a &= \left[\exists i, j \notin \{1,2\}, i \neq j : W^{\star n}(\mathbf{y}|\mathbf{X}_i, \mathbf{X}_j) \geq W^{\star n}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2)\right], \\ E_{b,1} &= \left[\exists i \notin \{1,2\} : W^{\star n}(\mathbf{y}|\mathbf{X}_i, \mathbf{x}_1) \geq W^{\star n}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2), (PW^\star)_1^n(\mathbf{y}|\mathbf{X}_i) \geq (PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)\right], \\ E_{b,2} &= \left[\exists i \notin \{1,2\} : W^{\star n}(\mathbf{y}|\mathbf{X}_i, \mathbf{x}_2) \geq W^{\star n}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2), (PW^\star)_1^n(\mathbf{y}|\mathbf{X}_i) \geq (PW^\star)_2^n(\mathbf{y}|\mathbf{x}_2)\right], \\ E_{c,1} &= \left[\exists j \notin \{1,2\} : W^{\star n}(\mathbf{y}|\mathbf{x}_1, \mathbf{X}_j) \geq W^{\star n}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2), (PW^\star)_2^n(\mathbf{y}|\mathbf{X}_j) \geq (PW^\star)_1^n(\mathbf{y}|\mathbf{x}_1)\right], \\ E_{c,2} &= \left[\exists j \notin \{1,2\} : W^{\star n}(\mathbf{y}|\mathbf{x}_2, \mathbf{X}_j) \geq W^{\star n}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2), (PW^\star)_2^n(\mathbf{y}|\mathbf{X}_j) \geq (PW^\star)_1^n(\mathbf{y}|\mathbf{x}_2)\right].\end{aligned}$$

Let us first look at the error event E_a. If

$$\frac{W^{\star n}(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2)}{(PPW^\star)^n(\mathbf{y})} < \frac{M^2}{\varepsilon^2},$$

we bound $\mathbf{P}\left[E_a\right]$ by 1. In the opposite case, since $\mathbf{E}[W^{\star n}(\mathbf{y}|\mathbf{X}_i, \mathbf{X}_j)] = (PPW^\star)^n(\mathbf{y})$, by the Markov inequality

$$\mathbf{P}\left[E_a\right] \leq \sum_{\substack{i \notin \{1,2\} \\ j \notin \{1,2\} \\ i \neq j}} \mathbf{P}\left[\frac{W^{\star n}(\mathbf{y}|\mathbf{X}_i, \mathbf{X}_j)}{(PPW^\star)^n(\mathbf{y})} \geq \frac{M^2}{\varepsilon^2}\right] \leq \varepsilon^2.$$

Using this in (4.28), we get

$$\begin{aligned}
&\sum_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}} P^n(\mathbf{x}_1)P^n(\mathbf{x}_2)W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)\mathbf{P}\left[E_a\right] \\
&\leq \sum_{\substack{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}: \\ \frac{W^{\star n}(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)}{(PPW^\star)^n(\mathbf{y})}<\frac{M^2}{\varepsilon^2}}} P^n(\mathbf{x}_1)P^n(\mathbf{x}_2)W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2) + \varepsilon^2 \\
&= \mathbf{P}\left[\frac{W^{\star n}(\mathbf{Y}|\mathbf{X}_1,\mathbf{X}_2)}{(PPW^\star)^n(\mathbf{Y})} < \frac{M^2}{\varepsilon^2}\right] + \varepsilon^2
\end{aligned}$$

where the independent r.v.'s $\mathbf{X}_1, \mathbf{X}_2$ have distribution P^n and $P_{\mathbf{Y}|\mathbf{X}_1\mathbf{X}_2} = W_{\mathbf{s}}^n$.

Lemma 4.10. *If $R \leq \frac{1}{2}I(P\times P, W^\star) - \varepsilon$, then for $n \geq \frac{2}{\varepsilon}\log\frac{1}{\varepsilon}$*

$$\mathbf{P}\left[\frac{W^{\star n}(\mathbf{Y}|\mathbf{X}_1,\mathbf{X}_2)}{(PPW^\star)^n(\mathbf{Y})} < \frac{M^2}{\varepsilon^2}\right] \leq c(\varepsilon)n^{-1} \tag{4.30}$$

where $c(\varepsilon) > 0$ is a constant which depends on ε.

The proof is provided in the Appendix. We now turn to the error event $E_{b,1}$.

$$\mathbf{P}\left[E_{b,1}\right] \leq \mathbf{P}\left[\exists i \notin \{1,2\}: \begin{array}{l} \frac{W^{\star n}(\mathbf{y}|\mathbf{X}_i,\mathbf{x}_1)}{(PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)} + \frac{(PW^\star)_1^n(\mathbf{y}|\mathbf{X}_i)}{(PPW^\star)^n(\mathbf{y})} \\ \geq \frac{W^{\star n}(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)}{(PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)} + \frac{(PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)}{(PPW^\star)^n(\mathbf{y})} \end{array}\right] \tag{4.31}$$

We will bound $\mathbf{P}\left[E_{b,1}\right]$ as follows. If

$$\frac{W^{\star n}(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)}{(PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)} + \frac{(PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)}{(PPW^\star)^n(\mathbf{y})} < \frac{M}{\varepsilon}$$

we bound $\mathbf{P}\left[E_{b,1}\right]$ by 1. In the opposite case, since $\mathbf{E}\left[W^{\star n}(\mathbf{y}|\mathbf{X}_i,\mathbf{x}_1)\right] = (PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)$ and $\mathbf{E}\left[(PW^\star)_1^n(\mathbf{y}|\mathbf{X}_i)\right] = (PPW^\star)^n(\mathbf{y})$,

$$\mathbf{P}\left[E_{b,1}\right] \leq \sum_{i\notin\{1,2\}} \mathbf{P}\left[\frac{W^{\star n}(\mathbf{y}|\mathbf{X}_i,\mathbf{x}_1)}{(PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)} + \frac{(PW^\star)_1^n(\mathbf{y}|\mathbf{X}_i)}{(PPW^\star)^n(\mathbf{y})} \geq \frac{M}{\varepsilon}\right] \leq 2\varepsilon$$

by the Markov inequality.

Using this in (4.28), we get

$$\begin{aligned}
&\sum_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}} P^n(\mathbf{x}_1)P^n(\mathbf{x}_2)W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)\mathbf{P}\left[E_{b,1}\right] \\
&\leq \sum_{\substack{\mathbf{x}_1,\mathbf{x}_2,\mathbf{y}: \\ \frac{W^{\star n}(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2)}{(PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)}+\frac{(PW^\star)_2^n(\mathbf{y}|\mathbf{x}_1)}{(PPW^\star)^n(\mathbf{y})}<\frac{M}{\varepsilon}}} P^n(\mathbf{x}_1)P^n(\mathbf{x}_2)W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{x}_1,\mathbf{x}_2) + 2\varepsilon \\
&= \mathbf{P}\left[\frac{W^{\star n}(\mathbf{Y}|\mathbf{X}_1,\mathbf{X}_2)}{(PW^\star)_2^n(\mathbf{Y}|\mathbf{X}_1)} + \frac{(PW^\star)_2^n(\mathbf{Y}|\mathbf{X}_1)}{(PPW^\star)^n(\mathbf{Y})} < \frac{M}{\varepsilon}\right] + 2\varepsilon \\
&\leq \mathbf{P}\left[\frac{W^{\star n}(\mathbf{Y}|\mathbf{X}_1,\mathbf{X}_2)}{(PW^\star)_2^n(\mathbf{Y}|\mathbf{X}_1)} < \frac{M}{\varepsilon}, \frac{(PW^\star)_2^n(\mathbf{Y}|\mathbf{X}_1)}{(PPW^\star)^n(\mathbf{Y})} < \frac{M}{\varepsilon}\right] + 2\varepsilon \\
&\leq \mathbf{P}\left[\frac{W^{\star n}(\mathbf{Y}|\mathbf{X}_1,\mathbf{X}_2)}{(PPW^\star)^n(\mathbf{Y})} < \frac{M^2}{\varepsilon^2}\right] + 2\varepsilon \qquad (4.32)
\end{aligned}$$

where $\mathbf{X}_1, \mathbf{X}_2$ have distribution P^n and $P_{\mathbf{Y}|\mathbf{X}_1\mathbf{X}_2} = W_{\mathbf{s}}^n$. The first probability term in (4.32) can again be bounded by Lemma 4.10. The remaining error events can also be bounded in the same way as $E_{b,1}$. Hence, if $R \leq \frac{1}{2}I(P \times P, W^\star) - \varepsilon$, then for any $\mathbf{s} \in \mathcal{S}^n$, when $n \geq \frac{2}{\varepsilon}\log\frac{1}{\varepsilon}$

$$e(\mathcal{C}, \mathcal{D}, (1,2), \mathbf{s}) \leq 5c(\varepsilon)n^{-1} + \varepsilon^2 + 8\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this establishes that the rate

$$\frac{1}{2}I(P \times P, W^\star) = \min_{W \in \overline{\mathcal{W}}} \frac{1}{2}I(X_1, X_2; Y)$$

claimed in the theorem is indeed achievable.

The remaining decoding options are handled as follows. We can focus on decoding only the message from the first terminal, treating the input from the other terminal as part of the channel. Specifically, for a fixed distribution P generating the codewords, we may select $W^\star \in \overline{\mathcal{W}}$ to be a channel minimizing $I(P, (PW)_1)$ for $W \in \overline{\mathcal{W}}$, where by $I(P, (PW)_1)$ we mean the quantity $I(X_1; Y)$. Now, performing standard (one-step) maximum likelihood decoding with respect to $(PW^\star)_1^n$, it is

possible to obtain codes with rate

$$I(P, (PW^\star)_1) = \min_{W \in \overline{\mathcal{W}}} I(X_1; Y)$$

essentially by using the results for the single-user (point-to-point) AVC [31, Lemma 6.10]. In a similar way, by decoding just the second message we can attain the rate $\min_{W \in \overline{\mathcal{W}}} I(X_2; Y)$. Since any of the above decoding choices may be invoked based on the distribution P, and P can be chosen to optimize the rate, this establishes the achievability of the stated lower bound. ■

Theorem 4.11 (Arbitrarily varying channel: Upper bounds)**.** *The capacity of the 2-input q-ary decode-one AVC over $\mathcal{W}$ satisfies*

$$\begin{aligned}
\text{(a)} \quad \mathsf{C}_q(\mathcal{W}) &\leq \min_{W \in \mathcal{W}} \max_{\substack{P_{X_1 X_2}: \\ X_1, X_2 \text{ i.i.d.}}} I(X_1, X_2; Y), \\
\text{(b)} \quad \mathsf{C}_q(\mathcal{W}) &\leq \min_{W \in \mathcal{W}} \max_{\substack{P_{X_1 X_2}: \\ X_1, X_2 \text{ indep.}}} \max\left(I(X_1; Y|X_2), I(X_2; Y|X_1)\right),
\end{aligned}$$

where X_1, X_2, Y are q-ary r.v.'s and $P_{Y|X_1X_2} = W$.

The proof is an extension of the ideas from the single known channel case (Theorem 4.5). We do not provide the proof here as we will face a similar, but considerably harder, task in the next chapter when we compute upper bounds for fingerprinting capacity. Please see Corollary 5.2 and Theorem 5.4.

4.3 Concluding Remarks

Our main objectives in this chapter were to enhance our understanding of the fingerprinting problem, and to develop techniques for computing capacity bounds by studying a similar, yet simpler, class of problems in multi-user information theory. In particular, we examined the decode-one MAC problem, where the input terminals

share the same codebook and the decoder is required to output only one of the transmitted messages. We calculated upper and lower bounds for the capacity of the decode-one MAC, both for the single channel and AVC case. However, the following fundamental question still remains unresolved.

Open Problem 4.12. Compute the exact value of capacity for a decode-one MAC.

4.4 Appendix

4.4.1 Proof of Lemma 4.10

Define

$$Z_l := \log \frac{W^\star(Y_l|X_{1l}, X_{2l})}{(PPW^\star)(Y_l)}, \quad l = 1, \ldots, n$$

where X_{1l}, X_{2l} are i.i.d. with distribution P, and $P_{Y_l|X_{1l}X_{2l}} = W_{s_l}$. As $(X_{1l}, X_{2l}, Y_l), l = 1, \ldots, n$ are mutually independent, the r.v.'s Z_l are also independent and

$$\mathbf{E}[Z_l] = \sum_{x_1,x_2,y} P(x_1)P(x_2)W_{s_l}(y|x_1, x_2) \log \frac{W^\star(y|x_1, x_2)}{(PPW^\star)(y)} \geq I(P \times P, W^\star) \quad (4.33)$$

where the last inequality is due to (4.27). Also, $|Z_l| \leq -\log m_{W^\star}$, where $m_{W^\star}$ is the smallest positive entry of $W^\star$. Hence

$$\mathrm{var}(Z_l) \leq (\log m_{W^\star})^2. \quad (4.34)$$

Now,

$$
\begin{aligned}
&\mathbf{P}\left[\frac{W^{\star n}(\mathbf{Y}|\mathbf{X}_1,\mathbf{X}_2)}{(PPW^\star)^n(\mathbf{Y})} < \frac{M^2}{\varepsilon^2}\right] \\
&\le \mathbf{P}\left[\sum_{l=1}^{n} Z_l < 2nR + 2\log\frac{1}{\varepsilon}\right] \\
&\overset{(a)}{\le} \mathbf{P}\left[\sum_{l=1}^{n} Z_l < n(I(P\times P, W^\star) - 2\varepsilon) + 2\log\frac{1}{\varepsilon}\right] \\
&\overset{(b)}{\le} \mathbf{P}\left[\left|\sum_{l=1}^{n}(Z_l - \mathbf{E}\left[Z_l\right])\right| > 3n\varepsilon\right] \\
&\overset{(c)}{\le} \frac{(\log m_{W^\star})^2}{9n\varepsilon^2}
\end{aligned}
$$

where inequalities (a) and (b) follow from the assumptions on R and n (resp.) in the lemma, and (c) is because of the Chebyshev inequality and (4.34). ■

Chapter 5

Upper Bounds on Fingerprinting Capacity

In this chapter we extend the information-theoretic techniques used for the decode-one multiple-access channel (MAC) from Chapter 4 to obtain ***new*** computable ***upper bound expressions*** for fingerprinting capacity. We consider several particular cases of the capacity problem for the binary fingerprint alphabet, evaluating numerically the upper bound for small-size coalitions as well for coalitions of arbitrary size.

Let us recall the equivalent communication problem associated with fingerprinting, described in Section 2.2. We identify the users of the fingerprinting system with the set of messages available for communication. Each message (user) is encoded into a codeword (fingerprint). Once a subset of messages (a coalition) is chosen, the corresponding codewords are transmitted over an *unknown* channel which represents the coalition strategy. The task of the decoder is to recover *one* of the transmitted messages (pirates) to have produced the channel output (the forgery).

The coalition can choose any channel from the admissible class $\mathcal{V}_t$, defined in (2.4) as its attack strategy. To make the problem of upper bounds tractable, we restrict the class of possible attacks investigated to the set of *memoryless* attack channels. By memoryless we mean that at any given coordinate, the symbol used in the forged fingerprint depends only on the coalition's observed symbols in that coordinate and not on any of the other coordinates. Notice that any upper bound on the capacity obtained with this restriction will be also valid in the original problem.

Consequently, we will be interested in the family $\mathcal{W}_t$ of memoryless channels $W : \mathcal{Q} \times \cdots \times \mathcal{Q} \to \mathcal{Q}$ with t inputs that satisfy the marking assumption for a single

letter, i.e.,

$$\mathcal{W}_t = \{W : W(y|x,\ldots,x) = 0 \text{ if } y \neq x, \forall x, y \in Q\}. \tag{5.1}$$

Note that $\mathcal{W}_t$ is a convex and compact set.

Observe that the above definition of $\mathcal{W}_t$ corresponds to the wide-sense attack rule $\mathcal{E}_W(\cdot)$ defined in (2.3). For the narrow-sense envelope $\mathcal{E}_N(\cdot)$ (2.2) and other variations of the problem it is possible to define similar communication channels and study their upper bounds on capacity. Our upper bounds extend to these situations by an appropriate restriction of the set $\mathcal{W}_t$ of permissible channels.

Unless stated otherwise, in the chapter all logarithms are to the base q.

5.1 Upper Bound I

We model the coalition's strategy by an arbitrarily varying channel (AVC) over $\mathcal{W}_t$ (cf. Section 4.2). The index $s \in \mathcal{S}_t$ that identifies the particular channel $W \in \mathcal{W}_t$ will be called a state below. We will write $W_s(y|x_1,\ldots,x_t)$ for channels in $\mathcal{W}_t$. For a given state sequence $\mathbf{s} \in \mathcal{S}_t^n$, the corresponding n-letter channel is given by

$$W_{\mathbf{s}}^n(\mathbf{y}|\mathbf{x}_1,\ldots,\mathbf{x}_t) = \prod_{l=1}^{n} W_{s_l}(y_l|x_{1l},\ldots,x_{tl}).$$

We denote the family of such channels $W_{\mathbf{s}}^n : \mathcal{Q}^n \times \cdots \times \mathcal{Q}^n \to \mathcal{Q}^n, \mathbf{s} \in \mathcal{S}_t^n$ by $\mathcal{W}_t^n$.

Since the state sequence $\mathbf{s}$ completely identifies the channel, we will use $e_{\text{avg}}(\mathcal{C}, \mathcal{D}, \mathbf{s})$ to denote the average error probability in (2.9). Below in this section we extend the techniques of Theorem 4.5(a) to cover channels with many inputs, AVCs and randomized codes, and obtain bounds on fingerprinting capacity. We begin with the case of general, arbitrarily sized alphabets, and then specialize the results obtained to the binary case.

5.1.1 The general case

Recall that a fingerprinting code is an r.v. taking values on a family $\{(\mathcal{C}_k, \mathcal{D}_k), k \in \mathcal{K}\}$ of codes, where $\mathcal{K}$ refers to the set of keys.

Theorem 5.1. *Let $(\mathcal{C}, \mathcal{D})$ be a q-ary t-fingerprinting code with ε-error ($0 < \varepsilon < 1$) of length n, rate R, and $|\mathcal{K}|$ keys, such that $\varepsilon q^{nR} \geq 2^t$. Then*

$$R \leq \frac{1}{1-2t\varepsilon}\left(\max_{P_{KX_1\cdots X_t}} \min_{W \in \mathcal{W}_t} I(X_1, \ldots, X_t; Y|K) + \xi_n\right)$$

where $\xi_n = t \log 2/n$, $X_1, \ldots, X_t, Y$ are q-ary r.v.'s, $P_{Y|X_1\ldots X_t} = W$, K is an r.v. taking values on a set of cardinality $|\mathcal{K}|$ and satisfying the Markov chain $K \leftrightarrow X_1, \ldots, X_t \leftrightarrow Y$, and the maximization is over joint distributions

$$\begin{aligned} &P_{KX_1\ldots X_t} = P_K \times P_{X_1|K} \times \cdots \times P_{X_t|K} \\ &\text{with } P_{X_1|K} = \cdots = P_{X_t|K}. \end{aligned} \tag{5.2}$$

Proof. Let $\mathcal{K}$ be a set of keys and let $\{(C_k, D_k), k \in \mathcal{K}\}$ be a family of codes with probability distribution $\pi(k)$ over $\mathcal{K}$. Since $(\mathcal{C}, \mathcal{D})$ is t-fingerprinting with ε-error, it satisfies

$$e_{\text{avg}}(\mathcal{C}, \mathcal{D}, \mathbf{s}) \leq \varepsilon \quad \text{for every } \mathbf{s} \in \mathcal{S}_t^n. \tag{5.3}$$

Let $U_1, \ldots, U_t$ be independent r.v.'s uniformly distributed over the message set $\{1, \ldots, q^{nR}\}$ and let K be an r.v. independent of $U_1, \ldots, U_t$, and with probability distribution $\pi(k)$ over $\mathcal{K}$. Also, let

$$\mathbf{X}_i := C_K(U_i), \quad i = 1, \ldots, t. \tag{5.4}$$

Fix some $\mathbf{s} \in \mathcal{S}_t^n$ and let $\mathbf{Y}$ be such that $P_{\mathbf{Y}|\mathbf{X}_1,\ldots,\mathbf{X}_t} = W_{\mathbf{s}}^n$. Then, we have

$$\mathbf{P}\left[D_K(\mathbf{Y}) \notin \{U_1, \ldots, U_t\}\right] \leq \varepsilon, \tag{5.5}$$

which follows from (5.3). We also have the following Markov chain

$$U_1, \ldots, U_t, K \leftrightarrow \mathbf{X}_1, \ldots, \mathbf{X}_t \overset{W_{\mathbf{s}}^n}{\longleftrightarrow} \mathbf{Y}. \tag{5.6}$$

Now,

$$I(U_1, \ldots, U_t; \mathbf{Y}|K) = tnR - H(U_1, \ldots, U_t|\mathbf{Y}, K), \tag{5.7}$$

because $U_1, \ldots, U_t$ are independent and uniformly distributed over $[q^{nR}]$. The second term in (5.7) can be bounded above as follows. Define $E_i = \mathbf{1}\left[D_K(\mathbf{Y}) \neq U_i\right]$, $i = 1, \ldots, t$. Let $p_i = \mathbf{P}\left[E_i = 0, E_j = 1, j = 1, \ldots, t, j \neq i\right]$, $i = 1, \ldots, t$. Since $D_K(\mathbf{Y}), E_1, \ldots, E_t$ are known given $K, \mathbf{Y}, U_1, \ldots, U_t$,

$$\begin{aligned}
H(U_1, \ldots, U_t|\mathbf{Y}, K) &= H(U_1, \ldots, U_t, E_1, \ldots, E_t|\mathbf{Y}, D_K(\mathbf{Y}), K) \\
&\overset{(a)}{\leq} t \log 2 + H(U_1, \ldots, U_t|\mathbf{Y}, D_K(\mathbf{Y}), K, E_1, \ldots, E_t) \\
&\overset{(b)}{\leq} t \log 2 + \varepsilon tnR + 2^t q^{-nR} tnR \\
&\quad + \sum_{i=1}^{t} p_i H(U_1^t \backslash U_i | U_i, \mathbf{Y}, K, E_i = 0, E_j = 1, j \neq i) \\
&\leq t \log 2 + (\varepsilon + 2^t q^{-nR}) tnR + (t-1)nR.
\end{aligned}$$

The inequality (a) holds true because $E_1, \ldots, E_t$ are binary r.v.'s, and (b) follows from the fact that U_i is determined under the event considered by p_i and there are at most 2^t remaining terms each of which can be bounded above by $q^{-nR}tnR$. Using this in (5.7), we obtain

$$nR(1 - (\varepsilon + 2^t q^{-nR})t) \leq I(U_1, \ldots, U_t; \mathbf{Y}|K) + t \log 2. \tag{5.8}$$

We now use the premise that $\varepsilon q^{nR} \geq 2^t$, together with (5.6) and the memoryless property of the channel, which results in

$$\begin{aligned} R &\leq \frac{1}{1-2t\varepsilon}\left(\frac{1}{n}I(U_1,\ldots,U_t;\mathbf{Y}|K)+\xi_n\right) \\ &\leq \frac{1}{1-2t\varepsilon}\left(\frac{1}{n}I(\mathbf{X}_1,\ldots,\mathbf{X}_t;\mathbf{Y}|K)+\xi_n\right) \\ &\leq \frac{1}{1-2t\varepsilon}\left(\frac{1}{n}\sum_{l=1}^{n}I(X_{1l},\ldots,X_{tl};Y_l|K)+\xi_n\right). \end{aligned}$$

Moreover, since the above bound applies for every $\mathbf{s} \in \mathcal{S}_t^n$, i.e., for every $W^n \in \mathcal{W}_t^n$,

$$\begin{aligned} R &\leq \frac{1}{1-2t\varepsilon}\left(\frac{1}{n}\min_{W^n\in\mathcal{W}_t^n}\sum_{l=1}^{n}I(X_{1l},\ldots,X_{tl};Y_l|K)+\xi_n\right) \\ &= \frac{1}{1-2t\varepsilon}\left(\frac{1}{n}\sum_{l=1}^{n}\min_{W\in\mathcal{W}_t}I(X_{1l},\ldots,X_{tl};Y_l|K)+\xi_n\right), \end{aligned}$$

because the minimization is over channels whose state may vary over $\mathcal{W}_t$ for every letter. Note that $\mathbf{X}_1,\ldots,\mathbf{X}_t$ are i.i.d. given K (by (5.4)). Therefore, given K, for every $l \in [n]$, $X_{1l},\ldots,X_{tl}$ are i.i.d. Hence,

$$R \leq \frac{1}{1-2t\varepsilon}\left(\max_{P_{KX_1\ldots X_t}}\min_{W\in\mathcal{W}_t}I(X_1,\ldots,X_t;Y|K)+\xi_n\right).$$

where the maximization is over joint distributions satisfying (5.2). ■

Corollary 5.2. *The capacity of t-fingerprinting over a q-ary alphabet satisfies*

$$\mathsf{C}_{t,q} \leq \min_{W\in\mathcal{W}_t}\max_{P_{X_1\ldots X_t}} I(X_1,\ldots,X_t;Y), \tag{5.9}$$

where $X_1,\ldots,X_t,Y$ are q-ary r.v.'s, $P_{Y|X_1\ldots X_t} = W$ and the maximization is over joint distributions such that $X_1,\ldots,X_t$ are i.i.d.

Proof. As we prove only a min-max type result, it becomes sufficient to consider only "fixed" memoryless coalition strategies, i.e., strategies that remain fixed at every symbol instead of varying arbitrarily. In the subsequent text, $W^n : \mathcal{Q}^n \times \mathcal{Q}^n \to \mathcal{Q}^n$ will denote the memoryless n-letter extension (4.1) of a channel $W : \mathcal{Q} \times \mathcal{Q} \to \mathcal{Q}$.

Consider any sequence of t-fingerprinting codes $(\mathcal{C}_n, \mathcal{D}_n), n = 1, 2, \ldots$ of rate R_n and error ε_n, where $\liminf_n R_n = R$ and ε_n approaches 0 as n increases. Then

$$e_{\text{avg}}(\mathcal{C}, \mathcal{D}, W^n) \leq \varepsilon_n \quad \text{for every } W \in \mathcal{W}_t. \tag{5.10}$$

Fix some $W \in \mathcal{W}_t$. We find that (5.8) holds for every n. Therefore by the arguments in Theorem 5.1

$$R_n \leq \frac{1}{1 - \varepsilon'_n} \left(\frac{1}{n} \sum_{l=1}^{n} I(X_{1l}, \ldots, X_{tl}; Y_l | K) + \xi_n \right), \tag{5.11}$$

where both $\varepsilon'_n = (\varepsilon_n + 2^t q^{-nR})t$ and ξ_n approach 0 as $n \to \infty$. Considering the inner term, we note that

$$\frac{1}{n} \sum_{l=1}^{n} I(X_{1l}, \ldots, X_{tl}; Y_l | K) \leq I(X_{1l^*}, \ldots, X_{tl^*}; Y_{l^*} | K = k^*),$$

where $l^* = l^*(W)$ and $k^* = k^*(W)$ are the coordinate and key which maximize the mutual information. The term on the r.-h.s. is a function of $(P_{X_{1l^*} \ldots X_{tl^*} | K = k^*}, W)$. For every $l \in [n]$, $X_{1l}, \ldots, X_{tl}$ are i.i.d. when conditioned on K. Therefore this term is at most

$$\max_{P_{X_1 \ldots X_t}} I(X_1, \ldots, X_t; Y),$$

where $X_1, \ldots, X_t, Y$ are q-ary r.v.'s with $P_{Y|X_1,\ldots,X_t} = W$, and the maximization is over i.i.d. r.v.'s. Finally, since (5.11) is true for every $W \in \mathcal{W}_t$, we obtain the stated result by taking $n \to \infty$. ■

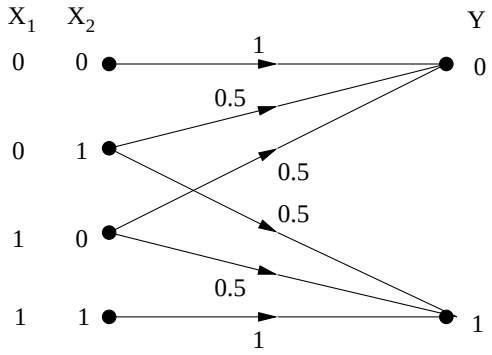


Figure 5.1: The uniform channel with 2 pirates

5.1.2 The binary case

Consider the case where $\mathcal{Q} = \{0, 1\}$. We would like to evaluate the upper bound on $\mathsf{C}_{t,2}$ given by Corollary 5.2. Computing the exact optimum in this formula is a difficult problem. Instead of attempting this, we will use a particular channel W in (5.9) and compute a maximum on the prior distribution $P_{X_1 \ldots X_t}$ for this channel. The resulting value of the rate gives an upper bound on capacity $\mathsf{C}_{t,2}$. Let W be the "uniform channel" defined by

$$W(1|x_1, \ldots, x_t) = \frac{w}{t}, \quad W(0|x_1, \ldots, x_t) = 1 - \frac{w}{t},$$

where w is the number of 1s among $x_1, \ldots, x_t$. Figure 5.1 shows the uniform channel for $t = 2$. Intuitively this choice is the worst strategy of the coalition from the distributor's perspective. If $X_1, \ldots, X_t$ are independent binary-valued r.v.'s with $\mathbf{P}[X_i = 1] = p, 0 \leq p \leq 1, i = 1, \ldots, t$, and Y is the output of the uniform channel with inputs $X_1, \ldots, X_t$, we have $\mathbf{P}[Y = 1] = p$ and

$$H(Y|X_1, \ldots, X_t) = \sum_{i=0}^{t} \binom{t}{i} p^i (1-p)^{t-i} h\left(\frac{i}{t}\right).$$

Evaluating the maximum mutual information in (5.9) for this channel gives a closed-form upper bound:

Theorem 5.3. *The capacity of t-fingerprinting over the binary alphabet satisfies*

$$\mathsf{C}_{t,2} \leq \max_{p\in[0,1]} \left\{ h(p) - \sum_{i=0}^{t} \binom{t}{i} p^i (1-p)^{t-i} h\left(\frac{i}{t}\right) \right\} \tag{5.12}$$

$$\leq \frac{1}{t \ln 2}. \tag{5.13}$$

A proof of the estimate (5.13) is given in the Appendix.

5.2 Upper Bound II

Next, we generalize the techniques of Theorem 4.5(b) to prove the following upper bound which is tighter compared to Corollary 5.2 for certain cases.

5.2.1 The general case

Theorem 5.4. *The capacity of t-fingerprinting over a q-ary alphabet satisfies*

$$\mathsf{C}_{t,q} \leq \min_{W\in\mathcal{W}_t} \max_{P_{X_1\ldots X_t}} \max_{i=1,\ldots,t} I(X_i; Y | X_1^{i-1}, X_{i+1}^t), \tag{5.14}$$

where $X_1, \ldots, X_t, Y$ are q-ary r.v.'s, $P_{Y|X_1\ldots X_t} = W$ and the maximization is over joint distributions such that $X_1, \ldots, X_t$ are independent.

Proof. The proof relies on Theorem 4.5(b). As before, we actually prove a "strong converse" theorem that the upper bound claimed is true even for a fixed average error probability $\varepsilon \in (0,1)$ (not necessarily decaying to 0 as the code length increases). For simplicity of presentation, the result is proved for the case $t = 2$.

Consider a sequence $(\mathcal{C}_n, \mathcal{D}_n), n = 1, 2, \ldots$ of $(n, q^{nR_n})_q$ randomized codes which are 2-fingerprinting with ε-error, where $\liminf_n R_n = R$ and $0 < \varepsilon < 1$.

In particular, this means that

$$e_{\text{avg}}(\mathcal{C}_n, \mathcal{D}_n, W^n) \leq \varepsilon \quad \text{for every } W \in \mathcal{W}_2.$$

For a given n and key k, let $\mathbf{x}_{i,n}^{(k)} = C_{k,n}(i)$ denote the fingerprints and $\mathbb{D}_{i,n}^{(k)} = \{\mathbf{y} : D_{k,n}(\mathbf{y}) = i\}$ denote the decoding regions for $i = 1, \ldots, M_n = q^{nR_n}$. Then the above error criterion can be written as follows: For every $W \in \mathcal{W}_2$,

$$\mathbf{E}_{K_n}\left[\frac{1}{M_n^2}\sum_{i,j=1}^{M_n} W^n(\mathbb{D}_{i,n}^{(K_n)} \cup \mathbb{D}_{j,n}^{(K_n)} | \mathbf{x}_{i,n}^{(K_n)}, \mathbf{x}_{j,n}^{(K_n)})\right] \geq 1 - \varepsilon.$$

Fix some $W \in \mathcal{W}_2$. There exists a sequence of keys $k_n^\star = k_n^\star(W)$ such that

$$\frac{1}{M_n^2}\sum_{i,j=1}^{M_n} W^n(\mathbb{D}_{i,n}^{(k_n^\star)} \cup \mathbb{D}_{j,n}^{(k_n^\star)} | \mathbf{x}_{i,n}^{(k_n^\star)}, \mathbf{x}_{j,n}^{(k_n^\star)}) \geq 1 - \varepsilon.$$

Now, applying Theorem 4.5(b) for the channel W and the code sequence indexed by $k_n^\star$, we obtain

$$R \leq \max_{\substack{P_{X_1X_2}: \\ X_1, X_2 \text{ indep.}}} \max\left(I(X_1; Y|X_2), I(X_2; Y|X_1)\right),$$

where X_1, X_2, Y are q-ary r.v.'s and $P_{Y|X_1X_2} = W$. Finally, we observe that the above inequality holds for every $W \in \mathcal{W}_2$ establishing the stated result for $t = 2$.

The general result for arbitrary t is obtained using the corresponding (straightforward) generalization of Theorem 4.5(b). ■

5.2.2 The binary case

Fix $\mathcal{Q} = \{0, 1\}$. For the case of $t = 2$ and $t = 3$, we again pick the uniform channel and obtain upper bounds on the expression in Theorem 5.4, which turn out to be stronger than the bounds resulting from (5.12). The calculations become quite

tedious for larger t. For $t = 2$, let X_1, X_2 be independent binary-valued r.v.'s with $\mathbf{P}[X_i = 1] = p_i, 0 \leq p_i \leq 1, i = 1, 2$, and let Y be the output of the uniform channel with inputs X_1 and X_2. We have

$$H(Y|X_2) = (1 - p_2)h\left(\frac{p_1}{2}\right) + p_2 h\left(\frac{1 - p_1}{2}\right)$$
$$H(Y|X_1, X_2) = (1 - p_1)p_2 + p_1(1 - p_2).$$

Computing the maximum conditional mutual information gives $\mathsf{C}_{2,2} \leq 0.322$. A similar computation for $t = 3$ yields

$$\begin{aligned} H(Y|X_2, X_3) = (1 - p_2)(1 - p_3)h\left(\frac{p_1}{3}\right) + (1 - p_2)p_3 h\left(\frac{1 + p_1}{3}\right) \\ + p_2(1 - p_3)h\left(\frac{1 + p_1}{3}\right) + p_2 p_3 h\left(\frac{1 - p_1}{3}\right), \\ H(Y|X_1, X_2, X_3) = \left(1 - p_1p_2p_3 - (1 - p_1)(1 - p_2)(1 - p_3)\right) h\left(\frac{1}{3}\right), \end{aligned}$$

and the maximization gives $\mathsf{C}_{3,2} \leq 0.199$. Combining these upper bounds with our lower bounds from Theorem 3.11 and Theorem 3.12 we obtain:

Theorem 5.5. *The capacity of* 2- *and* 3-*fingerprinting over the binary alphabet satisfy*

$$0.25 \leq \mathsf{C}_{2,2} \leq 0.322,$$

$$0.083 \leq \mathsf{C}_{3,2} \leq 0.199.$$

5.3 Summary and Recent Results

In this chapter, we proved several new upper bounds on fingerprinting capacity relying upon converse theorems for a class of channels which are similar to the multiple-access channel. For the binary case our results establish that $C_{t,2} \leq (t \ln 2)^{-1}$. Combined with the result of [79] this implies that $\Omega(1/t^2) \leq C_{t,2} \leq$

$O(1/t)$. For the general case with arbitrary alphabets, we established some upper bounds on the capacity involving only single-letter mutual information quantities.

The bounds presented here have been subsequently improved in recent works [4, 61, 54], relying in part on the ideas and results of this research.

Amiri and Tardos [4] provide tighter estimates of the upper bounds presented in this chapter. Moulin [61] provides an exact formula for fingerprinting capacity involving mutual information quantities. In Huang-Moulin's work [54], numerical estimates of the capacity formula in [61] are given for binary codes.

Publications: The results in this chapter have appeared in [9, 10].

5.4 Appendix

5.4.1 Proof of Theorem 5.3

Our goal is to estimate $\max_{p\in[0,1]} u(p,t)$, where we use the following notation

$$u(p,t) = h(p) - \sum_{i=0}^{t} \alpha_i h\left(\tfrac{i}{t}\right).$$

$$\alpha_i = \binom{t}{i} p^i (1-p)^{t-i}.$$

First, note that h is a concave function, and therefore $u(p,t)$ is non-negative for all $p \in [0,1]$. Bernstein proved that the sequence of polynomials $B_t(p) = \sum_{i=0}^{t} \alpha_i f(i/t), t = 1, 2, \ldots$, where f is a function continuous on $[0,1]$, provides a uniform approximation to f on $[0,1]$. His proof, found for instance in Feller [39] §7.2, relies on the weak law of large numbers. Refining the proof in the case of the function h, we show that for any $p \in [0,1]$ and any t,

$$u(p,t) \leq \frac{1}{t \ln 2}. \tag{5.15}$$

It suffices to consider the case $p \in (0, 1/2]$. Given some $x = i/t$, let us write a quadratic Taylor approximation for $h(x)$:

$$h(x) = h(p) + (x-p)\log_2 \frac{1-p}{p} + \frac{(x-p)^2}{2} a(x) \tag{5.16}$$

where the coefficient $a(x)$ depends on x, since $a(x) = h''(\gamma)$ for some $\gamma \in [x, p]$. We shall also consider the residual function

$$\tilde{g}(x) = h(x) - h(p) - (x-p)\log_2 \frac{1-p}{p}.$$

The main part of our proof is to show that for any $x \in [0, 1]$,

$$2p^{-2}\log_2(1-p) \le a(x) \le 0. \tag{5.17}$$

The right inequality is obvious since $h''(x) < 0$ for all $0 < x < 1$. The left inequality will be proven in two steps.

Let us take any point $x_0 \in [0, p]$. Then we compare $\tilde{g}(x)$ with the quadratic function

$$g_{x_0}(x) = a(x_0)\frac{(x-p)^2}{2}$$

on the entire interval $x \in [0, p]$. We first prove that functions $g_{x_0}(x)$ and $\tilde{g}(x)$ coincide at only two points, namely p and x_0. Indeed, let us assume that there exists a third such point x_1. Without loss of generality, let $x_0 < x_1 < p$. The functions $g_{x_0}(x)$ and $\tilde{g}(x)$ coincide at the ends of both intervals $[x_0, x_1]$ and $[x_1, p]$; therefore there exist two points $\theta' \in (x_0, x_1)$ and $\theta'' \in (x_1, p)$ where both functions have equal derivatives:

$$a(x_0)(\theta' - p) = \log_2 \frac{1-\theta'}{\theta'} - \log_2 \frac{1-p}{p};$$
$$a(x_0)(\theta'' - p) = \log_2 \frac{1-\theta''}{\theta''} - \log_2 \frac{1-p}{p}.$$

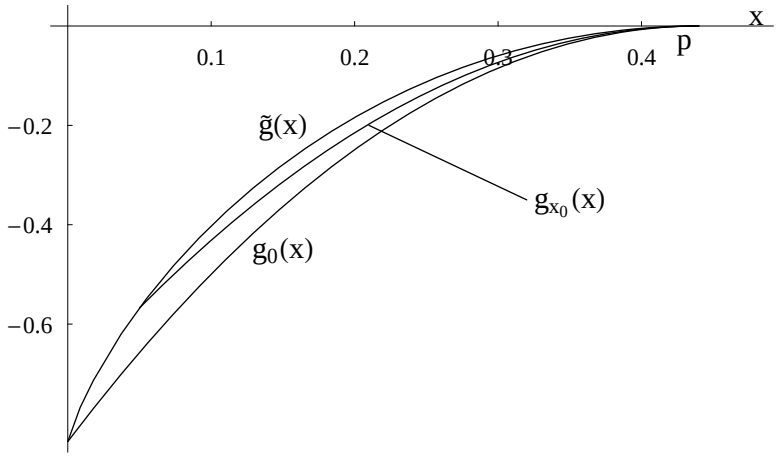


Figure 5.2: To the proof that $a(x_0) > a(0)$.

The left sides of both equalities represent a linear function of θ given by $a(x_0)(\theta - p)$ whereas the right sides represent a convex function $\log_2 \frac{1-\theta}{\theta} - \log_2 \frac{1-p}{p}$. A linear function can intersect a convex function at no more than two points. This leads to a contradiction, which shows that $x_0 = x_1$ and that the functions $g_{x_0}(x)$ and $\tilde{g}\,(x)$ intersect at two points p and x_0.

Our next step is to find the minimum $a(x) \leq 0$ for all $x \in [0, p]$. Compare the function $g_{x_0}(x)$ with $g_0(x)$ for any parameter $x_0 \in (0, p]$. Now we use the fact that both functions intersect $\tilde{g}\,(x)$ at only two points, one of which is $x = p$. However, $g_0(x)$ has its second intersection $x = 0$ to the left of x_0. Thus, $g_0(x) < g_{x_0}(x)$ for $0 \leq x < p$ and therefore, $a(x_0) > a(0)$ (see Figure 5.2). Now we conclude that

$$a(0) = \min_{x \in [0,p]} a(x).$$

Finally, we find $a(0)$ using the equality $g_0(0) = \tilde{g}\,(0)$, which gives $a(0) = 2p^{-2} \log_2(1-p)$.

The second interval $x \in [p, 1]$ can be considered similarly. Again, we use the same arguments and conclude that the end point $x = 1$ gives the minimum $a(1) = \min_{x \in [p,1]} a(x)$. Direct calculation also shows that the global minimum is

achieved at 0 as $a(0) < a(1)$ for all $p < 1/2$, and $a(1) = a(0)$ for $p = 1/2$. This gives us the left inequality in (5.17) and shows that for any $p \leq 1/2$ and any $x \in [0, 1]$,

$$h(x) \geq h(p) + (x - p) \log_2 \frac{1-p}{p} + (x - p)^2 \frac{\log_2 (1-p)}{p^2}.$$

Let us take $x = i/t, i = 0, 1, \ldots, t$ and substitute the above estimate into the expression for $u(p, t)$. In this substitution, we also use the first two moments of the binomial distribution $\{\alpha_i\}$, which gives

$$S_1 := \sum_{i=0}^{t} \alpha_i \left(\tfrac{i}{t} - p\right) = 0,$$

$$S_2 := \sum_{i=0}^{t} \alpha_i \left(\tfrac{i}{t} - p\right)^2 = \frac{p(1-p)}{t}.$$

Then

$$u(p, t) \leq -\log_2 \frac{1-p}{p} S_1 - \frac{\log_2 (1-p)}{p^2} S_2 \leq -\frac{(1-p)\ln(1-p)}{pt \ln 2}.$$

Finally, it is easy to verify that the function $-\frac{(1-p)\ln(1-p)}{p}$ monotonically decreases on the interval $[0, \frac{1}{2}]$ and achieves its maximum 1 at $p = 0$. This establishes (5.15) and hence the bound (5.13). ■

Chapter 6

Randomized Frameproof Codes

In this chapter, we consider a variation of the fingerprinting problem. In the modified system, each time a user accesses his fingerprinted copy, the fingerprint is validated to verify whether it is in fact permissible in the codebook being used and the execution continues only if the validation is successful. This limits the forgery possibilities for the pirates at the cost of an additional validation operation carried out every time a user accesses his copy. The purpose of this chapter is to design codes that permit a simple validation procedure, which does not lead to an excessive complexity overhead in the operation of the system.

In addition, since the pirates are limited to creating only a valid fingerprint and because we are interested in unique decoding, there is no additional tracing needed: instead, the distributor implicates as guilty the user corresponding to the fingerprint in the pirated copy.

Consequently, in this case, the coalition is successful if it is able to forge the fingerprint of an innocent user, thus "framing" him as the pirate. The distributor's objective is to design codes for which the probability that this error event occurs is small; thus the name *frameproof codes*.

In the deterministic case with zero error probability, frameproof codes (cf. Definition 2.14) arise as a special case of *separating codes*, which have been studied over many years since being introduced by Friedman et al. [42]. For further references on deterministic frameproof codes and separating codes, the interested reader may consult [67, 29, 78, 20]. To attain higher code rates in this chapter we introduce *randomized frameproof codes* with a small error probability.

Our specific ***objectives*** in this chapter are: (a) to characterize the theoretical lower limits on the rates of randomized frameproof codes (Section 6.2), and develop code constructions with polynomial-complexity validation performing close to these theoretical limits (Sections 6.3 and 6.4).

Remark 6.1. (a) The fingerprinting property of codes does not directly imply the frameproof property. Nevertheless, we demonstrate an improvement in the rates achievable by frameproof codes over fingerprinting constructions.

(b) Let us stress the difference between frameproof and fingerprinting codes. The former problem assumes that validation is performed *every time* the content is accessed by every user of the system. In the fingerprinting problem, no validation is performed until the distributor is alerted to a possible pirated copy of the data, at which point a tracing algorithm is executed. Thus the choice between these two scenarios depends on the application, in particular, the devices involved, complexity requirements, and the number of users to be supported.

6.1 Problem Definition

Following our previous notation, let $\mathcal{Q}$ denote an alphabet of finite size q. Also, let M and n denote the number of users in the system and the fingerprint length, respectively.

As before, the distributor constructs a randomized code $\mathcal{C}$ by choosing one of the $(n, M)_q$ codes $\{C_k, k \in \mathcal{K}\}$ according to some probability distribution $(\pi(k), k \in \mathcal{K})$. As remarked earlier, we will not require a decoding (tracing) procedure in this problem. The *rate* of this code is $R = n^{-1} \log_q M$.

Before a user executes his copy, his fingerprint is submitted to a *validation* algorithm that checks whether the fingerprint is a valid codeword in the current codebook. The execution continues only if the validation succeeds.

Once the key k is selected, each user is assigned a specific fingerprint. A *coalition* U of t users has access to the collection of fingerprints assigned to them, denoted $C_k(U)$. The goal of the pirates is to create a forged fingerprint different from theirs which is valid under the current key choice. In the following definition, the envelope $\mathcal{E}(\cdot)$ may correspond to either the narrow-sense rule (2.2) or the wide-sense rule (2.3), each leading to a different problem.

Definition 6.2. A randomized code $\mathcal{C}$ is said to be *t-frameproof with ε-error* if for all $U \subseteq [M]$ such that $|U| \leq t$, it holds that

$$\mathbf{P}\left[\mathcal{E}(\mathcal{C}(U)) \cap (\mathcal{C} \backslash \mathcal{C}(U)) \neq \emptyset\right] \leq \varepsilon, \tag{6.1}$$

where the probability is taken with respect to the distribution $\pi(\cdot)$.

Remark 6.3. Note again that a code which is t-fingerprinting with ε-error is not automatically t-frameproof with ε'-error, for any $0 \leq \varepsilon' < 1$.

In the subsequent text, we write $s_{\mathbf{z}}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ to denote the number of columns equal to $\mathbf{z}^T$ in the matrix formed with the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_t$ as the rows.

6.2 Lower Bounds for Binary Frameproof Codes

Code generation: Fix $R \in (0, 1]$, $M_n = \lfloor 2^{nR} \rfloor$, and $p \in [0, 1]$. Let us construct an $(n, M_n)_2$ randomized code $\mathcal{C}_n$ as follows. We pick each entry in the $M_n \times n$ matrix independently to be 1 with probability p.

Theorem 6.4. *The randomized (binary) code $\mathcal{C}_n$ is t-frameproof with error probability decaying exponentially in n for any rate*

$$R < -p^t \log_2 p - (1-p)^t \log_2 (1-p). \tag{6.2}$$

Proof. For $\varepsilon > 0$, define the set of t-tuples of vectors

$$\mathcal{T}_{t,n}^{\varepsilon} := \{(\mathbf{x}_1, \ldots, \mathbf{x}_t) : s_{\mathbf{1}}(\mathbf{x}_1, \ldots, \mathbf{x}_t) \in I_\varepsilon, s_{\mathbf{0}}(\mathbf{x}_1, \ldots, \mathbf{x}_t) \in J_\varepsilon\},$$

where $I_\varepsilon := [n(p^t - \varepsilon), n(p^t + \varepsilon)]$ and $J_\varepsilon := [n((1-p)^t - \varepsilon), n((1-p)^t + \varepsilon)]$. It is clear that for any coalition U of size t, the observed fingerprints $(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ belong to $\mathcal{T}_{t,n}^{\varepsilon}$ with high probability. Hence, we will refer to $\mathcal{T}_{t,n}^{\varepsilon}$ as the set of *typical* fingerprints. For any coalition U of size t

$$\begin{aligned}
&\mathbf{P}\left[\mathcal{E}(\mathcal{C}(U)) \cap (\mathcal{C} \backslash \mathcal{C}(U)) \neq \emptyset\right] \\
&\leq \mathbf{P}\left[\mathcal{C}(U) \notin \mathcal{T}_{t,n}^{\varepsilon}\right] + \mathbf{P}\left[\exists \mathbf{y} \in \mathcal{C} \backslash \mathcal{C}(U) : \mathbf{y} \in \mathcal{E}(\mathcal{C}(U)) | \mathcal{C}(U) \in \mathcal{T}_{t,n}^{\varepsilon}\right]. \qquad (6.3)
\end{aligned}$$

The first term in the above equation decays exponentially in n. It is left to prove that the second term is also exponentially decaying for R satisfying (6.2).

A codeword in $\mathcal{C} \backslash \mathcal{C}(U)$ is a part of $\mathcal{E}(\mathcal{C}(U))$ if it contains a 1 (resp., 0) in all $s_{\mathbf{1}}(\mathcal{C}(U))$ (resp., $s_{\mathbf{0}}(\mathcal{C}(U))$) positions. Since $\mathcal{C}(U) \in \mathcal{T}_{t,n}^{\varepsilon}$, by taking a union bound the second term in (6.3) is at most

$$2^{nR} p^{n(p^t - \varepsilon)} (1-p)^{n((1-p)^t - \varepsilon)},$$

which decays exponentially in n for

$$R < -(p^t - \varepsilon) \log_2 p - ((1-p)^t - \varepsilon) \log_2 (1-p).$$

The proof is completed by taking ε to be arbitrarily small. ■

The bias p in the construction of $\mathcal{C}_n$ can be chosen optimally for each value of t. Numerical values of the rate thus obtained are shown in Table 6.1, where they are compared with the existence bounds for deterministic zero-error frameproof codes

Table 6.1: Comparison of achievable rates of randomized frameproof codes with deterministic frameproof and fingerprinting codes

t	Rates		
	Randomized Frameproof	Deterministic Frameproof	Fingerprinting
2	0.5	0.2075	0.25
3	0.25	0.0693	0.098
4	0.1392	0.04	0.054
5	0.1066	0.026	0.034

(from [29]) and rates of fingerprinting codes (from [10, 4, 54]). Observe that there is a factor of t improvement compared to the rate of deterministic frameproof codes.

We remark that for large t, the optimizing value of p in

$$R_t = \max_{p\in[0,1]} \left[-p^t \log_2 p - (1-p)^t \log_2(1-p)\right] \tag{6.4}$$

is approximately $1/t$. Thus, for t large

$$R_t \approx \frac{1}{t^t}\log_2 t + \left(1-\frac{1}{t}\right)^t \log_2\left(\frac{t}{t-1}\right) = \Omega(t^{-t}\log t).$$

6.3 Linear Frameproof Codes

While randomized frameproof codes eliminate the need for a tracing algorithm, the fingerprints still need to be validated. Since the validation algorithm is executed every time a user accesses his copy, we require that this algorithm have an efficient running time. Although the codes designed in the previous section have high rates, they come at the price of an $\exp(n)$ complexity validation algorithm. *Linear* codes are an obvious first choice in trying to design efficient frameproof codes as they can be validated in $O(n^2)$ time by simply verifying the parity-check equations.

6.3.1 Linear Construction for $t = 2$

We now present a binary *linear* frameproof code for $t = 2$ which achieves the rate given by Theorem 6.4.

Code generation: Suppose we have $M_n = 2^{nR}$ users. We construct an $(n, M_n)_2$ randomized linear code $\mathcal{C}_n$ as follows. Pick a random $n(1-R)\times n$ parity-check matrix with each entry chosen independently to be 0 or 1 with equal probability. The null space of this matrix forms a linear code whose size is 2^{nR} with high probability. Each user is then assigned a unique codeword selected uniformly at random from this collection. In the few cases that the code size exceeds 2^{nR}, we simply ignore the remaining codewords during the assignment. However, note that since the validation algorithm simply verifies the parity-check equations, it will also identify the ignored vectors as valid.

Theorem 6.5. *The randomized (binary) linear code $\mathcal{C}_n$ is 2-frameproof with error probability decaying exponentially in n for any rate $R < 0.5$.*

Proof. As in the proof of Theorem 6.4, we begin by defining the set of typical pairs of fingerprints. For $\varepsilon > 0$, define

$$\mathcal{T}_n^{\varepsilon} := \Big\{(\mathbf{x}_1, \mathbf{x}_2) : s_{ij}(\mathbf{x}_1, \mathbf{x}_2) \in I_{\varepsilon}, \forall i, j \in \{0, 1\}\Big\},$$

where $I_{\varepsilon} := [n(1/4 - \varepsilon), n(1/4 + \varepsilon)]$. For any coalition U of two users

$$\begin{aligned}
&\mathbf{P}\left[\mathcal{E}(\mathcal{C}(U)) \cap (\mathcal{C}\backslash\mathcal{C}(U)) \neq \emptyset\right] \\
&\leq \mathbf{P}\left[\mathcal{C}(U) \notin \mathcal{T}_n^{\varepsilon}\right] + \sum_{(\mathbf{x}_1,\mathbf{x}_2)\in\mathcal{T}_n^{\varepsilon}} \mathbf{P}\left[\mathcal{C}(U) = (\mathbf{x}_1, \mathbf{x}_2)\right] \\
&\qquad\qquad \times \mathbf{P}\left[\exists \mathbf{y} \in \mathcal{C} : \mathbf{y} \in \mathcal{E}(\mathbf{x}_1, \mathbf{x}_2)\backslash\{\mathbf{x}_1, \mathbf{x}_2\} | \mathcal{C}(U) = (\mathbf{x}_1, \mathbf{x}_2)\right].
\end{aligned}$$

It can be seen that the first term again decays exponentially in n. We now consider the term inside the summation

$$\mathbf{P}\left[\exists \mathbf{y} \in \mathcal{C} : \mathbf{y} \in \mathcal{E}(\mathbf{x}_1, \mathbf{x}_2) \backslash \{\mathbf{x}_1, \mathbf{x}_2\} | \mathcal{C}(U) = (\mathbf{x}_1, \mathbf{x}_2)\right].$$

Observe that for any two binary vectors $(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{T}_n^{\varepsilon}$, the sum $\mathbf{x}_1 + \mathbf{x}_2 \notin \mathcal{E}(\mathbf{x}_1, \mathbf{x}_2)$ and also $\mathbf{0} \notin \mathcal{E}(\mathbf{x}_1, \mathbf{x}_2)$. Therefore, every vector in $\mathcal{E}(\mathbf{x}_1, \mathbf{x}_2) \backslash \{\mathbf{x}_1, \mathbf{x}_2\}$ is linearly independent from $\mathbf{x}_1, \mathbf{x}_2$. Thus for any $\mathbf{y} \in \mathcal{E}(\mathbf{x}_1, \mathbf{x}_2) \backslash \{\mathbf{x}_1, \mathbf{x}_2\}$,

$$\mathbf{P}\left[\mathbf{y} \in \mathcal{C} | \mathcal{C}(U) = (\mathbf{x}_1, \mathbf{x}_2)\right] = \mathbf{P}\left[\mathbf{y} \in \mathcal{C}\right] = 2^{-n(1-R)}.$$

Since $(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{T}_n^{\varepsilon}$, $|\mathcal{E}(\mathbf{x}_1, \mathbf{x}_2)| \leq 2^{n(1/2+2\varepsilon)}$. Using the union bound and taking ε to be arbitrarily small, we obtain the result. ■

6.3.2 Connection to Minimal Vectors

In this subsection, we show a connection between linear 2-frameproof codes and minimal vectors of linear codes. Let us first recall the definition of minimal vectors (see, e.g., [13]). Let C be an $[n, k]_q$ linear code. The *support* of a vector $\mathbf{c} \in C$ is defined as $\mathrm{supp}(\mathbf{c}) = \{i \in [n] : c_i \neq 0\}$. We write $\mathbf{c}' \preceq \mathbf{c}$ if $\mathrm{supp}(\mathbf{c}') \subseteq \mathrm{supp}(\mathbf{c})$.

Definition 6.6. A nonzero vector $\mathbf{c} \in C$ is called *minimal* if $\mathbf{0} \neq \mathbf{c}' \preceq \mathbf{c}$ implies $\mathbf{c}' = \alpha \mathbf{c}$, where $\mathbf{c}'$ is another code vector and α is a nonzero constant.

Proposition 6.7. *For any $\mathbf{x}_1, \mathbf{x}_2 \in C$, $\mathbf{x}_1 \neq \mathbf{x}_2$, if $\mathbf{x}_2 - \mathbf{x}_1$ is minimal then $\mathcal{E}_N(\mathbf{x}_1, \mathbf{x}_2) \cap (C \backslash \{\mathbf{x}_1, \mathbf{x}_2\}) = \emptyset$. If $q = 2$, the converse is also true.*

Proof. Consider any $\mathbf{y} \in \mathcal{Q}^n$ and define the translate $\mathbf{y}' := \mathbf{y} - \mathbf{x}_1$. It follows that

$$\mathbf{y} \in C \quad \Leftrightarrow \quad \mathbf{y}' \in C \tag{6.5}$$

$$\mathbf{y} \notin \{\mathbf{x}_1, \mathbf{x}_2\} \quad \Leftrightarrow \quad \mathbf{y}' \notin \{\mathbf{0}, \mathbf{x}_2 - \mathbf{x}_1\}. \tag{6.6}$$

Furthermore, if $y_i \in \{x_{1i}, x_{2i}\}$, then $y'_i \in \{0, x_{2i} - x_{1i}\}$ for all $i \in [n]$. Therefore,

$$\mathbf{y} \in \mathcal{E}_N(\mathbf{x}_1, \mathbf{x}_2) \Rightarrow \begin{cases} \mathbf{y}' \preceq \mathbf{x}_2 - \mathbf{x}_1, \\ \mathbf{y}' \neq \alpha(\mathbf{x}_2 - \mathbf{x}_1), \forall \alpha \notin \{0, 1\}. \end{cases} \tag{6.7}$$

Using (6.5), (6.6), (6.7), we obtain that $\mathcal{E}_N(\mathbf{x}_1, \mathbf{x}_2) \cap (C \backslash \{\mathbf{x}_1, \mathbf{x}_2\}) \neq \emptyset$ implies that $\mathbf{x}_2 - \mathbf{x}_1$ is non-minimal.

For $q = 2$, it is easily seen that the reverse statement also holds in (6.7) and thus the converse is also true. ■

Recall the random binary linear code constructed by generating a random $n(1 - R) \times n$ parity-check matrix in the previous subsection. With some abuse of notation, let us denote the (unordered) set of vectors satisfying the random parity-check matrix also by $\mathcal{C}_n$. Let $\mathcal{M}(\mathcal{C}_n)$ denote the set of minimal vectors in $\mathcal{C}_n$. We have the following companion result to Corollary 2.5 in [13].

Corollary 6.8. *As* $n \to \infty$,

$$\mathbf{E}\left[\frac{|\mathcal{M}(\mathcal{C}_n)|}{|\mathcal{C}_n|}\right] = \begin{cases} 1, & R < 1/2 \\ 0, & R > 1/2 \end{cases}$$

Proof. As a consequence of Proposition 6.7, for any two users $\{u_1, u_2\}$, we obtain

$$\begin{aligned} &\mathbf{P}\left[\mathcal{E}(\mathcal{C}_n(u_1, u_2)) \cap (\mathcal{C}_n \backslash \mathcal{C}_n(u_1, u_2)) \neq \emptyset\right] \\ &= \mathbf{P}\left[\mathcal{C}_n(u_2) - \mathcal{C}_n(u_1) \notin \mathcal{M}(\mathcal{C}_n)\right] \\ &= 1 - \mathbf{E}\left[\frac{|\mathcal{M}(\mathcal{C}_n)|}{|\mathcal{C}_n| - 1}\right]. \end{aligned}$$

The first part of the result is now true by Theorem 6.5. We skip the proof details of the second part which is easily established using Chernoff bounds. ■

6.3.3 Linear Codes for Larger t

In the light of Theorem 6.5, a natural question to ask is whether there exist randomized linear frameproof codes for $t > 2$, perhaps allowing even a larger alphabet. It turns out that, just as in the deterministic case, linear frameproof codes do not always exist in the randomized setting too.

Proposition 6.9. *There do not exist q-ary linear t-frameproof codes with ε-error, $0 \leq \varepsilon < 1$, which are secure with the wide-sense envelope if either $t > q$ or $q > 2$.*

Proof. Consider a coalition of $q + 1$ users. For any linear code realized from the family where the observed fingerprints are, say, $\mathbf{x}_1, \ldots, \mathbf{x}_{q+1}$, the forgery $\mathbf{y} = \mathbf{x}_1 + \cdots + \mathbf{x}_{q+1}$ is a part of $\mathcal{E}_W(\mathbf{x}_1, \ldots, \mathbf{x}_{q+1})$. In addition, it is also a valid fingerprint as the code is linear. This proves the first part of the proposition.

To prove the second part, consider an alphabet (a field) with $q > 2$. For any two pirates with fingerprints $\mathbf{x}_1$ and $\mathbf{x}_2$, the forgery $\mathbf{y} = \alpha\mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2$, where $\alpha \neq 0, 1$, is a valid codeword (by linearity) and is also a part of the wide-sense envelope. ■

Consequently, in considering *linear* frameproof codes which are *wide-sense* secure, we are limited to $t = 2, q = 2$.

6.4 Polynomial-time Validation for Larger t

We are mainly interested in constructing *binary* frameproof codes which have polynomial-time validation. With the binary alphabet, there is no distinction between wide-sense and narrow-sense envelopes. Therefore, there do not exist binary *linear* frameproof codes for $t > 2$ by Proposition 6.9. In this section, we use the idea of code concatenation [41] (described in Section 2.5) to construct a binary frameproof code with polynomial-time validation.

In the case of deterministic codes, it is known that if both the inner and outer codes are t-frameproof (with zero error), then the concatenated code is also t-frameproof. We will now establish a similar result when the inner code is a randomized t-frameproof code.

Construction: Let the outer code C_{out} be a (deterministic) linear $[N, K, \Delta]_q$ code with $K = R_{\text{out}}N$ and $\Delta = \delta N$. For each of the N coordinates of the outer code, generate an *independent* instance of an $(m, q)_2$ randomized code $\mathcal{C}_{\text{in}}$ which is t-frameproof with ε-error. Then the concatenated code $\mathcal{C}$ with outer code C_{out} and inner code independent instances of $\mathcal{C}_{\text{in}}$ is a randomized binary code of length $n = Nm$ and size q^K.

Validation Algorithm:

Given a fingerprint $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$, where each $\mathbf{y}_i$ is a binary word of length m, the algorithm operates in two steps.

1) For every outer coordinate $i = 1, \ldots, N$, validate $\mathbf{y}_i$ with the inner code. Specifically, verify whether $\mathbf{y}_i$ is a part of the realization of $\mathcal{C}_{\text{in}}$ at coordinate i by an exhaustive search over the q codewords. If the validation is successful at all N coordinates, map $\mathbf{y}$ into the corresponding q-ary vector $\hat{\mathbf{y}} = (\hat{y}_1, \ldots, \hat{y}_N)$.

2) Check whether $\hat{\mathbf{y}}$ is a member of the outer (linear) code by verifying its parity-check equations.

Theorem 6.10. *If the relative minimum distance of the outer code C_{out} satisfies*

$$\delta \geq 1 - \frac{1 - \xi}{t} \tag{6.8}$$

and the inner code $\mathcal{C}_{\text{in}}$ is t-frameproof with error probability $\varepsilon < \xi$, then the concatenated code $\mathcal{C}$ is t-frameproof with error probability $2^{-ND(\xi\|\varepsilon)}$.

Proof. In the proof, we will identify users with the codewords of C_{out}. Consider a coalition $U = \{\mathbf{x}_1, \ldots, \mathbf{x}_t\} \subseteq C_{\text{out}}$. In any coordinate $i \in [N]$ of the outer level, the coalition has the symbols $\{x_{1i}, \ldots, x_{ti}\}$ of which at most t are different. These t symbols correspond to t different codewords of the inner code. Thus if the t-frameproof property holds for the observed symbols for the realization of $\mathcal{C}_{\text{in}}$ in coordinate i, then at the outer level the coalition is restricted to output one of the symbols $\{x_{1i}, \ldots, x_{ti}\}$, i.e., the narrow-sense rule (2.2) holds. On the other hand, a failure of the t-frameproof property at the inner level code implies that the coalition is able to create a symbol different from $\{x_{1i}, \ldots, x_{ti}\}$.

Accordingly, let $\chi_i, i = 1, \ldots, N$, denote the indicator r.v.'s for failures at the inner level with $\mathbf{P}[\chi_i = 1] \leq \varepsilon$ since the inner code has ε-error. Note that χ_i are independent because we have an independent instance of the randomized code for every $i = 1, \ldots, N$. Then $Z = \sum_{i=1}^{N} \chi_i$ is a binomial r.v. denoting the number of coordinates where the narrow-sense rule fails at the outer level. For $0 \leq z \leq N$, let $\mathcal{E}_z(\cdot)$ denote the envelope when the narrow-sense rule is followed only in some $N - z$ outer-level coordinates, i.e.,

$$\mathcal{E}_z(\mathbf{x}_1, \ldots, \mathbf{x}_t) = \{\hat{\mathbf{y}} : s_H(\hat{\mathbf{y}}, \{\mathbf{x}_1, \ldots, \mathbf{x}_t\}) \geq N - z\}, \tag{6.9}$$

where

$$s_H(\hat{\mathbf{y}}, \{\mathbf{x}_1, \ldots, \mathbf{x}_t\}) := |\{i \in [N] : \hat{y}_i \in \{x_{1i}, \ldots, x_{ti}\}\}|.$$

For any $\hat{\mathbf{y}} \in \mathcal{E}_z(\mathbf{x}_1, \ldots, \mathbf{x}_t)$, there exists some $l \in [t]$ such that $s_H(\hat{\mathbf{y}}, \mathbf{x}_l) \geq (N - z)/t$, i.e., $d_H(\hat{\mathbf{y}}, \mathbf{x}_l) \leq N - (N - z)/t$. Therefore,

$$\mathcal{E}_z(\mathbf{x}_1, \ldots, \mathbf{x}_t) \subseteq \left\{\hat{\mathbf{y}} : d_H(\hat{\mathbf{y}}, \{\mathbf{x}_1, \ldots, \mathbf{x}_t\}) \leq N - \frac{N - z}{t}\right\}. \tag{6.10}$$

The coalition U succeeds when it creates a forgery which is valid in the outer code. Thus the probability of error is at most

$$\begin{aligned}
\mathbf{P}\left[\exists \hat{\mathbf{y}} \in C_{\text{out}} \backslash U : \hat{\mathbf{y}} \in \mathcal{E}_Z(U)\right] &\overset{(a)}{\leq} \mathbf{P}\left[\exists \mathbf{y} \in C_{\text{out}} \backslash U : d_H(\hat{\mathbf{y}}, U) \leq N - \frac{N-Z}{t}\right] \\
&\overset{(b)}{=} \mathbf{P}\left[N - \frac{N-Z}{t} \geq \Delta\right] \\
&\overset{(c)}{\leq} \mathbf{P}\left[Z \geq N\xi\right] \\
&\overset{(d)}{\leq} 2^{-ND(\xi\|\varepsilon)},
\end{aligned}$$

where inequality (a) follows from (6.10), (b) is because C_{out} is a linear code with minimum distance Δ, (c) is due to condition (6.8), and (d) is obtained using (1.1).

■

We now make specific choices for the outer and inner codes in Theorem 6.10 to arrive at explicit constructions. We take $\mathcal{C}_{\text{in}}$ to be the binary randomized t-frameproof code presented in Theorem 6.4 with length m growing. Thus we have the inner code rate as R_t, given by (6.4), and error probability $\varepsilon = 2^{-m\beta}$ for some $\beta > 0$. The outer code C_{out} is a $[q, K]$ extended Reed-Solomon (RS) code whose rate is at most $(1-\xi)/t$ so that the minimum distance of the code satisfies condition (6.8). Consequently, $n = m2^{mR_t}$ and therefore $m \approx (\log_2 n)/R_t$. Thus an exhaustive search over the inner code has only polynomial complexity in the overall code length n. Observe that for ε approaching 0 (for large m) and ξ fixed, $D(\xi\|\varepsilon) \sim \xi \log_2(1/\varepsilon)$. Therefore, with $\varepsilon = 2^{-m\beta}$, the error probability of the concatenated code is at most $2^{-n(\xi\beta + o(1))}$. We can take ξ arbitrarily small and m sufficiently large to satisfy $\varepsilon < \xi$ so that the effective rate of the concatenated code is close to R_t/t.

Alternately, take $\mathcal{C}_{\text{in}}$ to be a *fixed* $(m, q)_2$ randomized t-frameproof code with ε-error constructed using Theorem 6.4. Choose m to be sufficiently large, but fixed, so that the inner code rate is approximately R_t, and q is an even power of a prime.

In this case, the validation complexity for the inner code is constant. The outer code is an algebraic-geometry (AG) code with rate approaching $R_{\text{out}} = 1-\delta-1/(\sqrt{q}-1)$ for large N [80, 81]. Therefore, for large q and ξ arbitrarily small, we again achieve rates close to R_t/t. Moreover, the construction complexity for our outer code choice is polynomial in N [71] (it is constant for the inner code). As a consequence, we obtain the following result.

Corollary 6.11. *There exists a sequence of binary t-frameproof codes of length n with rate arbitrarily close to R_t/t, error probability* $\exp(-\Omega(n))$*, and having validation and construction complexity* $\text{poly}(n)$*.*

6.5 Concluding Remarks

In this chapter, we proposed a variation of the fingerprinting system where fingerprints are validated before the content can be accessed, leading to a formalization of randomized frameproof codes. We showed an improvement in the achievable rates with the proposed codes over traditional fingerprinting for coalitions of small size. For coalitions of size t, we proved the existence of randomized frameproof codes with rate R_t given in (6.4). Furthermore, we constructed concatenated frameproof codes where validation can be performed with complexity polynomial in the fingerprint length.

Open Problem 6.12. Find upper bounds on the maximum attainable rate of randomized frameproof codes.

Open Problem 6.13. Study frameproof code constructions which allow the validation algorithm to make an error with a small probability.

Publications: This chapter's contents have appeared in [6].

Chapter 7

Two-level Fingerprinting

7.1 Introduction

In the previous chapters, we have seen that fingerprinting codes enable the distributor to identify at least one of the pirates as long as the size of the coalition that produced the illegal copy does not exceed a certain threshold t. However, if the coalition size exceeds this threshold, the output of the tracing algorithm can be useless.

To overcome this weakness, we ***formalize*** the notion of ***multi-level fingerprinting codes***, which are inspired by error-correcting codes with unequal error protection used in communications problems (cf. for instance Bassalygo et al. [18]). We focus on the simplest case of two-level fingerprinting codes in this chapter, but the concepts introduced apply to an arbitrary number of protection levels.

In this setting, the users are organized in a hierarchical manner, for instance, according to geographical location. The distribution area is divided into several regions, and users from the same region are collected into one *group*. Fingerprinting systems for multimedia with a similar hierarchical structure were studied by Wang et al. [84].

The two-level fingerprinting codes studied in this chapter have the following property: As in traditional (one-level) codes, the tracing algorithm determines at least one of the guilty users if the coalition size is at most t. Moreover, even when a larger number s ($> t$) of pirates participate, the algorithm provides partial information by retrieving the index of a group that contains a member of the pirate coalition.

Formal definitions of two-level fingerprinting codes and related code families are given in Section 7.2. Our goal in this section is to devise analogs of the main concepts discussed for one-level codes for the present, more general context. In Section 7.3, we obtain sufficient conditions for two-level traceability and frameproof codes. Finally, we provide ***constructions for two-level fingerprinting codes*** and analyze the achievable rates in Section 7.4.

7.2 Problem Statement

Consider the problem where the content is to be distributed to $M_1 M_2$ users organized in M_1 groups, each of which contains M_2 users. Assume that there is some ordering of the groups, and of the users within each group. Thus, any user $\mathbf{u}$ is identified by a pair of indices $\mathbf{u} \equiv (u_1, u_2) \in [M_1] \times [M_2]$. For a user $\mathbf{u} = (u_1, u_2)$, let $\mathcal{G}(\mathbf{u})$ be its group index, i.e., $\mathcal{G}(\mathbf{u}) = u_1$.

The distributor hides a distinctive fingerprint in each legal copy. We carry the same notation as before, denoting the length of the fingerprints and the alphabet by n and $\mathcal{Q}$ respectively. We take $\mathcal{Q}$ to be $\{0, \ldots, q-1\}$ with modulo q addition (for some $q < \infty$).

As in the one-level case, the distributor's strategy of assigning fingerprints to the users can be either deterministic or randomized as explained in the following subsections. Randomization can potentially increase the number of users that can be supported for a given fingerprint length at the cost of a small error probability.

7.2.1 Deterministic Codes

An $(n, M_1, M_2)_q$ *two-level code* (C, D_1, D_2) is a triple consisting of one encoding and two decoding mappings

$$\begin{aligned} C &: [M_1] \times [M_2] \to \mathcal{Q}^n, \\ D_1 &: \mathcal{Q}^n \to [M_1] \cup \{0\}, \\ D_2 &: \mathcal{Q}^n \to ([M_1] \times [M_2]) \cup \{0\}, \end{aligned} \tag{7.1}$$

with 0 signifying a decoding failure. A two-level deterministic assignment of fingerprints is given by the encoding mapping C of such a two-level code. Using the same convention as before, we do not distinguish between the encoding mapping C and the collection of codewords in its range, using the same notation for both.

The *rate pair* of an $(n, M_1, M_2)_q$ two-level code is defined as

$$(R_1, R_2) := \left(\frac{1}{n} \log_q M_1, \frac{1}{n} \log_q M_2 \right).$$

It is assumed that the users have complete knowledge of the code (C, D_1, D_2).

A coalition $U \subseteq [M_1] \times [M_2]$ has access to the fingerprints $C(U)$ and attempts to produce an illegal copy. In order to conceal their identities from the distributor, the coalition creates a modified fingerprint $\mathbf{y}$ from the envelope $\mathcal{E}(C(U))$ consisting of all forgeries that follow the marking assumption (cf. Definition 2.2).

Given a pirated copy with a forged fingerprint, the distributor performs tracing based on D_1 and D_2 to locate one of the pirates. The decoder D_2 attempts to trace the exact identity of one of the pirates, while D_1 focuses only on locating a group containing at least one of the pirates. As in the one-level setting, with deterministic codes, the probability of decoding error is bounded away from zero if the pirates can output a forgery within the wide-sense envelope (2.3). Therefore, for the case

of fixed codes we investigate only the case of the narrow-sense rule defined in (2.2). Below, we write $\mathcal{E}(\cdot)$ for the narrow-sense envelope.

In order to extend the notion of traceability to two-level codes, let us consider the case where the tracing is accomplished using *minimum distance* (MD) decoding. Specifically, we take

$$\begin{aligned} D_2(\mathbf{y}) &= \underset{\mathbf{u}\in[M_1]\times[M_2]}{\arg\min} \; d_H(C(\mathbf{u}), \mathbf{y}), \\ D_1(\mathbf{y}) &= \mathcal{G}(D_2(\mathbf{y})). \end{aligned} \tag{7.2}$$

If the minimum distance above is attained for multiple users, the decoder D_2 outputs any one of the closest users. This leads us to the notion of two-level traceability codes in the deterministic setting.

Definition 7.1 ((t_1, t_2)-TA)**.** A two-level code C has (t_1, t_2)-*traceability* property (or is (t_1, t_2)-TA) where $t_1 > t_2$ if:

(a) For any coalition U of size at most t_2 and any $\mathbf{y} \in \mathcal{E}(C(U))$, the decoding result $D_2(\mathbf{y}) \in U$.

(b) For any coalition U of size at most t_1 and any $\mathbf{y} \in \mathcal{E}(C(U))$, the decoding result $D_1(\mathbf{y}) \in \mathcal{G}(U)$.

Comparing the above with Definition 2.13 for one-level traceability, we observe that a (t_1, t_2)-TA code is t_2-TA; moreover, for coalitions of the larger size t_1, one of the groups containing a pirate is closer to the forgery compared to the remaining groups.

Definition 7.2 ((t_1, t_2)-frameproof)**.** A two-level code C is (t_1, t_2)-*frameproof* where $t_1 > t_2$ if:

(a) For any coalition U of size at most t_2, $\mathcal{E}(C(U)) \cap (C \backslash C(U)) = \emptyset$.

(b) For any coalition U of size at most t_1, $\mathcal{E}(C(U)) \cap (C \backslash \tilde{C}(U)) = \emptyset$, where $\tilde{C}(U) = \{C(\mathbf{u}) : u_1 \in \mathcal{G}(U)\}$.

Consequently, a (t_1, t_2)-frameproof code satisfies the one-level t_2-frameproof property (Definition 2.14); furthermore, the larger size-t_1 coalitions cannot forge the fingerprint of an innocent user present in a different group from the pirates.

7.2.2 Randomized Codes

A randomized strategy to assign fingerprints in the two-level setting is defined as the following random experiment. The distributor has a family of $(n, M_1, M_2)_q$ two-level codes $\{(C_k, D_{1k}, D_{2k}), k \in \mathcal{K}\}$, where $\mathcal{K}$ is a finite set of "keys". The distributor chooses one of the keys according to a probability distribution $(\pi(k), k \in \mathcal{K})$. If the key k is selected, then fingerprints are assigned according to C_k and tracing is done using D_{1k} and D_{2k}. The code resulting from this random experiment is called a (two-level) *randomized code* and is denoted by $(\mathcal{C}, \mathcal{D}_1, \mathcal{D}_2)$.

Consider a coalition U of size t. As in the one-level case, any attack by the coalition can be modeled as a randomized strategy $V(\cdot|\cdot, \ldots, \cdot)$, where $V(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_t)$ gives the probability that the coalition creates $\mathbf{y}$ given that it observes the fingerprints $\mathbf{x}_1, \ldots, \mathbf{x}_t$. For tractability reasons, we limit the coalitions to create forgeries according to the narrow-sense rule (2.2) while considering two-level codes. Hereafter, $\mathcal{E}(\cdot)$ stands for the narrow-sense envelope. Accordingly, a strategy V is called *admissible* if

$$V(\mathbf{y}|\mathbf{x}_1, \ldots, \mathbf{x}_t) > 0 \text{ only if } \mathbf{y} \in \mathcal{E}(\mathbf{x}_1, \ldots, \mathbf{x}_t).$$

Let $\mathcal{V}_t$ denote the class of admissible strategies.

Denote the random forgery generated by U using the strategy V by $\mathbf{Y}_{\mathcal{C},U,V}$. The distributor, on observing the forged fingerprint, employs the decoders D_{1k} and D_{2k} while using the key k.

For a given coalition U and strategy V, we define the following error probabilities:

$$e_1(\mathcal{C}, \mathcal{D}_1, U, V) = \mathbf{P}\left[\mathcal{D}_1(\mathbf{Y}_{\mathcal{C},U,V}) \notin \mathcal{G}(U)\right] = \mathbf{E}_K \sum_{\mathbf{y}: D_{1K}(\mathbf{y}) \notin \mathcal{G}(U)} V(\mathbf{y}|C_K(U)),$$

$$e_2(\mathcal{C}, \mathcal{D}_2, U, V) = \mathbf{P}\left[\mathcal{D}_2(\mathbf{Y}_{\mathcal{C},U,V}) \notin U\right] = \mathbf{E}_K \sum_{\mathbf{y}: D_{2,K}(\mathbf{y}) \notin U} V(\mathbf{y}|C_K(U)),$$

where the expectation is taken with respect to the distribution $\pi(k)$.

Definition 7.3 ((t_1, t_2)-fingerprinting)**.** A randomized code $(\mathcal{C}, \mathcal{D}_1, \mathcal{D}_2)$ is said to be a (t_1, t_2)-*fingerprinting with ε-error* where $t_1 > t_2$ if:

(a) For any coalition U of size at most t_2 and any admissible strategy V, the error probability $e_2(\mathcal{C}, \mathcal{D}_2, U, V) \leq \varepsilon$.

(b) For any coalition U of size at most t_1 and any admissible strategy V, the error probability $e_1(\mathcal{C}, \mathcal{D}_1, U, V) \leq \varepsilon$.

Comparing this against Definition 2.6 for one-level fingerprinting, we find that a (t_1, t_2)-fingerprinting code is t_2-fingerprinting; in addition, when coalitions are of the larger size t_1, the tracing algorithm can locate a group containing one of the pirates with high probability.

Definition 7.4. A rate pair (R_1, R_2) is said to be *achievable for q-ary (t_1, t_2)-fingerprinting* if there exists a sequence of $(n, q^{nR_{1n}}, q^{nR_{2n}})_q$ randomized codes that are (t_1, t_2)-fingerprinting with error probability ε_n such that

$$\lim_{n\to\infty} \varepsilon_n = 0, \qquad \liminf_{n\to\infty} R_{in} = R_i, \quad i = 1, 2.$$

The goal of this chapter is to investigate randomized constructions of two-level fingerprinting codes and to characterize the corresponding set of achievable rate pairs.

Remark 7.5. 1. If an $(n, M_1, M_2)_q$ two-level code is (t_1, t_2)-fingerprinting (resp., TA, frameproof), then choosing any single user from every group forms an $(n, M_1)_q$ one-level code that is t_1-fingerprinting (resp., TA, frameproof).

2. If an $(n, M_1M_2)_q$ one-level code is t_1-fingerprinting (resp., TA, frameproof), then for any $t_2 < t_1$, it can also be treated as a $(n, M_1, M_2)_q$ two-level code that is (t_1, t_2)-fingerprinting (resp., TA, frameproof).

7.3 Traceability and Frameproof Codes: Simple Facts

For a given two-level code C, we define the following minimum distances:

$$d_1(C) := \min_{\substack{\mathbf{u},\mathbf{v}\in[M_1]\times[M_2]\\ u_1\neq v_1}} d_H(C(\mathbf{u}), C(\mathbf{v})), \tag{7.3}$$

$$d_2(C) := \min_{\substack{\mathbf{u},\mathbf{v}\in[M_1]\times[M_2]\\ u_2\neq v_2}} d_H(C(\mathbf{u}), C(\mathbf{v})). \tag{7.4}$$

Let $d(C) = \min(d_1(C), d_2(C))$. It is known [28] that a one-level code of length n is t-TA if the distance between any pair of fingerprints is strictly greater than $n(1-1/t^2)$ (recall Theorem 2.16(a)). In the case of two-level codes, we obtain the following analogous result.

Proposition 7.6. *Suppose $t_1 > t_2$ and C is a two-level code of length n with*

$$d_1(C) > n(1-1/t_1^2) \quad \textit{and} \quad d_2(C) > n(1-1/t_2^2).$$

Then C is (t_1, t_2)-TA.

Proof. It is straightforward to see that the assumptions in the proposition imply that $d(C) > n(1-1/t_2^2)$. Therefore, property (a) in Definition 7.1 follows directly from the result for one-level codes.

Next, we show that property (b) is a consequence of $d_1(C) > n(1 - 1/t_1^2)$. Let U be a coalition of size at most t_1 and $\mathbf{y} \in \mathcal{E}(C(U))$. Then, there exists some user $\mathbf{u} \in U$ who coincides with $\mathbf{y}$ in at least n/t_1 coordinates. For any user $\mathbf{u}'$ such that $\mathcal{G}(\mathbf{u}') \notin \mathcal{G}(U)$, the number of agreements with $\mathbf{y}$ is at most $t_1(n - d_1(C)) < \frac{n}{t_1}$, thus establishing property (b). ■

We now establish a similar sufficient condition for two-level frameproof codes, which is the analog of Theorem 2.16(b) for one-level codes. The proof is a straightforward extension of the one-level proof and is along the same lines as above.

Proposition 7.7. *Suppose $t_1 > t_2$ and C is a two-level code of length n with*

$$d_1(C) > n(1 - 1/t_1) \quad \textit{and} \quad d_2(C) > n(1 - 1/t_2).$$

Then C is (t_1, t_2)-frameproof.

As a consequence, ideas used for constructing unequal error protection codes can be also employed to construct two-level TA and frameproof codes. As in the one-level case, to be able to construct large-size codes based on the above sufficient conditions, one needs large alphabets.

7.4 Fingerprinting Codes

We denote the Hamming *weight* of $\mathbf{x} \in \mathcal{Q}^n$ by $w_H(\mathbf{x})$. For $w \in [n]$, define $\mathcal{S}_{w,n} = \{\mathbf{x} \in \mathcal{Q}^n : w_H(\mathbf{x}) = w\}$.

7.4.1 Code Generation

For $R_1, R_2 \in [0, 1]$, define $M_{1n} = \lfloor q^{nR_1} \rfloor$, $M_{2n} = \lfloor q^{nR_2} \rfloor$. Fix $\omega \in [0, 1]$. We take n such that $w = \omega n$ is an integer and construct an $(n, M_{1n}, M_{2n})_q$ two-level randomized code $(\mathcal{C}_n^\omega, \mathcal{D}_{1n}^\omega, \mathcal{D}_{2n}^\omega)$ as follows.

For $i \in [M_{1n}]$, pick random vectors $\mathbf{R}_i$ independently and uniformly at random from $\mathcal{Q}^n$. We will refer to the $\mathbf{R}_i$'s as "centers". Choose $\mathbf{S}_{ij}, (i,j) \in [M_{1n}] \times [M_{2n}]$, independently and uniformly at random from $\mathcal{S}_{w,n}$. Generate $M_{1n}M_{2n}$ fingerprints

$$\mathbf{X}_{ij} = \mathbf{R}_i + \mathbf{S}_{ij}, \quad (i,j) \in [M_{1n}] \times [M_{2n}]$$

and assign $\mathbf{X}_{ij}$ as the fingerprint for user (i,j).

Once the fingerprint assignments are fixed, tracing is based on the MD decoder (7.2). The MD decoder may be sub-optimal in general; however, it is amenable for analysis in our code construction.

7.4.2 Useful Facts

In the following subsections, we analyze the error probability and characterize the achievable rate pairs for two-level fingerprinting using the above construction. The lemmas below will be useful in the analysis.

Lemma 7.8. *Let* $\mathbf{S}$ *have a uniform distribution on* $\mathcal{S}_{w,n}$. *Then, for* $l \in [n]$ *and* $a \in \mathcal{Q}\backslash\{0\}$, $\mathbf{P}[S_l = a] = \omega/(q-1)$. *Moreover, the r.v.'s* $\{S_l, l \in [n]\}$ *are asymptotically pairwise independent.*

For $p \in [0,1]$ and $\varepsilon > 0$, define

$$I_n(p,\varepsilon) := [n(p-\varepsilon), n(p+\varepsilon)].$$

Lemma 7.9. *Fix* $p \in [0,1]$ *and* $\varepsilon > 0$. *For* $l \in [n]$, *let* Z_l *be a Bernoulli r.v. with* $\mathbf{P}[Z_l = 1] = p$, *and let* $\{Z_l, l \in [n]\}$ *be pairwise independent. Then, with* $Z := \sum_{l\in[n]} Z_l$, *we have*

$$\mathbf{P}[Z \notin I_n(p,\varepsilon)] \leq \frac{p(1-p)}{\varepsilon^2 n}.$$

Notation: For a coalition $U = \{\mathbf{u}^1, \ldots, \mathbf{u}^t\}$, we denote the realizations of $\mathbf{X}_{\mathbf{u}^i}, \mathbf{R}_{u_1^i}, \mathbf{S}_{\mathbf{u}^i}$ by $\mathbf{x}_i, \mathbf{r}_i, \mathbf{s}_i$ respectively, with $\mathbf{x}_i = \mathbf{r}_i + \mathbf{s}_i, i \in [t]$. Let $\mathbf{z} \in \mathcal{Q}^t$ be a vector. Denote by $s_{\mathbf{z}}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ the number of columns equal to $\mathbf{z}^T$ in the matrix whose rows are $\mathbf{x}_1, \ldots, \mathbf{x}_t$. We will denote the q-ary entropy function by $h(x) = -x \log_q x/(q-1) - (1-x)\log_q(1-x)$.

7.4.3 $(t, 1)$-fingerprinting

We begin by considering the $(2, 1)$-fingerprinting property. This is the simplest case of two-level fingerprinting that goes beyond the known techniques for one-level codes. Although coalitions of size 1 are trivial to handle for one-level fingerprinting, it is still non-trivial to construct a $(2, 1)$-fingerprinting code.

Theorem 7.10. *For any $\omega \in [0, (q-1)/2q]$, the randomized code $(\mathcal{C}_n^\omega, \mathcal{D}_{1n}^\omega, \mathcal{D}_{2n}^\omega)$ is $(2, 1)$-fingerprinting with error probability decaying to 0 if*

$$R_1 < 1 - h\left(\frac{1}{2}\left(1 - \frac{1}{q}\right) + \omega\right), \tag{7.5}$$

$$R_2 < h(\omega). \tag{7.6}$$

Discussion: The above theorem provides a set of achievable rate pairs for q-ary $(2, 1)$-fingerprinting. Let us fix $\mathcal{Q} = \{0, 1\}$ and put the result in the perspective of bounds available for one-level fingerprinting (see Figure 7.1).

- *Outer bound*: Since the $(2, 1)$-fingerprinting property implies one-level 1-fingerprinting, we should have $R_1 + R_2 \leq 1$. Moreover, R_1 cannot exceed the rate of a one-level 2-fingerprinting code (by part (1) of Remark 7.5); thus, any upper bound for it also applies to R_2. In particular, by [54] $R_2 \leq 0.25$.

- *Inner bound*: By part (2) of Remark 7.5, the rate pairs (R_1, R_2) such that $R_1 + R_2 < 0.188$ are achievable (with MD decoding) using the 2-fingerprinting

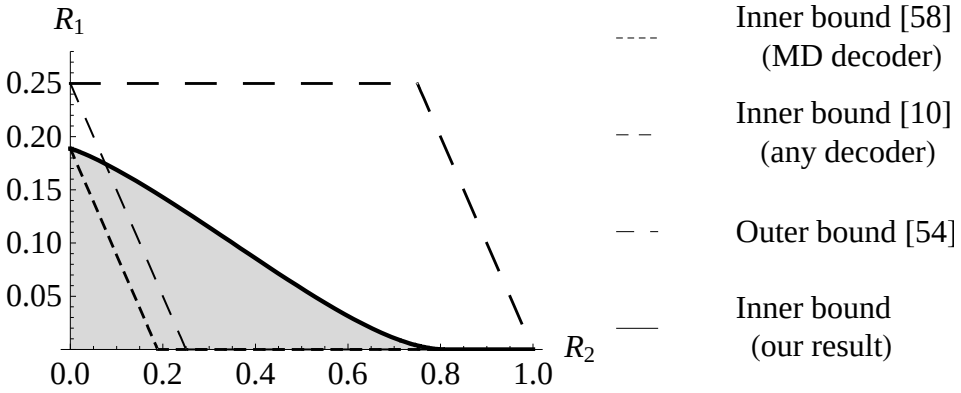


Figure 7.1: Achievable rate region for binary $(2, 1)$-fingerprinting. The bounds from previous works follow by using one-level fingerprinting schemes.

code given in [58]. In fact, by allowing other decoders, we can do better, achieving $R_1 + R_2 < 0.25$ through the 2-fingerprinting construction in [10].

Proof of Theorem 7.10.

Size-1 coalitions: Let $\mathbf{u} = (u_1, u_2)$ be the pirate. For size-1 coalitions, the envelope is degenerate as it consists of only the user's own fingerprint. Now,

$$\begin{aligned}
e_2(\mathcal{C}_n^\omega, \mathcal{D}_{2n}^\omega, \mathbf{u}) &= \mathbf{P}\left[\exists \mathbf{u}' \neq \mathbf{u} : \mathbf{X}_{\mathbf{u}'} = \mathbf{X}_{\mathbf{u}}\right] \\
&\leq \mathbf{P}\left[\exists \mathbf{u}' \neq \mathbf{u} : u_1' = u_1, \mathbf{X}_{\mathbf{u}'} = \mathbf{X}_{\mathbf{u}}\right] + \mathbf{P}\left[\exists \mathbf{u}' \neq \mathbf{u} : u_1' \neq u_1, \mathbf{X}_{\mathbf{u}'} = \mathbf{X}_{\mathbf{u}}\right] \\
&\overset{(a)}{\leq} \mathbf{P}\left[\exists u_2' \neq u_2 : \mathbf{S}_{u_1 u_2'} = \mathbf{S}_{u_1 u_2}\right] + \mathbf{P}\left[\exists u_1' \neq u_1 : d_H(\mathbf{R}_{u_1'}, \mathbf{X}_{\mathbf{u}}) \leq w\right] \\
&\overset{(b)}{\leq} q^{nR_2}\mathbf{P}\left[\mathbf{S}_{u_1 u_2'} = \mathbf{S}_{u_1 u_2}\right] + q^{nR_1}\mathbf{P}\left[d_H(\mathbf{R}_{u_1'}, \mathbf{X}_{\mathbf{u}}) \leq w\right] \\
&\doteq q^{-n(h(\omega)-R_2)} + q^{-n(1-h(\omega)-R_1)},
\end{aligned}$$

where (a) is due to the fact that if the fingerprint of another user matches with the pirate's fingerprint, then the corresponding center is within distance w from the pirate's fingerprint, and (b) follows from the union bound. Consequently, the error probability for size-1 coalitions approaches 0 if $R_2 < h(\omega)$ and $R_1 < 1 - h(\omega)$.

Size-2 coalitions: There are two possibilities: either both users are in the same group or they are in different groups. It turns out that the latter case is the dominant one. Since the analysis for the two cases is similar, we only consider the latter case below.

Let $U = \{\mathbf{u}^1, \mathbf{u}^2\}$ be such a coalition. For any strategy $V \in \mathcal{V}_2$, we have

$$\begin{aligned} & e_1(\mathcal{C}_n^\omega, \mathcal{D}_{1n}^\omega, U, V) \\ &= \sum_{\mathbf{r}_1,\mathbf{r}_2,\mathbf{s}_1,\mathbf{s}_2} \mathbf{P}\left[\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2\right] \sum_{\mathbf{y}} V(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2) \mathbf{P}\left[\mathcal{D}_{1n}^\omega(\mathbf{y}) \notin \mathcal{G}(U) \middle| \mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2\right]. \end{aligned} \tag{7.7}$$

Consider the inner probability term

$$\begin{aligned} & \mathbf{P}\left[\mathcal{D}_{1n}^\omega(\mathbf{y}) \notin \mathcal{G}(U) \middle| \mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2\right] \\ & \overset{(a)}{=} \mathbf{P}\left[\exists \mathbf{u}' \notin U : u_1' \notin \mathcal{G}(U), d_H(\mathbf{X}_{\mathbf{u}'}, \mathbf{y}) \leq d_H(\{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{y})\right] \\ & \overset{(b)}{\leq} \mathbf{P}\left[\exists u_1' \notin \mathcal{G}(U) : d_H(\mathbf{R}_{u_1'}, \mathbf{y}) \leq d_H(\{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{y}) + w\right] \\ & \leq q^{nR_1} \mathbf{P}\left[d_H(\mathbf{R}_{u_1'}, \mathbf{y}) \leq d_H(\{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{y}) + w\right], \end{aligned}$$

where we have exploited the independence in the construction in (a), and (b) follows because if the fingerprint of another user is within distance d from $\mathbf{y}$, then the corresponding center is within $d + w$ from $\mathbf{y}$. For $\varepsilon > 0$, define

$$\mathcal{T}_n^\varepsilon := \left\{(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2) : s_{(a,a)}(\mathbf{x}_1, \mathbf{x}_2) \in I_n(1/q^2, \varepsilon/q), \forall a \in \mathcal{Q}\right\}.$$

Observe that $\mathbf{X}_{\mathbf{u}^1}$ and $\mathbf{X}_{\mathbf{u}^2}$ are independent and uniformly distributed over $\mathcal{Q}^n$. Therefore, using Lemma 7.9, it is a simple matter to show that

$$\mathbf{P}\left[(\mathbf{R}_{u_1^1}, \mathbf{R}_{u_2^1}, \mathbf{S}_{\mathbf{u}^1}, \mathbf{S}_{\mathbf{u}^2}) \notin \mathcal{T}_n^\varepsilon\right] \to 0 \quad \text{as} \quad n \to \infty.$$

Now, take any $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2) \in \mathcal{T}_n^{\varepsilon}$ and $\mathbf{y} \in \mathcal{E}(\mathbf{x}_1, \mathbf{x}_2)$. The number of undetectable positions in $\{\mathbf{x}_1, \mathbf{x}_2\}$ is at least $n(1/q - \varepsilon)$, implying that $d_H(\{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{y}) \leq \frac{n}{2}\left(1 - \frac{1}{q} + \varepsilon\right)$. Thus, in this case

$$\begin{aligned}
&q^{nR_1}\mathbf{P}\left[d_H(\mathbf{R}_{u_1'}, \mathbf{y}) \leq d_H(\{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{y}) + w\right] \\
&\leq q^{nR_1}\mathbf{P}\left[d_H(\mathbf{R}_{u_1'}, \mathbf{y}) \leq \frac{n}{2}\left(1 - \frac{1}{q} + \varepsilon\right) + w\right] \\
&\doteq q^{-n\left(1 - h\left(\frac{1}{2}\left(1 - \frac{1}{q} + \varepsilon\right) + \omega\right) - R_1\right)}.
\end{aligned}$$

Substituting the above in (7.7) and taking $\varepsilon \to 0$, we conclude that the error probability for size-2 coalitions approaches 0 if (7.5) holds. ■

We now extend the techniques to larger coalitions.

Theorem 7.11. *For any ω such that $\frac{t-1}{t}\left(1 - \frac{1}{q^{t-1}}\right) + \omega \leq \frac{q-1}{q}$, the randomized code $(\mathcal{C}_n^{\omega}, \mathcal{D}_{1n}^{\omega}, \mathcal{D}_{2n}^{\omega})$ is $(t, 1)$-fingerprinting with error probability decaying to 0 if*

$$R_1 < 1 - h\left(\frac{t-1}{t}\left(1 - \frac{1}{q^{t-1}}\right) + \omega\right), \tag{7.8}$$

$$R_2 < h(\omega). \tag{7.9}$$

Proof. *Size-1 coalitions:* For a single pirate $\mathbf{u}$, the analysis in Theorem 7.10 proves that the probability of decoding error approaches 0 if $R_2 < h(\omega)$ and $R_1 < 1 - h(\omega)$.

Size-t coalitions: It can be shown that the case where the t pirates are in distinct groups is the dominant one. Once this is shown, we use exactly the same arguments as in the case of size-2 coalitions in Theorem 7.10. We finally obtain that the error probability for coalitions of size t approaches 0 if (7.8) holds. ■

Remark 7.12. A sufficiently large alphabet is required in order for an ω satisfying $\frac{t-1}{t}\left(1 - \frac{1}{q^{t-1}}\right) + \omega \leq \frac{q-1}{q}$ to exist. For instance, it suffices to take $q \geq t + 1$.

7.4.4 $(t, 2)$-fingerprinting

Let $q \geq 3$. For $\omega, \gamma, \alpha, \beta \in [0, 1]$, with $\alpha \leq 1-\gamma$, $\beta \leq \gamma$, $\alpha+\beta \leq \omega$, $\omega-\alpha \leq \gamma$, let

$$\begin{aligned}\varphi(\omega, \gamma, \alpha, \beta) := (1-\gamma)h\left(\frac{\alpha}{1-\gamma}\right) + (\gamma-\beta)h\left(\frac{\omega-\alpha-\beta}{\gamma-\beta}\right) \\ + \gamma h\left(\frac{\beta}{\gamma}\right) + (\omega-\alpha)\log_q\left(\frac{q-2}{q-1}\right) - \beta\log_q(q-2).\end{aligned}$$

Let

$$\begin{aligned}\delta_1(\omega) &= \frac{1}{2}\left(1-(1-\omega)^2 - \frac{\omega^2}{q-1}\right), \\ \delta_2(\omega) &= \frac{1}{2}\left(1-\frac{1}{q}\right), \\ f_1(\omega) &= \max_{\substack{\gamma,\alpha,\beta: \\ \omega^2 \leq \gamma \leq 1-(1-\omega)^2, \gamma-\beta+\alpha \leq \delta_1(\omega)}} \varphi(\omega, \gamma, \alpha, \beta), \\ f_2(\omega) &= \max_{\substack{\gamma,\alpha,\beta: \\ \omega\left(\frac{q-1}{q}\right) \leq \gamma \leq 1-\frac{1-\omega}{q}, \gamma-\beta+\alpha \leq \delta_2(\omega)}} \varphi(\omega, \gamma, \alpha, \beta).\end{aligned}$$

Theorem 7.13. *Let $q \geq 3$. For any ω such that $\frac{t-1}{t}\left(1-\frac{1}{q^{t-1}}\right)+\omega \leq \frac{q-1}{q}$, the randomized code $(\mathcal{C}_n^\omega, \mathcal{D}_{1n}^\omega, \mathcal{D}_{2n}^\omega)$ is $(t, 2)$-fingerprinting with error probability decaying to 0 if*

$$R_1 < 1 - h\left(\frac{t-1}{t}\left(1-\frac{1}{q^{t-1}}\right)+\omega\right), \tag{7.10}$$

$$R_2 < h(\omega) - \max(f_1(\omega), f_2(\omega)). \tag{7.11}$$

Proof. Size-t coalitions are handled in the same way as in Theorem 7.11.

Size-2 coalitions: There are two possibilities depending on whether the pirates belong to the same group or not. We sketch the case where they are in different groups below. The other case is analyzed similarly.

Consider a coalition $U = \{\mathbf{u}^1, \mathbf{u}^2\}$, where the users are in different groups, and let $V \in \mathcal{V}_2$ be an admissible strategy. We have

$$\begin{aligned} & e_2(\mathcal{C}_n^\omega, \mathcal{D}_{2n}^\omega, U, V) \\ &= \sum_{\mathbf{r}_1,\mathbf{r}_2,\mathbf{s}_1,\mathbf{s}_2} \mathbf{P}\left[\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2\right] \sum_{\mathbf{y}} V(\mathbf{y}|\mathbf{x}_1, \mathbf{x}_2) \mathbf{P}\left[\mathcal{D}_{2n}^\omega(\mathbf{y}) \notin U | \mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2\right]. \end{aligned} \tag{7.12}$$

Now,

$$[\mathcal{D}_{2n}^\omega(\mathbf{y}) \notin U] = E_1 \cup E_2 \cup E_3,$$

where, the events E_1, E_2, E_3 are formed of those $\mathbf{u}' \notin U$ that satisfy $d_H(\mathbf{X}_{\mathbf{u}'}, \mathbf{y}) \leq d_H(\{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{y})$ and the conditions $u_1' = u_1^1$, $u_1' = u_1^2$, $u_1' \notin \mathcal{G}(U)$, respectively. The error event E_3 was already analyzed in Theorem 7.10 and its conditional probability approaches 0 if (7.5) holds. We consider E_1 below. The analysis for E_2 is identical by symmetry.

$$\begin{aligned} & \mathbf{P}\left[E_1 \middle| \mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2\right] \\ &= \mathbf{P}\left[\exists u_2' \neq u_2^1 : d_H(\mathbf{r}_1 + \mathbf{S}_{u_1^1 u_2'}, \mathbf{y}) \leq d_H(\{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{y})\right] \\ &\leq q^{nR_2} \mathbf{P}\left[d_H(\mathbf{r}_1 + \mathbf{S}_{u_1^1 u_2'}, \mathbf{y}) \leq d_H(\{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{y})\right] \\ &= q^{nR_2} \mathbf{P}\left[d_H(\mathbf{S}_{u_1^1 u_2'}, \mathbf{y}') \leq d_H(\{\mathbf{s}_1, \mathbf{r}_1 + \mathbf{x}_2\}, \mathbf{y}')\right], \end{aligned} \tag{7.13}$$

where $\mathbf{y}' = \mathbf{y} + \mathbf{r}_1 \in \mathcal{E}(\mathbf{s}_1, \mathbf{r}_1 + \mathbf{x}_2)$. In this case, we use Lemmas 7.8 and 7.9 to show that

$$\mathcal{T}_n^\varepsilon := \left\{ (\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2) : \begin{array}{c} s_{(0,0)}(\mathbf{s}_1, \mathbf{r}_1 + \mathbf{x}_2) \simeq n\frac{1-\omega}{q} \\ s_{(a,a')}(\mathbf{s}_1, \mathbf{r}_1 + \mathbf{x}_2) \simeq n\frac{\omega}{(q-1)q} \\ \forall a, a' \in \mathcal{Q}\backslash\{0\} \end{array} \right\}.$$

is the typical set. For simplicity, we have omitted ε and will use the approximate relations $\simeq, \lesssim, \gtrsim$ in its place. Now, take any $(\mathbf{r}_1, \mathbf{r}_2, \mathbf{s}_1, \mathbf{s}_2) \in \mathcal{T}_n^\varepsilon$ and $\mathbf{y}' \in \mathcal{E}(\mathbf{s}_1, \mathbf{r}_1 + \mathbf{x}_2)$. The number of undetectable positions in $\{\mathbf{s}_1, \mathbf{r}_1 + \mathbf{x}_2\}$ is $\simeq n/q$, while the

number of coordinates where both symbols are non-zero is $\simeq n\omega(q-1)/q$. This implies $d_H(\{\mathbf{s}_1, \mathbf{r}_1+\mathbf{x}_2\}, \mathbf{y}') \lesssim n\delta_2(\omega)$ and $n\omega(q-1)/q \lesssim w_H(\mathbf{y}') \lesssim n(1-(1-\omega)/q)$.

Let $w_H(\mathbf{y}') = \gamma n$, where $\gamma \in [0,1]$. Then

$$\mathbf{P}\left[d_H(\mathbf{S}_{u_1^1 u_2'}, \mathbf{y}') \le n\delta_2(\omega)\right] \doteq q^{-nE(\omega,\gamma)},$$

where

$$E(\omega,\gamma) = h(\omega) - \max_{\substack{\alpha,\beta: \\ \gamma-\beta+\alpha\le\delta_2(\omega)}} \varphi(\omega,\gamma,\alpha,\beta).$$

Since γ can be chosen by the pirates such that $\omega\frac{q-1}{q} \lesssim \gamma \lesssim 1-\frac{1-\omega}{q}$, by substituting the above in (7.13), we conclude that the conditional probability of E_1 (and E_2) approaches 0 if $R_2 < h(\omega) - f_2(\omega)$. Similarly, we obtain $R_2 < h(\omega) - f_1(\omega)$ when the pirates are in the same group. ■

Let us show that the rate region thus defined is nontrivial. Given ω and γ, the maximizing values of the other arguments of φ are $\alpha = \omega(1-\gamma)$ and $\beta = \omega\gamma/(q-1)$, so

$$\varphi(\omega,\gamma,\alpha,\beta) \le h(\omega) - \gamma\omega\Big(\log_q\frac{q-1}{q-2} + \frac{\log_q(q-2)}{q-1}\Big).$$

Consequently, we get $\max(f_1(\omega), f_2(\omega)) \le h(\omega) - D$, where $D = D(\omega) = \omega^3\big(\log_q\frac{q-1}{q-2} + \frac{\log_q(q-2)}{q-1}\big)$ and $D(\omega) > 0$ for all $\omega > 0$. This shows that the r.-h.s. of (7.11) is positive. By Remark 7.12, the r.-h.s. of (7.10) is also positive if $q \ge t+1$ and $\frac{t-1}{t}\Big(1-\frac{1}{q^{t-1}}\Big) + \omega < \frac{q-1}{q}$. This calculation can be further refined because of the additional constraints on the parameters α, β, γ mentioned above.

7.5 Concluding Remarks

We introduced a new class of problems involving two-level codes, where the licensed users are organized in several groups. The main advantage of two-level codes is their ability to partially tolerate coalitions larger than the designed threshold. For instance, in the case of two-level fingerprinting codes, if the coalition size is beyond the designed limit, then up to a certain larger threshold, the tracing algorithm can identify a group containing a pirate. We presented constructions of codes with the two-level fingerprinting property. Our main focus was on the narrow-sense rule (2.3) and minimum distance based decoding.

In the next chapter, we investigate concatenated constructions with the objective of designing two-level codes with polynomial-time tracing. The concept of two-level codes raises several new questions. A few are identified below.

Open Problem 7.14. Construct two-level fingerprinting codes under the wide-sense rule (2.3) for generating forgeries. In particular, constructions for binary codes and other tracing algorithms besides MD decoding are especially of interest.

Open Problem 7.15. Find upper bounds on the rates of two-level fingerprinting (also TA, frameproof) codes.

Publications: The results of this chapter appear in [7]. An expanded version [8] is presently being prepared for publication.

Chapter 8

Fingerprinting Codes with Polynomial-time Tracing

In this chapter, we construct fingerprinting codes (both one-level and two-level) with a tracing algorithm of complexity polynomial in the fingerprint length and code rates better than previously known in the literature. We employ the technique of code concatenation (explained in Section 2.5, Eqns. (2.11)-(2.12)) in our constructions.

The use of concatenated codes in (one-level) fingerprinting has been demonstrated in previous constructions [26, 15]. The paper of Barg et al. [15] proposed the idea of using list decoding of algebraic codes [48, 47] to develop fingerprinting codes with polynomial-time tracing algorithms. Similar applications of list decoding to IPP and TA codes have been studied in [17] and [72] respectively. Yet, the best known rate of binary t-fingerprinting codes with polynomial-time decoding is approximately $\Omega(1/t^4 2^{2t})$ [15]. In comparison, with no complexity restraints the overall best available rate of binary t-fingerprinting codes is $\Omega(1/t^2)$ [79].

The ***main contribution*** of this chapter is an adaptation of the idea of concatenation to two-level fingerprinting codes. The definition of these codes was given in Chapter 7 where we provided some constructions which employ the minimum distance (MD) decoder (7.2). Attempting at improving the tracing complexity from exponential estimates for the MD decoder, we employ multilevel concatenation for the construction and analysis of efficient fingerprinting codes.

We also advance constructions of single-level efficient fingerprinting codes. Specifically, referring to earlier works [15, 40, 72] we suggest to use a high-rate, but relatively short fingerprinting code in the inner level of the concatenated scheme. As

before, the tracing complexity of the inner-level code will not have a major impact on the overall complexity of the scheme because of this code's small length. In this way we can combine powerful earlier constructions such as Tardos' codes [79] with algebraic outer codes arriving at one-level binary t-fingerprinting code with rate $\Omega(1/t^4)$ and polynomial-time tracing.

8.1 One-level Codes

Before proceeding to describe our construction, let us recall the list decoding result of Guruswami-Sudan (GS) [48, 47]. For $\mathbf{x}, \mathbf{y} \in \mathcal{Q}^N$, the notation $s_H(\mathbf{x}, \mathbf{y})$ stands for the number of coordinates i where $x_i = y_i$.

Theorem 8.1. *[48, 47] Let C be an $[N, K, \delta N]_q$ RS code (or one-point AG code) over the alphabet Q of size q. Then for any given $\mathbf{y} \in \mathcal{Q}^n$, the number of codewords $\mathbf{x} \in C$ such that*

$$s_H(\mathbf{x}, \mathbf{y}) \geq N\sqrt{1-\delta}$$

is polynomial in N. Moreover, there exists an algorithm with complexity polynomial in N which outputs the list of all such codewords.

8.1.1 Code Construction

In this section we employ the concatenated code construction to design t-fingerprinting codes.

Let the outer code C_{out} be an $[N, K, \Delta]_q$ RS (or one-point AG) code with $K = R_{\text{out}}N$ and $\Delta = \delta N$. Let $(\mathcal{C}_{\text{in}}, \mathcal{D}_{\text{in}})$ denote an $(m, q)_2$ randomized code which is t-fingerprinting with ε-error. For every outer coordinate $i = 1, \ldots, N$, generate an *independent* instance of $\mathcal{C}_{\text{in}}$ for the inner encoding. In this way we obtain a randomized binary concatenated code $\mathcal{C}$ of length $n = Nm$ and size q^K.

Assumptions: W.l.o.g. we assume the tracing strategy of the inner fingerprinting code always outputs exactly one user. In practice, sometimes the decoder can suggest several pirate candidates or fail to provide even one such candidate. We assume that in the former case the decoding output is chosen randomly from the candidate list, while in the latter case the decoder outputs a user chosen uniformly out of the set of all users.

Secondly, we assume that the inner fingerprinting code is "symmetric" across the users, meaning that the fingerprints of different users are identically distributed random variables. Indeed, most constructions, and in particular codes that we have in mind for this chapter's applications, satisfy this condition. (In fact, even if the fingerprinting code is asymmetric, we can artificially introduce symmetry by randomly permuting the inner codewords, similarly to [15]).

The following statement is clearly true.

Fact 8.2. *If the inner t-fingerprinting code of size q with ε-error satisfies the above assumptions, then for any coalition U of size at most t, the probability that a given innocent user $u' \notin U$ is accused during tracing is at most $\varepsilon/(q-t)$.*

In the subsequent text, we find it convenient to identify the users with the codewords of C_{out}.

Decoding Algorithm:

Given a forged fingerprint $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$, where each $\mathbf{y}_i \in \{0,1\}^m$, the algorithm operates as follows.

1) For every $i = 1, \ldots, N$, apply the tracing strategy $\mathcal{D}_{\text{in}}$ of the inner fingerprinting code to $\mathbf{y}_i$ to obtain a q-ary symbol $\hat{y}_i$. Completing this procedure for all N outer coordinates produces a q-ary vector $\hat{\mathbf{y}} = (\hat{y}_1, \ldots, \hat{y}_N)$.

2) Let $\xi > \varepsilon$. Using the GS list decoding algorithm of C_{out} for the vector $\hat{\mathbf{y}}$, output the list of all outer codewords (users) $\mathbf{x} \in C_{\text{out}}$ such that

$$s_H(\mathbf{x}, \hat{\mathbf{y}}) \geq \frac{1-\xi}{t} N. \tag{8.1}$$

The concatenated code above together with the decoding algorithm described is written as $(\mathcal{C}, \mathcal{D})$ below.

8.1.2 Main Result

Theorem 8.3. *Let $0 < \varepsilon < \xi$ and $\sigma = \frac{1-\xi}{t} - t(1-\delta)$. Suppose that the relative minimum distance of C_{out} satisfies*

$$\delta \geq 1 - \left(\frac{1-\xi}{t}\right)^2 + \frac{\varepsilon}{t(q-t)} \tag{8.2}$$

and the inner code $(\mathcal{C}_{\text{in}}, \mathcal{D}_{\text{in}})$ is t-fingerprinting with ε-error. Then the concatenated code $(\mathcal{C}, \mathcal{D})$ is t-fingerprinting with error probability at most

$$2^{-ND(\xi\|\varepsilon)} + q^{NR_{\text{out}}} 2^{-ND(\sigma\|\frac{\varepsilon}{q-t})} \tag{8.3}$$

and tracing complexity $\text{poly}(n)$.

Proof. Consider a coalition $U = \{\mathbf{x}_1, \ldots, \mathbf{x}_t\} \subseteq C_{\text{out}}$. Let $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_N)$, where $\mathbf{Y}_i$ are binary length-m vectors, be a random forgery generated by U using an admissible strategy. In any outer coordinate $i \in [N]$, the coalition observes at most t distinct q-ary symbols among $\{x_{1i}, \ldots, x_{ti}\}$. At the inner level, this can be viewed as a virtual coalition among at most t symbols, and correspondingly, $\mathbf{Y}_i$ is generated by an admissible strategy for these symbols. Thus if the output $\hat{Y}_i$ of the inner decoding $\mathcal{D}_{\text{in}}$ is correct and matches one of the symbols $\{x_{1i}, \ldots, x_{ti}\}$, then the narrow-sense rule (2.2) applies at the outer level in producing the q-ary vector

$\hat{\mathbf{Y}}$. This observation will be used in the analysis of the error probability which is split into two cases corresponding to the probability of missed detection and of identifying an innocent user.

Let $\chi_i, i = 1, \ldots, N$, be an r.v. that indicates an inner decoding error in the ith coordinate. Then for all i we have $\mathbf{P}[\chi_i = 1] \leq \varepsilon$ since the inner code has ε-error, and $\chi_i, i = 1, \ldots, N$, are mutually independent because independent instances of $\mathcal{C}_{\text{in}}$ are taken for every $i \in [N]$. Thus $Z = \sum_{i=1}^{N} \chi_i$ is a binomial r.v. denoting the number of coordinates where a decoding error occurs at the inner level. From our earlier observation, Z corresponds to the number of outer coordinates where the narrow-sense rule fails. Therefore, $\hat{\mathbf{Y}} \in \mathcal{E}_Z(\mathbf{x}_1, \ldots, \mathbf{x}_t)$, where the notation $\mathcal{E}_z(\cdot)$, introduced in (6.9), denotes the envelope of the outer codevectors when the narrow-sense rule holds only in some $N - z$ positions.

Let us begin by examining the probability that none of the members of U are output by the decoding algorithm.

$$\begin{aligned}
\mathbf{P}[\mathcal{D}(\mathbf{Y}) \cap U = \emptyset] &= \mathbf{P}\left[\max_{l \in [t]} s_H(\mathbf{x}_l, \hat{\mathbf{Y}}) < \frac{1-\xi}{t} N\right] \\
&\overset{(a)}{\leq} \mathbf{P}\left[\frac{N - Z}{t} < \frac{1-\xi}{t} N\right] \\
&= \mathbf{P}[Z > N\xi] \overset{(b)}{\leq} 2^{-ND(\xi\|\varepsilon)}.
\end{aligned}$$

The inequality (a) follows from the fact that if $\hat{\mathbf{y}} \in \mathcal{E}_z(\mathbf{x}_1, \ldots, \mathbf{x}_t)$, then there exists $l \in [t]$ such that $s_H(\mathbf{x}_l, \hat{\mathbf{y}}) \geq (N - z)/t$, while (b) is true by (1.1).

Next, consider the false-positive probability. Let $\mathbf{x}' \notin U$ be an innocent user. In any coordinate $i \in [N]$, there are two possible ways for the inner decoder to output the symbol x'_i. Namely, either x'_i is one of the symbols $\{x_{1i}, \ldots, x_{ti}\}$ of the actual coalition. The number of such positions is at most $t(1 - \delta)N$. Alternately, if x'_i is different from the coalition's symbols, it may be output when the inner decoder commits an error. By Fact 8.2, the probability of this event is at most $\varepsilon/(q - t)$.

Let $\tilde{Z}$ be a binomial r.v. that represents the number of coordinates where the latter error event occurs. Then

$$s_H(\mathbf{x}', \hat{\mathbf{Y}}) \leq \tilde{Z} + t(1-\delta)N$$

and we get

$$\begin{aligned}
\mathbf{P}\left[\mathbf{x}' \in \mathcal{D}(\mathbf{Y})\right] &= \mathbf{P}\left[s_H(\mathbf{x}', \hat{\mathbf{Y}}) \geq \frac{1-\xi}{t}N\right] \\
&\leq \mathbf{P}\left[\tilde{Z} + t(1-\delta)N \geq \frac{1-\xi}{t}N\right] \\
&= \mathbf{P}\left[\tilde{Z} \geq N\sigma\right] \leq 2^{-ND(\sigma\|\frac{\varepsilon}{q-t})},
\end{aligned}$$

where the last inequality again holds because of (1.1). Taking the union bound completes the proof of the error probability estimate.

Lastly, the tracing complexity is polynomial in the code length because under the condition (8.2) we have $(1-\xi)/t > \sqrt{1-\delta}$, and thus the GS algorithm (Theorem 8.1) succeeds in finding the list of codewords satisfying (8.1) in time poly(N). ■

We now proceed to make explicit code choices in Theorem 8.3. Let $(\mathcal{C}_{\text{in}}, \mathcal{D}_{\text{in}})$ be a sequence of binary t-fingerprinting codes with error probability $\varepsilon = 2^{-m\beta}$ for some $\beta > 0$ obtained from Tardos' construction [79]. The rate of $\mathcal{C}_{\text{in}}$ is close to $R_t = 1/(100t^2)$ for arbitrarily small β. The outer code C_{out} is a $[q, K]$ (extended) RS code satisfying the condition (8.2). We have $m \approx O(\log_2 n)$ since $n = m2^{mR_t}$, and so the tracing algorithm for the inner code has only polynomial complexity in the overall code length n. The first term in (8.3) is easily shown to be exponentially decaying in n. With ξ, t fixed and m growing, we have

$$D\Big(\sigma\Big\|\frac{\varepsilon}{q-t}\Big) \sim N\sigma \log_2 \frac{q}{\varepsilon} = n\sigma(R_t + \beta).$$

Therefore, with $R = R_{\text{out}} R_t$ representing the rate of the concatenated code, and using $1 - \delta \sim R_{\text{out}}$ for RS codes, the error probability (8.3) approaches 0 exponentially if

$$R < \left(\frac{1 - \xi}{t} - t R_{\text{out}} \right) R_t,$$

$$\text{i.e., } R < \frac{1 - \xi}{t(t+1)} R_t.$$

We can take ξ arbitrarily small and m sufficiently large to satisfy $\varepsilon < \xi$ obtaining

Corollary 8.4. *There exists a sequence of binary t-fingerprinting codes of length n with rate $\Omega(1/t^4)$, error probability $\exp(-\Omega(n))$, and having decoding complexity $\text{poly}(n)$.*

8.2 Two-level Codes

In this section we extend the concatenated construction of single-level codes of the previous section to the two-level scenario.

Following the definition of two-level codes earlier in Chapter 7, let $t_1 > t_2$ be the two threshold values of the coalition size. Let $\mathcal{Q}_1$ and $\mathcal{Q}_2$ denote finite alphabets of size Q_1 and Q_2 respectively. We introduce the operation $*$ which is the n-fold extension of the direct product operation on the alphabets. For $\mathbf{x} \in \mathcal{Q}_1^N$, $\mathbf{y} \in \mathcal{Q}_2^N$,

$$\mathbf{x} * \mathbf{y} = ((x_1, y_1), \ldots, (x_N, y_N)) \in (\mathcal{Q}_1 \times \mathcal{Q}_2)^n.$$

Let C_1 be an $[N, K_1, \Delta_1]_{Q_1}$ linear code and C_2 be an $[N, K_2, \Delta_2]_{Q_2}$ linear code. The $*$ product extends to the codes C_1 and C_2 as follows:

$$C_1 * C_2 = \{\mathbf{x}_1 * \mathbf{x}_2 : \mathbf{x}_1 \in C_1, \mathbf{x}_2 \in C_2\}.$$

To obtain a two-level code, we associate C_1 with groups and C_2 with users within the group. Thus $C_1 * C_2$ can be viewed as an $(N, Q_1^{K_1}, Q_2^{K_2})_{Q_1Q_2}$ two-level code over the alphabet $\mathcal{Q}_1 \times \mathcal{Q}_2$. Obviously, it is true that

$$d_1(C_1 * C_2) \geq \Delta_1 \quad \text{and} \quad d_2(C_1 * C_2) \geq \Delta_2,$$

where the quantities d_1 and d_2 are defined in (7.3), (7.4). Therefore, choosing $\Delta_1 > N(1 - 1/t_1^2)$ and $\Delta_2 > N(1 - 1/t_2^2)$ makes the resulting code $C_1 * C_2$ into a (t_1, t_2)-TA code by Proposition 7.6. This observation forms the motivation for the code choices in our concatenated scheme described below.

8.2.1 Code Construction

In this section we construct a (t_1, t_2)-fingerprinting code by adapting the concatenation technique for two-level codes.

Let C_1 and C_2 both be RS (or one-point AG) codes with parameters $[N, K_1, \Delta_1]_{Q_1}$ and $[N, K_2, \Delta_2]_{Q_2}$ respectively, where $K_i = R_{i,\text{out}}N$ and $\Delta_i = \delta_i N$ for $i = 1, 2$. Each codeword from C_1 corresponds to a particular group, while the codewords of C_2 are associated with the user indices within a group. Then the outer code $C_{\text{out}} = C_1 * C_2$ is an $(N, Q_1^{K_1}, Q_2^{K_2})_{Q_1Q_2}$ deterministic two-level code. Let $(\mathcal{C}_{\text{in}}, \mathcal{D}_{1,\text{in}}, \mathcal{D}_{2,\text{in}})$ denote an $(m, Q_1, Q_2)_q$ randomized code which is (t_1, t_2)-fingerprinting with ε-error under exhaustive search decoding. For every outer coordinate $i = 1, \ldots, N$, we generate an *independent* instance of $(\mathcal{C}_{\text{in}}, \mathcal{D}_{1,\text{in}}, \mathcal{D}_{2,\text{in}})$ for the inner level.

For a given user $\mathbf{u} \equiv (u_1, u_2)$, the fingerprint is assigned as follows. At the outer level, pick $\mathbf{x}_1 \in C_1$ and $\mathbf{x}_2 \in C_2$ corresponding to u_1 and u_2 respectively, and construct $\mathbf{x} = \mathbf{x}_1 * \mathbf{x}_2$. Next, for each $i = 1, \ldots, N$, encode $(x_{1i}, x_{2i}) \in \mathcal{Q}_1 \times \mathcal{Q}_2$ using the realization of the two-level code $\mathcal{C}_{\text{in}}$. This procedure results in a concatenated code $\mathcal{C}$ which is a randomized $(n, Q_1^{K_1}, Q_2^{K_2})_q$ two-level code with $n = Nm$.

Remark 8.5. This code construction is an adaptation of *generalized concatenated codes* which are known in coding theory literature [24] for problems dealing with error correction.

Below, we state some assumptions on the inner fingerprinting code used in our construction which are analogous to the assumptions made for the one-level construction.

Assumptions: Suppose that the tracing strategy $\mathcal{D}_{2,\text{in}}$ of the inner fingerprinting code always outputs exactly one user. As before in Section 8.1, sometimes the decoder can suggest several pirate candidates or fail to provide even one such candidate. We assume that in the former case the decoding output is chosen randomly from the candidate list, while in the latter case the decoder outputs a user chosen uniformly out of the set of all users. Secondly, it is assumed that for any given forged fingerprint $\mathbf{y}$, $\mathcal{D}_{1,\text{in}}(\mathbf{y}) = \mathcal{G}(\mathcal{D}_{2,\text{in}}(\mathbf{y}))$ as in the case of MD decoding.

We also assume that the inner fingerprinting code is "symmetric" across the users, meaning that the fingerprints of different users are identically distributed random variables. We also assume that this applies to different groups as a whole.

Note that all the specified conditions are satisfied for the codes presented in Section 7.4, and are quite reasonable to expect in other general constructions.

Fact 8.6. *If the inner* $(m, Q_1, Q_2)_q$ *code* $(\mathcal{C}_{\text{in}}, \mathcal{D}_{1,\text{in}}, \mathcal{D}_{2,\text{in}})$ *which is* (t_1, t_2)*-fingerprinting code with* ε*-error satisfies the above assumptions, then*

(a) For any coalition U *of size at most* t_2*, the probability that a given innocent user* $\mathbf{u}' \notin U$ *such that* $\mathcal{G}(\mathbf{u}') \in \mathcal{G}(U)$ *is accused by* $\mathcal{D}_{2,\text{in}}$ *is at most* $\varepsilon/(Q_2 - t_2)$*.*

(b) For any coalition U *of size at most* t_1*, the probability that a given group* $u' \notin \mathcal{G}(U)$ *is accused by* $\mathcal{D}_{1,\text{in}}$ *is at most* $\varepsilon/(Q_1 - t_1)$*.*

In the subsequent text, the users are identified with the codewords of C_{out}. For $\mathbf{x} = \mathbf{x}_1 * \mathbf{x}_2 \in C_{\text{out}}$, with some abuse of notation we write $\mathcal{G}(\mathbf{x}) = \mathbf{x}_1$. We again make use of GS list decoding (Theorem 8.1) in the tracing algorithm below.

Decoding Algorithm:

Given a forged fingerprint $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$, where each $\mathbf{y}_i$ is a q-ary length-m vector, the algorithm operates as follows.

1) For every $i = 1, \ldots, N$, apply the tracing strategy $\mathcal{D}_{2,\text{in}}$ of the inner fingerprinting code to $\mathbf{y}_i$ to obtain a $\mathcal{Q}_1 \times \mathcal{Q}_2$ symbol $(\hat{y}_{1i}, \hat{y}_{2i})$. Completing this procedure for all N outer coordinates produces a vector $\hat{\mathbf{y}} = \hat{\mathbf{y}}_1 * \hat{\mathbf{y}}_2$ over the alphabet $\mathcal{Q}_1 \times \mathcal{Q}_2$.

2) Let $\xi > \varepsilon$. Run the GS list decoding algorithm for C_1 and C_2 to compute the lists
$$L_j(\hat{\mathbf{y}}_j) = \left\{ \mathbf{x}_j \in C_j : s_H(\mathbf{x}_j, \hat{\mathbf{y}}_j) \geq \frac{1-\xi}{t_j} N \right\}, \qquad j = 1, 2. \tag{8.4}$$

3) Let $L_1(\hat{\mathbf{y}}_1)$ be the output of the decoder $\mathcal{D}_1$. The decoder $\mathcal{D}_2$ outputs the list
$$L(\hat{\mathbf{y}}) = \left\{ \mathbf{x} \in C_{\text{out}} : s_H(\mathbf{x}, \hat{\mathbf{y}}) \geq \frac{1-\xi}{t_2} N \right\}. \tag{8.5}$$
computed as follows. First we find $L_1(\hat{\mathbf{y}}_1) * L_2(\hat{\mathbf{y}}_2)$ and then construct the subset list $L(\hat{\mathbf{y}})$ by eliminating codewords that appear in the product list above but do not satisfy the condition for $L(\hat{\mathbf{y}})$.

The concatenated code thus defined together with the decoding algorithm described is denoted by $(\mathcal{C}, \mathcal{D}_1, \mathcal{D}_2)$ below.

8.2.2 Main Result

Theorem 8.7. *Let* $0 < \varepsilon < \xi$ *and* $\sigma_i = \frac{1-\xi}{t_i} - t_i(1-\delta_i)$, $i = 1, 2$. *Suppose that the relative minimum distances of* C_1 *and* C_2 *satisfy*

$$\delta_i \geq 1 - \left(\frac{1-\xi}{t_i}\right)^2 + \frac{\varepsilon}{t_i(Q_i - t_i)}, \quad i = 1, 2, \tag{8.6}$$

and the inner code $(\mathcal{C}_{\text{in}}, \mathcal{D}_{1,\text{in}}, \mathcal{D}_{2,\text{in}})$ *is* (t_1, t_2)*-fingerprinting with* ε*-error. Then the concatenated code* $(\mathcal{C}, \mathcal{D}_1, \mathcal{D}_2)$ *is* (t_1, t_2)*-fingerprinting with error probability at most*

$$q^{-ND(\xi\|\varepsilon)} + Q_1^{K_1} q^{-ND(\sigma_1\|\frac{\varepsilon}{Q_1 - t_1})} + Q_2^{K_2} q^{-ND(\sigma_2\|\frac{\varepsilon}{Q_2 - t_2})} \tag{8.7}$$

and decoding complexity $\text{poly}(n)$.

Proof. We begin by outlining a basic argument used repeatedly in the proof. Throughout the proof, we write a coalition U of size t as a subset $\{\mathbf{x}^1, \ldots, \mathbf{x}^t\} \subseteq C_{\text{out}}$ where $\mathbf{x}^i = \mathbf{x}_1^i * \mathbf{x}_2^i$, $i = 1, \ldots, t$. Let $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_N)$, where $\mathbf{Y}_i$ are q-ary vectors of length m, be a random forgery generated by U using any admissible strategy. In any outer coordinate $i \in [N]$, the coalition observes at most t distinct $\mathcal{Q}_1 \times \mathcal{Q}_2$ symbols among $\{(x_{1i}^1, x_{2i}^1), \ldots, (x_{1i}^t, x_{2i}^t)\}$. At the inner level this is equivalent to the action of a virtual coalition of size at most t, and correspondingly $\mathbf{Y}_i$ is generated by an admissible strategy from these symbols. This enables us to utilize the fingerprinting property of the inner code to derive some properties for $\hat{\mathbf{Y}} = \hat{\mathbf{Y}}_1 * \hat{\mathbf{Y}}_2 \in (\mathcal{Q}_1 \times \mathcal{Q}_2)^N$ which is the result of inner level decoding performed in Step 1 of the algorithm.

We use the above argument to establish the (t_1, t_2)-fingerprinting property of the concatenated code by analyzing separately t_1- and t_2-sized coalitions. For each of these two cases, we analyze the probability of missed detection and of identifying an innocent user.

Size-t_1 coalitions: Consider a coalition $U = \{\mathbf{x}^1, \ldots, \mathbf{x}^{t_1}\}$ and a random forgery $\mathbf{Y}$ generated by U. As mentioned above, this amounts to a virtual coalition of at most t_1 symbols acting in every outer coordinate. Therefore, for every $i \in [N]$, the probability that the output $\hat{Y}_{1i}$ of the inner decoder $\mathcal{D}_{1,\text{in}}$ (i.e., the group index output by $\mathcal{D}_{2,\text{in}}$) does not match one of the symbols $\{x_{1i}^1, \ldots, x_{1i}^{t_1}\}$ is at most ε. Let Z_1 be a binomial r.v. denoting the number of coordinates where this error event occurs. Then

$$\max_{l \in [t_1]} s_H(\mathbf{x}_1^l, \hat{\mathbf{Y}}_1) \geq \frac{N - Z_1}{t_1},$$

and so the probability that none of the guilty groups are output

$$\begin{aligned}
\mathbf{P}\left[\mathcal{D}_1(\mathbf{Y}) \cap \mathcal{G}(U) = \emptyset\right] &= \mathbf{P}\left[L_1(\hat{\mathbf{Y}}_1) \cap \mathcal{G}(U) = \emptyset\right] \\
&\leq \mathbf{P}\left[\frac{N - Z_1}{t_1} < \frac{1-\xi}{t_1} N\right] \\
&= \mathbf{P}\left[Z_1 > N\xi\right] \leq q^{-ND(\xi \| \varepsilon)},
\end{aligned}$$

where the last inequality holds due to (1.1). This concludes the analysis of the missed group detection case.

Next, consider an innocent group $\mathbf{x}_1' \notin \mathcal{G}(U)$. In any coordinate $i \in [N]$, there are two possible ways for $\mathcal{G}(\mathcal{D}_{2,\text{in}})$ to output the symbol x_{1i}'. The first possibility is that $x_{1i}' \in \{x_{1i}^1, \ldots, x_{1i}^{t_1}\}$, and the number of such positions is at most $t_1(1-\delta_1)N$. Otherwise, if x_{1i}' is different from the above symbols, it can be output when the inner decoder makes an error. By Fact 8.6(b) the probability of this event is at most $\varepsilon/(Q_1 - t_1)$. Let $\tilde{Z}_1$ be a binomial r.v. counting the number of coordinates where the latter error event occurs. Then

$$s_H(\mathbf{x}_1', \hat{\mathbf{Y}}_1) \leq \tilde{Z}_1 + t_1(1-\delta_1)N$$

and we obtain

$$\begin{aligned}\mathbf{P}\left[\mathbf{x}'_1 \in \mathcal{D}_1(\mathbf{Y})\right] &= \mathbf{P}\left[s_H(\mathbf{x}'_1, \hat{\mathbf{Y}}_1) \geq \frac{1-\xi}{t_1}N\right] \\ &\leq \mathbf{P}\left[\tilde{Z}_1 + t_1(1-\delta_1)N \geq \frac{1-\xi}{t_1}N\right] \\ &= \mathbf{P}\left[\tilde{Z}_1 \geq N\sigma_1\right] \leq q^{-ND(\sigma_1 \| \frac{\varepsilon}{Q_1 - t_1})},\end{aligned}$$

because by (8.6), $\sigma_1 > \varepsilon/(Q_1 - t_1)$. Applying the union bound, we conclude that the error probability is less than the estimate (8.7).

Size-t_2 coalitions: Consider the coalition $U = \{\mathbf{x}^1, \ldots, \mathbf{x}^{t_2}\}$ and let $\mathbf{Y}$ be a forged fingerprint generated by it. For every $i \in [N]$, we employ the inner decoder $\mathcal{D}_{2,\text{in}}$ to obtain $(\hat{Y}_{1i}, \hat{Y}_{2i})$. The probability that $(\hat{Y}_{1i}, \hat{Y}_{2i})$ is not one of the symbols $\{(x^1_{1i}, x^1_{2i}), \ldots, (x^{t_2}_{1i}, x^{t_2}_{2i})\}$ is at most ε. Let the binomial r.v. Z_2 denote the number of coordinates where this error event occurs. Then

$$\max_{l \in [t_2]} s_H(\mathbf{x}^l, \hat{\mathbf{Y}}) \geq \frac{N - Z_2}{t_2},$$

and so the probability that none of the users in U are output

$$\begin{aligned}\mathbf{P}\left[\mathcal{D}_2(\mathbf{Y}) \cap U = \emptyset\right] &= \mathbf{P}\left[L(\hat{\mathbf{Y}}) \cap U = \emptyset\right] \\ &\leq \mathbf{P}\left[\frac{N - Z_2}{t_2} < \frac{1-\xi}{t_2}N\right] \\ &= \mathbf{P}\left[Z_2 > N\xi\right] \leq q^{-ND(\xi \| \varepsilon)}.\end{aligned}$$

This concludes the analysis of the missed detection case.

Next, consider an innocent user $\mathbf{x}' \notin U$ contained in one of the guilty groups, i.e., $\mathbf{x}'_1 \in \mathcal{G}(U)$. In any coordinate $i \in [N]$, there are two possible ways for $\mathcal{D}_{2,\text{in}}$ to output the symbol (x'_{1i}, x'_{2i}). One possibility is that $(x'_{1i}, x'_{2i}) \in \{(x^1_{1i}, x^1_{2i}), \ldots, (x^{t_2}_{1i}, x^{t_2}_{2i})\}$, and there are at most $t_2(1-\delta_2)N$ such positions. Secondly, if (x'_{1i}, x'_{2i}) is different

from the actual coalition's symbols, it may be output when $\mathcal{D}_{2,\text{in}}$ makes an error. The probability of this event is at most $\varepsilon/(Q_2 - t_2)$ by Fact 8.6(a). Let $\tilde{Z}_2$ be a binomial r.v. denoting the number of coordinates where the second error event occurs. Then

$$s_H(\mathbf{x}', \hat{\mathbf{Y}}) \le \tilde{Z}_2 + t_2(1 - \delta_2)N,$$

and we get

$$\begin{aligned}
\mathbf{P}\left[\mathbf{x}' \in \mathcal{D}_2(\mathbf{Y})\right] &= \mathbf{P}\left[s_H(\mathbf{x}', \hat{\mathbf{Y}}) \ge \frac{1-\xi}{t_2}N\right] \\
&\le \mathbf{P}\left[\tilde{Z}_2 + t_2(1-\delta_2)N \ge \frac{1-\xi}{t_2}N\right] \\
&= \mathbf{P}\left[\tilde{Z}_2 \ge N\sigma_2\right] \le q^{-ND(\sigma_2 \| \frac{\varepsilon}{Q_2 - t_2})},
\end{aligned}$$

since $\sigma_2 > \varepsilon/(Q_2 - t_2)$ from (8.6).

The only remaining case is to consider an innocent group $\mathbf{x}'_1 \notin \mathcal{G}(U)$. The analysis is similar to the corresponding case for coalitions of size t_1, so we omit some details. We have

$$\begin{aligned}
\mathbf{P}\left[\mathbf{x}'_1 \in \mathcal{G}(\mathcal{D}_2(\mathbf{Y}))\right] &\le \mathbf{P}\left[s_H(\mathbf{x}'_1, \hat{\mathbf{Y}}_1) \ge \frac{1-\xi}{t_2}N\right] \\
&\le q^{-ND(\sigma_1 \| \frac{\varepsilon}{Q_1 - t_2})}.
\end{aligned}$$

Using the union bound, we see that the error probability is less than the estimate (8.7).

Finally, let us show that the tracing complexity is polynomial in n. It is straightforward to see that $(1-\xi)/t_i > \sqrt{1-\delta_i}$, $i = 1, 2$, under the condition (8.6). Therefore, the GS list decoding algorithm (Theorem 8.1) succeeds in finding the lists L_1 and L_2, defined in (8.4), with polynomial complexity. Moreover, the list

L (see (8.5)) can also be computed in polynomial time since both lists L_1 and L_2 are of size polynomial in N. ■

Let us analyze the rates attained by Theorem 8.7 by fixing our code choices. Let C_1 and C_2 be extended RS codes with parameters $[Q, K_1]$ and $[Q, K_2]$, respectively, satisfying the condition (8.6). At the inner level, consider a sequence of q-ary (t_1, t_2)-fingerprinting codes with error probability $\varepsilon = o(1)$ and rate $(R_{\text{in}}, R_{\text{in}})$. The tracing procedure of the inner code will be performed by exhaustive search; for instance, the codes in Section 7.4 can be used at the inner level. We have $m \approx O(\log_q n)$ since $n = mq^{mR_{\text{in}}}$. Hence, the tracing for the inner code has only polynomial complexity in the code length n. With ξ, t_1, t_2 fixed and m growing, we have

$$D\Big(\sigma_i \Big|\Big| \frac{\varepsilon}{Q - t_i}\Big) \sim N\sigma_i \log_q \frac{Q}{\varepsilon} \geq n\sigma_i R_{\text{in}}, \quad i = 1, 2.$$

Let $R_i = R_{i,\text{out}} R_{\text{in}}$, $i = 1, 2$ denote the rate pair of the concatenated code. Since for RS codes we have $1 - \delta_i \sim R_{i,\text{out}}$, the error probability (8.7) for the concatenated code approaches 0 if

$$R_i < \left(\frac{1 - \xi}{t_i} - t_i R_{i,\text{out}}\right) R_{\text{in}}, \quad i = 1, 2,$$

$$\text{i.e., } R_i < \frac{1 - \xi}{t_i(t_i + 1)} R_{\text{in}}, \quad i = 1, 2.$$

Finally, taking ξ arbitrarily small and m sufficiently large to satisfy $\varepsilon < \xi$ we obtain the following result.

Corollary 8.8. *There exists a sequence of q-ary (t_1, t_2)-fingerprinting codes of length n with error probability decaying with n, having decoding complexity* $\text{poly}(n)$ *and rate pair $R_i = \Omega(R_{\text{in}}/t_i^2)$, $i = 1, 2$.*

The material of this chapter is intended for publication as a part of [8].

Chapter 9

Beyond the Disk Model of Wireless Networks

In this chapter we describe the part of our research that deals with connectivity analysis of WSNs using random graphs. We study two such applications introduced informally in Section 1.2. In both the applications of interest, the scarce energy resources of the sensors translate into a limited range for communication. For this reason we assume that any two nodes can establish a link only if they are located within the communication range. This geometric constraint is captured by random graphs referred to as *disk models* in the literature, see Gupta and Kumar [46], Han and Makowski [50], or Penrose's book [66]. In the networks considered in this chapter, further constraints need to be taken into account by the random graph model as explained below.

The first problem that we consider concerns a WSN where the link between each pair of nodes can experience an outage independently of the other links with a certain probability. Thus, two nodes can communicate directly if they are located within the range, and in addition the link between them is active. We study a simple network topology described by the sensors being placed on a circle, making the first step of the analysis of the exact threshold parameters for connectivity of the class of geometric random graphs with link failures studied in this chapter. As our ***main result*** we provide a complete characterization of the ***zero-one law***, thereby establishing the exact critical parameter scaling which guarantees that with high probability none of the sensors are isolated. This result strengthens similar theorems proved earlier in the works of Yi et al. [88] by removing assumptions on the outage probability taken there.

The second problem that we study addresses node isolation in a WSN where a probabilistic scheme is used to assign secret keys to the sensors for establishing secure links. In this case, a secure link is possible between two sensors if they are within the range and possess a shared secret key. We study both secure WSNs and networks with random link failures with nodes located on a sphere or on a torus in $\mathbb{R}^d$. We prove a one-law for such networks, which establishes sufficient conditions for the scalings so that with high probability, the WSN does not contain isolated nodes.

9.1 Model and Assumptions

We begin with a formal definition of the random graph models studied in our research. Throughout this chapter, we are only concerned with undirected graphs. Following standard terminology, we say a node is *isolated* if no edge exists between the node and any of the remaining nodes. Throughout this chapter we use natural logarithms.

Let n denote the number of nodes in the WSN. Each node is identified by a label in $\{1, \ldots, n\}$. We first describe the geometric component of the model which captures the range constraints associated with both kinds of random graphs studied below. Suppose that the n nodes are placed in a compact region $\mathbb{D} \subseteq \mathbb{R}^d$ $(d > 0)$ (the only regions that we consider in this chapter are: the interval $[0, 1]$, a circle, a sphere, and a torus in $\mathbb{R}^d$). Let the r.v.'s $X_1, \ldots, X_n$ represent the node locations. We assume that $X_1, \ldots, X_n$ are i.i.d. r.v.'s which are distributed uniformly over $\mathbb{D}$. Once deployed, the sensors do not change their location. Let $d(\cdot, \cdot)$ denote an appropriate notion of distance on $\mathbb{D}$. All nodes have the same transmission range $r > 0$. Thus, nodes i and j are within the range if $d(X_i, X_j) \leq r$. Let $\chi_{ij}(r) := \mathbf{1}[d(X_i, X_j) \leq r]$ denote the indicator r.v. of this event.

The 1-dimensional case $d = 1$ will be our main focus. Specifically, we take $\mathbb{D} = [0, 1]$ as well as $[0, 1]$ whose ends are identified (a circle). Correspondingly, the distance is measured either as

$$d(x, y) = |x - y|, \quad x, y \in [0, 1]$$

or as

$$d(x, y) = \min(|x - y|, 1 - |x - y|), \quad x, y \in [0, 1].$$

The unit circle is a simpler model to study because it eliminates the boundary effects in the unit interval. For any pair of nodes $i, j \in [n]$ on the unit circle, observe that $\mathbf{P}[\chi_{ij}(r) = 1] = \mathbf{P}[d(X_i, X_j) \leq r] = \min(1, 2r)$. For convenience, we use the shorthand

$$\ell(r) := \min(1, 2r), \quad r \geq 0$$

in the subsequent text.

To distinguish between the cases of the unit circle and unit interval, we use the superscripts (C) and (L) respectively in our notation. As a way to lighten the notation, we omit the superscripts (C) and (L) when the discussion applies equally well to both cases. We now proceed to describe each of the random graphs studied below.

9.1.1 WSNs with random link failures

Let $p \in [0, 1]$ be the probability that a link is active (i.e., not in outage). Let $\{B_{ij}(p),\ 1 \leq i < j \leq n\}$ be a collection of i.i.d. $\{0, 1\}$-valued r.v.'s with success probability p. The link between nodes i and j is active if $B_{ij}(p) = 1$. Throughout we always assume that the r.v.'s $\{X_i,\ i = 1, \ldots, n\}$ representing the node locations, and $\{B_{ij}(p),\ 1 \leq i < j \leq n\}$ are *mutually independent*.

The random graph model for WSNs with link failures is parametrized by the number n of nodes, the transmission range r and the probability p. To lighten the notation we often group the parameters r and p into the ordered pair $\boldsymbol{\theta} \equiv (r, p)$ and denote the random graph by $\mathbb{G}(n; \boldsymbol{\theta})$.

Two nodes in the WSN can communicate if and only if they are located within range and the pairwise link between them is active. Therefore, the indicator r.v. $\chi_{ij}(\boldsymbol{\theta})$ that an edge is present between nodes i and j in $\mathbb{G}(n; \boldsymbol{\theta})$ is given by

$$\chi_{ij}(\boldsymbol{\theta}) = \begin{cases} \chi_{ij}(r)B_{ij}(p) & \text{if } \ i < j \\ \\ \chi_{ij}(r)B_{ji}(p) & \text{if } \ j < i. \end{cases}$$

For each $i = 1, \ldots, n$, node i is isolated in $\mathbb{G}(n; \boldsymbol{\theta})$ if it is either not within transmission range from each of the $(n-1)$ remaining nodes, or within range from some nodes, but the corresponding links all are inactive. The indicator r.v. $\chi_{n,i}(\boldsymbol{\theta})$ that node i is an isolated node in $\mathbb{G}(n; \boldsymbol{\theta})$ can be expressed as

$$\chi_{n,i}(\boldsymbol{\theta}) = \prod_{i=1,\ j \neq i}^{n} \left(1 - \chi_{ij}(\boldsymbol{\theta})\right). \tag{9.1}$$

The number of isolated nodes in $\mathbb{G}(n; \boldsymbol{\theta})$ is equal to

$$I_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \chi_{n,i}(\boldsymbol{\theta}). \tag{9.2}$$

9.1.2 Secure WSNs

Next, we formally describe the randomized key distribution scheme proposed by Eschenauer and Gligor [38] and define a random graph model for a WSN employing this scheme. Suppose the distributor has a total of P secret keys available in a key pool. Each node i is assigned a size-K subset $S_i(K, P)$ of keys chosen uniformly

at random from the key pool. The procedure is repeated *independently* for each node $i \in [n]$.

The parameters for the random graph model are the number n of nodes, the transmission range r, and the numbers K and P. For convenience, the parameters r, K and P are grouped, and written as $\boldsymbol{\omega} \equiv (r, K, P)$. The corresponding graph model is denoted by $\mathbb{G}(n; \boldsymbol{\omega})$.

In this case, two sensors can establish a secure link if and only if they are located within range and they have a common secret key. Therefore, the indicator r.v. $\chi_{ij}(\boldsymbol{\omega})$ for the edge between nodes i and j in $\mathbb{G}(n; \boldsymbol{\omega})$ is given by

$$\chi_{ij}(\boldsymbol{\omega}) = \chi_{ij}(r)\mathbf{1}\left[S_i(K, P) \cap S_j(K, P) \neq \emptyset\right].$$

Accordingly, the indicator r.v. $\chi_{n,i}(\boldsymbol{\omega})$ that node i is isolated, and the number of isolated nodes in $\mathbb{G}(n; \boldsymbol{\omega})$ are defined by substituting $\boldsymbol{\omega}$ for $\boldsymbol{\theta}$ in (9.1) and (9.2), respectively. For any two nodes, $i, j \in [n]$, we note that

$$\mathbf{P}\left[S_i(K, P) \cap S_j(K, P) = \emptyset\right] = \frac{\binom{P-K}{K}}{\binom{P}{K}}.$$

We find it convenient to denote the r.-h.s. above by

$$q(K, P) := \frac{\binom{P-K}{K}}{\binom{P}{K}}.$$

Remark 9.1. For each of the models described above, it will be convenient to view the specific cases on the circle and interval, $\mathbb{G}^{(C)}(n; \cdot)$ and $\mathbb{G}^{(L)}(n; \cdot)$, as *coupled* in that they are constructed from the *same* r.v.'s $X_1, \ldots, X_n$ defined on the *same* probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

9.1.3 Objectives

Some terminology: A *scaling* for the WSN with link failures is defined as a mapping $\boldsymbol{\theta} : \mathbb{N}_0 \to \mathbb{R}_+ \times [0, 1]$. Similarly, in the case of a secure WSN, the *scaling* is a mapping $\boldsymbol{\omega} : \mathbb{N}_0 \to \mathbb{R}_+ \times \mathbb{N} \times \mathbb{N}$.

The main objective of this chapter can be stated as follows: For each of the random graph models $\mathbb{G}(n; \boldsymbol{\theta})$ and $\mathbb{G}(n; \boldsymbol{\omega})$ defined above, we wish to identify the conditions for the parameter scalings such that the probability that the random graph contains no isolated nodes is either 0 or 1.

Specifically, we would like to establish conditions on the scalings $\boldsymbol{\theta}_n$ and $\boldsymbol{\omega}_n$ to ensure that

$$\lim_{n\to\infty} \mathbf{P}\left[I_n(\boldsymbol{\theta}_n) = 0\right] = 1 \quad \text{(resp., 0)},$$

$$\lim_{n\to\infty} \mathbf{P}\left[I_n(\boldsymbol{\omega}_n) = 0\right] = 1 \quad \text{(resp., 0)}.$$

In the literature such results are known as *zero-one laws*. Interest in them stems from their ability to capture the threshold behavior of the underlying random graphs.

9.2 Previous Work

Some previous results related to the random graph models studied in this chapter are mentioned in Section 1.2.2. In particular, the best characterization of the zero-one laws for $\mathbb{G}(n; \boldsymbol{\theta})$ and $\mathbb{G}(n; \boldsymbol{\omega})$ is given by Yi et al. [88], where the 2-dimensional case with the nodes located on a unit-area disk or square is considered. The authors prove that the asymptotic distribution of the number of isolated nodes in both random graph models is Poisson under certain conditions on the parameter scalings. For the case of WSNs with random link failures, a zero-one law (stated in Theorem 9.2 below) can be obtained as a direct consequence of their results.

Let us associate the sequence α_n with a scaling $\boldsymbol{\theta}_n$ for the WSN with random link failures through

$$p_n \tilde{\ell}(r_n) = \frac{\log n + \alpha_n}{n}, \quad n = 1, 2, \ldots \tag{9.3}$$

where $\tilde{\ell}(r) := \pi r^2$.

Theorem 9.2. *[88] If $\boldsymbol{\theta}_n$ is a scaling for the WSN with random link failures such that*

$$\lim_{n\to\infty} p_n \log n = \infty, \tag{9.4}$$

then we have the zero-one law

$$\lim_{n\to\infty} \mathbf{P}\left[I_n(\boldsymbol{\theta}_n) = 0\right] = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = -\infty \\ \\ 1 & \text{if } \lim_{n\to\infty} \alpha_n = +\infty, \end{cases}$$

where the sequence α_n is determined through (9.3).

We draw the reader's attention to the technical *assumption* (9.4) used in the above zero-one law. For instance, if the link outage probability for the given application scales as $p_n = \sqrt{(\log n)/n}$, then Theorem 9.2 is unable to provide the transmission range (i.e., power) needed to guarantee the absence of isolated nodes. Therefore, the above result provides only a partial characterization of the zero-one law for WSNs with random link failures. Similar remarks also apply for the analogue result obtained from [88] for secure WSNs.

Our goal is to eliminate the need for additional assumptions and thereby provide a complete zero-one law. We establish such a result for WSNs with random link failures on the circle.

9.3 Main Results

WSN with link failures on the unit circle. With a scaling $\boldsymbol{\theta}_n$ for the WSN with link failures we associate the sequence α_n through

$$p_n \ell(r_n) = \frac{\log n + \alpha_n}{n}, \quad n = 1, 2, \ldots \tag{9.5}$$

In the case of the random graph $\mathbb{G}^{(C)}(n; \boldsymbol{\theta})$ defined on the unit circle, we get a complete zero-one law.

Theorem 9.3 (Unit circle)**.** *For any scaling $\boldsymbol{\theta}_n$ for the WSN with random link failures, we have the zero-one law*

$$\lim_{n\to\infty} \mathbf{P}\left[I_n^{(C)}(\boldsymbol{\theta}_n) = 0\right] = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = -\infty \\ \\ 1 & \text{if } \lim_{n\to\infty} \alpha_n = +\infty \end{cases}$$

where the sequence α_n is determined through (9.5).

WSN with link failures on the unit interval. With a scaling $\boldsymbol{\theta}_n$, we also associate the sequence α'_n through

$$p_n \ell(r_n) = \frac{2(\log n - \log\log n) + \alpha'_n}{n}, \quad n = 1, 2, \ldots \tag{9.6}$$

For the random graph $\mathbb{G}^{(L)}(n; \boldsymbol{\theta})$ defined on the unit interval there is a gap between the zero and one laws that we are able to prove.

Theorem 9.4 (Unit interval)**.** *For any scaling $\boldsymbol{\theta}_n$ for the WSN with random link failures, we have the zero-one law*

$$\lim_{n\to\infty} \mathbf{P}\left[I_n^{(L)}(\boldsymbol{\theta}_n) = 0\right] = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = -\infty \\ \\ 1 & \text{if } \lim_{n\to\infty} \alpha_n' = +\infty, \end{cases}$$

where the sequences α_n, α_n' are determined through (9.5) and (9.6), respectively.

Remark 9.5. An elementary coupling argument shows that for any particular realization of the r.v.'s $\{X_i,\ i = 1, \ldots, n\}$ and $\{B_{ij}(p),\ 1 \le i < j \le n\}$, the graph on the circle contains more edges than the graph on the interval. As a result, the zero law for the unit circle automatically implies the zero law for the unit interval, and so only the former needs to be established.

General results. Consider the case where the nodes are located in a region $\mathbb{D}$ which is a *sphere* or a *torus* in $\mathbb{R}^d$. Let us define

$$\ell_{\mathbb{D}}(r) := \mathbf{P}\left[d(x, Y) \le r\right],$$

where x is an arbitrary point in $\mathbb{D}$, the r.v. Y is uniformly distributed over $\mathbb{D}$, and $d(\cdot, \cdot)$ is the appropriate notion of distance. The quantity $\ell_{\mathbb{D}}(r)$ gives the probability that any two nodes are within the transmission range r.

For a scaling $\boldsymbol{\theta}_n$ for the WSN with random link failures, define the sequence α_n through

$$p_n \ell_{\mathbb{D}}(r_n) = \frac{\log n + \alpha_n}{n}, \quad n = 1, 2, \ldots.$$

We obtain the one law

Theorem 9.6 (WSN with link failures)**.** *For any scaling* $\boldsymbol{\theta}_n$

$$\lim_{n\to\infty} \mathbf{P}\left[I_n(\boldsymbol{\theta}_n) = 0\right] = 1 \quad \textit{if} \ \lim_{n\to\infty} \alpha_n = +\infty.$$

Similarly, in the case of a secure WSN, for a scaling $\boldsymbol{\omega}_n$ with $K_n \leq P_n$, defining the sequence α_n through

$$(1 - q(K_n, P_n))\ell_{\mathbb{D}}(r_n) = \frac{\log n + \alpha_n}{n}, \quad n = 1, 2, \ldots$$

we obtain the one law

Theorem 9.7 (Secure WSN)**.** *For any scaling* $\boldsymbol{\omega}_n$ *with* $K_n \leq P_n$

$$\lim_{n\to\infty} \mathbf{P}\left[I_n(\boldsymbol{\omega}_n) = 0\right] = 1 \quad \textit{if} \ \lim_{n\to\infty} \alpha_n = +\infty.$$

9.4 Method of First and Second Moments

The proofs rely on the method of first and second moments [55, p. 55], an approach widely used in the theory of Erdős-Rényi graphs: Let Z denote an $\mathbb{N}$-valued r.v. with finite second moment. The method of first moments [55, Eqn. (3.10), p. 55] relies on the inequality

$$1 - \mathbf{E}\left[Z\right] \leq \mathbf{P}\left[Z = 0\right], \tag{9.7}$$

while the method of second moments [55, Remark 3.1, p. 55] uses the bound

$$\mathbf{P}\left[Z = 0\right] \leq 1 - \frac{\mathbf{E}\left[Z\right]^2}{\mathbf{E}\left[Z^2\right]}. \tag{9.8}$$

We describe applications of this technique to the model with link failures. Note that the discussion in this section also applies for secure WSNs. Pick any scaling

$\boldsymbol{\theta}_n$. From (9.7) we see that the one law

$$\lim_{n\to\infty} \mathbf{P}\left[I_n(\boldsymbol{\theta}_n) = 0\right] = 1$$

is established if we show that

$$\lim_{n\to\infty} \mathbf{E}\left[I_n(\boldsymbol{\theta}_n)\right] = 0. \tag{9.9}$$

On the other hand, it is plain from (9.8) that

$$\lim_{n\to\infty} \mathbf{P}\left[I_n(\boldsymbol{\theta}_n) = 0\right] = 0$$

if

$$\liminf_{n\to\infty} \left(\frac{\mathbf{E}\left[I_n(\boldsymbol{\theta}_n)\right]^2}{\mathbf{E}\left[I_n(\boldsymbol{\theta}_n)^2\right]} \right) \geq 1. \tag{9.10}$$

Upon using the exchangeability and the binary nature of the r.v.'s involved in the count variables of interest, we can obtain simpler characterizations of the convergence statements (9.9) and (9.10). Indeed, for all $n = 2, 3, \ldots$ and every $\boldsymbol{\theta}$, the calculations

$$\mathbf{E}\left[I_n(\boldsymbol{\theta})\right] = \sum_{i=1}^{n} \mathbf{E}\left[\chi_{n,i}(\boldsymbol{\theta})\right] = n\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\right]$$

and

$$\begin{aligned} \mathbf{E}\left[I_n(\boldsymbol{\theta})^2\right] &= \sum_{i=1}^{n} \mathbf{E}\left[\chi_{n,i}(\boldsymbol{\theta})\right] + \sum_{i,j=1,\ i\neq j}^{n} \mathbf{E}\left[\chi_{n,i}(\boldsymbol{\theta})\chi_{n,j}(\boldsymbol{\theta})\right] \\ &= n\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\right] + n(n-1)\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right] \end{aligned}$$

are straightforward, so that

$$\frac{\mathbf{E}\left[I_n(\boldsymbol{\theta})^2\right]}{\mathbf{E}\left[I_n(\boldsymbol{\theta})\right]^2} = \frac{1}{n\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\right]} + \frac{n-1}{n} \cdot \frac{\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right]}{\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\right]^2}.$$

Thus, for the given scaling $\boldsymbol{\theta}_n$, we obtain the one law by showing that

$$\lim_{n\to\infty} n\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta}_n)\right] = 0, \tag{9.11}$$

while the zero law will follow if we show that

$$\lim_{n\to\infty} n\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta}_n)\right] = \infty \tag{9.12}$$

and

$$\limsup_{n\to\infty}\left(\frac{\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta}_n)\chi_{n,2}(\boldsymbol{\theta}_n)\right]}{\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta}_n)\right]^2}\right) \leq 1. \tag{9.13}$$

The bulk of the technical discussion therefore amounts to establishing (9.11), (9.12) and (9.13) under the appropriate conditions on the scaling $\boldsymbol{\theta}_n$.

9.5 Calculation of First Moments

Let us consider the model with random link failures. Let X and Y be mutually independent r.v.'s with a uniform distribution on $[0,1]$. Fix $n = 2, 3, \ldots$ and $\boldsymbol{\theta}$ in $\mathbb{R}_+ \times [0,1]$. For both the unit circle and unit interval, the enforced independence assumptions readily imply

$$\begin{aligned}
\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\right] &= \mathbf{E}\left[\prod_{i=1,\ j\neq i}^{n}\left(1-\chi_{ij}(\boldsymbol{\theta})\right)\right] \\
&= \mathbf{E}\left[\left(1-pa(X;r)\right)^{n-1}\right] \\
&= \int_0^1 \left(1-pa(x;r)\right)^{n-1} dx
\end{aligned} \tag{9.14}$$

where we have set

$$a(x;r) := \mathbf{P}\left[d(x,Y)\leq r\right], \quad 0\leq x\leq 1, r>0. \tag{9.15}$$

Closed-form expressions for (9.15) depend on the geometry of the region where the nodes are located.

The unit circle. As there are no border effects, we get

$$a^{(C)}(x;r) = \ell(r), \quad 0 \le x \le 1, r > 0 \tag{9.16}$$

and with the help of (9.14) this yields

$$\mathbf{E}\left[\chi_{n,1}^{(C)}(\boldsymbol{\theta})\right] = (1 - p\ell(r))^{n-1}, \quad r > 0, p \in [0,1]. \tag{9.17}$$

The unit interval. For $r \ge 1$, it is plain that

$$a^{(L)}(x;r) = 1, \quad 0 \le x \le 1.$$

On the other hand, when $0 < r < 1$, elementary calculations show that

$$a^{(L)}(x;r) \;=\; \begin{cases} x + r & \text{if } 0 < r \le 0.5, 0 \le x \le r \\ & \text{or } 0.5 < r < 1, 0 \le x \le 1 - r \\ \ell(r) & \text{if } 0 < r \le 0.5, r \le x \le 1 - r \\ & \text{or } 0.5 < r < 1, 1 - r \le x \le r \\ 1 - x + r & \text{if } 0 < r \le 0.5, 1 - r \le x \le 1 \\ & \text{or } 0.5 < r < 1, r \le x \le 1. \end{cases}$$

Reporting this information into (9.14), we obtain the following upper bound for any fixed $n = 2, 3, \ldots$, and $\boldsymbol{\theta}$ in $\mathbb{R}_+ \times [0, 1]$:

$$\mathbf{E}\left[\chi_{n,1}^{(L)}(\boldsymbol{\theta})\right] \leq (1 - p\ell(r))^{n-1} + \frac{2}{np}\left(1 - \frac{1}{2}p\ell(r)\right)^n. \tag{9.18}$$

The general case. In the general case when the nodes are located on a *torus* or *sphere* in higher dimensions, the analysis is a direct extension of that in the unit circle. Because there are no boundary effects, we get

$$a(x; r) = \ell_{\mathbb{D}}(r), \quad 0 \leq x \leq 1, r > 0,$$

and using (9.14) we obtain

$$\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\right] = (1 - p\ell_{\mathbb{D}}(r))^{n-1}, \quad r > 0, p \in [0, 1]. \tag{9.19}$$

Similarly, in the case of a secure WSN, we get

$$\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\omega})\right] = (1 - (1 - q(K, P))\ell_{\mathbb{D}}(r))^{n-1}, \quad r > 0, 0 \leq K \leq P. \tag{9.20}$$

9.6 Proof of the One Laws

As discussed in Section 9.4, the one law will be established if we show that (9.11) holds. Below we consider separately the unit circle and the unit interval. In that discussion we repeatedly use the elementary bound

$$1 - x \leq e^{-x}, \quad x \geq 0. \tag{9.21}$$

One law for the unit circle and torus. The one law over the unit circle reduces to showing the following convergence.

Lemma 9.8. *For any scaling $\boldsymbol{\theta}_n$, we have*

$$\lim_{n\to\infty} n\mathbf{E}\left[\chi_{n,1}^{(C)}(\boldsymbol{\theta}_n)\right] = 0 \quad \textit{if} \quad \lim_{n\to\infty}\alpha_n = +\infty$$

where the sequence α_n is determined through (9.5).

Proof. Fix $n = 1, 2, \ldots$ and in the expression (9.17) substitute (r, p) by (r_n, p_n) according to the scaling $\boldsymbol{\theta}_n$. We get

$$\begin{aligned} n\mathbf{E}\left[\chi_{n,1}^{(C)}(\boldsymbol{\theta}_n)\right] &= n\left(1 - p_n\ell(r_n)\right)^{n-1} \\ &= n\left(1 - \frac{\log n + \alpha_n}{n}\right)^{n-1} \\ &\leq n^{\frac{1}{n}} e^{-\frac{n-1}{n}\alpha_n} \end{aligned}$$

where the bound (9.21) was used. Letting n go to infinity we get the desired conclusion since $\lim_{n\to\infty}\alpha_n = \infty$. ■

Following the same steps as above, the one laws given in Theorem 9.6 and Theorem 9.7 are established using the expressions (9.19) and (9.20), respectively.

One law for the unit interval. A similar approach is taken for the random graphs over the unit interval.

Lemma 9.9. *For any scaling $\boldsymbol{\theta}_n$, we have*

$$\lim_{n\to\infty} n\mathbf{E}\left[\chi_{n,1}^{(L)}(\boldsymbol{\theta}_n)\right] = 0 \quad \textit{if} \quad \lim_{n\to\infty}\alpha'_n = +\infty$$

where the sequence α'_n is determined through (9.6).

Proof. Fix $n = 1, 2, \ldots$ and in the upper bound (9.18) substitute (r, p) by (r_n, p_n) according to the scaling $\boldsymbol{\theta}_n$. We get

$$n\mathbf{E}\left[\chi_{n,1}^{(L)}(\boldsymbol{\theta}_n)\right] \leq n\left(1 - p_n\ell(r_n)\right)^{n-1} + \frac{2}{p_n}\left(1 - \frac{1}{2}p_n\ell(r_n)\right)^n.$$

As in the proof of Lemma 9.8, we can show that

$$\lim_{n\to\infty} n\left(1 - p_n\ell(r_n)\right)^{n-1} = 0$$

under the condition $\lim_{n\to\infty} \alpha_n' = \infty$. The desired conclusion will be established as soon as we show that

$$\lim_{n\to\infty} \frac{2}{p_n}\left(1 - \frac{1}{2}p_n\ell(r_n)\right)^n = 0 \tag{9.22}$$

under the same condition $\lim_{n\to\infty} \alpha_n' = \infty$.

To do so, fix $n = 1, 2, \ldots$ sufficiently large so that $\alpha_n' \geq 0$ – This is always possible under the condition $\lim_{n\to\infty} \alpha_n' = \infty$. On that range we note that

$$\frac{1}{p_n}\left(1 - \frac{1}{2}p_n\ell(r_n)\right)^n \leq \frac{1}{p_n\ell(r_n)}\left(1 - \frac{1}{2}p_n\ell(r_n)\right)^n \leq \frac{\log n}{2(\log n - \log\log n)}e^{-\frac{1}{2}\alpha_n'}$$

upon using the fact $\ell(r_n) \leq 1$ and the bound (9.21). Letting n go to infinity we obtain (9.22). ■

9.7 Calculation of Second Moments

The calculation of second moments is needed in the proof of the zero laws in Theorems 9.3 and 9.4 using (9.13). By Remark 9.5, we need only consider the unit circle as we do from now on. We drop the superscript (C) for simplicity of notation.

Throughout we denote by X, Y and Z three mutually independent r.v.'s which are uniformly distributed on $[0, 1]$, and by B, B' and B'' three mutually independent

$\{0,1\}$-valued r.v.'s with success probability p. The two groups of r.v.'s are assumed to be independent.

Again fix $n = 2, 3, \ldots$ and $\boldsymbol{\theta}$ in $\mathbb{R}_+ \times [0,1]$. The same arguments apply for both the unit circle and unit interval: For x, y in $[0,1]$, write

$$\begin{aligned} b(x,y;\boldsymbol{\theta}) &:= \mathbf{E}\left[\left(1 - B'\mathbf{1}\left[d(x,Z) \le r\right]\right)\left(1 - B''\mathbf{1}\left[d(y,Z) \le r\right]\right)\right] \\ &= 1 - pa(x;r) - pa(y;r) + p^2 u(x,y;r) \end{aligned}$$

with

$$u(x,y;r) := \mathbf{P}\left[d(x,Z) \le r, d(y,Z) \le r\right].$$

We then proceed with the decomposition

$$\begin{aligned} \chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta}) &= \prod_{j=2}^{n}\left(1 - \chi_{1j}(\boldsymbol{\theta})\right) . \prod_{k=1,k\neq 2}^{n}\left(1 - \chi_{2k}(\boldsymbol{\theta})\right) \\ &= \left(1 - \chi_{12}(\boldsymbol{\theta})\right)\prod_{j=3}^{n}\left(1 - \chi_{1j}(\boldsymbol{\theta})\right)\left(1 - \chi_{2j}(\boldsymbol{\theta})\right). \end{aligned}$$

Under the enforced independence assumptions, an easy conditioning argument (with respect to the triple X_1, X_2 and B_{12}) based on this decomposition now gives

$$\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right] = \mathbf{E}\left[\left(1 - B\mathbf{1}\left[d(X,Y) \le r\right]\right) b(X,Y;\boldsymbol{\theta})^{n-2}\right].$$

From (9.16) it is plain that

$$b(x,y;\boldsymbol{\theta}) = 1 - 2p\ell(r) + p^2 u(x,y;r)$$

for all x, y in $[0,1]$.

We note that

$$\begin{aligned} u(x,y;r) &= \mathbf{P}\left[d(x,Z) \le r, d(y,Z) \le r\right] \\ &= u(0,d(x,y);r) \end{aligned}$$

by translation invariance. Thus, writing

$$\tilde{b}(z;\boldsymbol{\theta}) := 1 - 2p\ell(r) + p^2\tilde{u}(z;r), \quad z \in [0,0.5] \tag{9.23}$$

with

$$\tilde{u}(z;r) := u(0,z;r),$$

we get

$$b(x,y;\boldsymbol{\theta}) = \tilde{b}(d(x,y);\boldsymbol{\theta}), \quad x,y \in [0,1].$$

Taking advantage of these facts we now find

$$\begin{aligned} \mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right] &= \mathbf{E}\left[\left(1 - p\mathbf{1}\left[d(X,Y) \le r\right]\right)\tilde{b}(d(X,Y);\boldsymbol{\theta})^{n-2}\right] \\ &= 2\int_0^{0.5}\left(1 - p\mathbf{1}\left[z \le r\right]\right)\tilde{b}(z;\boldsymbol{\theta})^{n-2}dz \end{aligned}$$

by a straightforward evaluation of the double integral

$$\int_0^1 dx \int_0^1 dy \left(1 - p\mathbf{1}\left[d(x,y) \le r\right]\right)\tilde{b}(d(x,y);\boldsymbol{\theta})^{n-2}.$$

Consequently,

$$\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right] \le 2\int_0^{0.5}\tilde{b}(z;\boldsymbol{\theta})^{n-2}dz. \tag{9.24}$$

It is possible to compute the value of $\tilde{u}(z;r)$ for various values for z, r (details have been omitted). For $0 < r < 0.5$, we find

$$\tilde{u}(z;r) = \begin{cases} 2r - z & \text{if } 0 < r < 0.25, 0 \leq z \leq 2r \\ 0 & \text{if } 0 < r < 0.25, 2r < z \leq 0.5 \\ 2r - z & \text{if } 0.25 \leq r < 0.5, 0 \leq z \leq 1 - 2r \\ 4r - 1 & \text{if } 0.25 \leq r < 0.5, 1 - 2r < z \leq 0.5. \end{cases}$$

Obviously, if $r \geq 0.5$, then $\tilde{u}(z;r) = 1$ for every z in $[0, 0.5]$. Thus, for $0 \leq p \leq 1$, through (9.23) we obtain

$$\tilde{b}(z;\boldsymbol{\theta}) = \begin{cases} 1 - 4pr + p^2(2r - z) & \text{if } 0 < r < 0.25, 0 \leq z \leq 2r \\ 1 - 4pr & \text{if } 0 < r < 0.25, 2r < z \leq 0.5 \\ 1 - 4pr + p^2(2r - z) & \text{if } 0.25 \leq r < 0.5, \\ & \quad 0 \leq z \leq 1 - 2r \\ 1 - 4pr + p^2(4r - 1) & \text{if } 0.25 \leq r < 0.5, \\ & \quad 1 - 2r < z \leq 0.5. \end{cases}$$

Using this fact in (9.24) and evaluating the integral, we obtain the following upper bounds:

(i) For $0 < r < 0.25$ and $0 < p \leq 1$,

$$\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right] \leq (1-4r)(1-4pr)^{n-2} + \frac{2(1-4pr)^{n-1}}{(n-1)p^2}\left(\left(1+\frac{2p^2r}{1-4pr}\right)^{n-1} - 1\right).$$

(ii) For $0.25 \leq r < 0.5$ and $0 < p \leq 1$,

$$\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right] \leq (4r-1)(1-2pr)^{2(n-2)} + (2-4r)(1-4pr+2p^2r)^{n-2}.$$

(iii) For $r \geq 0.5$ and $0 < p \leq 1$,

$$\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right] = (1-p)^{2n-3}.$$

(iv) For $r > 0$ and $p = 0$,

$$\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right] = 1.$$

Furthermore, combining these bounds with (9.17), we obtain the following upper bound on

$$R_n(\boldsymbol{\theta}) := \frac{\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\chi_{n,2}(\boldsymbol{\theta})\right]}{\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta})\right]^2}$$

in the various cases listed below.

(i) For $0 < r < 0.25$ and $0 < p \leq 1$,

$$R_n(\boldsymbol{\theta}) \leq \frac{1-4r}{1-4pr} + \frac{2}{(n-1)p^2}\left(\left(1+\frac{2p^2r}{1-4pr}\right)^{n-1} - 1\right). \tag{9.25}$$

(ii) For $0.25 \leq r < 0.5$ and $0 < p \leq 1$,

$$R_n(\boldsymbol{\theta}) \leq \frac{4r-1}{(1-2pr)^2} + (2-4r)\frac{(1-4pr+2p^2r)^{n-2}}{(1-2pr)^{2(n-1)}}. \tag{9.26}$$

(iii) For $r \geq 0.5$ and $0 < p \leq 1$,

$$R_n(\boldsymbol{\theta}) = \frac{1}{1-p}. \tag{9.27}$$

(iv) For $r > 0$ and $p = 0$,

$$R_n(\boldsymbol{\theta}) = 1. \tag{9.28}$$

9.8 Proof of the Zero Laws

As observed in Remark 9.5, when dealing with the zero law we need only concern ourselves with the unit circle case. Therefore, the superscript (C) is ignored for simplicity.

Throughout this section, we take $\boldsymbol{\theta}_n$ and associate with it the sequence α_n through (9.5). We now show (9.12) and (9.13) under the condition $\lim_{n\to\infty} \alpha_n = -\infty$. This will complete the proof of the zero laws.

In the discussion we shall make use of the following elementary fact: For any sequence $a : \mathbb{N}_0 \to \mathbb{R}_+$, the asymptotic equivalence

$$(1 - a_n)^n \sim e^{-na_n} \tag{9.29}$$

holds provided $\lim_{n\to\infty} a_n = \lim_{n\to\infty} na_n^2 = 0$.

Establishing (9.12). The first step is contained in the following zero law complement of Lemma 9.8.

Lemma 9.10. *For any scaling $\boldsymbol{\theta}_n$, we have*

$$\lim_{n\to\infty} n\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta}_n)\right] = \infty \quad \textit{if} \quad \lim_{n\to\infty}\alpha_n = -\infty$$

where the sequence α_n is determined through (9.5).

Proof. Fix $n = 1, 2, \ldots$ and in the expression (9.17) substitute (r, p) by (r_n, p_n) according to the scaling $\boldsymbol{\theta}_n$. As in the proof of Lemma 9.8 we start with the expression

$$n\mathbf{E}\left[\chi_{n,1}(\boldsymbol{\theta}_n)\right] = n\left(1 - p_n\ell(r_n)\right)^{n-1}. \tag{9.30}$$

Under the condition $\lim_{n\to\infty}\alpha_n = -\infty$ we note that $\alpha_n = -|\alpha_n|$ for all n sufficiently large, say for all $n \geq n^\star$ for some finite integer $n^\star$. Using (9.5) we get $|\alpha_n| \leq \log n$ on that range by the non-negativity condition $p_n\ell(r_n) \geq 0$. Therefore,

$$p_n\ell(r_n) \leq \frac{\log n}{n} \quad \text{and} \quad n\left(p_n\ell(r_n)\right)^2 \leq \frac{(\log n)^2}{n} \tag{9.31}$$

for all $n \geq n^\star$, and the equivalence (9.29) (with $a_n = p_n\ell(r_n)$) now yields

$$n\left(1 - p_n\ell(r_n)\right)^{n-1} \sim ne^{-np_n\ell(r_n)} \tag{9.32}$$

with

$$ne^{-np_n\ell(r_n)} = ne^{-(\log n + \alpha_n)} = e^{-\alpha_n}, \quad n = 1, 2, \ldots \tag{9.33}$$

Finally, letting n go to infinity in (9.30) and using (9.32)-(9.33), we find

$$\lim_{n\to\infty} n\left(1 - p_n\ell(r_n)\right)^{n-1} = \lim_{n\to\infty} e^{-\alpha_n} = \infty$$

as desired under the condition $\lim_{n\to\infty}\alpha_n = -\infty$. ■

Establishing (9.13). The proof of the one law will be completed if we establish the next result.

Proposition 9.11. *For any scaling $\boldsymbol{\theta}_n$, with the sequence α_n determined through (9.5), we have*

$$\limsup_{n\to\infty} R_n(\boldsymbol{\theta}_n) \leq 1 \quad \textit{if} \quad \lim_{n\to\infty}\alpha_n = -\infty.$$

The proof of Proposition 9.11 is organized around the following simple observation: Consider a sequence $a : \mathbb{N}_0 \to \mathbb{R}$ and let $N_1, \ldots, N_K$ constitute a partition of $\mathbb{N}_0$ into K subsets, i.e., $N_k \cap N_\ell = \emptyset$ for distinct $k, \ell = 1, \ldots, K$, and $\cup_{k=1}^K N_k = \mathbb{N}_0$. In principle, some of the subsets $N_1, \ldots, N_K$ may be either empty or finite. For each $k = 1, \ldots, K$ such that N_k is *non*-empty, we set

$$\alpha_k := \limsup_{\substack{n\to\infty \\ n\in N_k}} a_n = \inf_{n\in N_k} \left(\sup_{m\in N_k:\ m\geq n} a_m \right)$$

with the natural convention that $\alpha_k = -\infty$ when N_k is finite. In other words, α_k is the limsup for the subsequence $\{a_n,\ n \in N_k\}$. It is a simple matter to check that

$$\limsup_{n\to\infty} a_n = \max{}^\star \left(\alpha_k,\ k = 1, \ldots, K\right)$$

with $\max^\star$ denoting the maximum operation over all indices k such that N_k is non-empty.

Proof. As we plan to make use of this fact with $K = 4$, we write

$$R_k := \limsup_{\substack{n\to\infty \\ n\in N_k}} R_n(\boldsymbol{\theta}_n), \quad k = 1, \ldots, 4$$

with

$$\begin{aligned}
N_1 :=& \{n \in \mathbb{N}_0 :\ 0 < r_n < 0.25,\ 0 < p_n \leq 1\}, \\
N_2 :=& \{n \in \mathbb{N}_0 :\ 0.25 \leq r_n < 0.5,\ 0 < p_n \leq 1\}, \\
N_3 :=& \{n \in \mathbb{N}_0 :\ 0.5 \leq r_n,\ 0 < p_n \leq 1\}
\end{aligned}$$

and

$$N_4 := \{n \in \mathbb{N}_0 : \ r_n > 0, \ p_n = 0\}.$$

Therefore, we have

$$\limsup_{n\to\infty} R_n(\boldsymbol{\theta}_n) = \max{}^{\star}(R_k, \ k = 1, \ldots, 4)$$

and the result will be established if we show that

$$R_k \leq 1, \quad k = 1, \ldots, 4.$$

In view of the convention made earlier, we need only discuss for each $k = 1, \ldots, 4$, the case when N_k is countably infinite, as we do from now on.

The easy cases are handled first: From (9.28) it is obvious that $R_4 = 1$. Next as observed before, (9.31) holds for all n sufficiently large under the condition $\lim_{n\to\infty} \alpha_n = -\infty$. Since $\ell(r_n) = 1$ for all n in N_3, we conclude that

$$\lim_{\substack{n\to\infty \\ n\in N_3}} p_n = 0$$

and the conclusion $R_3 = 1$ is now immediate from (9.27). We complete the proof by invoking Lemmas 9.12 and 9.13 given next which establish $R_1 \leq 1$ and $R_2 \leq 1$, respectively. ■

Lemma 9.12. *Under the assumptions of Proposition 9.11, with N_1 countably infinite, we have $R_1 \leq 1$.*

Proof. Fix $n = 2, 3, \ldots$ and pick (r, p) such that $0 < r < 0.25$ and $0 < p \leq 1$. With (9.25) in mind, we note that

$$\frac{2}{(n-1)p^2}\left(\left(1+\frac{2p^2r}{1-4pr}\right)^{n-1}-1\right) = \frac{2}{(n-1)p^2}\sum_{k=1}^{n-1}\binom{n-1}{k}\left(\frac{2p^2r}{1-4pr}\right)^k$$
$$= \frac{4r}{1-4pr} + \frac{2}{(n-1)p^2}\sum_{k=2}^{n-1}\binom{n-1}{k}\left(\frac{2p^2r}{1-4pr}\right)^k$$

and we can rewrite the r.-h.s of (9.25) as

$$\frac{1-4r}{1-4pr} + \frac{2}{(n-1)p^2}\left(\left(1+\frac{2p^2r}{1-4pr}\right)^{n-1}-1\right)$$
$$= \frac{1}{1-4pr} + \frac{2}{(n-1)p^2}\sum_{k=2}^{n-1}\binom{n-1}{k}\left(\frac{2p^2r}{1-4pr}\right)^k$$
$$\leq \frac{1}{1-4pr} + \frac{2}{(n-1)}\sum_{k=2}^{n-1}\binom{n-1}{k}\left(\frac{2pr}{1-4pr}\right)^k$$

since $p^k \leq p^2$ for $k = 2, \ldots, n-1$. Therefore,

$$R_n(\boldsymbol{\theta}) \leq \frac{1}{1-4pr} + \frac{2}{(n-1)}\left(1+\frac{2pr}{1-4pr}\right)^{n-1}.$$

In this last bound, fix n in N_1 and substitute (r, p) by (r_n, p_n) according to the scaling $\boldsymbol{\theta}_n$. Standard properties of the limsup operation yield

$$R_1 \leq \limsup_{\substack{n\to\infty \\ n\in N_1}}\left(\frac{1}{1-4p_nr_n}\right) + \limsup_{\substack{n\to\infty \\ n\in N_1}}\left(\frac{2}{(n-1)}\left(1+\frac{2p_nr_n}{1-4p_nr_n}\right)^{n-1}\right)$$

and the desired result $R_1 \leq 1$ will follow if we show that

$$\limsup_{\substack{n\to\infty \\ n\in N_1}}\left(\frac{1}{1-4p_nr_n}\right) = 1 \tag{9.34}$$

and

$$\limsup_{\substack{n\to\infty \\ n\in N_1}} \left(\frac{2}{(n-1)} \left(1 + \frac{2p_n r_n}{1-4p_n r_n} \right)^{n-1} \right) = 0. \tag{9.35}$$

To do so, under the condition $\lim_{n\to\infty} \alpha_n = -\infty$ we once again use the fact that (9.31) holds for large n with $p_n \ell(r_n) = 2p_n r_n$ for all n in N_1. Thus,

$$\lim_{\substack{n\to\infty \\ n\in N_1}} p_n r_n = 0$$

and the convergence (9.34) follows.

Next, since $1 + x \le e^x$ for all x in $\mathbb{R}$, we note for all n in N_1 that

$$\frac{2}{n-1} \left(1 + \frac{2p_n r_n}{1-4p_n r_n} \right)^{n-1} \le 2e^{\beta_n}$$

with

$$\beta_n := (n-1) \frac{p_n \ell(r_n)}{1 - 2p_n \ell(r_n)} - \log(n-1).$$

Thus, (9.35) follows if we show that

$$\lim_{\substack{n\to\infty \\ n\in N_1}} \beta_n = -\infty. \tag{9.36}$$

From (9.31) we get

$$\beta_n \le \left(\frac{n-1}{n} \right) \frac{\log n + \alpha_n}{1 - 2\frac{\log n}{n}} - \log(n-1)$$

for large n. It is now a simple exercise to check that

$$\lim_{n\to\infty} \left(\frac{n-1}{n} \right) \frac{\log n}{1 - 2\frac{\log n}{n}} - \log(n-1) = 0$$

and the conclusion (9.36) is obtained under the assumption $\lim_{n\to\infty} \alpha_n = -\infty$. ■

Lemma 9.13. *Under the assumptions of Proposition 9.11, with N_2 countably infinite, we have $R_2 \leq 1$.*

Proof. Fix $n = 2, 3, \ldots$ and pick (r, p) such that $0.25 < r \leq 0.5$ and $0 < p \leq 1$. From (9.26) we get

$$\begin{aligned} R_n(\boldsymbol{\theta}) &\leq \frac{4r-1}{(1-2pr)^2} + \frac{2-4r}{(1-2pr)^2}\frac{(1-4pr+2p^2r)^{n-2}}{(1-2pr)^{2(n-2)}} \\ &= \frac{4r}{(1-2pr)^2}\left(1 - \frac{(1-4pr+2p^2r)^{n-2}}{(1-2pr)^{2(n-2)}}\right) \\ &\quad + \frac{1}{(1-2pr)^2}\left(2\frac{(1-4pr+2p^2r)^{n-2}}{(1-2pr)^{2(n-2)}} - 1\right). \end{aligned}$$

Now fix n in N_2 and substitute (r, p) by (r_n, p_n) according to the scaling $\boldsymbol{\theta}_n$ in (9.26). As before, properties of the limsup operation yield

$$R_2 \leq R_{2c}\,(R_{2a} + R_{2b}) \tag{9.37}$$

with

$$R_{2a} := \limsup_{\substack{n\to\infty \\ n\in N_2}} \left(4r_n\left(1 - \frac{(1-4p_nr_n+2p_n^2r_n)^{n-2}}{(1-2p_nr_n)^{2(n-2)}}\right)\right),$$

$$R_{2b} := \limsup_{\substack{n\to\infty \\ n\in N_2}} \left(2\frac{(1-4p_nr_n+2p_n^2r_n)^{n-2}}{(1-2p_nr_n)^{2(n-2)}} - 1\right)$$

and

$$R_{2c} := \limsup_{\substack{n\to\infty \\ n\in N_2}} \frac{1}{(1-2p_nr_n)^2}.$$

As in the proof of Lemma 9.12, it is also the case here that R_{2c} exists as a limit and is given by

$$R_{2c} = \lim_{\substack{n\to\infty \\ n\in N_2}} \frac{1}{(1-2p_nr_n)^2} = 1.$$

Next, we show that

$$\lim_{\substack{n\to\infty \\ n\in N_2}} \frac{(1-4p_nr_n+2p_n^2r_n)^{n-2}}{(1-2p_nr_n)^{2(n-2)}} = 1. \tag{9.38}$$

Once this is done, we see from their definitions that $R_{2a} = 0$ and $R_{2b} = 1$, and the conclusion $R_2 \leq 1$ follows from (9.37).

To establish (9.38) we note that

$$4p_nr_n - 2p_n^2r_n = p_n\ell(r_n)(2-p_n) \leq 2p_n\ell(r_n)$$

and

$$2p_nr_n = p_n\ell(r_n)$$

for all n in N_2. Now making use of (9.31) we conclude that

$$\lim_{\substack{n\to\infty \\ n\in N_2}} \left(4p_nr_n - 2p_n^2r_n\right) = \lim_{\substack{n\to\infty \\ n\in N_2}} (n-2)\left(4p_nr_n - 2p_n^2r_n\right)^2 = 0$$

while

$$\lim_{\substack{n\to\infty \\ n\in N_2}} 2p_nr_n = \lim_{\substack{n\to\infty \\ n\in N_2}} (n-2)\left(2p_nr_n\right)^2 = 0.$$

By the equivalence (9.29) used with $a_n = 4p_nr_n - 2p_n^2r_n$ and $a_n = 2p_nr_n$, respectively, we now conclude that

$$\frac{(1-4p_nr_n+2p_n^2r_n)^{n-2}}{(1-2p_nr_n)^{2(n-2)}} \sim \frac{e^{-(n-2)\left(4p_nr_n-2p_n^2r_n\right)}}{\left(e^{-(n-2)(2p_nr_n)}\right)^2} = e^{2(n-2)\left(p_n^2r_n\right)} \tag{9.39}$$

as n goes to infinity in N_2.

Finally, for n in N_2, because $\ell(r_n) = 2r_n \geq 0.5$, we get

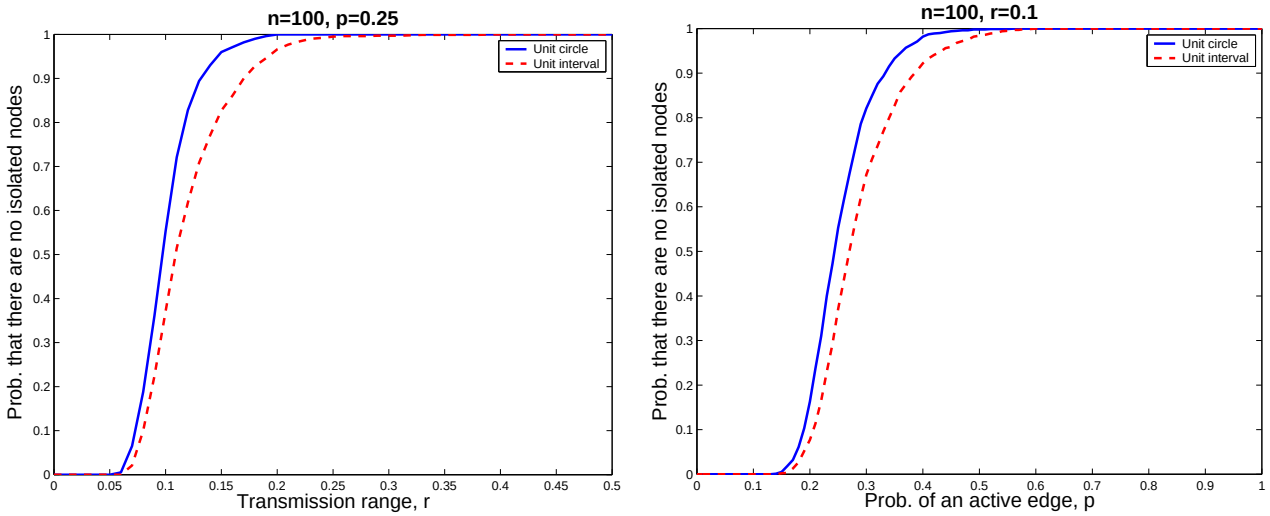
$$2(n-2)\left(p_n^2r_n\right) = (n-2)\frac{\left(p_n\ell(r_n)\right)^2}{\ell(r_n)} \leq \frac{2(n-2)}{n} \cdot n\left(p_n\ell(r_n)\right)^2$$

so that

$$\lim_{\substack{n\to\infty \\ n\in N_2}} 2(n-2)\left(p_n^2 r_n\right) = 0$$

with the help of (9.31). The conclusion (9.38) now follows from (9.39), and the proof of Lemma 9.13 is complete. ■

9.9 Simulation Results



(a) Fix $n = 100$, $p = 0.25$ and vary r

(b) Fix $n = 100$, $r = 0.1$ and vary p

Figure 9.1: Simulation results for WSNs with random link failures

In this section, we present some plots from simulations in Matlab which confirm the results in Theorem 9.3 and Theorem 9.4. For given n, p and r, we estimate the probability that there are no isolated nodes by averaging over 1000 instances of the random graphs $\mathbb{G}^{(C)}(n;\boldsymbol{\theta})$ and $\mathbb{G}^{(L)}(n;\boldsymbol{\theta})$.

In Figure 9.1(a), we have taken $n = 100$ and $p = 0.25$, and examine the threshold behavior of the probability that there are no isolated nodes by varying r. Theorem 9.3 suggests that the critical range for the graph over the unit circle when $n = 100$ and $p = 0.25$ should be $r^\star = 0.09$. This is confirmed by the simulation

results. In the case of the unit interval, we expect from Theorem 9.4 that the critical range will be between $r^\star = 0.09$ and $r^{\star\star} = 0.12$; this is in agreement with the plot.

In Figure 9.1(b), we have taken $n = 100$ and $r = 0.1$, and repeat the analysis by choosing various values for p. As expected from Theorem 9.3, the critical edge probability for the unit circle is found to occur at $p^\star = 0.23$. It is also clear that for the unit interval, the critical edge probability is between $p^\star = 0.23$ and $p^{\star\star} = 0.31$ as predicted by Theorem 9.4.

9.10 Concluding Remarks

We have proved that the critical scaling for the link failure model on the unit circle is given by

$$p_n^\star \ell(r_n^\star) = \frac{\log n}{n}, \quad n = 1, 2, \ldots$$

This is an example of the general phenomenon that critical scalings for the absence of isolated nodes for many random graph models in the literature are determined through the requirement

$$\mathbf{P}\left[\text{Edge exists between two nodes}\right] = \frac{\log n}{n}. \tag{9.40}$$

The analogous result for the unit interval established in Theorem 9.4 still shows a gap between the scalings for the zero and one laws. While we believe that this gap can be bridged, the technique employed in this chapter is apparently insufficient to accomplish this.

Open Problem 9.14. Study the asymptotic distribution of the number of isolated nodes for WSNs with link failures on the unit circle with no assumptions on the link outage probability.

Open Problem 9.15. Examine the critical scalings for connectedness of the random graphs studied in this chapter.

The results of this chapter are intended for publication as a part of [11].

Bibliography

[1] R. Ahlswede, "An elementary proof of the strong converse theorem for the multiple-access channel," *Journal of Combinatorics, Information and System Sciences*, Vol. 7, No. 3, pp. 216–230, 1982.

[2] R. Ahlswede and N. Cai, "Codes with the identifiable parent property and the multiple-access channel," *General Theory of Information Transfer*, R. Ahlswede at al., Eds., Lecture Notes Comput. Sci., Vol. 4123, Berlin: Springer Verlag, pp. 249–257, 2006.

[3] N. Alon and U. Stav, "New bounds on parent-identifying codes: the case of multiple parents," *Combinatorics, Probability and Computing*, Vol. 13, No. 6, pp. 795–807, 2004.

[4] E. Amiri and G. Tardos, "High rate fingerprinting codes and the fingerprinting capacity," *Proc. 20th Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA 2009), pp. 336–345, 2009.

[5] N. P. Anthapadmanabhan and A. Barg, "Random binary fingerprinting codes for arbitrarily sized coalitions," *Proc. IEEE Internat. Sympos. Information Theory* (ISIT 2006), pp. 351–355, 2006.

[6] N. P. Anthapadmanabhan and A. Barg, "Randomized frameproof codes: Fingerprinting plus validation minus tracing," *Proc. Conf. on Information Sciences and Systems* (CISS 2008), pp. 365–369, 2008.

[7] N. P. Anthapadmanabhan and A. Barg, "Two-level fingerpinting codes," to appear in *Proc. IEEE Internat. Sympos. Information Theory* (ISIT 2009), Seoul, Korea, Jun. 28 - Jul. 3, 2009.

[8] N. P. Anthapadmanabhan and A. Barg, "Two-level fingerpinting," manuscript in preparation.

[9] N. P. Anthapadmanabhan, A. Barg and I. Dumer, "Fingerprinting capacity under the marking assumption," *Proc. IEEE Internat. Sympos. Information Theory* (ISIT 2007), pp. 706–710, 2007.

[10] N. P. Anthapadmanabhan, A. Barg and I. Dumer, "On the fingerprinting capacity under the marking assumption," *IEEE Trans. Information Theory - Special issue of Information-theoretic security*, Vol. 54, No. 6, pp. 2678–2689, Jun. 2008.

[11] N. P. Anthapadmanabhan and A. M. Makowski, "Beyond the disk model of wireless networks," manuscript in preparation.

[12] M. J. B. Appel and R. P. Russo, " The connectivity of a graph on uniform points on $[0,1]^d$," *Statistics & Probability Letters*, Vol. 60, pp. 351–357, 2002.

[13] A. Ashikhmin and A. Barg, "Minimal vectors in linear codes," *IEEE Trans. Information Theory*, Vol. 44, No. 5, pp. 2010–2017, Sep. 1998.

[14] U. Augustin, "Gedächtnisfreie Kanäle für diskrete Zeit," *Z. Wahrscheinlichkeitstheorie u. verw. Gebiete*, Vol. 6, pp. 10–61, 1966.

[15] A. Barg, G. R. Blakley and G. Kabatiansky, "Digital fingerprinting codes: Problem statements, constructions, identification of traitors," *IEEE Trans. Information Theory*, Vol. 49, No. 4, pp. 852–865, Apr. 2003.

[16] A. Barg, G. Cohen, S. Encheva, G. Kabatiansky, and G. Zémor, "A hypergraph approach to the identifying parent property: The case of multiple parents," *SIAM Journal Discrete Math.*, Vol. 14, No.3, pp. 423–431, 2001.

[17] A. Barg and G. Kabatiansky, "A class of i.p.p. codes with efficient identification," *Journal of Complexity*, Vol. 20, Nos. 2–3, pp. 137–147, 2004.

[18] L. A. Bassalygo, V. A. Zinov'ev, V. V. Zyablov, M. S. Pinsker and G. Sh. Poltyrev, "Bounds for codes with unequal error protection of two sets of messages," *Probl. Information Transmission*, Vol. 15, No. 3, pp. 190–197, Jul.–Sep. 1979.

[19] S. R. Blackburn, "An upper bound on the size of a code with the k-identifiable property," *Journal of Combinatorial Theory Ser. A*, Vol. 102, pp. 179–185, 2003.

[20] S. R. Blackburn, "Combinatorial schemes for protecting digital content," *Surveys in Combinatorics*, 2003 (Bangor), London Math. Soc. Lecture Note Ser., Vol. 307, pp. 43–78, Cambridge Univ. Press, Cambridge, 2003.

[21] G. R. Blakley and G. Kabatiansky, "Random coding technique for digital fingerprinting codes: fighting two pirates revisited," *Proc. IEEE Internat. Sympos. Information Theory* (ISIT 2004), p. 203, 2004.

[22] G. R. Blakley, C. Meadows, and G. Purdy, "Fingerprinting long forgiving messages," *Proc. CRYPTO 1985*, pp. 180–189.

[23] O. Blayer and T. Tassa, "Improved versions of Tardos' fingerprinting scheme," *Designs, Codes and Cryptography*, Vol. 48, No. 1, pp. 79–103, Jul. 2008.

[24] E. L. Blokh and V. V. Zyablov, *Linear Concatenated Codes*, Moscow: Nauka, 1982 (in Russian).

[25] B. Bollobás, *Random Graphs*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge (UK), 2001.

[26] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," *IEEE Trans. Information Theory*, Vol. 44, No. 5, pp. 1897–1905, Sep. 1998.

[27] B. Chor, A. Fiat and M. Naor, "Tracing traitors," *Advances in Cryptology - Proc. Crypto '94*, Lecture Notes in Comput. Sci., Vol. 839, pp. 257–270, 1994.

[28] B. Chor, A. Fiat, M. Naor and B. Pinkas, "Tracing traitors," *IEEE Trans. Information Theory*, Vol. 46, No. 3, pp. 893–910, May 2000.

[29] G. Cohen and H. G. Schaathun, "Asymptotic overview of separating codes," Report no. 248, Department of Informatics, University of Bergen, 52pp., May 2003. Available at www.ii.uib.no.

[30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.

[31] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York: Academic, 1981.

[32] R. Di Pietro, L. V. Mancini, A. Mei, A. Panconesi and J. Radhakrishnan, "Connectivity properties of secure wireless sensor networks," *Proc. 2nd ACM Workshop on Security of Ad Hoc And Sensor Networks* (SASN 2004), Oct. 2004.

[33] R. Di Pietro, L. V. Mancini, A. Mei, A. Panconesi and J. Radhakrishnan, "Sensor networks that are provably resilient," *Proc. SecureComm 2006, the 2nd IEEE/CreateNet Internat. Conf. on Security and Privacy in Communication Networks*, Aug. 2006.

[34] O. Dousse, F. Baccelli and P. Thiran, "Impact of interferences on connectivity in ad hoc networks," *Proc. IEEE INFOCOM 2003*, Apr. 2003.

[35] O. Dousse, M. Franceschetti, N. Macris, R. Meester and P. Thiran, "Percolation in the signal-to-interference-ratio graph," *Journal of Applied Probability*, Vol. 43, pp. 552–562, 2006.

[36] I. Dumer, "Equal-weight fingerprinting codes," manuscript, Jan. 2009.

[37] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci.*, Vol. 5, pp. 17–61, 1960.

[38] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," *Proc. 9th ACM Conf. on Computer and Communications Security* (CCS 2002), pp. 41–47, Nov. 2002.

[39] W. Feller, *An Introduction to Probability Theory and Its Applications,* Vol. II, New York, N.Y.: John Wiley & Sons (1971).

[40] M. Fernandez and J. Cotrina, "Obtaining asymptotic fingerprint codes through a new analysis of the Boneh-Shaw codes," *Proc. of Inscrypt 2006*, H. Lipmaa, M. Yung amd D. Lin, Eds., Lecture Notes in Computer Science, Vol. 4318, pp. 289–303, 2006.

[41] G. D. Forney, Jr., *Concatenated Codes*, Cambridge, MA: MIT Press, 1966.

[42] A. D. Friedman, R. L. Graham, and J. D. Ullman, "Universal single transition time asynchronous state assignments," *IEEE Trans. Computers*, Vol. C-18, pp. 541–547, 1969.

[43] F. Galand, "Practical construction against theoretical approach in digital fingerprinting," *Proc. IEEE Internat. Sympos. Information Theory* (ISIT 2006), pp. 2603–2606, 2006.

[44] E. Godehardt, *Graphs as Structural Models: The Application of Graphs and Multigraphs in Cluster Analysis*, Vieweg, Braunschweig and Wiesbaden, 1990.

[45] E. Godehardt and J. Jaworski, "On the connectivity of a random interval graph," *Random Structures and Algorithms*, Vol. 9, pp. 137–161, 1996.

[46] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," Chapter in *Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, Edited by W. M. McEneany, G. Yin and Q. Zhang, Birkhäuser, Boston (MA), 1998.

[47] V. Guruswami, *List decoding of error-correcting codes*, Lecture Notes in Computer Science, Vol. 3282, Springer, 2005.

[48] V. Guruswami and M. Sudan, "Improved decoding of Reed-Solomon and algebraic-geometry codes," *IEEE Trans. Inform. Theory*, Vol. 45, pp. 1757–1767, Sep. 1999.

[49] G. Han, "Connectivity Analysis of Wireless Ad-Hoc Networks," Ph.D. disseration, University of Maryland, College Park, MD, Apr. 2007.

[50] G. Han and A. M. Makowski, "A very strong zero-one law for connectivity in one-dimensional geometric random graphs," *IEEE Communications Letters*, Vol. 11, pp. 55–57, 2007.

[51] G. Han and A. M. Makowski, "Connectivity in one-dimensional geometric random graphs: Poisson approximations, zero-one laws and phase transitions," submitted to *IEEE Trans. Information Theory*, 2008.

[52] S. He and M. Wu, "Joint coding and embedding techniques for multimedia fingerprinting," *IEEE Trans. Information Forensics and Security*, Vol. 1, No. 2, pp. 231–247, Jun. 2006.

[53] H. D. L. Hollmann, J. H. van Lint, J.-P. Linnartz and L. M. G. M. Tolhuizen, "On codes with the identifiable parent property," *Journal of Combinatorial Theory Ser. A*, Vol. 82, pp. 121–133, 1998.

[54] Y.-W. Huang and P. Moulin, "Saddle-point solution of the fingerprinting capacity game under the marking assumption," submitted to *IEEE Internat. Sympos. Information Theory* (ISIT 2009), Jan. 2009.

[55] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, 2000.

[56] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Information Theory*, vol. 44, no. 6, pp. 2148–2177, Oct. 1998.

[57] T. V. Le, M. Burmester, and J. Hu, "Short c-secure fingerprinting codes," *Proc. 6th Information Security Conference* (ISC 2003), pp. 422–428, Oct. 2003.

[58] S-C. Lin, M. Shahmohammadi and H. El Gamal, "Fingerprinting with minimum distance decoding," preprint, arxiv:0710.2705, Oct. 2007.

[59] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*, North-Holland, Amsterdam, New York, Oxford, 1977.

[60] H. Maehara, " On the intersection graph of random arcles on a circle," *Random Graphs*, pp.159–173, 1987.

[61] P. Moulin, "Universal Fingerprinting: Capacity and Random-Coding Exponents," preprint, Jan. 2008, revised Dec. 2008. Available at arXiv:0801.3837v2 [cs.IT].

[62] P. Moulin and R. Koetter, "Data-hiding codes," *Proc. IEEE*, Vol. 93, No. 12, pp. 2083–2126, Dec. 2005.

[63] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Information Theory*, Vol. 49, No. 3, pp. 563–593, Mar. 2003.

[64] K. Nuida, S. Fujitsu, M. Hagiwara, T. Kitagawa, H. Watanabe, K. Ogawa and H. Imai, "An improvement of the discrete Tardos fingerprinting codes," *Cryptology ePrint Archive*, Report 2008/338.

[65] K. Nuida, M. Hagiwara, H. Watanabe and H. Imai, "Optimization of memory usage in Tardos's fingerprinting codes," Available at arXiv:cs/0610036.

[66] M. D. Penrose, *Random Geometric Graphs*, Oxford Studies in Probability, Vol. 5, Oxford University Press, New York (NY), 2003.

[67] Y. L. Sagalovich, "Separating systems," *Probl. Information Transmission*, Vol. 30, No. 2, pp. 105–123, 1994.

[68] H. G. Schaathun, "Fighting three pirates with scattering codes," *Proc. IEEE Internat. Sympos. Information Theory* (ISIT 2004), p. 202, Jun. 2004.

[69] H. G. Schaathun and M. Fernandez, "Boneh-Shaw fingerprinting and soft decision decoding," *Proc. 2005 IEEE ISOC ITW on Coding and Complexity*, M. J. Dinneen (Ed.), pp. 183–186, 2005.

[70] F. Sebe and J. Domingo-Ferrer, "Short 3-secure fingerprinting codes for copyright protection," *Proc. ACISP 2002*, L. Batten and J. Seberry (Eds.), Lect. Notes Comput. Science, Vol. 2384, pp. 316–327, 2002.

[71] K. W. Shum, I. Aleshnikov, P. V. Kumar, H. Stichtenoth and V. Deolalikar, "A low-complexity algorithm for the construction of algebraic-geometric codes better than the Gilbert-Varshamov bound," *IEEE Trans. Inform. Theory*, Vol. 47, pp. 2225–2241, Sep. 2001.

[72] A. Silverberg, J. Staddon and J. L. Walker, "Applications of list decoding to tracing traitors," *IEEE Trans. Information Theory*, Vol. 49, No. 5, pp. 1312–1318, May 2003.

[73] B. Škorić, S. Katzenbeisser and M. U. Celik, "Symmetric Tardos fingerprinting codes for arbitrary alphabet sizes," *Designs, Codes and Cryptography*, Vol. 46, No. 2, pp. 137–166, Feb. 2008.

[74] B. Škorić, T. U. Vladimirova, M. Celik and J. C. Talstra, "Tardos fingerprinting is better than we thought," *IEEE Trans. Information Theory*, Vol. 54, No. 8, pp. 3663–3676, Aug. 2008.

[75] A. Somekh-Baruch and N. Merhav, "On the capacity game of private fingerprinting systems under collusion attacks," *IEEE Trans. Information Theory*, Vol. 51, No. 3, pp. 884–899, Mar. 2005.

[76] A. Somekh-Baruch and N. Merhav, "Achievable Error Exponents for the Private Fingerprinting Game," *IEEE Trans. Information Theory*, Vol. 53, No. 5, pp. 1827–1838, May 2007.

[77] A. Somekh-Baruch and N. Merhav, "Correction to "On the Capacity Game of Private Fingerprinting Systems Under Collusion Attacks" [Mar 05 884–899]," *IEEE Trans. Information Theory*, Vol. 54, No. 11, pp. 5263–5264, Nov. 2008.

[78] J. N. Staddon, D. R. Stinson and R. Wei, "Combinatorial properties of frameproof and traceability codes," *IEEE Trans. Information Theory*, Vol. 47, No. 3, pp. 1042–1049, Mar. 2001.

[79] G. Tardos, "Optimal probabilistic fingerprint codes," *Journal of the ACM*, Vol. 55, No. 2, Art. 10, 24pp., May 2008, Preliminary version in *Proc. 35th Annual ACM Symposium on Theory of Computing* (STOC 2003), pp. 116–125, 2003.

[80] M. A. Tsfasman and S. G. Vlăduţ, *Algebraic-Geometric Codes*, Dordrecht, The Netherlands: Kluwer, 1991.

[81] M. A. Tsfasman, S. G. Vlăduţ and T. Zink, "Modular curves, Shimura curves and Goppa codes better than the Varshamov-Gilbert bound," *Math. Nachrichtentech.*, Vol. 109, pp. 21–28, 1982.

[82] W. Trappe, M. Wu, Z. Wang and K.J.R. Liu, "Anti-collusion fingerprinting for multimedia," *IEEE Trans. on Signal Processing*, Vol. 51, No. 4, pp. 1069–1087, Apr. 2003.

[83] N. Wagner, "Fingerprinting," *Proc. IEEE Sympos. Security and Privacy*, pp. 18–22, Apr. 1983.

[84] Z. J. Wang, M. Wu, W. Trappe, and K.J.R. Liu, "Group-Oriented Fingerprinting for Multimedia Forensics," *EURASIP Journal on Applied Signal Processing*, Vol. 2004, No. 14, pp. 2153–2173, 2004.

[85] M. Wu, W. Trappe, Z. Wang and K.J.R. Liu, "Collusion resistant fingerprinting for multimedia," *IEEE Signal Processing Magazine - Special Issue on Digital Rights Management*, pp. 15–27, Mar. 2004.

[86] O. Yağan and A. M. Makowski, "On the random graph induced by a random key predistribution scheme under full visibility," *Proc. IEEE Internat. Sympos. Information Theory* (ISIT 2008), pp. 544–548, 2008.

[87] C.-W. Yi, P.-J. Wan, X.-Y. Li and O. Frieder, "Asymptotic distribution of the number of isolated nodes in wireless ad hoc networks with Bernoulli nodes," *Proc. IEEE Wireless Communications and Networking Conference* (WCNC 2003), pp. 1585–1590, 2003.

[88] C.-W. Yi, P.-J. Wan, K.-W. Lin and C.-H. Huang, "Asymptotic distribution of the number of isolated nodes in wireless ad hoc networks with unreliable nodes and links," *Proc. IEEE Global Telecommunications Conference* (GLOBECOM 2006), Nov. 2006.