Performance Analysis of an Authentication Method relying on Graphical Codes

Bao An Mai Hoang

N° d'ordre : 254

ÉCOLE CENTRALE DE LILLE

THÈSE

pour obtenir le grade de

Docteur

En

Spécialité : **Automatique, Génie informatique, Traitement du signal et images**

présentée et soutenue publiquement
par

Bao An MAI HOANG

DOCTORAT DELIVRE PAR L'ÉCOLE CENTRALE DE LILLE

Titre:

Analyse de performance d'un système d'authentification utilisant des codes graphiques

Performance Analysis of an Authentication Method relying on Graphical Codes

Soutenue le 01/12/2014 devant le jury d'examen:

Jury

M. Philippe VanHeeghe, Prof. à l'Ecole Centrale de Lille,	Président
M. Lionel Fillatre, Prof. à l'Université de Nice Sophia-Antipolis,	Rapporteur
M. William Puech, Prof. à l'Université de Montpellier,	Rapporteur
M. François Cayre, MCF à l'Université de Grenoble,	Examinateur
M. Patrick Bas, Chercheur CNRS,	Directeur de thèse
M. Wadih Sawaya, MCF à Télecom Lille,	Co-Encadrant de thèse

Thèse préparée dans le Laboratoire d'automatique, génie informatique et signal **(LAGIS)**

dans le cadre de l'École Doctorale **SPI 072 (Lille I, Lille III, Artois, ULCO, UVHC, EC Lille)**

PRES Université Lille Nord-de-France

Analyse de performance d'un système d'authentification utilisant des codes graphiques

Bao An Mai Hoang

January 19, 2015

I dedicate this thesis to my family, my girlfriend and friends whom I luckily have in my life.

Contents

Acknowledgments

Firstly, I want to express my sincere gratitude to the everybody in the jury: two reviewers Prof. Lionel Fillatre and Prof. William Puech, the President of the jury Prof. Philippe VanHeeghe, the examiner Dr. François Cayre, and my two advisors Dr. Patrick Bas and Dr. Wadih Sawaya, who have accepted to read my manuscript and will give me invaluable comment and advice, from which I believe that I would have not only a nice thesis but also progress in my future research.

Especially, I would like to use this opportunity to express my thanks to my PhD advisors, Dr. Patrick Bas and Dr. Wadih Sawaya, for supporting me during these past three years in my PhD life. I am thankful for their helpful guidance, endlessly constructive criticism and friendly advice. Without their help, I cannot imagine where I would be in my research. Pedagogically, carefully and patiently, they shared their experiences and illuminating views with me, put me in the right direction and taught me how to write a good scientific paper. Working with them was a great chance for learning and professional development.

I am sincerely grateful to my family who always give me unconditional love, constant encouragement and continuous support. If I didn't have them in my life, I would be sure that I cannot have chance to sit here and write these words of gratitude. Therefore, I would like to dedicate this thesis to my beloved family.

Last but not least, I also express my warm thanks to my colleagues Thi Le Thu Nguyen, Quoc Thong Nguyen and Anh Thu Phan Ho and all of my friends in France for every wonderful and unforgettable time that we have spent together. I especially thank to all nice people in Telecom Lille and Ecole Centrale de Lille whose appearance in my life made me feel happier and more joyful.

Bao An Mai Hoang

Liste des symboles

The list of notations and their interpretation: See below table

Notations	Interpretation
α	Probability of type I error.
β	Probability of type II error.
P_{FA}	Probability of false alarm.
P_{ND}	Probability of non-detection.
ROC curve	Receiver Operating Characteristic curve.
LLR	Log-likelihood ratio.
GC	Graphical code.
GLRT	Generalized likelihood ratio test.
UMP Test	Uniformly Most Powerful Test.
BIBO	Binary input binary output.
CLT	Central limit theorem.
EM algorithm	Expectation Maximization algorithm.
NP test	Neyman-Pearson test.
MLE	Maximum Likelihood Estimation.
ρ	The quadratic form of the error (the variation of the estimation).
$\mu(s)$	Cumulant generating function.
X^N	Original source of the GC of the main (legitimate) channel.
Y^N	Printed source from the main (legitimate) channel.
$\hat{X}^N$	Estimator of X^N.
Z^N	Reprinted from the opponent channel after thresholding.
θ	Parameters of the main channel.
θ	Parameters of the opponent channel.
$\hat{\bar{\theta}}$	Estimated parameters of the opponent channel.
$P_{Y\|X,\theta}$	Probability distribution of the main channel.
$P_{Z\|X,\bar{\theta}}$	Probability distribution of the opponent channel.
$D_{KL}(\cdot \parallel \cdot)$	Kullback-Leibler divergence.
p_{s_i}	Boltzmann's distribution indexed by parameter s_i.

Chapter 1

Introduction

"*Books aren't written, they're rewritten. Including your own. It is one of the hardest things to accept, especially after the seventh rewrite hasn't quite done it...*"

Michael Crichton

1.1 Motivations.

Authentication of genuine goods is a problem which is nowadays more and more concerning in our Society. According to the news channel CNN [10]: "*The global trade in counterfeit goods is booming, and it's shifting from relatively innocuous items like shoes and handbags to things like medicine and pesticides that can carry serious health and safety implications.*"

The economic impact of counterfeiting industry is also significant. A report of The Organization for Economic Co-operation and Development (OECD) written in 2009 states that the profit from counterfeit goods was responsible for more than $250 billion in total profit of the world trade [5, 2]. This amount can be compared with the profit

of international trade in illegal drugs such as heroin, cocaine, methamphetamine or ecstasy,...and is even more important than other underworld economies such as weapons smuggling, money laundering and human trafficking [10].

According to other sources, the figures are different but still impressive: for example the Commercial Crime Services (CCS), a specialized division of the world business organization International Chamber of Commerce (ICC) declares that the profits of counterfeiting could reach an estimated $600 billion a year [1] on average and accounts for between $5-7\%$ of international trade around the world. For example, in 2008 CCS approximated counterfeit goods were worth up to $650 billion.

From the assessment of CCS, specialists of ICC state that the revenue of counterfeit goods across the world market could surpass $1.7 trillion and contribute over 2% of the world's total output by 2015 [10]. It gives a huge profit for those who produce and distribute counterfeit goods but it's also a disaster for both the consumers and the current economy.

Several examples are presented below to show the aftermath of counterfeiting industry to the global consumers and world economy.

According to CNBC [9], among all counterfeit items imported to the United States in 2009, there was $260 million coming from clothing and accessories. Especially, in New York the counterfeit market benefited approximately $34 billion for imitations, robbing $1.6 billion in tax revenue. Disappointingly in the anti-counterfeiting fight, the results have been poor so far despite many efforts done by the governments and security companies. For example in 2002 the U.S Federal officials only confiscated $138 million knockoffs in which counterfeit clothing is the most popular product. This accounts 18% of all counterfeit items (see Fig. 1.1). As a matter of fact, this amount of money was much smaller than the revenue of counterfeit goods.

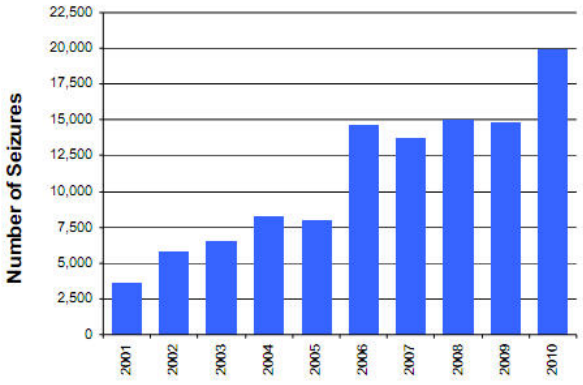


Figure 1.1: Growth in seizures of counterfeit goods by U.S. Taken from [13].

The problem of counterfeiting is even more worrying within the pharmaceutical industry: with $206.2 billion pharmaceutical sales in Europe per year the profits based

on counterfeit products account for 16%, while in Asia this rate reaches 39% [9]. In 2005, the World Health Organization (WHO) reported that among the medicines produced in developing countries, there was nearly 25% of forgery [4]. Interpol [3] warned that such products nowadays are threatening public health at an international alert level and day by day the consumption of fake drugs, medicines and medical devices endangers the health and life of patients across the world, especially across developing countries. Why?

According to Interpol [3]: *"Illicit drugs can contain the wrong dose of active ingredient, or none at all, or a different ingredient. They are associated with a number of dangers and, at worst, can result in heart attack, coma or death*". Also declared in [3], due to the growing of number of Internet users as well as untrusted sale-online web pages, patients can buy medicines easily, cheaply without prescription of the doctor and this fact makes the fight against counterfeit medicines even more difficult.

The Nato Office on Drugs and Crime stated in a recent report titled "Transnational Organized Crime in East Asia and the Pacific", in which there was nearly 70% of all counterfeits confiscated coming from China in the period 2008-2010, while in U.S, according to the Customs, this percentage was 87% for the same period [12]. We can see in the chart 1.2 the ten countries with the most important seizure values of products violating the Intellectual Property Rights (IPR):

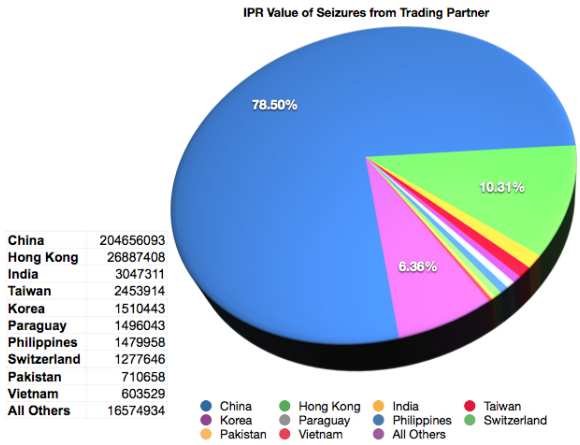


Figure 1.2: Ten largest contributors of value of IPR seizures, taken from [11].

In another report [8], the U.S International Trade Commission investigated U.S businesses and they observed that there was approximately $48 billion lost by the infringement of IPR by China in 2009 (see Fig. 1.3).

We can see that the problem of counterfeit prevention and authentication of physical products such as documents, goods, drugs, jewels, ... becomes a major concern in a world of global exchanges, and one important task is to protect the legal manufacturers and consumers. This is the main motivation of the ANR "Estampille" project which is presented in the next section and with which this thesis is related.

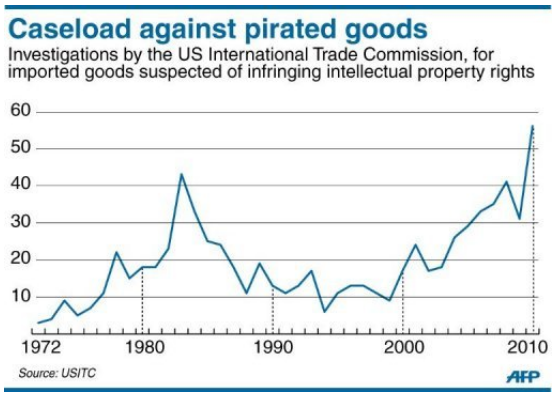


Figure 1.3: Chart showing the rising number of complaints lodged in the US against imported goods suspected of infringing intellectual property rights [8].

1.2 The Estampille project

The general framework of the Estampille project is to fight against forged printed documents and counterfeited products by protecting their packages [7]. In order to do so, the project proposes to insert Graphical Codes (GC) on the document or the package of the commercial product (see an example of such a code in Figure 1.4). The use of GC in security framework enables both to perform integrity check of the printed document (detecting that a document has not been tempered) and to perform authentication (detecting which document is a counterfeit). In fact, CG have already been used by different companies, among which the company Advanced Track and Trace belongs to, on millions of commercial products in the pharmaceutical industry, cosmetics, wines and spirits, valuable documents and parts

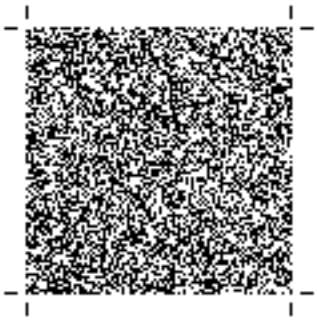


Figure 1.4: A simulated Graphical Code before being printed.

1.2.1 Participant organizations

This project is a four years industrial research project financed by the French National Research Agency (ANR) and led by Dr Patrick Bas in a collaboration of 6 partners:

- ATT (Advanced Track & Trace) is an industrial company working on authentication, protection against counterfeit and products tracking. In Estampille, ATT provides technical part in 2D graphical code.
- LAGIS (Laboratoire d'Automatique, Génie Informatique et Signal) is a scientific laboratory in Ecole Centrale de Lille working on automatic systems, computer engineering and signal processing. LAGIS provides the expertise about authentication and stochastic modeling of printing processes. The work presented in this memoir has been conducted within this laboratory.
- GIPSA (Grenoble Images Parole Signal Automatique) is a joint laboratory between CNRS and university of Grenoble working on theoretical and applied research on signals and systems. GIPSA provides expertise about security analysis and integrity control.
- LGP2 (Laboratoire Génie des Procédés Papetiers) is a laboratory in university of Grenoble working on intelligent processes, materials chemistry, solid mechanics, mechanics of materials and printing processes. LGP2 provides expertise about description and analysis of printing processes at the microscopic level.
- LATA is an industrial company working on printing technologies. LATA provides expertise and Data from various printing processes.
- CERDI (Centre d'Etudes et de Recherche en Droit de l'Immatériel) is belongs to university Paris 11 working on the law in intellectual property of new technologies. CERDI provides legal basis for the use of graphical code.

1.2.2 The authentication setup in Estampille

The general framework of Estampille project can be depicted in the Figure (1.5). We recall here the main step of the authentication process:

- In step (1), we generate a simulated GC from a random source and model the legitimate printing channel to print this GC out, called printed original GC, then we insert it into the legal product in step (2).
- The opponent observes the printed original GC, and tries to process it in order to be able to print it by his printing channel in step (3) creating a reprinted forged GC.
- He then inserts it into his illegal product in step (4).
- Both printed original and reprinted forged GC are observed by the receiver, and in order to detect the fake product, firstly the receiver has to process these GCs in step (5) and then to perform the authentication test in step (6).

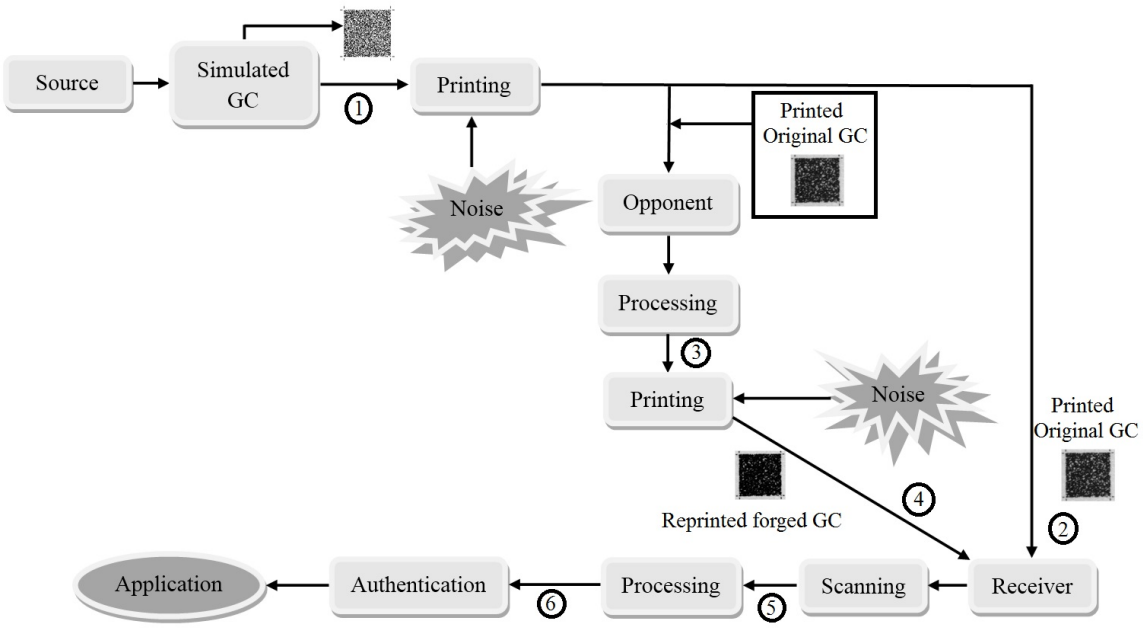


Figure 1.5: Description of the works of Estampille project

Generally speaking, the Estampille's setup can be summarized briefly in three main tasks as follow:

- Analysis and modeling of the printing processes from a physical and signal processing approaches.
- Achieving global security of the authentication system.
- Design of efficient GC for authentication.

The company ATT has developed the technology for making 2D GC in order to maximize the quantity of information lost by the forgers. This design is based on the fact that the printing process in the real environment comes from complex phenomena. For example, it can be governed by the intrinsic features of the printers, the physical properties of the ink drop, the randomness of the paper's fiber, etc... Viewed under a microscope, we can see in Figure (1.6), the surface of a sheet of paper is not perfectly flat but is tangled. In fact, it is like a mixture of wood fibers which is highly random and difficult to reproduce [28]. Because of this randomness, a scanner will produce a different image depending on the orientation of the page and the printer will cause a stochastic non-invertible noise when conducting a printing process in the paper to print a GC out.

Randomness affects the authentic process and may degrade the accuracy of authentication performance. Consequently, the opponent can try to take advantage of this objective factor to entrap the detector (the receiver). It is consequently important

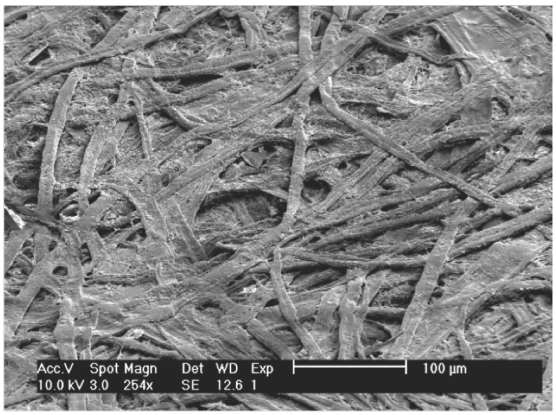


Figure 1.6: An ordinary piece of paper viewed under a microscope [6, 28].

to evaluate both the security and authentication performances of this complex system more accurately.

Our team in LAGIS proposes two directions of development to Estampille project: **1)** the modeling of the printing processes as well as the characterization of the parameters for the printer and **2)** the analysis of the authentication performances for different possible types of attacks that the opponent develop to forge the graphical codes and the impact of coding theory in authentication.

– The first direction is carried out by my colleague Quoc Thong Nguyen and his advisors Prof. Yves Delignon and Lionel Chagas. Their works mainly consist in characterizing the intrinsic features of the printer. Motivated from the microscopic analysis of paper printing, in [76, 77] they propose to model the micrometric scan of document printing by a binary response model whose the parameters depend on the location and the shape of ink dots. Ink dots viewed under microscope are shown in Figure (1.7). They provide a maximum likelihood identification algorithm, its performance is assessed through simulations and true data.

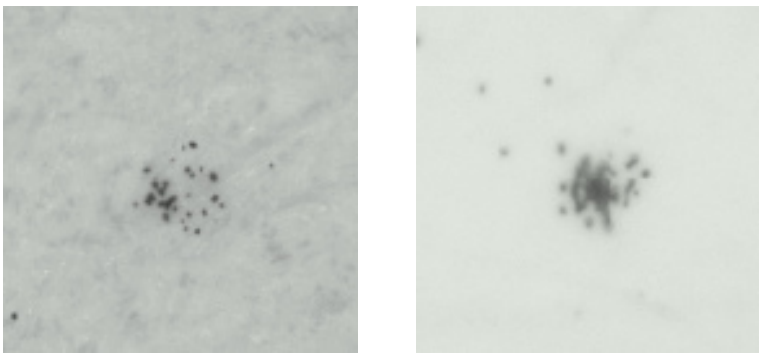


Figure 1.7: Left: Ink dots in uncoated paper printed in Laser printer (600dpi). Right: Ink dots in coated paper printed in Laser printer (600dpi).

Furthermore, they illustrate the benefit of a such model and estimation algorithm in the case of authentication of printer from micro-tag made of one dot in [76] and multiple dots in [77].

– The second direction is mostly taken into account by my colleague Anh Thu Phan Ho and me with our advisors Dr. Wadih Sawaya and Dr. Patrick Bas. While the work of Anh Thu is to consider the benefit of information theory, channel coding and coding schemes for authentication, my researches focus on the use of signal processing, statistical estimation and hypothesis testing to improve authentication performance while guaranteeing security.

1.3 Sketch of the thesis

This thesis is a part of my researches that aims to answer the second direction of the Estampille project. We study an authentication system on a 2D GC, which will be presented in details within the next chapter, based on the fact that a printing process at very high resolution can be seen as a stochastic process. This is due to the nature of different elements such as the paper fibers, the ink heterogeneity, or the dot addressability of the printer as mentioned above. Our solution to perform authentication is to use the hypothesis testing on the observed memoryless sequences of a printed GC, by considering the assumption that we are able to perfectly model the printing process [81, 53] and by deriving an optimal test for which the probabilities of error can be accurately approximated. Moreover, when looking for a more practical scenario, we take into account the estimation of the printing process used to generate the GC of the opponent and we see how it impacts the performance of authentication [66]. We also try to optimize the printing channel controlled by the legitimate manufacturers in order to maximize the ability of detecting a forged GC [53, 54].

The main context of the thesis begins with Chapter 2 in order to present the necessary theoretical backgrounds and state of the art for the thesis.

In Chapter 3, the theoretical analysis of the authentication test and associated error probabilities, together with numerical implementations are proposed.

In Chapter 4, security constraints are taken into account by the fact that the receiver tries to optimize of the parameters of the original printing channel while the adversary tries to optimize his own channel. This is modeled as a min-max game which is solved by using optimization tools.

The thesis is ended in Chapter 5 with the overall conclusions, the existing drawbacks and perspectives as well as the indication for the directions of future researches.

1.4 Publications related to the thesis

A part of works presented in this thesis has been published in one journal paper and three conferences papers:

- *Document Authentication Using Graphical Codes: Reliable Performance Analysis and Channel Optimization*,
  Anh Thu Phan Ho, **Bao An Hoang Mai**, Wadih Sawaya and Patrick Bas,
  EURASIP Journal on Information Security, 2014, pp. 10.1186/1687

- *Image Model and Printed Document Authentication: a Theoretical Analysis,*
  Bao An Hoang Mai, Wadih Sawaya and Patrick Bas
  IEEE International Conference on Image Processing, Oct 2014, France. IEEE International Conference on Image Processing, IEEE ICIP

- *Authentication using Graphical Codes: Optimisation of the Print and Scan Channels,*
  Anh Thu Phan Ho, **Bao An Hoang Mai**, Wadih Sawaya and Patrick Bas
  EUSIPCO 2014, Sep 2014, Portugal.

- *Document Authentication Using Graphical Codes: Impacts of the Channel Model,*
  Anh Thu Phan Ho, **Bao An Hoang Mai**, Wadih Sawaya and Patrick Bas
  ACM Workshop on Information Hiding and Multimedia Security, Jun 2013, Montpellier, France. ACM IH-MMSEC

Chapter 2

Fundamental backgrounds and Related works

> *"We can only see a short distance ahead, but we can see plenty there that needs to be done."*
>
> Alan Turing

In this chapter we presents two important aspects in mathematical statistics which are used in the entire thesis: Parameter estimation and hypothesis testing. They are essential and very useful in a vast area of research field in signal processing. From the practical point of view, we have to use estimation theory in order to extract information about the printing channel and hypothesis testing to derive a test to perform authentication.

We also introduce some aspects of authentication, especially the authentication of printed objects, and we remind several advances in this field and in modeling the printing-scanning process.

2.1 Fundamental backgrounds

2.1.1 Hypothesis testing for known parameters

In the entire thesis we define "Authentication" as the problem of classifying between two groups: one contains authentic objects and the other contains inauthentic objects. To solve it, it is common to use machine learning or hypothesis testing.

Hypothesis testing problem arises in many context (statistical signal processing, communication, life sciences, social sciences....) and is an active topic in statistics. The primary task of hypothesis testing is to use observed data to take decisions by distinguishing the true hypothesis among the set of M surveyed hypotheses.

Classical binary hypothesis testing

We can consider two different testing approaches: classical and Bayesian hypothesis testing. The Bayesian approach considers that the prior distributions of the hypothesis are concerned while for other approach the prior distributions are assumed to be equiprobable.

Because it is difficult to gather information on the prior distribution, we focus our work on the classical binary hypothesis testing (see Fig. 2.1) to test two kinds of hypotheses: 1) the null hypothesis H_0 and 2) the alternative hypothesis H_1. A decision d is derived based on the observed data in order to decide if H_0 is true or not. For any decision, there are two types of error called type I and type II error (Fig. 2.2).

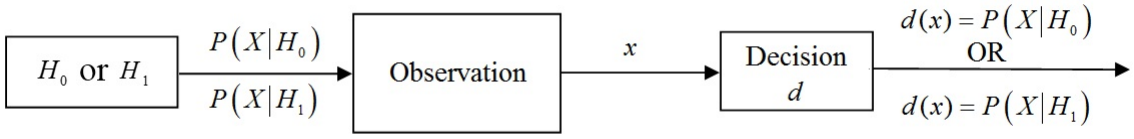


Figure 2.1: Classical (non-Bayesian) binary hypothesis testing.

A type I error is equivalent to rejecting the null hypothesis H_0 while H_0 is true. In hypothesis testing, the probability of type I error α is often referred as the significance level of the test or the probability of false alarm (P_{FA}). On the other hand, accepting H_0 when H_1 is true will cause a type II error whose probability β is often called the

probability of type II error or probability of non-detection (P_{ND}). We define the power of a test $P_D = 1 - \beta$ as the probability of rejecting H_0 while H_1 is indeed true.

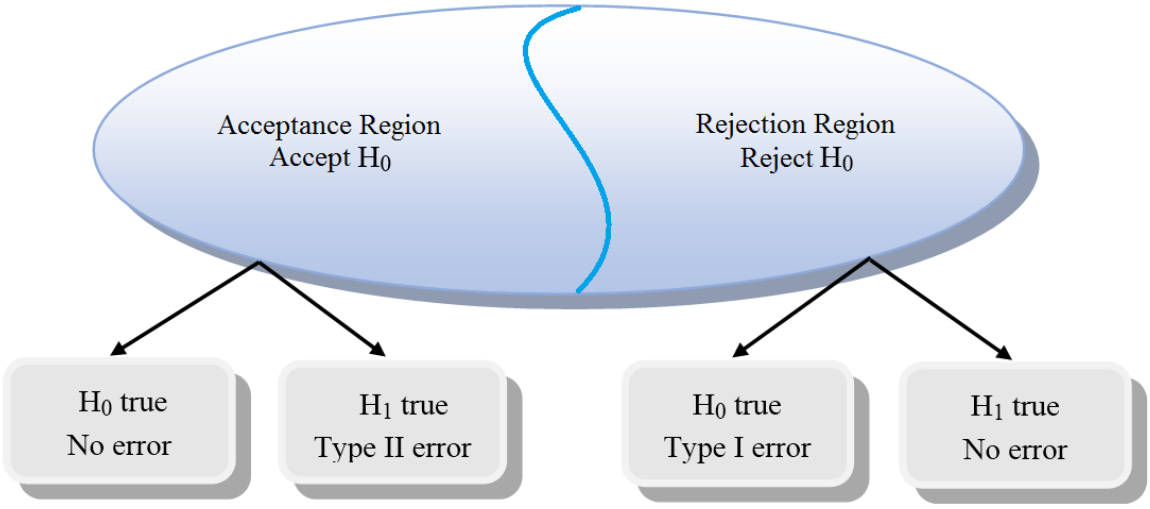


Figure 2.2: Graphical description of a decision rule for classical binary hypothesis testing.

In hypothesis testing, it is desired to make α and β as small as possible. Nevertheless, there is an interplay between α and β and both cannot be negligible at the same time. It means that when α is increased, β decreases and vice versa. Hence we have to accept "a trade-off" between α and β, for example by using Receiver Operating Characteristic (ROC) curves which show the evolution of β (or power P_D) w.r.t α. A ROC curve has several important properties:

1. (α, β) moves continuously along the ROC curve.
2. All points on a ROC curve satisfy $P_D \geq \alpha$.

An example of the ROC curves is shown in Fig. 2.3, in which we consider the null hypothesis $H_0 : P(X|H_0) \equiv BSC(p)$ and alternative hypothesis $H_1 : P(X|H_1) \equiv BSC(2p(1-p))$ (BSC means the Binary Symmetric Channel with transition probability p), and we use Monte Carlo simulation to compute α and β.

In classical binary testing, the *Neyman-Pearson theorem* plays an essential role for getting an optimum decision rule. This theorem states that the optimum decision rule that minimizes β for a given α (see in [63]) is given by comparing the likelihood ratio between the two hypothesis with a threshold. We recall it below:

Neyman-Pearson theorem : Suppose we have random variables x^N distributed by an unknown probability density in a sample space in $\mathcal{X} \subseteq \mathbb{R}^N$. Among all the procedures

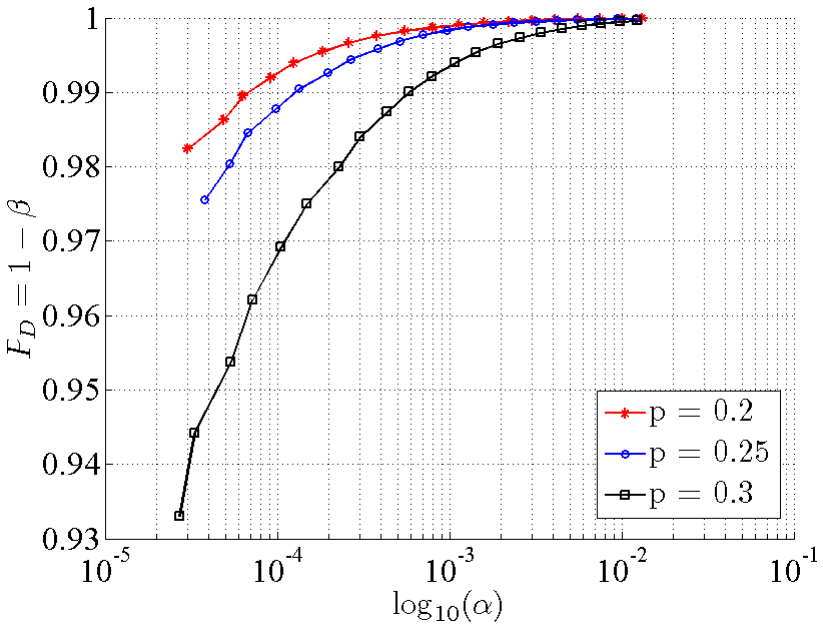


Figure 2.3: ROC curves of test between two hypothesis: $H_0 : P(X|H_0) \equiv BSC(p)$ and $H_1 : P(X|H_1) \equiv BSC(2p(1-p))$, the length of $i.i.d$ sequence X is 500, p is equal to 0.2, 0.25 and 0.3 respectively. The number of trials is 10^6.

applied to x^N to test if the distribution of x^N comes from the hypothesis $H_0 : \theta = \theta_0$ or $H_1\, \theta = \theta_1$, the likelihood-ratio test between H_0 and H_1 with a threshold λ satisfying:

$$\Lambda(x) = \frac{L(\theta_1 \mid x^N, H_1)}{L(\theta_0 \mid x^N, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda$$

is the most powerful test of a significance level α given by $Pr[\Lambda(x) > \lambda \mid H_0]$. Thus, β or the probability of non-detection P_{ND} is defined as $Pr[\Lambda(x) < \lambda \mid H_1]$. Herein, $L(\theta_i \mid x^N, H_i)$ is likelihood function based on hypothesis H_i $(i = 0, 1)$ and the most powerful test means the test with the largest power (1-β) for a given significant level α. If the logarithm of the likelihood ratio is used, the test is known as a log-likelihood ratio test (LLR test).

The optimum solution coming from the Neyman-Pearson problem requires to select the threshold λ to obtain the smallest possible β while keeping $\alpha \leq \alpha^*$, α^* is fixed. The ROC curve is then used to analyse the performances of the test for different thresholds and the area under the curve (AUC):

$$\int_0^1 \beta d\alpha,$$

can be used to measure its average performance.

If now we want to apply hypothesis testing using the LLR test for authentication, we assume that the receiver observes an unknown $i.i.d$ sequence $v^N = (v_1, v_2, ..., v_N)$ of length N in the observation space $\mathcal{V}^N$. In order to perform authentication he assumes that H_0 is the hypothesis that v^N comes from the legitimate source with probability distribution Q_0 while H_1 is the hypothesis that v^N is sent by the opponent with probability distribution Q_1, then the receiver can use the LLR test between H_0 and H_1 to map the N dimensional problem into a one dimensional problem (see also Fig. 2.4):

$$L_R = \log \frac{Q_1\left(v^N \middle| H_1\right)}{Q_0\left(v^N \middle| H_0\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \tag{2.1}$$

If two densities Q_0 and Q_1 are known, the probability of type I and type II error are given by:

$$\begin{array}{rcl} \alpha & = & \int_{\lambda}^{+\infty} P_{L_R|H_0}(l)dl \\ \beta & = & \int_{-\infty}^{\lambda} P_{L_R|H_1}(l)dl. \end{array} \tag{2.2}$$

Here, in Fig. 2.4, we define:

$$\mathcal{H}_0 = \left\{v^N \in \mathcal{V} \subseteq \mathbb{R}^N \,:\, L_R(v^N) < \lambda\right\} \tag{2.3}$$

and

$$H_0 = \left\{L_R \in \mathbb{R} \,:\, L_R < \lambda\right\}. \tag{2.4}$$

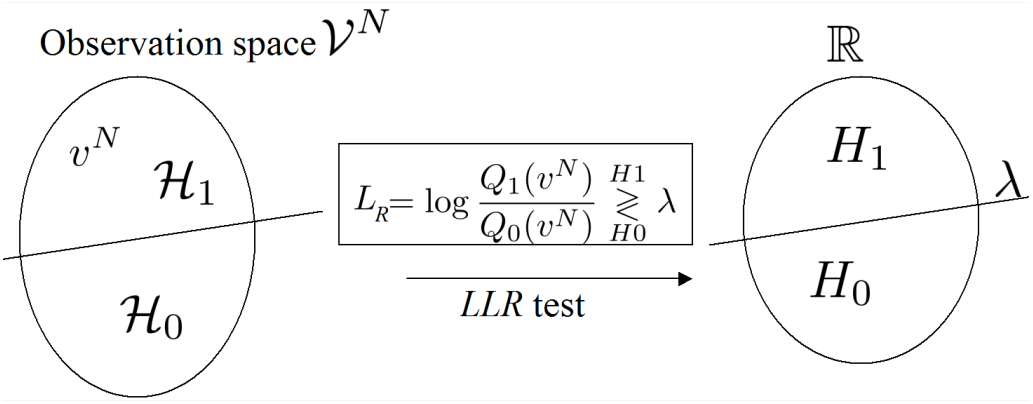


Figure 2.4: Classical (non-Bayesian) binary hypothesis testing using LLR test.

For example, a signal sequence $v = (v_1, v_2, \ldots, v_n)$ with $\{v_i\}_1^n$ is n i.i.d random variables is sent to the receiver. We assume that the receiver knows that v_i is only distributed by $Q_0 \sim \mathcal{N}(\mu_1, \sigma^2)$ or $Q_1 \sim \mathcal{N}(\mu_2, \sigma^2)$, and he has to determine where x comes from by considering a test:

$$\begin{aligned} H_0 : v_i &\sim \mathcal{N}(\mu_1, \sigma^2), \\ H_1 : v_i &\sim \mathcal{N}(\mu_2, \sigma^2), \end{aligned}$$

where $\mu_1, \mu_2 > 0$. The receiver can use an LLR test given in this case by:

$$L_R(v) = \log \frac{Q_1(v^n \mid H_1)}{Q_0(v^n \mid H_0)} = \frac{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(v_i-\mu_2)^2}}{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(v_i-\mu_1)^2}} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \tag{2.5}$$

It leads to

$$\sum_{i=1}^{n} v_i \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma^2}{\mu_2 - \mu_1} log(\lambda) + \frac{n}{2}(\mu_2 + \mu_1) \equiv \zeta. \tag{2.6}$$

The probability of type I error is given by:

$$\alpha = Pr\left[\sum_{i=1}^{n} v_i > \zeta \mid H_0\right] = Q\left(\frac{\zeta - n\mu_1}{\sigma\sqrt{n}}\right). \tag{2.7}$$

and probability of type II error is given by:

$$\beta = Pr\left[\sum_{i=1}^{n} v_i < \zeta \mid H_1\right] = 1 - Q\left(\frac{\zeta - n\mu_2}{\sigma\sqrt{n}}\right). \tag{2.8}$$

where $Q(x)$ is error function. In this particular case, we can rewrite β as the function of α as follow:

$$\beta = 1 - Q\left(Q^{-1}(\alpha) + \sqrt{n}\frac{\mu_1 - \mu_2}{\sigma}\right). \tag{2.9}$$

However, generally it is difficult to compute α and β because the densities of $P_{L_R|H_1}(l)$ and $P_{L_R|H_0}(l)$ are unknown and both the probability of error and the likelihoods have to be approximated numerically. In this case, because of the approximations and the numerical computation, the test is not optimal anymore.

Asymptotic properties of *LLR* test

In order to derive asymptotic properties for the LLR, if we set $q_k = \log \frac{Q_1(v_k|H_1)}{Q_0(v_k|H_0)}$ and $S_N = \frac{1}{N}\Sigma_{k=1}^{N} q_k$. From strong law of large number, when $N \to +\infty$, we consequently have:

$$\begin{aligned} H_0 : \quad S_N &\overset{a.s}{\to} E\left[q_k \middle| H_0\right] = \int \log \tfrac{Q_1(x|H_1)}{Q_0(x|H_0)} Q_0\left(x \middle| H_0\right) dx, \\ H_1 : \quad S_N &\overset{a.s}{\to} E\left[q_k \middle| H_1\right] = \int \log \tfrac{Q_1(x|H_1)}{Q_0(x|H_0)} Q_1\left(x \middle| H_1\right) dx. \end{aligned} \tag{2.10}$$

It is equivalent to write, using $D_{KL}(P\|Q) = \int P(x)\log\frac{P(x)}{Q(x)}dx$ the Kullback - Leibler divergence between two densities $P(x)$ and $Q(x)$, that:

$$\begin{array}{rrl} H_0: & S_N \overset{a.s}{\to} & -D_{KL}(Q_0\|Q_1), \\ H_1: & S_N \overset{a.s}{\to} & D_{KL}(Q_1\|Q_0). \end{array} \tag{2.11}$$

Therefore as long as we can collect an arbitrarily large number of $i.i.d$ observations, we can separate perfectly H_0 and H_1 (Q_0 and Q_1 are completely different).

If one of the probabilities of error goes to zero arbitrarily slowly, the Stein's lemma [30] provides the best exponent bound to minimize the other probability of error. It means that when α is very close to zero then

$$\frac{1}{N}\lim_{N\to\infty}\log\beta = -D_{KL}(Q_0\|Q_1), \tag{2.12}$$

or similarly when β tends to zero, it gives

$$\frac{1}{N}\lim_{N\to\infty}\log\alpha = -D_{KL}(Q_1\|Q_0). \tag{2.13}$$

However, in a realistic manner, the Stein's lemma cannot be considered as an approximation of the probabilities of error due to the fact that N is limited and every practical detector has to cope with a value of α or β that may be small but not very close to zero.

Uniformly Most Powerful Test

In statistics, the classical binary hypothesis testing (or NP-test) based on likelihood ratio test statistic can be seen as a specific case of a more general testing problem called the uniformly most powerful test (UMP test) which can be used to test between simple hypothesis $H_0: \theta \in \Theta_0$ and $H_1: \theta \in \Theta_1$ with $\Theta_0 \cup \Theta_1 = \Theta$ and $\Theta_0 \cap \Theta_1 = \emptyset$.

An hypothesis test is called an UMP test of a significance level α^* with any test statistic $T^*(x)$ satisfying:

$$\sup_{\theta\in\Theta_0}\mathbb{E}_\theta[T^*(X)] = \alpha^*, \tag{2.14}$$

if for any other test statistic $T(x)$ of size α such that:

$$\sup_{\theta\in\Theta_0}\mathbb{E}_\theta[T(X)] = \alpha \leq \alpha^*, \tag{2.15}$$

we always suffer a loss in power, i.e.,

$$\mathbb{E}_\theta[T(X)] \leq \mathbb{E}_\theta[T^*(X)] \qquad \forall\theta\in\Theta_1. \tag{2.16}$$

An UMP test in general does not always exist. For example, we can let $X \sim Binom(n,\theta)$ and suppose we want to test:

$$H_0 : \theta = \theta_0 \quad v.s \quad H_1 : \theta \neq \theta_0 \tag{2.17}$$

at some level α. There is no UMP test in this case [63].

However when the test exists, it can be found by two methods. The first one comes from Neyman-Pearson theorem as we have stated above. The second method, used in case of scalar parameter and based on the concept of monotone likelihood ratio, can be understood as an extension of Neyman-Pearson theorem and it is called Karlin-Rubin theorem [60].

Another difficulty with the optimality in NP-test or UMP test (if it exists) is that the density of the population must be assumed to be known, except for a finite number of parameters. This assumption makes the testing problem easier to solve, but in the real scenario it will rarely be true.

2.1.2 Hypothesis testing for unknown parameters

As we know from the previous subsection, the main requirement for a NP-test of an observed sequence v^N

$$\begin{aligned} H_0 &: v^N \overset{i.i.d}{\sim} f_0(x|\,\theta_0)\ \ \theta_0 \in \Theta_0 \\ H_1 &: v^N \overset{i.i.d}{\sim} f_1(x|\,\theta_1)\ \ \theta_1 \in \Theta_1 \end{aligned} \tag{2.18}$$

is the knowledge of the underlying distributions $f_0(x|\,\theta_0)$ and $f_1(x|\,\theta_1)$ as well as the specific parameters θ_0 and θ_1, where

$$\begin{aligned} \Theta_0 \cup \Theta_1 &= \Theta \\ \Theta_0 \cap \Theta_1 &= \emptyset \end{aligned} \tag{2.19}$$

However, in a realistic situation we do not know exactly the densities of f_0 and f_1 as well as the true parameters θ_0 and θ_1. In this case, we can use a generalized likelihood ratio test (GLRT).

The GLRT is a general procedure for composite testing problems. The fundamental idea is to compare the maximum likelihood of the model in class H_1 to the maximum likelihood of the model in class H_0 . The test statistic based on the observation v^N is

$$\hat{\Lambda}(v^N) = \frac{\sup\limits_{\theta_1 \in \Theta_1} L(\theta_1 \mid v^N, H_1)}{\sup\limits_{\theta_0 \in \Theta_0} L(\theta_0 \mid v^N, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \tag{2.20}$$

or equivalently

$$\log \hat{\Lambda}(v^N) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \tag{2.21}$$

It can be supposed in practice that the null hypothesis H_0 is completely known, i.e. θ_0 is fixed, the expression can be written as

$$\hat{\Lambda}(v^N) = \frac{\sup_{\theta_1 \in \Theta_1} L(\theta_1 \mid v^N, H_1)}{L(\theta_0 \mid v^N, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \tag{2.22}$$

If we want to perform a GLRT, we have to solve first with a maximum likelihood estimation (MLE) problem (which will be discussed more clearly in the next subsection) because the test statistics $\log\hat{\Lambda}(v^N)$ cannot be expressed explicitly unless we estimate θ_i $(i = 1, 2)$. Instead, if we find $\hat{\theta}_i$ which maximizes the corresponding likelihood $L(\theta_i \mid v^N, H_i)$, i.e.,

$$\hat{\theta}_i = \underset{\theta_i \in \Theta_i}{\operatorname{argmax}} L(\theta_i \mid v^N, H_i), \tag{2.23}$$

then we may write

$$\hat{\Lambda}(v^N) = \frac{L(\hat{\theta}_1 \mid v^N, H_1)}{L(\hat{\theta}_0 \mid v^N, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \tag{2.24}$$

The quantity $\hat{\theta}_i$ is called the restricted maximum likelihood estimate of θ_i under H_i.

Although the law of $\hat{\Lambda}(v^N)$ is unknown, the following theorem hopefully unveils the method to approximate the threshold given a significance level α.

Wilks's theorem : [98] Let $\theta_0 = (\theta_{0,1}, \theta_{0,2}, ..., \theta_{0,m}) \in \Theta_0 \subset \mathbb{R}^m$ be a vector of parameters of a density family $p(x \mid \theta_0)$ in which $\theta_{0,1}, ..., \theta_{0,l} \in \mathbb{R}$ are free parameters that need to be estimated using MLE, and $\theta_{0,l+1} = t_{l+1}, ..., \theta_{0,m} = t_m$ are fixed at the real values $t_{l+1}, ..., t_m$. Assume that $p(x \mid \theta_1)$ is a density family parametrized by $\theta_1 \in \Theta_1 \subset \mathbb{R}^m$ with θ_1 includes all free parameters. Consider a composite testing problem

$$\begin{aligned} H_0 &: v^N \overset{i.i.d}{\sim} p\left(x \middle| \theta_0\right) \\ H_1 &: v^N \overset{i.i.d}{\sim} p\left(x \middle| \theta_1\right) \end{aligned} \tag{2.25}$$

where the parametric density has the same form in each hypothesis. If the 1^{st} and 2^{nd} order derivatives of $p(x \mid \theta_i)$ w.r.t θ_i exist, then the test statistic

$$\hat{W}(v^N) = \frac{\sup_{\theta_1 \in \Theta_1} p\left(v^N \middle| \theta_1\right)}{\sup_{\theta_0 \in \Theta_0} p\left(v^N \middle| \theta_0\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \tag{2.26}$$

has the following asymptotic distribution when H_0 is true and the sample size $N \to \infty$

$$2\log\hat{W}(v^N) \xrightarrow{d} \chi^2_{m-l}. \tag{2.27}$$

Thus, for large N

$$\alpha = \Pr\left\{ \chi^2_{m-l}(x) \geq 2\log(\lambda) \middle| H_0 \right\}. \tag{2.28}$$

Sometimes it is easy to compute only one of the true parameters $\hat{\theta}_0$ or $\hat{\theta}_1$ but difficult or impossible to compute the other one. This motivates the birth of two other tests that are asymptotically equivalent to the Wilks test. The first one is the Rao test [32] with test statistic:

$$\hat{R}(v^N) = \left(\nabla L(\hat{\theta}_0 \mid v^N, H_0)\right)^T J_N^{-1}(\hat{\theta}_0) \left(\nabla L(\hat{\theta}_0 \mid v^N, H_0)\right), \tag{2.29}$$

where

$$J_N(\theta) = -\nabla^2 L(\theta \mid v^N, H_0), \tag{2.30}$$

is the observed Fisher information matrix for sample size N. The Rao test statistic is asymptotically equivalent to the Wilks test statistic in the order $o_p(1)$ under the same conditions for Wilks's theorem, i.e.,

$$\hat{R}(v^N) = \hat{W}(v^N) + o_p(1). \tag{2.31}$$

The most important property of Rao test is that the test statistic depends only on the MLE for the null hypothesis H_0.

The second one is Wald test [32, 62] with test statistic is

$$\hat{U}(v^N) = \left(g(\hat{\theta}_1)\right)^T \left[\nabla g(\hat{\theta}_1) J_N^{-1}(\hat{\theta}_1) \left(\nabla g(\hat{\theta}_1)\right)^T\right]^{-1} g(\hat{\theta}_1), \tag{2.32}$$

where $g : \mathbb{R}^N \to \mathbb{R}^N$ is a constrain function on the set of θ_0 such that $g(\theta_0) = 0$. Like the Rao test, the Wald test under the same conditions for Wilks test is asymptotically equivalent to the Wilks test in the order $o_p(1)$.

$$\hat{U}(v^N) = \hat{W}(v^N) + o_p(1). \tag{2.33}$$

Contrarily, the important point about Wald test statistic is that it depends only on the MLE for the alternative hypothesis H_1.

All these tests above asymptotically may give us a tool to estimate the probability of type I error α but nevertheless it is still far from being enough to propose a general solution for authentication problems:

- Firstly, most of GLRT fails in obtaining optimality [43, 101].
- The second important point is that even if we accept a arbitrarily large sample size, when α is very small, i.e. when the threshold is far from the means (it can happen for a highly accurate detector), the calculation of the tail probability of chi-squared distribution may be incorrect [33].

– The asymptotic property of the test statistic only occurs in case when the parametric density has the same form in each hypothesis.

– Moreover we need also to compute the asymptotic distribution of the test statistic when the alternative hypothesis happens in order to approximate the probability of type II error β.

This is why another approach proposed to overcome these drawbacks is discussed in the main content of Chapter 3 and Chapter 4.

2.1.3 Parameter estimation

When the receiver observes two GCs how could he certifies that a GC is original or is forged? A proposed solution is to use statistical decision to perform authentication. In our scenario, we assume that the physical process that generate the original GC is known while the other one used by the opponent is different and unknown. From the previous subsection, we know that one way for the receiver to perform authentication is firstly to estimate the generating process of the GC and secondly to use hypothesis testing.

This section presents consequently the theoretical background related to parameter estimation.

Parameter estimation is a branch of statistics in which the parameters, describing the whole underlying physical setting of a population, are supposed to be unknown and need to be estimated based on empirical measured data that are supposed to be outcomes of random variables. Although our proposed analysis does not depend on the estimation method, we consider mostly in this thesis Maximum Likelihood Estimation in order to achieve optimal estimation.

Maximum likelihood estimation

Given a statistical model, Maximum Likelihood Estimation (MLE) relates to a popular class of methods in statistical estimation that use sample of empirical data to estimate the model's parameters [89]. The aim of MLE is to find an estimated parameter $\hat{\theta}$, given the sample x, that maximizes the likelihood function $L(\theta \mid x)$.

We consider a n-length sample $x = (x_1, x_2, ..., x_n)$ of n $i.i.d$ observations coming from a distribution in which its density function $f(x|\theta)$ is supposed to belong to a certain parametric family indexed by the unknown parameters $\theta \in \Theta \subset \mathbb{R}^m$, called the parametric model. The likelihood of the whole sample is defined as the product of the individual likelihoods:

$$\begin{aligned} L(\theta \mid x) &= \prod_{i=1}^{n} L(\theta \mid x_i), \\ &= \prod_{i=1}^{n} f(x_i \mid \theta). \end{aligned} \tag{2.34}$$

Instead of working directly with the likelihood function, we can use the log-likelihood $\mathcal{L}(\theta) = \log L(\theta \mid x)$. Since all individual likelihoods are always positive, the likelihood function and its log version achieve the maxima at the same point

$$\hat{\theta}_{ML} = \underset{\theta \in \Theta}{\operatorname{argmax}}\, \mathcal{L}(\theta), \tag{2.35}$$

which is the solution for a MLE and called maximum likelihood estimator (ML estimator) for θ.

Let us denote $H_m(\theta)$ the Hessian matrix of the log-likelihood with respect to the parameters:

$$H_m(\theta) = \left(\frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta_i \partial \theta_j} \right)_{i,j=1,..,m}, \tag{2.36}$$

and let $I_m(\theta)$ be the Fisher information matrix [40] defined as:

$$\begin{aligned} I_m(\theta) &= \mathbb{E}\left[\left(\frac{\partial \mathcal{L}(\theta)}{\partial \theta} \right)^2 \middle| \theta \right] \\ &= -\mathbb{E}_{f(x|\theta)} \left[H_m(\theta) \right], \end{aligned} \tag{2.37}$$

where $\mathbb{E}_{f(x|\theta)}$ or simply $\mathbb{E}_\theta$ is the expectation taken w.r.t $f(x \mid \theta)$.

From the *Cramer-Rao theorem* [88] for a large sample size, it can be stated that the covariance matrix of any unbiased ML estimator $\hat{\theta}$ of a parameter θ_0 satisfies:

$$\operatorname{Cov}(\hat{\theta}) \simeq I_m^{-1}(\theta_0). \tag{2.38}$$

Since $\hat{\theta}_{ML}$ is an unbiased estimator, we have:

$$\operatorname{Cov}(\hat{\theta}_{ML}) \simeq I_m^{-1}(\theta_0), \tag{2.39}$$

i.e., the covariance of a unbiased ML estimator can be approximated by the inverse of the Fisher information matrix at the true parameters θ_0, and $\hat{\theta}_{ML}$ is an estimator yielding the smallest variance. The asymptotic distribution of $\hat{\theta}_{ML}$ is then given by:

$$\hat{\theta}_{ML} \overset{\text{asym}}{\sim} \mathcal{N}(\theta, I_m^{-1}(\theta_0)). \tag{2.40}$$

The normality of $\hat{\theta}_{ML}$ helps us to provide a measure of how the estimated parameters spread w.r.t the true value. The quadratic form of the error (the variation of the estimation) is chi-squared distributed:

$$\rho(\hat{\theta}_{ML}) = \left(\hat{\theta}_{ML} - \theta_0 \right)^T \operatorname{Cov}^{-1}(\hat{\theta}_{ML}) \left(\hat{\theta}_{ML} - \theta_0 \right) \overset{\text{asym}}{\sim} \chi^2_\kappa, \tag{2.41}$$

or equivalently

$$\rho(\hat{\theta}_{ML}) = \left(\hat{\theta}_{ML} - \theta_0 \right)^T I_m(\theta) \left(\hat{\theta}_{ML} - \theta_0 \right) \overset{\text{asym}}{\sim} \chi^2_\kappa, \tag{2.42}$$

where χ^2_κ is the chi-squared distribution with κ degree of freedom. Here, κ is the number of free parameters that govern the proposed model. One may observe that $\rho(\hat{\theta}_{ML}) = cte$ is an ellipsoid in the κ-dimensional space. In practice, the true parameters are always unknown, hence the Fisher information need to be estimated, for example by the observed Fisher information matrix [40, 65] at $\theta = \hat{\theta}_{ML}$

$$\begin{aligned} J_m(\hat{\theta}_{ML}) &= -H_m(\hat{\theta}_{ML}) \\ &= -\left(\frac{\partial^2 \mathcal{L}(\theta)}{\partial\theta_i \partial\theta_j}\right)\Big|_{\theta=\hat{\theta}_{ML}}. \end{aligned} \tag{2.43}$$

Another important property of MLE is the invariant property, i.e. if g is a continuous function w.r.t θ, then $\eta_{ML} = g(\hat{\theta}_{ML})$ is a ML estimator of $\eta = g(\theta)$. Thenceforth, rather than estimating directly a parameter θ, we can first estimate some function $g(\theta)$ using MLE and then recover an estimate of θ from $g(\theta)$.

Expectation Maximization

The Expectation Maximization (EM) algorithm [34, 69, 22] is an iterative method for finding maximum likelihood in cases where the equations in a MLE problem cannot be solved directly. Suppose that we have a given statistical model with a set of observation x^N, a set of unobserved latent or missing data u^N, and a vector of unknown parameters θ. The EM algorithm is used to find a ML estimator for θ by iteratively applying two steps:

1. The Expectation step (E step) computes the expected value of the log likelihood function w.r.t the conditional distribution of U^N given X^N and the estimate $\theta^{(t)}$ of the parameters θ at iteration (t):

$$Q(\theta|\,\theta^{(t)}) = \mathbb{E}_{U^N|X^N,\theta^{(t)}}\left[\mathcal{L}\left(\theta|\,x^N, u^N\right)\right]. \tag{2.44}$$

2. The Maximization steps (M step) finds the estimated parameters at iteration $(t+1)$ that maximizes this quantity:

$$\theta^{(t+1)} = \underset{\theta}{\operatorname{argmax}}\, Q(\theta|\,\theta^{(t)}). \tag{2.45}$$

The algorithm repeats these two steps and assigns, at each iteration T, the estimated $\hat{\theta} = \theta^{(T)}$ of θ until convergence.

In some certain cases, it is more convenient to express EM algorithm under an alternative form. Let us denote

$$\begin{aligned} F(q,\theta) &= \mathbb{E}_g\left[\mathcal{L}\left(\theta|\,x^N, u^N\right)\right] + H(q), \\ &= -D_{KL}\left(q\|\,p_{U^N|X^N}\left(u^N\big|\,x^N,\theta\right)\right) + \mathcal{L}\left(\theta|\,x^N\right), \end{aligned} \tag{2.46}$$

where $p_{U^N|X^N}\left(u^N\big|\,x^N,\theta\right)$ is the conditional distribution of the unobserved or missing data u^N given the observation x^N; q is an arbitrary probability density over $u^N$and H

is the entropy function of q. Then the EM algorithm can be reformulated by the two following steps:

1. Expectation step: select q satisfying

$$q^{(t)} = \underset{q}{\operatorname{argmax}}\, F(q, \theta^{(t)}). \tag{2.47}$$

2. Maximization step: choose θ satisfying

$$\theta^{(t+1)} = \underset{\theta}{\operatorname{argmax}}\, F(q^{(t)}, \theta). \tag{2.48}$$

It should be noted that an EM algorithm only yields to a local solution for the estimation, so that it requires an initialization value that is close enough to the true model's parameters to run the algorithm and to insure the convergence of the algorithm.

The EM algorithm has very large applications in many research directions. It can be used for data clustering in data mining [27, 57, 71, 78], in signal processing [45], in computer vision [72] or even psychology and social researches [14, 24], etc. One of the most important application of EM algorithm is to fit the mixtures of distributions by using what we might call "pseudo missing data" [50, 70, 73], i.e. the data that we never obtain but they can be considered as missing in order to facilitate the computation of ML estimators.

In chapter 4, we propose a modification of the EM algorithm for a mixture of truncated Gaussian distributions. This algorithm is used to estimated the parameters of the opponent channel.

2.2 Previous works related to authentication of GC

In this section, we present the connections between this thesis and previous works that:

- have the same general goals of authenticating items,
- use directly Graphical Codes to perform authentication,
- use the same methodology, i.e. hypothesis testing, for security related applications,
- are related with modeling the print and scan channel.

2.2.1 Overview of authentication processes

Authentication processes are used in a lot of different fields, often related to computer sciences, such as:

- Access control on a network/intelligent systems to check the authority of the right users [87, 46].
- Password protection based on RSA system [35, 41], considered as a digital signature. It is used, for example, to access to bank accounts.
- Authentication between objects based on the interaction between all objects which are belongs to the same user [39, 47].
- Digital image authentication based on watermarking. Note that "image authentication" in this case is related to the problem of integrity check and not to authentication per se since a digital plain copy is always authentic. One examples for this approach can be found, for instance in Ho *et al* [55].
- Authentication of physical materials where the goal is to distinguish genuine products from forged ones. This approach is extremely active nowadays due to the explosion of counterfeiting industry.

The following section proposes more examples of this last application.

2.2.2 Authentication of physical products

As introduced briefly in Chapter 1, the authentication of physical products can generally be obtained by using the stochastic structure of either the materials that compose the product or the stochasticity of a physical process that is associated to the generation of the product.

Authentication of physical objects can be for example performed by recording the random patterns of the fiber of a paper: in [51], the authors combine the optical detection method with recorded digital signatures based on public key codes in order to protect not only banknotes but also credit- and chip-cards, checks, contracts, etc., against counterfeiting. The general idea, depicted in Fig. 2.5, is separated into two parts:

1. Protection: the image is detected then compressed and a digital signature is attached. The result is encoded.
2. Verification: the fiber structure of the banknote to be verified is extracted and then compared with the fiber structure of the original object stored in a available database.

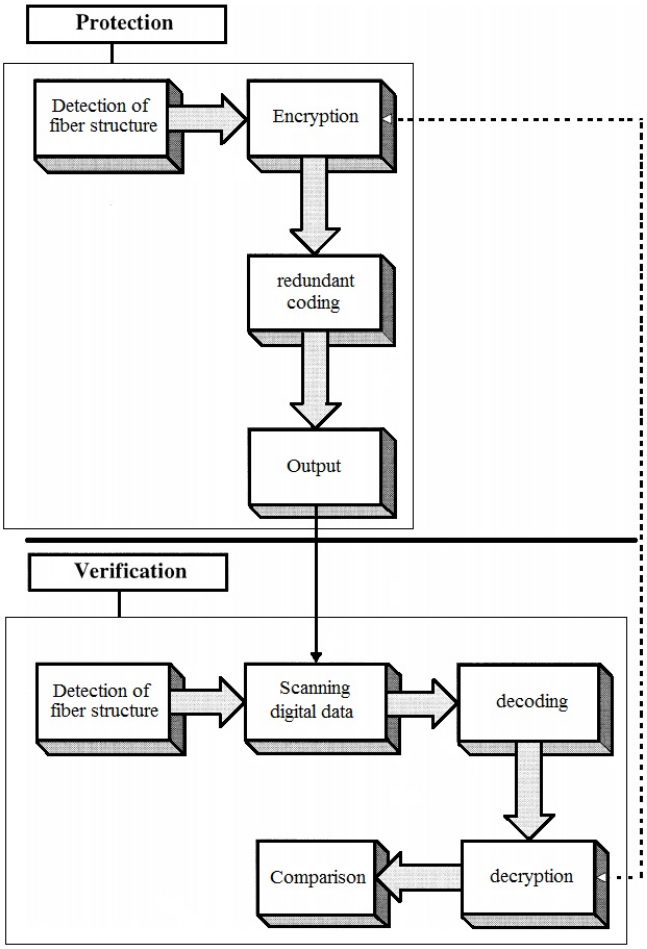


Figure 2.5: Protection and Verification of banknote (taken from [51])

The authors claim that they can obtain a high security without using any expensive production techniques due to the fact that the complicated features, used for verification, are already included in the object. Instead of using random features of objects, they can also detect, code and store the characteristics such as fingerprints, speech, faces... of those who are authorized to access the objects. However, such a system is practically heavy to deploy since each product needs to be linked to its high definition capture stored in a database.

Another solution is to rely on the degradation induced by the interaction between the product and a physical process such as printing, marking, embossing, carving ...

Because of both the defaults of the physical process and the stochastic nature of the matter, this interaction can be considered as a Physically Unclonable Function (PUF) [94] that cannot be reproduced by the forger and can consequently be used to perform authentication. According to [94], rigorously a PUF is defined as *a "random"*

assignment that maps a set of challenges to a set of responses based on an intractably complex physical system. The authors indicate that PUF can be originally described as a innovative circuit primitives that enable significantly higher physical security with low-cost authentication of individual integrated circuits (ICs) by deriving secrets from complex physical characteristics of ICs rather than storing them in non-volatile digital memory.

According to Fig. 2.6, an authentic device A includes the pairs Challenges-Response stored in the database of A for future authentication operations. To check the authenticity of an unknown device, firstly a challenge that had been recorded but has never been used is selected, and the corresponding response is obtained by PUF. This response is then compared with the one already stored in the database of A for authentication.

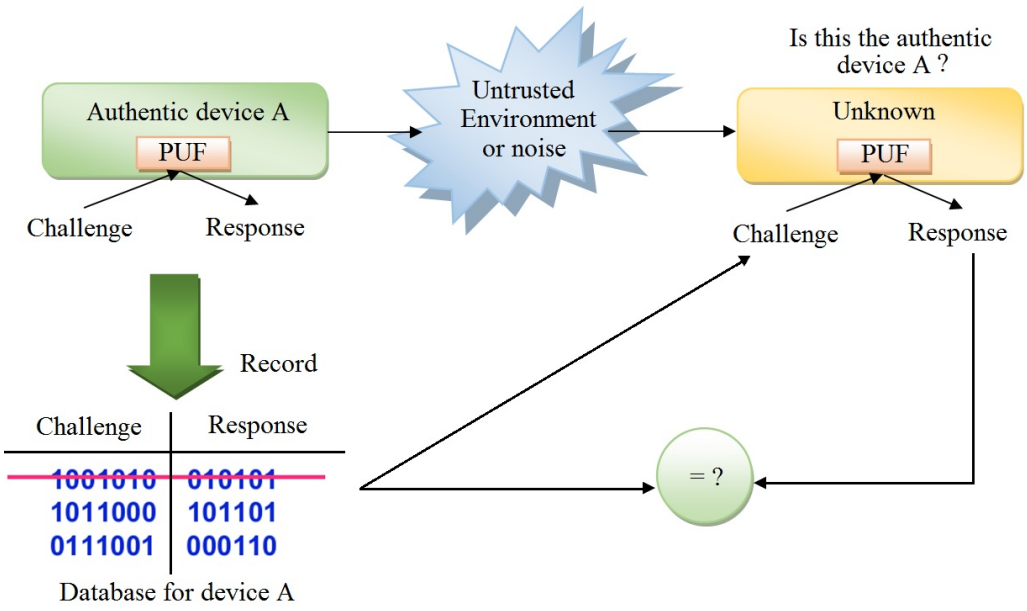


Figure 2.6: Description of PUF-based authentication (taken from [94])

Connections with authentication using GC

It is important to note that the Integrated Circuits presented above act in a similar way than the features characterizing the physical printing process in this thesis. The particularity of our system is the fact that this mapping may be understood as a "one time PUF" or a physical unclonable process because it can be called only once for each object. There are no challenge-response processes in this case but the authentication still relies on a physical unclonable process.

A similar example to our problem can be found in [91]. The authors propose another marking technique, called Laser-written PUF or simply LPUF, in order to characterize the random profiles of laser marks on materials such as metals. From a technological

achievement of TOMO3D project [15], LPUF is generated based on 3-D profile of laser marks using a diameter of 60μm and carved on the surface of a physical object. They also provide an anti-counterfeiting system based on LPUF and use them for authentication. This scheme is split in two parts:

1. The registration is executed once on the object before it is released to the market.
2. The verification can be executed whenever someone wants to check whether the product is a genuine one or not.

Their approach can be used to protect small objects or to be linked with other authentication methods.

The unclonable process can also be represented solely by the interaction between paper and ink. In [49], the authors measure the degradation of the inks within printed color-tiles, and use the discrepancy between the statistics of the authentic and print-and-scan tiles to perform authentication. They try to estimate whether a printed impediment will authenticate without imploring the actual authentication process. Several algorithms are also proposed to predict the result of the authentication process. The used AOC (area over the curve) statistic shows that there are two metrics, called *no-reference metric* and *full-reference metric* respectively, which are particularly useful for estimating authentication performance in the existence of distortions caused via different choices of print-and-scan systems. Surprisingly, by using AOC, they show that a *no-reference metric* gives the best performance for authentication.

Another effective authentication system has been proposed by Picard et al. [82, 83] and uses 2D pseudo random binary codes that are printed at the native resolution of the printer (2400 dpi on a standard offset printer or 812 dpi on a digital HP Indigo printer). It is important to notice that this system is very similar to the one that will be studied in the next chapter.

At the authentication step, in order to perform authentication the receiver computes a test on the observed scanned code, being either the scanned version of the original printed code or the scanned version of the reprinted forged code.

One advantage of this system over previously cited ones is that it is easy to deploy since the authentication process needs only a scan of the graphical code under scrutiny and the seed used to generate the original one: no fingerprint database is required in this case.

Security analyses

The security of this system relies on the fact that the opponent is not able to accurately estimate the original binary code due to solely relying on the use of a PUF. Different security analysis have already been performed w.r.t. this authentication system, or to very similar ones.

The authors have studied in [18] the impact of multiple printed observations of the same graphical codes and the authors have shown that the influence of the noise due to the printing process can be reduced in this particular setup, but not completely removed. Here, the authors consider a batch attack where the product manufacturer (Bob) generates a batch of printed GCs coming from the same original code. Eve can access a number N_c of printed versions of a genuine 2D-GC and tries to estimate the original one. The goal of Eve is to convince the receiver that her reproduced GC is a genuine one. This paper shows also that the original code cannot be totally removed even with a large number of observations, which leaves room for secure authentication even under batch attacks.

In [38], a print and scan model is proposed to be adapted to graphical code and a smart attack based on the model is analyzed to show that it can be used to corrupt the effectiveness of authentication (precisely authentication using GC). To handle this type of attack, the authors propose four new detection metrics which are sensitive to print-scan distortions. Through experimental analysis, they show that their proposed features can be employed to improve significantly the authentication accuracy.

In [37], the authors propose to study the security of authentication considering GCs by using machine learning techniques in order to extract the original code from an observation of the printed code. Their results show that the estimation accuracy can be improved without recovering perfectly the original code. A "black box" strategy is employed to analyze the security. They propose to use a set of observations and try to invert the printing system by inferring a linear classifier based on these observations.

In [19], the considered security analysis is quite similar to the setup of passive fingerprinting using binary fingerprints, which is similar to binary GCs, under informed attacks (the channel between the original code and the copied code is assumed to be a BSC). In this case, the security is shown to increase w.r.t the code length and the authors propose a practical threshold when type I error (original detected as a forgery) and type II error (forgery detected as an original) are equal. A information-theoretic analysis is also derived based on the assumption of the code length.

2.2.3 Hypothesis testing in authentication and forensics

A lot of research papers related to hypothesis testing and information security have been written over last several decades. In this thesis, we deal with authentication considering hypothesis testing technique for detecting the genuineness of products.

Although hypothesis testing is a classical methodology, it is still very effective in many application researches of authentication and forensic. For instance, in [96], the fundamental idea is also the use of hypothesis testing, albeit the authors deal with the identification of camera models. They are successful to design a camera model considering only two camera parameters (a, b). They then develop an estimation of these parameters based on the weighted least squares estimation. A binary statistical

test based on GLRT is developed to analyze the performance of identification problem using heteroscedastic noise which is stated to describe more precisely the acquisition noise of a natural raw image. Numerical experiments are carried out based on both simulated and real images taken from *Nikon D70* and *Nikon D200* show that

- The performance of proposed tests depends on the discriminability of camera parameters (a, b).
- Only a small number of pixels is required to achieve a perfect detection performance which proves the sharpness of the tests.

The authors claim that their proposed method is the only one which employs raw images to identify camera model. However, because the main limitation is that raw images may not be available in practice, they consider to extend their approach to other image formats that are related to the post-acquisition and compression processes.

Hypothesis testing is also used in steganalysis, a branch in computer sciences used to detect hidden information in the cover media such as image, audio, video, etc,.. using steganography. In [102], the authors propose to use the classical binary hypothesis testing and show how it is useful for detecting hidden information. Both simple and composite hypothesis testing schemes based on likelihood ratio tests are used to analyze the performance of hidden information detection. Their approach is strictly based on the parametric statistical model of the media object. Both theoretical and numerical results show not only the impact of observation quantization on the probabilities of type I and type II error but also the benefits of using statistical decision on hidden information detection.

In [68], the authors propose to interpret message authentication as a hypothesis testing problem coming from an information-theoretic point of view based on the concept of discrimination whose expression is related to mutual information in channel coding theory [23]. They provide a generalized scheme to evaluate the information-theoretic lower bounds on an opponent's probability of fooling the receiver by forging one of the messages in the sequence shared between sender and receiver. Two types of cheating, impersonation and substitution attacks, are analyzed and lower bounds on cheating probability are also obtained for any authentication system.

2.2.4 Printing-scanning models

Because the principle of our authentication system relies on the degradation induced by the print and scan process, we draw here an overview of the different works in this domain.

We recall first a list of the most important printing techniques (see Fig. 2.7 for examples of each type)

- Offset printing in which the inked image is spread on a metal plate then transferred to a surface of rubber blanket and finally pressed to the paper (see in [61]).
- Laser printing [61] in which the printer uses a laser beam to carve an image on a charged drum. The drum is then rolled through a reservoir of ink. The ink is then transferred to the paper using a combination of heat and pressure.
- Inkjet printing in which the printer spray droplets of ionized ink on a paper. A inkjet printer can print at a moderately high resolution (300 dots per inch or more). Because of both its outstanding properties and low cost, inkjet printing nowadays is used widely as a printing tool [93].

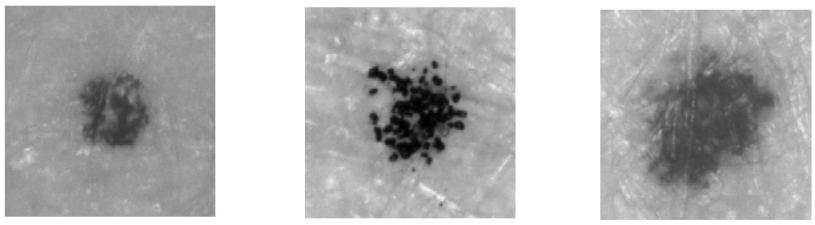


Figure 2.7: From left to right: offset, laser and inkjet printing dots on uncoated papers.

As a matter of fact, the printed images are always degraded by the sequence of printing, scanning, copying,... Even if the humans eyes cannot distinguish the difference between the printed images and the digital, the quality of printed images can be evaluated using text recognition systems [16] because the microscopic view contains more details that can be used to measure the quality of the printed image. Many methods are consequently based on observations from a microscope to estimate the parameters of the printer.

From the aspect of modeling, in [79] the authors propose to use a probabilistic model to generate the images having the toner location (the toner is a special ink used in laser printing technology) that is similar with the one of actual printed images. Using a so called "geometric probability" described by measuring the area filled by toner and its complementary area, they can develop a deterministic model to describe the average amount of paper to be covered by toner. Simulations show that the results of their model fit well with the average coverage of printed images.

Another feature which needs to be modeled is the ink spreading, a phenomenon of dot gain that produces high color deviations in ink jet printing. In [42], the authors propose a new model of ink spreading by extending the drop impact w.r.t the shape of its neighbors and the condition of the surface to improve the prediction of the reflection spectra of halftoned samples printed on various inkjet printers. By using Pólya's

counting theory in combinatorics, they can reduce the number of all possible ink drop configurations to a significantly small number of cases. For instance, in a three-ink-color printing, they show that only 30 important cases must be considered instead of $3 \times 4^6 = 12288$ cases.

Printer modeling can be also carried out using purely the techniques from signal processing [97] in which a signal processing model is proposed to model multi-level halftoning and resolution enhancement, as well as traditional halftoning.

Recently, in [76, 77], using statistical signal processing techniques, the authors provide a model for the scanning and printing process through a binary response scheme based on the shape and the location of the ink dots. They also develop an algorithm called maximum likelihood unsupervised identification to show the accuracy of printing process at the microscopic scale. The algorithm's performance is evaluated through simulation using the true data collected by microscope of various types of papers and printing modes.

2.3 Conclusions of Chapter 2

This chapter has presented elements of the theoretical background needed in order to conduct this thesis. Hypothesis testing will be used in the next chapter to derive authentication test, and its combination with parameter estimation in chapter 4 to evaluate the impact of estimation on authentication.

The related works also show that this thesis is connected to multiple domains: authentication schemes, physical unclonable functions, forensics and printing modeling. The different security analyses presented here are also connected to the one presented in chapter 5.

Chapter 3

Authentication using hypothesis testing

"*Essentially, all models are wrong, but some are useful.* "

George E. P. Box

The goal of this chapter is first to present an authentication system for GC that relies on hypothesis testing, and then to provide accurate computations of the error probabilities of this system. The authentication system is defined in the first section of this chapter, then the print and scan model is presented. We afterwards present two possible strategies for the receivers, which consist in thresholding or not the observed code before applying the hypothesis test and we show that the authentication is more performant without thresholding. Finally we end this chapter by presenting reliable computations of the errors probabilities for this setup. These computations are based

on an asymptotic expression and they will be further extended in the next chapter to take into account the estimation of the opponent channel.

3.1 The authentication system

As stated in Chapter 1 with the general framework described in Figure 1.5, we focus here on the authentication aspect of the GC , so that the principle of the studied system in this thesis can be depicted more accurately by Figure 3.1 .

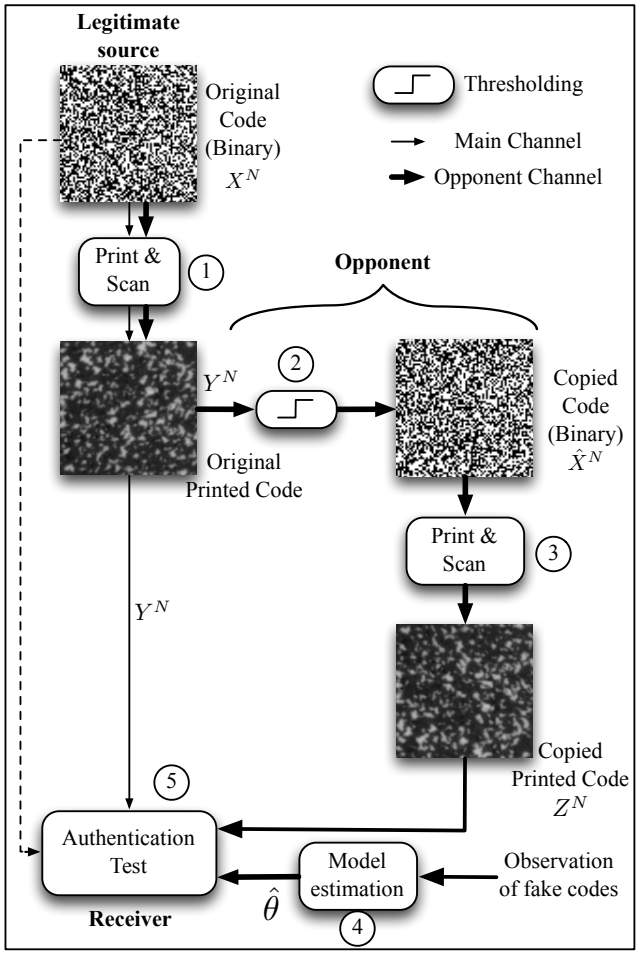


Figure 3.1: Principle of authentication using graphical codes.

Based on Fig. 3.1, let us introduce our setup for authentication. Our authentication system is based on printed graphical 2D codes using very high resolution printers (2400dpi). Each printed and scanned set of dots (a dot being a binary element) suffers from a stochastic non-invertible noise which makes the reproduction of the original GC impossible (see Fig. 3.2 for an example of real the GCs after printing and reprinting).

The opponent's goal is then to reproduce a printed and scanned code similar to the original printed one using a printer that will also generate a non-invertible noise.

In the whole of thesis the authentication model involves two channels $\mathcal{X} \to (\mathcal{Y}, \mathcal{Z})$, we define the main channel as the channel between the original source and the receiver, while the opponent channel as the channel between the original source and the receiver but passing through the counterfeiter (or opponent) channel (see in Fig. 3.1).

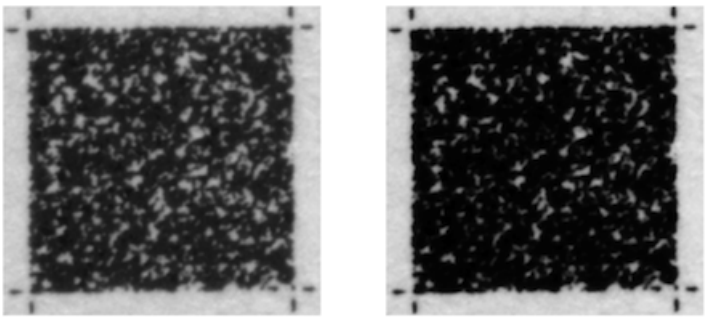


Figure 3.2: Left: an original printed and scanned GC. Right: a re-printed and scanned (forged) GC.

The authentication system works as follow: A binary graphical code can be considered as an authentication sequence x^N chosen randomly from the message set $\mathcal{X}^N$ ($\mathcal{X}^N$ mostly is $\{0,1\}^N$) by the legitimate sender and shared secretly with the legitimate receiver. In our authentication model, x^N is published as a noisy version y^N, taking values in the set of points $\mathcal{V}^N$, modeling the original printed and scanned graphical code (see Fig. 3.2 on the left). An opponent may observe y^N and, naturally, tries to retrieve the original authentication sequence using his skills in data estimation. After his processing, he obtains an estimated sequence $\hat{x}^N$ which is supposed practically to be different from the original x^N (see (2) in Fig. 3.1). He then prints it using his printing process to create a forged observable noisy image z^N taking values in the same set of points $\mathcal{V}^N$. He publishes z^N hoping that it will be accepted by the receiver as coming from the original source (see Fig. 3.2 on the right). The observed images y^N and z^N are 8 bits grey level images. In practice, this attack will be used to create false documents or fake packages that could be considered as authentic.

The whole physical process, precisely printing and scanning devices used by the legitimate parts (see (1) in Fig. 3.1) and by the counterfeiter (see (3) in Fig. 3.1), are respectively modeled by probability distributions conditioned to the original data $P_{Y|X,\theta}$ and $P_{Z|X,\bar{\theta}}$ in which θ and $\bar{\theta}$ are set of parameters, taken in Θ, specifying the devices in each case.

When observing a sequence v^N, which may be one of the two possible sequences y^N or z^N, the detector has to determine whether this observed sequence comes from the legitimate source or not (see (5) in Fig. 3.1) supposed that the models $P_{Y|X,\theta}$ is known.

The print and scan process in this particular setup for example can be modeled by an AWGN channel or an additive i.i.d. lognormal noise as in [17].

As mentioned before, our authentication is based on NP-test in which the receiver considers two hypothesis H_0 and H_1. The former hypothesis attests authenticity, i.e. that the received sequence is generated by $P_{Y|X,\theta}$ and the latter one unveils a fake code, i.e. that the observed sequence is driven from $P_{Z|X,\bar{\theta}}$. Performances are evaluated via computing accurately the probability of type I error and the probability of type II error.

3.2 Proposed models for print and scan

The two channels $\mathcal{X} \to (\mathcal{Y}, \mathcal{Z})$ are considered being discrete and memoryless with conditional probability distribution $P_{Y\,Z|X,\theta,\bar{\theta}}(y\,z|\,x,\theta,\bar{\theta})$. The marginal main and opponent channels $P_{Y|X,\theta}(y|\,x,\theta)$ and $P_{Z|X,\bar{\theta}}(z|\,x,\bar{\theta})$ are defined by the transition probability matrices of the main channel and the opponent channel, respectively.

Let $T_{V|X,\theta}$ be the generic transition matrix modeling the printing and scanning devices. The entries of this matrix are conditional probabilities $T_{V|X,\theta}(v|\,x,\theta)$ or simply $T_{V|X}(v|\,x)$ relating an input alphabet $\mathcal{X}$ and the output alphabet $\mathcal{V}$. In practice, $\mathcal{X}$ is a digital value, i.e., a binary alphabet standing for black bit (0) and white bit (1), and the channel output set $\mathcal{V}$ stands for the set of grey level values in the set $\{0, 1, ..., 255\}$.

The marginal distribution of the main channel $P_{Y|X,\theta}$ is equivalent to one print and scan process, represents a grey level distribution of the authentic image conditioned to the knowledge of both the authentication dots and the parameters governing the legitimate print and scan process. On the other hand, $P_{Z|X,\bar{\theta}}$ depends on the opponent processing while he tries to recover the original sequence X^N by the estimated sequence $\hat{X}^N$ before reprinting it, hoping that $\Pr(\hat{X}^N \neq X^N) = \varepsilon$ with $\varepsilon > 0$ is arbitrarily small.

Practically, when performing a detection to obtain an estimated sequence $\hat{X}^N$, the opponent always undergoes errors coming from the realistic fact that he is not able to infer the original code. It is important to note that the opponent will have to print a binary version of its observation because an industrial printer at this very high resolution can only print binary images. These errors yield to the probability $P_{e,W}$ for the confusion between an original white dot with a black one and to the probability$P_{e,B}$ for the confusion between an original black dot with a white one. This distinction is due to the fact that the distribution $T_{V|X,\theta}$ of the physical devices is arbitrary and not necessarily symmetric. Let $\mathcal{D}_{\mathcal{W}}$ be the optimal decision region for guessed white dots using thresholding:

$$\mathcal{D}_{\mathcal{W}} = \left\{ v \in \mathcal{V} :\, P_{Y|X,\theta}(v|\,x=1,\theta) > P_{Y|X,\theta}(v|\,x=0,\theta) \right\}, \tag{3.1}$$

and $\mathcal{D}_{\mathcal{W}}^{c}$ is the complementary region in the set $\mathcal{V}$. Error probabilities $P_{e,B}$ and $P_{e,W}$ are then equal to

$$P_{e,B} = \sum_{v \in \mathcal{D}_{\mathcal{W}}} P_{Y|X,\theta}(v|\, x = 0, \theta), \tag{3.2}$$

and

$$P_{e,W} = \sum_{v \in \mathcal{D}^c_{\mathcal{W}}} P_{Y|X,\theta}(v|\, x = 1, \theta). \tag{3.3}$$

The channel $X^N \to \hat{X}^N$ can be modeled as a binary input binary output (BIBO) channel with transition probability matrix $P_{\hat{X}|X}$:

$$\begin{bmatrix} P_{\hat{X}|X}(\hat{x} = 0|\, x = 0) & P_{\hat{X}|X}(\hat{x} = 1|\, x = 0) \\ P_{\hat{X}|X}(\hat{x} = 0|\, x = 1) & P_{\hat{X}|X}(\hat{x} = 1|\, x = 1) \end{bmatrix} = \begin{bmatrix} 1 - P_{e,B} & P_{e,B} \\ P_{e,W} & 1 - P_{e,W} \end{bmatrix} \tag{3.4}$$

As depicted in Fig. 3.1, because the opponent channel $X^N \to Z^N$ is a physically degraded version of the main channel, $X^N \to \hat{X}^N \to Z^N$ forms a Markov chain with the relation $P_{\hat{X}\,Z|X}(\hat{x}\, z|\, x) = P_{\hat{X}|X}(\hat{x}|\, x) T_{Z|\hat{X}}(z|\, \hat{x})$ where $T_{Z|\hat{X}} \equiv T_{Z|\hat{X},\bar{\theta}}$ is the transition matrix of the counterfeiter physical device. Therefore, given $P_{Z|X,\bar{\theta}}$ a grey level distribution of the forged image conditioned to the knowledge of both the authentication dots and the parameters governing opponent print and scan process, we have:

$$\begin{aligned} &P_{Z|X,\bar{\theta}}(\, Z = v|\, X = 0, \bar{\theta}) \\ &\qquad = (1 - P_{e,B}) P_{Z|\hat{X},\bar{\theta}}(v|\, \hat{X} = 0, \bar{\theta}) \\ &\qquad + P_{e,B} P_{Z|\hat{X},\bar{\theta}}(v|\, \hat{X} = 1, \bar{\theta}), \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} &P_{Z|X,\bar{\theta}}(\, Z = v|\, X = 1, \bar{\theta}) \\ &\qquad = (1 - P_{e,W}) P_{Z|\hat{X},\bar{\theta}}(v|\, \hat{X} = 1, \bar{\theta}) \\ &\qquad + P_{e,W} P_{Z|\hat{X},\bar{\theta}}(v|\, \hat{X} = 0, \bar{\theta}). \end{aligned} \tag{3.6}$$

Without loss of generality, a generalized symmetric exponential family (or generalized Gaussian distributions) can be used to model the physical device in our analysis, i.e., the association of a printer with a scanner, used by the legitimate source $T_{Y|X}(v|\, x)$ and by the opponent $T_{Z|\hat{X}}(v|\, \hat{x})$ which may be expressed as follow:

$$p(v|\, x) = \frac{b}{2a\Gamma(1/b)} e^{-\left(\frac{|v - \mu(x)|}{a}\right)^b} \tag{3.7}$$

where $\mu(x)$ is the mean and the parameter a can be derived from the variance $\sigma^2 = \mathrm{Var}(V)$ by using below formula

$$a = \sqrt{\sigma \Gamma(1/b) \Gamma(3/b)}. \tag{3.8}$$

The parameter b is used to control the sparsity of the distribution, for example, when $b = 1$ the distribution is Laplacian, $b = 2$ the distribution is Gaussian, and $b \to +\infty$ the distribution is uniform. The resulting distributions $P_{Y|X,\theta}$ and $P_{Z|X,\bar{\theta}}$ are first discretized then truncated to provide values within the finite set $[0, 1, ..., 255]$ to model a reasonable scanning process. Each channel is defined by four parameters, two per each type of dots, $\mu_b = \mu(0)$ and σ_b for black dots and $\mu_w = \mu(1)$ and σ_w for white dots.

Fig. 3.3 illustrates the different types of GCs using generalized Gaussian distributions on the main and the opponent channels of same mean and variance and with $b = 1$, $b = 2$ and $b = 6$ (a distribution which is close to uniform).

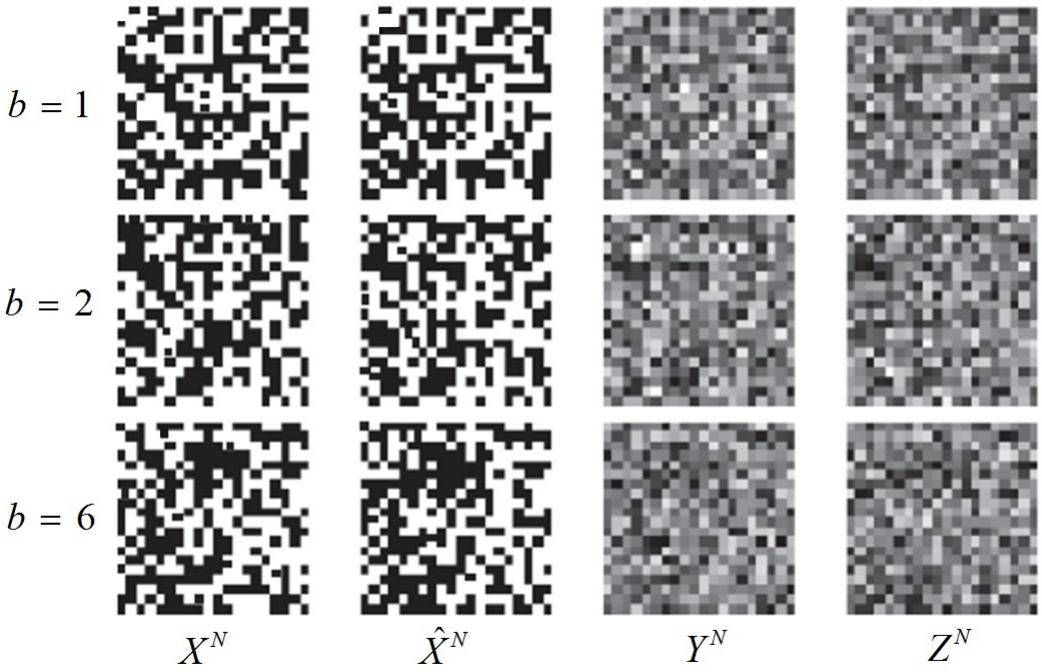


Figure 3.3: Examples of 20×20 code which are generated (X^N) and printed (Y^N) by the main channel, then estimated ($\hat{X^N}$) and reprinted (Z^N) by an opponent using generalized Gaussian distributions in case $b = 1$, $b = 2$ and $b = 6$. Main and opponent channels are identical, $\mu_b = 50$, $\mu_w = 150$, $\sigma_b = 42$ and $\sigma_w = 42$.

Another choice to model the print and scan channel, mentioned in [17], is the use of Lognormal distribution:

$$p(v \mid x) = \frac{1}{vs(x)\sqrt{2\pi}} e^{-\frac{(\log v - \mu(x))^2}{2s^2(x)}} \tag{3.9}$$

with the mode of the distribution is defined as $M = e^{\mu(x)-s^2(x)}$, and the variance is

given by $\sigma^2 = \left(e^{s^2(x)} - 1\right) e^{2\mu(x)+s^2(x)}$. In our case, the Lognormal distribution can be parametrized by the standard deviations σ_b, σ_w and the modes M_b, M_w respectively for black and white dots. Fig. 3.4 depicts truncated Lognormal distributions having same modes but different standard deviations.

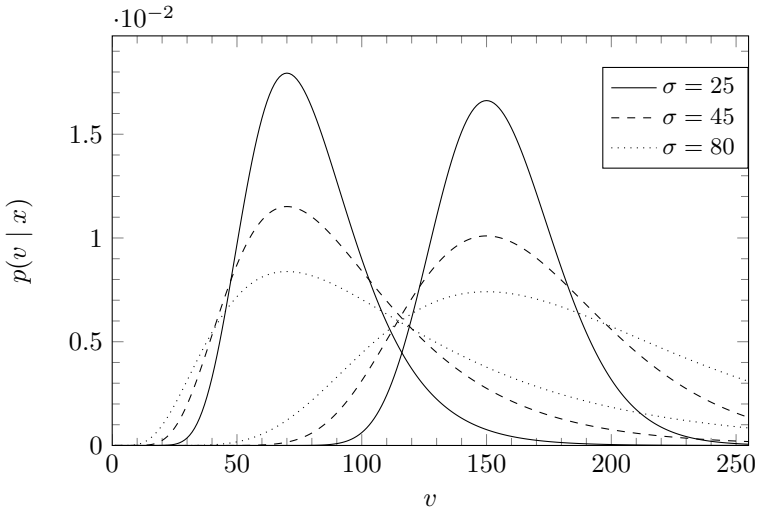


Figure 3.4: Representation of the print and scan model for the black dots (on the left) and the white parts of the paper (on the right) for different standard deviations $\sigma_b = \sigma_w = \sigma$ with $M_b = 70$ and $M_w = 150$ for the Lognormal distribution.

Note that other print and scan models that deal with the gamma transfer function or additive noise with input dependent variance can be found in [64], but the general methodology considered in this thesis is not dependent on the model and can still be applied.

3.3 Receiver strategies

In this subsection we introduce the testing strategies to check whether, for a given fixed codeword $(x_1, ..., x_N)$ in $\{0,1\}^N$, an observed $i.i.d$ sequence $(o_1, ..., o_N | x_1, ..., x_N)$ (with $(o_i | x_i)$ belonging to a discrete finite set $\mathcal{V}$) is generated from a given distribution $P_{Y|X,\theta}$ of the main channel or if it comes from an alternative hypothesis associated to distribution $P_{Z|X,\bar{\theta}}$ of the opponent channel. Generally, we are interested in performing authentication after observing a sequence of N samples $(o_i | x_i)$ checking whether this sequence comes from a original source or from a counterfeiter. Similar to the example presented in subsection (2.1.1), the strategy of receiver is to establish a decision based on binary testing problems between two hypothesis H_0 and H_1 corresponding respectively to each of the former cases. As a matter of fact, the sample space $\mathcal{V}^N$ will be partitioned into two regions $\mathcal{H}_0$ and $\mathcal{H}_1$ and it leads to two kinds of errors as being introduced from the previous chapter: type I error with occurred probability α and type II error with

occurred probability β. According to Neyman-Pearson theorem, under the constraint $\alpha \leq \alpha^*$, β is minimized if and only if the following log-likelihood test deduces the choice of H_1:

$$L = \log \frac{P_{Z^N|X^N,\bar{\theta}}\left(o^N \middle| x^N, H_1\right)}{P_{Y^N|X^N,\theta}\left(o^N \middle| x^N, H_0\right)} \geq \lambda \tag{3.10}$$

where λ is a threshold verifying the constraint $\alpha \leq \alpha^*$.

However, the test statistic (3.10) is just a general expression while in practice, the receiver does not know exactly the true parameters $\bar{\theta}$ related to the print and scan process of the opponent or even the distribution $P_{Z|X,\bar{\theta}}$ of this process. Therefore, the receiver who want to perform authentication can use two possible strategies:

- Firstly, it is assumed that he does not know anything about $P_{Z|X,\bar{\theta}}$, in this case the receiver can use a threshold to count the number of errors between y^N and x^N or z^N andx^N respectively. He can after build a test based on the distributions of the number of errors in y^N and z^N to perform authentication. We call this strategy authentication via binary thresholding.

- Secondly, if the underlying distribution $P_{Z|X,\bar{\theta}}$ is known or may be guessed, the receiver can use the knowledge of the true parameters $\bar{\theta}$ to establish a test statistic (see subsection 2.1.2) . This strategy is called authentication via grey level observations (see section 3.3.2).

- The last scenario is indeed similar with the second strategy but in the case where the receiver has to first estimate the opponent channel parameters before designing the authentication test. The impact of the estimation of these parameters on the performance of the authentication system is detailed in the next chapter.

3.3.1 Authentication via binary thresholding

The legitimate receiver first observes sequence o^N and uses a threshold based on the main channel marginal distribution $P_{Y|X,\theta}$ to restore a binary version $\tilde{x}^N$, called the decoded sequence of the original message x^N using the same decision region as defined by (3.1), which naturally generates errors.

- In the main channel, i.e., when $O^N = Y^N$, error probabilities are equivalent to (3.2) and (3.3).

- In the opponent channel, i.e., when $O^N = Z^N$, we make use of (3.5) and (3.6) to express the corresponding error probabilities:

$$\tilde{P}_{e,B} = \sum_{v \in \mathcal{D}_{\mathcal{W}}} P_{Z|X,\bar{\theta}}(v| X = 0, \bar{\theta}) \tag{3.11}$$

hence

$$\begin{aligned}\tilde{P}_{e,B} &= (1-P_{e,B}) \sum_{v\in\mathcal{D}_{\mathcal{W}}} T_{Z|\hat{X},\bar{\theta}}(v|\,\hat{X}=0,\bar{\theta}) \\ &\quad + P_{e,B} \sum_{v\in\mathcal{D}_{\mathcal{W}}} T_{Z|\hat{X},\bar{\theta}}(v|\,\hat{X}=1,\bar{\theta}) \\ &= (1-P_{e,B})P'_{e,B} + P_{e,B}(1-P'_{e,W})\end{aligned} \tag{3.12}$$

where $P'_{e,B} = \sum_{v\in\mathcal{D}_{\mathcal{W}}} T_{Z|\hat{X},\bar{\theta}}(v|\,\hat{X}=0,\bar{\theta})$ and $P'_{e,W} = \sum_{v\in\mathcal{D}^c_{\mathcal{W}}} T_{Z|\hat{X},\bar{\theta}}(v|\,\hat{X}=1,\bar{\theta})$. The same development yields:

$$\tilde{P}_{e,W} = (1-P_{e,W})P'_{e,W} + P_{e,W}(1-P'_{e,B}) \tag{3.13}$$

For this first strategy, the opponent channel may be viewed as the cascade of two binary input/binary output channels:

$$\begin{bmatrix} 1-\tilde{P}_{e,B} & \tilde{P}_{e,B} \\ \tilde{P}_{e,W} & 1-\tilde{P}_{e,W} \end{bmatrix} = \begin{bmatrix} 1-P_{e,B} & P_{e,B} \\ P_{e,W} & 1-P_{e,W} \end{bmatrix} \times \begin{bmatrix} 1-P'_{e,B} & P'_{e,B} \\ P'_{e,W} & 1-P'_{e,W} \end{bmatrix} \tag{3.14}$$

We realize that the test used to decide whether the observed decoded sequence $\tilde{x}^N$ comes from the original source or not is equivalent to counting the number of erroneous decoded dots. It should be noted that the conditional distribution of each random component $(\tilde{X}_i \mid x_i)$ of the random $i.i.d$ sequence $(\tilde{X}^N \mid x^N)$ is the same for each given type. We compute then the probabilities describing the random $i.i.d$ sequence $(\tilde{X}^N \mid x^N)$ for each of the two possible hypothesis H_0 and H_1, and we derive the corresponding test based on the general expression (3.10). Under hypothesis H_j, $j \in \{1,2\}$, these probabilities are formulated conditionally to the known original code x^N. Let $\mathcal{N}_B = \{i : x_i = 0\}$ and $\mathcal{N}_W = \{i : x_i = 1\}$ with $N_B = |\mathcal{N}_B|$, $N_W = |\mathcal{N}_W|$ and $N = N_B + N_W$. From the property of $i.i.d$ sequences we have:

$$\begin{aligned} P\left(\tilde{x}^N \middle|\, x^N, H_j\right) &= \prod_{i=1}^{N} P\left(\tilde{x}_i|\, x_i, H_j\right) \\ &= \prod_{i\in\mathcal{N}_B} P\left(\tilde{x}_i|\, 0, H_j\right) \times \prod_{i\in\mathcal{N}_W} P\left(\tilde{x}_i|\, 1, H_j\right) \end{aligned} \tag{3.15}$$

Particularly,

– Under hypothesis H_0, the channel $X \to \tilde{X}$ has distribution given by (3.2) and (3.3) and we have:

$$\begin{aligned} P\left(\tilde{x}^N \middle|\, x^N, H_0\right) &= (P_{e,B})^{n_{e,B}} (1-P_{e,B})^{N_B - n_{e,B}} \\ &\times (P_{e,W})^{n_{e,W}} (1-P_{e,W})^{N_W - n_{e,W}}, \end{aligned} \tag{3.16}$$

where $n_{e,B}$ and $n_{e,W}$ are the number of error $(\tilde{x}_i \neq x_i)$ when black is decoded into white and when white is decoded into black respectively.

– Under hypothesis H_1, the channel $X \to \tilde{X}$ has distribution given by (3.12) and (3.13) and we have:

$$\begin{aligned} P\left(\tilde{x}^N \middle| x^N, H_1\right) &= \left(\tilde{P}_{e,B}\right)^{n_{e,B}} \left(1-\tilde{P}_{e,B}\right)^{N_B - n_{e,B}} \\ &\times \left(\tilde{P}_{e,W}\right)^{n_{e,W}} \left(1-\tilde{P}_{e,W}\right)^{N_W - n_{e,W}}, \end{aligned} \tag{3.17}$$

Applying now the Neyman-Pearson decision (3.10) the test is expressed as:

$$L_1 = \log \frac{P\left(\tilde{x}^N \middle| x^N, H_1\right)}{P(\tilde{x}^N | x^N, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \tag{3.18}$$

or

$$L_1 = n_{e,B} \log \frac{\left(\tilde{P}_{e,B}\left(1-P_{e,B}\right)\right)}{\left(P_{e,B}\left(1-\tilde{P}_{e,B}\right)\right)} + n_{e,W} \log \frac{\left(\tilde{P}_{e,W}\left(1-P_{e,W}\right)\right)}{\left(P_{e,W}\left(1-\tilde{P}_{e,W}\right)\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_1, \tag{3.19}$$

where $\lambda_1 = \lambda - N_B \log \frac{\left(1-\tilde{P}_{e,B}\right)}{\left(1-P_{e,B}\right)} - N_W \log \frac{\left(1-\tilde{P}_{e,W}\right)}{\left(1-P_{e,W}\right)}$. This expression has the practical advantage to only count the number of errors in order to perform the authentication task but at a cost of a loss of optimality.

3.3.2 Authentication via grey level observations

In the second strategy, the receiver performs his test directly on the received sequence o^N without using any given threshold or decoding. We will see in the next subsection (3.4) that this strategy is better than the previous one. Here again, the conditional distribution of each random component $(O_i \mid x_i)$ of the random $i.i.d$ sequence $(O^N \mid x^N)$ is the same for each type of data of X. The Neyman-Pearson test is expressed as:

$$L_2 = \log \frac{P\left(o^N \middle| x^N, H_1\right)}{P(o^N | x^N, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_2, \tag{3.20}$$

which can be developed as

$$L_2 = \sum_{i \in \mathcal{N}_B} \log \frac{P_{Z|X,\bar{\theta}}(o_i|0)}{P_{Y|X,\theta}(o_i|0)} + \sum_{i \in \mathcal{N}_W} \log \frac{P_{Z|X,\bar{\theta}}(o_i|1)}{P_{Y|X,\theta}(o_i|1)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_2, \tag{3.21}$$

or more specifically,

$$\begin{aligned} L_2 = &\sum_{i \in \mathcal{N}_B} \log \left[(1-P_{e,W}) \frac{T_{Z|\hat{X},\bar{\theta}}(o_i|0)}{T_{Y|X,\theta}(o_i|0)} + P_{e,W} \frac{T_{Z|\hat{X},\bar{\theta}}(o_i|1)}{T_{Y|X,\theta}(o_i|0)} \right] + \\ &\sum_{i \in \mathcal{N}_W} \log \left[(1-P_{e,B}) \frac{T_{Z|\hat{X},\bar{\theta}}(o_i|1)}{T_{Y|X,\theta}(o_i|1)} + P_{e,B} \frac{T_{Z|\hat{X},\bar{\theta}}(o_i|0)}{T_{Y|X,\theta}(o_i|1)} \right] \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_2. \end{aligned} \tag{3.22}$$

Note that the expressions of the transition matrix modeling the physical processes $T_{Y|X,\theta}$ and $T_{Z|\hat{X},\bar{\theta}}$ are required in order to perform the optimal test, i.e., the receiver needs to know or to estimate the print and scan processes of the opponent.

3.3.3 Comparison between the two strategies

For this comparison and without loss of generality, we consider only the Gaussian model with variance σ^2 for the physical devices $T_{Y|X,\theta}$ and $T_{Z|\hat{X},\bar{\theta}}$. and we compare the receiver operating characteristic (ROC) curves associated with the two different strategies. Note that the error probabilities are computed using the results given in the next subsection (see 3.4). We can notice that the gap between the two strategies is significant in the variation of magnitude. This is not surprising since the binary thresholding removes information from the gray-level observation, yet this has a practical impact because one practitioner can be tempted to count the number of errors as given in (3.19) as an authentication score for its easy implementation or in case when he cannot estimate the opponent channel. The information theoretical analysis presented in [53] confirms also that authentication is more accurate without thresholding.

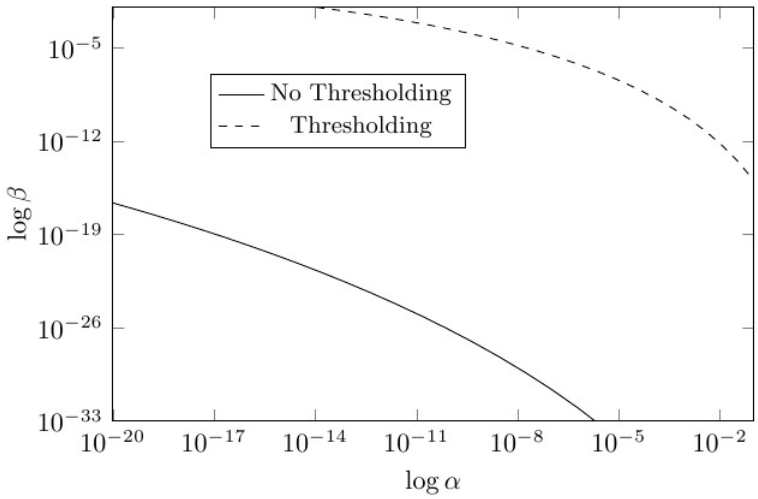


Figure 3.5: ROC curves for two different strategies in case $N = 2.10^3$, $\sigma_b = \sigma_w = 52$

3.4 Reliable computation for error probabilities

In the previous subsection we have expressed explicitly the Neyman-Pearson test for the two proposed receiver's strategies summarized by (3.19) and (3.20). These tests may then be practically performed on the observed sequence of size N coming from an observed GC in order to decide about its authenticity. We aim now at expressing the error probabilities of types I and II and comparing the authentication performance of two possible strategies described previously. It should be reminded that throughout the thesis we deal with discrete case for the main and the opponent densities although the results can be extended to the continuous case. So if we let $m = 1, 2$ be the index denoting the strategy, a straightforward calculation gives exactly

$$\alpha_m = \sum_{l>\lambda_m} P_{L_m}(l \mid H_0) \tag{3.23}$$

$$\beta_m = \sum_{l<\lambda_m} P_{L_m}(l \mid H_1) \tag{3.24}$$

where $P_{L_m}(l \mid H_j)$ is the density of the log-likelihood ratio L_m under hypothesis H_j.

3.4.1 Gaussian approximation

As the length N of the sequence is generally large enough, we commonly use the central limit theorem (CLT) to approximate the distributions P_{L_m}, $m = 1, 2$ (a similar strategy was considered in [84]).

– For the binary thresholding strategy, we know that $n_{e,B}$ and $n_{e,W}$ in (3.19) are binomial random variables depending on the origin of the observed sequence. Let N_x stands for the number of data of type x in the original code, $n_{e,x}$ stands for $n_{e,B}$ or $n_{e,W}$, and $P_{e,x}$ stands for the cross-over probabilities emerging from type x in the BIBO channels (3.4) or (3.14). When N is large enough, using CLT the binomial random variables can be computed from a Gaussian distribution. We have thus:

$$n_{e,x} \sim \mathcal{N}(N_x P_{e,x}, N_x P_{e,x}(1 - P_{e,x})) \tag{3.25}$$

From (3.19), L_1 is a weighted sum of Gaussian random variables and one can obviously infer the parameters of the normal approximation describing the log-likelihood L_1.

– For authentication via grey level observations, , i.e. when the receiver applies the test directly on the observed gray-level sequence, the log-likelihood L_2 in Eq. (3.22) may be expressed as two sums of $i.i.d$ variables and becomes:

$$L_2 = \sum_{i \in \mathcal{N}_B} l(o_i \mid 0) + \sum_{i \in \mathcal{N}_W} l(o_i \mid 1) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_2, \tag{3.26}$$

where $l(v \mid t)$ is a function $l : \mathcal{V} \mid \mathcal{X} \to \mathbb{R}$ having a distribution with mean and variance respectively equal to:

$$\mu_x = \mathbb{E}\left[l(V \mid x) \mid H_j\right] = \sum_{v \in \mathcal{V}} l(v \mid x) P(v \mid x, H_j), \tag{3.27}$$

and

$$\sigma_x^2 = \mathrm{Var}\left[l(V \mid x) \mid H_j\right] = \sum_{v \in \mathcal{V}} (l(v \mid x) - \mu_x)^2 P(v \mid x, H_j), \tag{3.28}$$

with $P = P_{Y|X}$ (respectively $P = P_{Z|X}$) for $j = 0$ (respectively 1). The CLT can then be used again to approximate the distribution of L_2 and compute type I and type II error probabilities.

3.4.2 Asymptotic expression

In this part, for the shake of simplicity we drop the subscript m denoting the strategy as all the subsequent analysis is common for both of them.

One important problem is the fact that the Gaussian approximation proposed previously provides inaccurate error probability values when the threshold λ in (3.23) and (3.24) is far from the mean of the log-likelihood random variable L. Instead Chernoff bound and large deviation theory [33] can be employed in this context as very small error probabilities of types I and II may be desired [48]. Given a real number s, the Chernoff bound on type I and type II errors may be expressed as:

$$\alpha = \Pr\left(L \geq \lambda \mid H_0\right) \leq e^{-s\lambda} g_L(s \mid H_0) \text{ for any } s > 0, \tag{3.29}$$

$$\beta = \Pr\left(L \leq \lambda \mid H_1\right) \leq e^{-s\lambda} g_L(s \mid H_1) \text{ for any } s < 0, \tag{3.30}$$

where the function $g_L(s \mid H_j)$, $j = 0, 1$ is the moment generating function of L defined as:

$$g_L(s \mid H_j) = \mathbb{E}_{P_L(L|H_j)}\left[e^{sL}\right], \tag{3.31}$$

and the expectation is calculated w.r.t distribution $P_L(L \mid H_j)$.

Because L is a sum of N independent random variables, asymptotic analysis in probability theory (when N is large enough) shows that bounds similar to (3.29) and (3.30) are much more appropriate for estimating α and β than the Gaussian approximation especially when λ is far from $\mathbb{E}[L]$, namely when bounding the tails of a distribution [33, 48]. The tightest bound is obtained by finding the value of s that provides the minimum of the right hand side (RHS) of (3.29) and (3.30), i.e. for the minimum of $e^{-s\lambda} g_L(s \mid H_j)$ for each $j = 0, 1$. Taking the derivative, the value s that provides the tightest bound under each hypothesis is such that[1]:

$$\begin{aligned} \lambda = \left.\frac{\frac{dg_L(s|H_j)}{\mathrm{d}s}}{g_L(s|H_j)}\right|_{s=s_j} &= \left.\frac{d}{ds}\log g_L(s \mid H_j)\right|_{s=s_j} \\ &= \left.\frac{d}{ds}\mu_L(s \mid H_j)\right|_{s=s_j} \end{aligned} \tag{3.32}$$

where

[1] one can show that $e^{-s\lambda} g_L(s \mid H_j)$ is a convex function of s

$$\mu_L(s \mid H_j) = \log g_L(s \mid H_j), \tag{3.33}$$

is the semi-invariant moment generating function or cumulant generating function. This function has many interesting properties that ease the extraction of an asymptotic expression for (3.29) and (3.30) [48]. For instance, this function is additive for the sum of independent random variables, and we have

$$\mu_L(s \mid H_j) = \sum_{i \in \mathcal{N}_B} \mu_{i|0}(s \mid H_j) + \sum_{i \in \mathcal{N}_W} \mu_{i|1}(s \mid H_j) \tag{3.34}$$

where $\mu_{i|x}(s \mid H_j)$ is the cumulant generating function of the random variable $l(O_i \mid x)$ when the observed sequence comes from the distribution associated to hypothesis H_j. Additionally, the relation (3.32) may be interpreted as the sum of the derivatives at the value s_j optimizing the bounds of α and β as:

$$\lambda = \sum_{i \in \mathcal{N}_B} \mu'_{i|0}(s_j \mid H_j) + \sum_{i \in \mathcal{N}_W} \mu'_{i|1}(s_j \mid H_j). \tag{3.35}$$

The Chernoff bounds on α and β in (3.29) and (3.30) may thus be expressed as:

$$\begin{aligned} \alpha &= \Pr(L \geq \lambda \mid H_0) \\ &\leq \exp\left[\sum_{i \in \mathcal{N}_B} \Big(\mu_{i|0}(s_0 \mid H_0) - s_0 \mu'_{i|0}(s_0 \mid H_0) \Big) \right. \\ &+ \left. \sum_{i \in \mathcal{N}_W} \Big(\mu_{i|1}(s_0 \mid H_0) - s_0 \mu'_{i|1}(s_0 \mid H_0) \Big) \right]. \end{aligned} \tag{3.36}$$

and

$$\begin{aligned} \beta &= \Pr(L \leq \lambda \mid H_1) \\ &\leq \exp\left[\sum_{i \in \mathcal{N}_B} \Big(\mu_{i|0}(s_1 \mid H_1) - s_1 \mu'_{i|0}(s_1 \mid H_1) \Big) \right. \\ &+ \left. \sum_{i \in \mathcal{N}_W} \Big(\mu_{i|1}(s_1 \mid H_1) - s_1 \mu'_{i|1}(s_1 \mid H_1) \Big) \right]. \end{aligned} \tag{3.37}$$

From our assumption, the distribution of each random component $(O_i \mid x_i)$ in the $i.i.d$ sequence $(O^N \mid x^N)$ is the same for each type of data X, and consequently, $\mu_{i|x}(s \mid H_j) = \mu_x(s \mid H_j)$, i.e. $\mu_{i|x}(s \mid H_j)$ is independent from i for each type of data x. The RHS in (3.36) and (3.37) can be simplified as

$$\exp\left[N_B \left(\mu_0(s_j \mid H_j) - s_j \mu'_0(s_j \mid H_j) \right) + N_W \left(\mu_1(s_j \mid H_j) - s_j \mu'_1(s_j \mid H_j) \right) \right]. \tag{3.38}$$

Roughly speaking, Cramér's theorem [33] states that for sufficiently large N, the upper bounds expressed for $j = 0, 1$ in (3.38) are also lower bounds for α and β respectively. Thus without loss of generality, one can suppose that $N_B = N_W = N/2$; we have then:

$$\lim_{N\to\infty}\frac{2}{N}\log\alpha=\left[\mu(s_0\mid H_0)-s_0\mu'(s_0\mid H_0)\right], \tag{3.39}$$

$$\lim_{N\to\infty}\frac{2}{N}\log\beta=\left[\mu(s_1\mid H_1)-s_1\mu'(s_1\mid H_1)\right], \tag{3.40}$$

where $s_0>0$, $s_1<0$, $\mu(s_j\mid H_j)=\mu_0(s_j\mid H_j)+\mu_1(s_j\mid H_j)$ and $\mu'(s_j\mid H_j)=\mu'_0(s_j\mid H_j)+\mu'_1(s_j\mid H_j)$. One can show also that $s_1=s_0-1$. A modified asymptotic expression including a correction factor is evaluated for the sum of an $i.i.d$ random sequence (see [48], Appendix 5A), and for large N we have:

$$\begin{aligned}\alpha &= \Pr(L\geq\lambda\mid H_0)\\ &\underset{N\to\infty}{\to} \frac{1}{|s_0|\sqrt{N\pi\mu''(s_0|H_0)}}\exp\left\{\tfrac{N}{2}\left[\mu(s_0\mid H_0)-s_0\mu'(s_0\mid H_0)\right]\right\},\end{aligned} \tag{3.41}$$

and

$$\begin{aligned}\beta &= \Pr(L\leq\lambda\mid H_1)\\ &\underset{N\to\infty}{\to} \frac{1}{|s_1|\sqrt{N\pi\mu''(s_1|H_1)}}\exp\left\{\tfrac{N}{2}\left[\mu(s_1\mid H_1)-s_1\mu'(s_1\mid H_1)\right]\right\}.\end{aligned} \tag{3.42}$$

where $\mu''(s_j\mid H_j)=\mu''_0(s_j\mid H_j)+\mu''_1(s_j\mid H_j)$ is the second derivative of cumulant generating function $l(V\mid x)$ defined by:

$$l(v\mid 0)=\log\left((1-P_{e,B})\frac{T_{Z|\hat{X},\bar{\theta}}(v\mid 0)}{T_{Y|X,\theta}(v\mid 0)}+P_{e,B}\frac{T_{Z|\hat{X},\bar{\theta}}(v\mid 1)}{T_{Y|X,\theta}(v\mid 0)}\right), \tag{3.43}$$

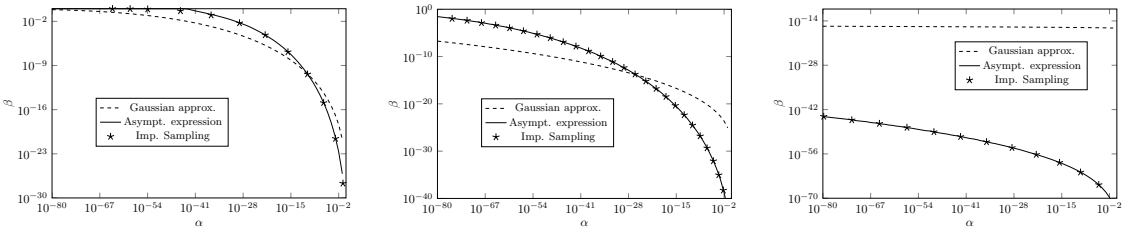
$$l(v\mid 1)=\log\left((1-P_{e,W})\frac{T_{Z|\hat{X},\bar{\theta}}(v\mid 1)}{T_{Y|X,\theta}(v\mid 1)}+P_{e,W}\frac{T_{Z|\hat{X},\bar{\theta}}(v\mid 0)}{T_{Y|X,\theta}(v\mid 1)}\right). \tag{3.44}$$

We give below the numerical results for the difference between Gaussian approximation and Asymptotic Expression and for the comparison with Monte-Carlo simulation in order to see which is the best choice for authentication performance.

In order to assess the accuracy of the computations of α and β using either the Gaussian approximation given by (3.23) and (3.24), the Asymptotic Expression given by (3.39) and (3.42) and the Monte-Carlo simulations using importance sampling given in our paper [53], we respectively derive ROC curves for generalized Gaussian distributions and $b=\{1,2,6\}$. The ROCs are practically computed by first setting a threshold λ and then deriving the probabilities associated to this threshold.

Fig. 3.6 illustrates the gap between the estimation of α and β using the Gaussian approximation and the asymptotic expression or the Monte-Carlo simulations. The Monte-Carlo simulations confirm the fact that the derived Cramér Chernoff bounds are tight, and the difference between the results obtained with the Gaussian approximation are very important especially for close to uniform channels. This sheds light on

an important conclusion that using Chernoff bounds are more reliable to perform authentication than using Gaussian approximation when the code length is large enough. From Fig. 3.6 we can also notice that for a same channel power, the authentication performances are better for $b = 6$ then for $b = 2$ and $b = 1$.



$b = 1$ $b = 2$ $b = 6$

Figure 3.6: Comparison between the Gaussian approximation, the Asymptotic expression and Monte-Carlo simulations via importance sampling for $b = 1$, $b = 2$ and $b = 6$. Main and opponent channels are identical, $\mu_b = 50$, $\mu_w = 150$, $\sigma_b = 40$, $\sigma_w = 40$.

3.5 Conclusion of Chapter 3

In this chapter, rigorously we have introduced the general model for print and scan channels. We have also proposed the receiver's strategies to perform authentication relying on the classical binary hypothesis testing. We have made the comparisons for all strategies and indicated the best strategy for the receiver. A reliable computation of the authentication performance based on Asymptotic Expression method has been derived and compared with the Monte-Carlo simulation (see also in [81, 53]).

From the practical point of view, the next chapter studies the impact of the estimation on authentication performance. Another expression of Asymptotic Expression, based on a well-known distribution in statistical physic, will be proposed to fulfill this study.

Chapter 4

Impact of estimation on authentication performances

" *There is no end to education. It is not that you read a book, pass an examination, and finish with education. The whole of life, from the moment you are born to the moment you die, is a process of learning.* "

Jiddu Krishnamurti

In this chapter we extend our analysis to the case where the receiver does not know the true parameter $\bar{\theta}$ related to the opponent print and scan process, but establishes a test statistic using estimated ones obtained by a maximum likelihood based algorithm. The estimated parameters are computed from several codes identified previously as fake codes which represent a set of printed and scanned dots driven from $P_{Z|X,\bar{\theta}}$ (see (4) in Fig. 3.1).

4.1 Asymptotic expression employing Boltzmann's distribution

As we have mentioned in the setup, practically the receiver have only a partial knowledge about the opponent channel. We assume indeed that the receiver has an a-priori knowledge about the families of distributions that govern the opponent channel. Furthermore we assume that the receiver observes several identified fake GCs, he then uses these observed data to estimate the parameters of the opponent channel. Because every estimation yields noise, the loss in authentication performance is computed by comparing the error probabilities obtained for the true and estimated channel. A rigorous study of this loss is consequently needed due to the natural fact that the receiver in this chapter performs authentication based on the estimated opponent channel instead of the true one that gives the optimal performance.

It is important to compare our study with a direct use of the Generalized Likelihood Ratio Tests (GLRT, see 2.1.2). Or analysis is motived by the following points:

- In our setup, the opponent channel parameters are not estimated directly from the received code (as for the GLRT) but they are estimated beforehand from a batch of codes which are known to be fake.
- Because we are interested in computing an accurate estimation of the error probabilities, we cannot invoke a variation of the Wilks' theorem (see eq. (2.26)), since it is only accurate for large N or small false alarm probabilities.
- Moreover, our setup is different from the assumptions of the Wilks's theorem since the parametric definition of densities under H_0 and H_1 can be different in our case.
- There is one connection however: in the case where the observation used by the GLRT comes from only fake codes, then the test that we use in this chapter and the GLRT are equivalent.

In this section, we develop an essential tool to fulfill this study, which is originally related to the more general concepts of f-divergence in information theory and Boltzmann's distribution in statistical physics.

There exists different expressions for the asymptotic tail probabilities of sum of $i.i.d$ random variables, for example, in the part "Asymptotic expression" of section (3.4) or in [30] on Sanov's theorem that uses information theoretical arguments. Here, we present a unified tool employing the Boltzmann's distribution, that allows to derive these expressions and to use them to link the probability errors of the tests obtained for true parameters and for estimated ones. All these expressions or proofs use a twisted distribution (see in [31, 26]) which is centered on the desired threshold when optimized. This desired threshold is bounded from above and below by the quantities related to the Kullback-Leibler divergences between the distribution of the main and opponent channels. By the use of Boltzmann's distribution, we can generalize the concept of this twisted distribution. The most important fact is that the generalized tool helps us approximate accurately the losses on β via a very simple expression when the opponent channel is estimated.

The goal of this part is to show the relation between the Boltzmann's distribution and the Chernoff bounds mentioned in the previous part and to deduce several interesting properties. We start with generating a property related to the f-divergence. For simplicity, we denote:

$$p_0 \equiv P_{V|H_0}, \tag{4.1}$$

the parametric density of the sequence coming from the null hypothesis H_0, and

$$p_1 \equiv P_{V|H_1}, \tag{4.2}$$

for the one coming from the alternative hypothesis H_1.

Generally, we can consider two densities $p_0(v)$ and $p_1(v)$ defined on a space Ω. The requirement is that the n-moment of the log-likelihood ratio w.r.t p_0 and p_1 are finite, i.e.,

$$\int_\Omega \left(\log \frac{p_1(v)}{p_0(v)}\right)^n p_1(v)d(v) < \infty, \tag{4.3}$$

$$\int_\Omega \left(\log \frac{p_0(v)}{p_1(v)}\right)^n p_0(v)d(v) < \infty, \tag{4.4}$$

which implies that $D_{KL}(p_1||p_0) < \infty$ and $D_{KL}(p_0||p_1) < \infty$ by taking $n = 1$.Hence $p_0(v)$ and $p_1(v)$ have common support on Ω, i.e., both p_0 and p_1 are absolutely continuous w.r.t a common Lebesgue measure on Ω. So the following parametric density, called a Boltzmann distribution, is well-defined:

$$p_t(v) = \frac{[p_0(v)]^{1-t}\,[p_1(v)]^t}{N_t}, \tag{4.5}$$

on Ω with N_t is normalized constant $N_t = \int_\Omega [p_0(v)]^{1-t}[p_1(v)]^t dv$ $(0 \le t \le 1)$. We are then able to define the f-divergence between p_t and p_u as:

$$D_f(p_t||p_u) = \int_\Omega f\left(\frac{p_t(v)}{p_u(v)}\right) p_u(v)dv, \tag{4.6}$$

with f is a convex function.

Equations (A.2) and (A.1) developed in Appendix A.1 are interesting from the mathematical point of view because each of them merges the cumulant generating function and the Kullback-Leibler divergence in a same formula. Expressions for the tail probabilities α and β of the sum of N_r $i.i.d$ random variables expressed in (3.39) and (3.42) may be expressed then by, after pointing that $s_1 = s_0 - 1$:

$$\begin{array}{rcl} \alpha & \simeq & \exp\left\{N_r\left[\mu(s) - s\mu'(s)\right]\right\}, \\ \beta & \simeq & \exp\left\{N_r\left[\mu(s) + (1-s)\mu'(s)\right]\right\}. \end{array} \qquad 0 < s < 1 \tag{4.7}$$

or equivalently:

$$\log\alpha \simeq N_r(\mu(s) - s\mu'(s)) = -N_r D_{KL}(p_s||p_0), \tag{4.8}$$

and

$$\log\beta \simeq N_r(\mu(s) + (1-s)\mu'(s)) = -N_r D_{KL}(p_s||p_1). \tag{4.9}$$

Now we apply this general development in our case in which $D_{KL}(p_s \parallel p_0)$ and $D_{KL}(p_s \parallel p_1)$ $(0 < s < 1)$ are defined as:

$$D_{KL}(p_s \parallel p_0) = D_{KL}(p_s(\cdot \mid 0) \parallel p_0(\cdot \mid 0)) + D_{KL}(p_s(\cdot \mid 1) \parallel p_0(\cdot \mid 1)), \tag{4.10}$$

and

$$D_{KL}(p_s \parallel p_1) = D_{KL}(p_s(\cdot \mid 0) \parallel p_1(\cdot \mid 0)) + D_{KL}(p_s(\cdot \mid 1) \parallel p_1(\cdot \mid 1)), \tag{4.11}$$

where

$$\begin{array}{rcl} p_0(v \mid x) & = & P_{Y|X,\theta}\left(v|\, x, H_0\right) \\ p_1(v \mid x) & = & P_{Z|X,\bar{\theta}}\left(v|\, x, H_1\right) \end{array} \qquad x = 0, 1 \tag{4.12}$$

and

$$p_s(v \mid x) = \frac{[p_0(v|x)]^{1-s}[p_1(v|x)]^s}{\sum\limits_{v\in\mathcal{V}} [p_0(v|x)]^{1-s}[p_1(v|x)]^s} \qquad x = 0, 1. \tag{4.13}$$

As we have mentioned above, we can consider without loss of generality that $N_B = N_W = N/2$ (N is the sample size), the formulas (4.8) and (4.9) are then:

$$\log\alpha \simeq -\frac{N}{2} D_{KL}(p_s||p_0) \tag{4.14}$$

$$\log \beta \simeq -\frac{N}{2} D_{KL}(p_s||p_1), \tag{4.15}$$

which gives us an explicit approximation of the error probabilities w.r.t. the Boltzmann's distribution p_s.

4.2 Approximation of authentication performance up to the second order

In this section we analyze how the set of estimated parameters impacts the performance of the probability of type II error $\beta(\alpha, \hat{\bar{\theta}})$ for a fixed value of α. Precisely, we provide a relation between the variation on $\log \beta(\alpha, \hat{\bar{\theta}})$ and the variation on $\hat{\bar{\theta}}$. To do this, we assume that the proposed estimation scheme is able to provide unbiased estimated parameters $\hat{\bar{\theta}}$ that are close to the true parameters $\bar{\theta}$ of the opponent channel.

4.2.1 Second order approximation

Here, we suppose that $\hat{\bar{\theta}}$ is a vector of m unknown parameters, i.e., $\hat{\bar{\theta}} = (\hat{\bar{\theta}}_1, ..., \hat{\bar{\theta}}_m)$. Without loss of generality, we can take into account our analysis on the codes with the same number of bits 0 (N_b) and bits 1 (N_w), i.e., $N_b = N_w = \frac{N}{2}$. Furthermore, for large enough N, the changes of correcting factors $\frac{1}{|s_j|\sqrt{N\pi\mu''(s_j|H_j)}}$ $(j = 0, 1)$ in (3.41) and (3.42) are negligible and we drop their analysis.

According to the Taylor expansion we can write:

$$\beta^*(\hat{\bar{\theta}}) \cong \beta^*(\bar{\theta}) + \nabla\beta^*(\hat{\bar{\theta}})\Big|_{\hat{\bar{\theta}}=\bar{\theta}} (\hat{\bar{\theta}} - \bar{\theta}) + \frac{1}{2}(\hat{\bar{\theta}} - \bar{\theta})^T \nabla^2\beta^*(\hat{\bar{\theta}})\Big|_{\hat{\bar{\theta}}=\bar{\theta}} (\hat{\bar{\theta}} - \bar{\theta}) + ... \tag{4.16}$$

where

$$\beta^*(\hat{\bar{\theta}}) \equiv \beta^*(\alpha, \hat{\bar{\theta}}) = \frac{2}{N} \log \beta(\alpha, \hat{\bar{\theta}}) = \mu(s_1(\hat{\bar{\theta}}) \mid H_1) - s_1(\hat{\bar{\theta}})\mu'(s_1(\hat{\bar{\theta}}) \mid H_1),$$

and $\nabla\beta^*(\hat{\bar{\theta}})$ and $\nabla^2\beta^*(\hat{\bar{\theta}})$ are the gradient vector and Hessian matrix of $\beta^*(\hat{\bar{\theta}})$ respectively. For simplicity, we denote $\mathcal{D} = \nabla\beta^*(\hat{\bar{\theta}})$ and $\mathcal{H} = \nabla^2\beta^*(\hat{\bar{\theta}})$.

Let's denote also:

$$\alpha^*(\hat{\bar{\theta}}) \equiv \frac{2}{N} \log \alpha(\hat{\bar{\theta}}) = \mu(s_0(\hat{\bar{\theta}}) \mid H_0) - s_0(\hat{\bar{\theta}})\mu'(s_0(\hat{\bar{\theta}}) \mid H_0),$$

We have for each $\hat{\bar{\theta}}_i$ the first partial derivative of $\alpha^*(\hat{\bar{\theta}})$ w.r.t $\hat{\bar{\theta}}_i$ as:

$$\frac{\partial \alpha^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} = \frac{\mathbb{E}_{p_0}\left[s_0(\hat{\bar{\theta}}) l_i'(\hat{\bar{\theta}}) e^{s_0(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]}{\mathbb{E}_{p_0}\left[e^{s_0(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]} - s_0(\hat{\bar{\theta}}) \frac{\frac{\partial \lambda_0(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i}}{N/2}, \tag{4.17}$$

where $l(\hat{\bar{\theta}}) = \log \frac{p_1(v|\hat{\bar{\theta}})}{p_0(v|\hat{\bar{\theta}})}$ and $l_i'(\hat{\bar{\theta}}) = \frac{\partial l(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i}$. Because α is fixed, $\frac{\partial \alpha^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} = 0$ $(i = 1, ..., m)$, and hence:

$$\frac{\frac{\partial \lambda_0(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i}}{N/2} = \frac{\mathbb{E}_{p_0}\left[l_i'(\hat{\bar{\theta}}) e^{s_0(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]}{\mathbb{E}_{p_0}\left[e^{s_0(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]}. \tag{4.18}$$

Similarly, the first partial derivative of $\beta^*(\hat{\bar{\theta}})$ w.r.t $\hat{\bar{\theta}}_i$ is

$$\frac{\partial \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} = \frac{\mathbb{E}_{p_1}\left[s_1(\hat{\bar{\theta}}) l_i'(\hat{\bar{\theta}}) e^{s_1(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]}{\mathbb{E}_{p_1}\left[e^{s_1(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]} - s_1(\hat{\bar{\theta}}) \frac{\frac{\partial \lambda_1(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i}}{N/2}. \tag{4.19}$$

We always choose the same threshold for the test, so $\lambda_0(\hat{\bar{\theta}}) = \lambda_1(\hat{\bar{\theta}})\ \forall \hat{\bar{\theta}}$ and then

$$\frac{\partial \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} = s_1(\hat{\bar{\theta}}) \left\{ \frac{\mathbb{E}_{p_1}\left[l_i'(\hat{\bar{\theta}}) e^{s_1(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]}{\mathbb{E}_{p_1}\left[e^{s_1(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]} - \frac{\mathbb{E}_{p_0}\left[l_i'(\hat{\bar{\theta}}) e^{s_0(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]}{\mathbb{E}_{p_0}\left[e^{s_0(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]} \right\}. \tag{4.20}$$

By using two Boltzmann's distributions $p_{s_0}(\hat{\bar{\theta}}) \equiv p_{s_0}(v \mid \hat{\bar{\theta}})$ and $p_{s_1}(\hat{\bar{\theta}}) \equiv p_{s_1}(v \mid \hat{\bar{\theta}})$ defined as follow:

$$\begin{aligned} p_{s_0}(\hat{\bar{\theta}}) &= \frac{e^{s_0(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})} p_0(\hat{\bar{\theta}})}{\mathbb{E}_{p_0}\left[e^{s_0(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]}, \\ p_{s_1}(\hat{\bar{\theta}}) &= \frac{e^{s_1(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})} p_1(\hat{\bar{\theta}})}{\mathbb{E}_{p_1}\left[e^{s_1(\hat{\bar{\theta}}) l(\hat{\bar{\theta}})}\right]}. \end{aligned} \tag{4.21}$$

and with the approximation that $p_1(\hat{\bar{\theta}}) \approx p_1(\bar{\theta})$, we can simplify Eq. 4.20 with:

$$\frac{\partial \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} \approx s_1(\hat{\bar{\theta}}) \left[\mathbb{E}_{p_{s_1}}\left(l_i'(\hat{\bar{\theta}})\right) - \mathbb{E}_{p_{s_0}}\left(l_i'(\hat{\bar{\theta}})\right)\right]. \tag{4.22}$$

Note that this approximation does not impair our results since in the following, we compute the expectation of all derivatives for $\hat{\bar{\theta}} = \bar{\theta}$. As pointed out previously, at the true parameters $\bar{\theta}$, we know that :

$$s_1(\bar{\theta}) = s_0(\bar{\theta}) - 1. \tag{4.23}$$

This property leads to the following lemma:

Lemma 1. *For every integrable function $f(\theta)$, whenever the Chernoff's bounds are optimized we have:*

$$\mathbb{E}_{p_{s_1}}\left(f(\bar{\theta})\right) = \mathbb{E}_{p_{s_0}}\left(f(\bar{\theta})\right) \tag{4.24}$$

at the true parameters $\bar{\theta}$.

Proof. The proof is directly obtained when plugging property (4.23) in (4.24). □

Now we respectively compute the values of $\mathcal{D}$ and $\mathcal{H}$ at the actual model parameters $\bar{\theta}$ of opponent's printing process.

Value of $\mathcal{D}$ at $\bar{\theta}$:

Applying this lemma we can directly imply the following equality for the gradient of β^* at the true parameter $\bar{\theta}$:

$$\nabla\beta^*(\hat{\bar{\theta}})\Big|_{\hat{\bar{\theta}}=\bar{\theta}} = 0, \tag{4.25}$$

hence $\mathcal{D} = 0$. The equation (4.25) is not surprising since the NP-test is known to reach the optimum when applied on the true parameter.

Value of $\mathcal{H}$ at $\bar{\theta}$:

In order to compute the value of $\mathcal{H}$ at $\bar{\theta}$, we respectively derive the analytical formulas for $\frac{\partial^2\beta^*(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i^2}\Big|_{\hat{\bar{\theta}}=\bar{\theta}}$ and $\frac{\partial^2\beta^*(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i\partial\hat{\bar{\theta}}_k}\Big|_{\hat{\bar{\theta}}=\bar{\theta}}$. Now let us first denote:

$$\begin{aligned}
\operatorname{cov}(l, l'_i) &= \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})l'_i(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right]\mathbb{E}_{p_{s_0}}\left[l'_i(\bar{\theta})\right] \\
\operatorname{cov}(l'_i, l'_k) &= \mathbb{E}_{p_{s_0}}\left[l'_i(\bar{\theta})l'_k(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l'_i(\bar{\theta})\right]\mathbb{E}_{p_{s_0}}\left[l'_k(\bar{\theta})\right] \\
\operatorname{Var}(l) &= \mathbb{E}_{p_{s_0}}\left[l^2(\bar{\theta})\right] - \mathbb{E}^2_{p_{s_0}}\left[l(\bar{\theta})\right] \\
\operatorname{Var}(l'_i) &= \mathbb{E}_{p_{s_0}}\left[\left(l'_i(\bar{\theta})\right)^2\right] - \mathbb{E}^2_{p_{s_0}}\left[l'_i(\bar{\theta})\right].
\end{aligned}$$

Then we can prove the two lemmas presented below that give expression of the the second partial derivatives of $\beta^*(\hat{\bar{\theta}})$:

Lemma 2. *At $\hat{\bar{\theta}} = \bar{\theta}$ we have, for each parameter $\hat{\bar{\theta}}_i$:*

$$\frac{\partial^2\beta^*(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i^2}\Big|_{\hat{\bar{\theta}}=\bar{\theta}} = s_1(\bar{\theta})\left[\frac{\operatorname{cov}^2(l,l'_i)}{\operatorname{Var}(l)} - \operatorname{Var}(l'_i)\right] \tag{4.26}$$

Proof. See A.2 in Appendix. □

Similar to the above lemma, we can prove the following result for the second partial derivatives of $\beta^*(\hat{\bar{\theta}})$ w.r.t the parameters $\hat{\bar{\theta}}_i$ and $\hat{\bar{\theta}}_k$:

Lemma 3. *At $\hat{\bar{\theta}} = \bar{\theta}$ we get, for each couple of parameter $(\hat{\bar{\theta}}_i, \hat{\bar{\theta}}_k)$, the following result:*

$$\left.\frac{\partial^2 \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i \partial \hat{\bar{\theta}}_k}\right|_{\hat{\bar{\theta}}=\bar{\theta}} = s_1(\bar{\theta}) \left[\frac{\mathrm{cov}(l,l'_i)\mathrm{cov}(l,l'_k)}{\mathrm{Var}(l)} - \mathrm{cov}(l'_i, l'_k)\right] \tag{4.27}$$

Proof. Taking the partial derivative on both sides of (4.22) w.r.t $\hat{\bar{\theta}}_k$ and follow the same way as in lemma (2). □

From Taylor expansion, the property (4.23), the lemma (2), (3) and Eq. 4.25, we get the close form formula for the Hessian matrix $\mathcal{H}$ (i.e., we do not have to use any sample to compute $\mathcal{H}$) and obtain the important theorem below:

Theorem 4. *The expansion for the* log *of the probability of type II error w.r.t the estimated parameters close to the true parameters $\bar{\theta}$, can be expressed as a quadratic form:*

$$\log \beta(\hat{\bar{\theta}}) \cong \log \beta(\bar{\theta}) + \frac{N}{4}(\hat{\bar{\theta}} - \bar{\theta})^T \mathcal{H}(\bar{\theta})(\hat{\bar{\theta}} - \bar{\theta}) \tag{4.28}$$

where

$$\begin{aligned} \mathcal{H}_{i,i}(\bar{\theta}) &= s_1(\bar{\theta}) \left[\frac{\mathrm{cov}^2(l,l'_i)}{\mathrm{Var}(l)} - \mathrm{Var}(l'_i)\right] \\ \mathcal{H}_{i,k}(\bar{\theta}) &= s_1(\bar{\theta}) \left[\frac{\mathrm{cov}(l,l'_i)\mathrm{cov}(l,l'_k)}{\mathrm{Var}(l)} - \mathrm{cov}(l'_i, l'_k)\right]. \end{aligned} \tag{4.29}$$

All quantities cov *and* Var *in (4.29) are taken w.r.t the Boltzmann density p_{s_0} defined in (4.21).*

It should be reminded that in order to analyze the authentication performance, we try to analyze the variation of $\log \beta(\hat{\bar{\theta}})$ w.r.t the set of estimated parameters $\hat{\bar{\theta}}$. It means that we want to approximate the distribution of $\log \beta(\hat{\bar{\theta}})$ considering the distribution of $\hat{\bar{\theta}}$. Because we consider an unbiased MLE in order to achieve optimal estimation, $\hat{\bar{\theta}}$ asymptotically tends to a normal distribution and consequently the distribution of $\log \beta(\hat{\bar{\theta}})$ can be determined.

4.2.2 Distribution of $\log \beta(\hat{\bar{\theta}})$

Distribution for one estimated parameter

For the sake of simplicity, we can start with the assumption that there is only one parameter $\bar{\theta}_{i_0}$ that needs to be estimated. In this case the formula (4.28) in Theorem (4) becomes:

$$\log \beta(\hat{\bar{\theta}}) \cong \log \beta(\bar{\theta}) + \frac{N}{4} \left. \frac{\partial^2 \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} \right|_{\hat{\bar{\theta}}=\bar{\theta}} \left(\triangle\bar{\theta}\right)^2, \tag{4.30}$$

where $\triangle\bar{\theta} = (\hat{\bar{\theta}}_{i_0} - \bar{\theta}_{i_0})$ and

$$\left. \frac{\partial^2 \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} \right|_{\hat{\bar{\theta}}=\bar{\theta}} = s_1(\bar{\theta}) \left[\frac{\mathrm{cov}^2(l, l'_{i_0})}{\mathrm{Var}(l)} - \mathrm{Var}(l'_{i_0}) \right].$$

Let

$$\gamma(\alpha, \bar{\theta}) = \frac{N}{4} \left. \frac{\partial^2 \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} \right|_{\hat{\bar{\theta}}=\bar{\theta}} \times \mathrm{Var}(\hat{\bar{\theta}}_{i_0}), \tag{4.31}$$

and the normalized module be:

$$\rho(\hat{\bar{\theta}}_{i_0}) = \frac{\left(\triangle\bar{\theta}\right)^2}{\mathrm{Var}(\hat{\bar{\theta}}_{i_0})}, \tag{4.32}$$

where $\mathrm{Var}(\hat{\bar{\theta}}_{i_0})$ is the variance of the estimated parameter $\hat{\bar{\theta}}_{i_0}$. is the variation of the estimation . The expression (4.30) is then:

$$\log \beta(\hat{\bar{\theta}}) \cong \log \beta(\bar{\theta}) + \gamma(\alpha, \bar{\theta}) \rho(\hat{\bar{\theta}}_{i_0}). \tag{4.33}$$

From the property of NP-test, we always have $\log \beta(\hat{\bar{\theta}}) \geq \log \beta(\bar{\theta})$ for all $\hat{\bar{\theta}}$ so $\gamma(\alpha, \bar{\theta})$ is always nonnegative. By using (2.41), we can see $\rho(\hat{\bar{\theta}}_{i_0})$ has normalized chi-squared distribution with 1 degree of freedom. Therefore, from (4.33) we show that $\log \beta(\hat{\bar{\theta}})$ follows a shifted and scaled chi-squared distribution. The mean and variance of $\log \beta(\hat{\bar{\theta}})$ for an unbiased estimator $\hat{\bar{\theta}}$ are then:

$$\begin{aligned} \mathbb{E}\left[\log \beta(\hat{\bar{\theta}})\right] &= \log \beta(\bar{\theta}) + \gamma(\alpha, \bar{\theta}), \\ \mathrm{Var}\left[\log \beta(\hat{\bar{\theta}})\right] &= 2\gamma^2(\alpha, \bar{\theta}). \end{aligned} \tag{4.34}$$

Now if we take $100(1-\eta)\%$ of an error spread region $\left[\chi^2_{1,\frac{\eta}{2}}, \chi^2_{1,1-\frac{\eta}{2}}\right]$ for $\bar{\theta}_{i_0}$ satisfying:

$$\begin{aligned} \Pr\left[\rho(\hat{\bar{\theta}}_{i_0}) \leq \chi^2_{1,\frac{\eta}{2}}\right] &= \frac{\eta}{2} \\ \Pr\left[\rho(\hat{\bar{\theta}}_{i_0}) \leq \chi^2_{1,1-\frac{\eta}{2}}\right] &= 1 - \frac{\eta}{2}. \end{aligned}$$

We thus have a corresponding $100(1-\eta)\%$ error spread region for $\log \beta(\hat{\bar{\theta}})$:

$$\left[\log \beta(\bar{\theta}) + \gamma(\alpha, \bar{\theta}) \chi^2_{1,\frac{\eta}{2}} \,,\, \log \beta(\bar{\theta}) \gamma(\alpha, \bar{\theta}) \chi^2_{1,1-\frac{\eta}{2}}\right], \tag{4.35}$$

and derive two critical ROC curves that bound $100(1-\eta)\%$ error spread region for the losses in authentication performance.

Distribution for several estimated parameters

The extension of this analysis for vectors of estimated parameters of size m is more practical and needs to be taken into account. Hopefully, in this case if $\hat{\bar{\theta}}$ is unbiased ML estimator and if we define:

$$\mathcal{H}^* \equiv \frac{N}{2}\mathcal{H} \tag{4.36}$$

then

$$\log \beta(\hat{\bar{\theta}}) \cong \log \beta(\bar{\theta}) + \frac{1}{2}(\hat{\bar{\theta}} - \bar{\theta})^T \mathcal{H}^*(\bar{\theta})(\hat{\bar{\theta}} - \bar{\theta}) \tag{4.37}$$

is called a quadratic form of normal distribution and sometimes its distribution is called generalized chi-squared distribution, say $\mathcal{G}\chi^2_m$, in statistical literature [58, 59, 67, 86, 92]. Generally, its density has no explicit form but it can be approximated numerically in [86, 92]. For such quadratic models, which are extremely popular in financial risk calculation [58], the quantiles can be estimated in [58]. Based on these developments, we can deduce the confidence regions for $\log \beta(\hat{\bar{\theta}})$ which are important to analyze the losses in authentication performance.

We can obtain the explicit formulas for the expectation and variance of the quadratic form (4.37) by using the following proposition:

Proposition 5. *If $\hat{\bar{\theta}}$ is unbiased ML estimator with mean $\bar{\theta}$ and covariance matrix $\Sigma_{\hat{\bar{\theta}}}$, $\mathcal{H}^*$ is a symmetric matrix with constant terms in $\mathbb{R}$, then the expectation and variance of $\log \beta(\hat{\bar{\theta}})$ are*

$$\begin{aligned} \mathbb{E}\left[\log \beta(\hat{\bar{\theta}})\right] &= \log \beta(\bar{\theta}) + \tfrac{1}{2}\mathrm{tr}\left(\mathcal{H}^*\Sigma_{\hat{\bar{\theta}}}\right), \\ \mathrm{Var}\left[\log \beta(\hat{\bar{\theta}})\right] &= \tfrac{1}{2}\mathrm{tr}\left[\left(\mathcal{H}^*\Sigma_{\hat{\bar{\theta}}}\right)^2\right]. \end{aligned} \tag{4.38}$$

Proof. See in Appendix A.3. □

In our case, the $100(1-\eta)\%$ confidence region for $\log \beta(\hat{\bar{\theta}})$ is $\left[\mathcal{G}\chi^2_{m,\frac{\eta}{2}}, \mathcal{G}\chi^2_{m,1-\frac{\eta}{2}}\right]$, where two critical points (or quantiles) $\mathcal{G}\chi^2_{m,\frac{\eta}{2}}$ and $\mathcal{G}\chi^2_{m,1-\frac{\eta}{2}}$ can be computed numerically in [58]. The approximations for the probabilities of $\log \beta(\hat{\bar{\theta}})$ can be provided using the method described in [92].

4.3 Approximation of authentication performance up to the third order

Note that for authentication purposes only, our goal is to obtain a good approximation of the distribution of the error probability as we have done in the previous section.

We are also interested here in analyzing the loss of accuracy due to the rest of the Taylor expansion of $\log \beta(\hat{\bar{\theta}})$ w.r.t. the estimation error, especially the most important term after the second derivative, i.e. the third order derivative. We show in this section that if we take into account the third derivative, we are able to obtain a better match with the asymptotic expression, however its impact is rather marginal compared with the influence of the second order derivative. As a perspective we foresee that the use of the third derivative might lead to a better approximation of the distribution of $\log \beta(\hat{\bar{\theta}})$.

This development help us obtain such a more accurate behavior of $\log \beta(\hat{\bar{\theta}})$ w.r.t to $\rho(\hat{\bar{\theta}})$.

For the sake of simplicity, we start this analysis for only one estimated parameter first and then, as above, we provide an extension for the case of multiple estimated parameters. Precisely, we suppose that there is only unknown parameter $\bar{\theta}_{i_0}$ and the aim is to compute analytically the third partial derivative of $\log \beta(\hat{\bar{\theta}})$ w.r.t $\hat{\bar{\theta}}_{i_0}$.

From (A.12) we have:

$$\frac{\partial^3 \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^3} = \ddot{s}_1^{(i_0)}(\hat{\bar{\theta}}) A(\hat{\bar{\theta}}) + 2\dot{s}_1^{(i_0)}(\hat{\bar{\theta}}) \left[\frac{\partial A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right](\hat{\bar{\theta}}) + s_1(\hat{\bar{\theta}}) \left[\frac{\partial^2 A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right](\hat{\bar{\theta}}),$$

with $\ddot{s}_j^{(i_0)}(\hat{\bar{\theta}}) = \left[\frac{\partial s_j^2(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right](\hat{\bar{\theta}})$ $(j = 0, 1)$. By using the the lemma (1), the third partial derivative of $\beta^*(\hat{\bar{\theta}})$ w.r.t $\hat{\bar{\theta}}_{i_0}$ at the true parameter $\bar{\theta}$ is then

$$\left.\frac{\partial^3 \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^3}\right|_{\hat{\bar{\theta}}=\bar{\theta}} = 2\dot{s}_1^{(i_0)}(\bar{\theta}) \left.\frac{\partial A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right|_{\hat{\bar{\theta}}=\bar{\theta}} + s_1(\bar{\theta}) \left.\frac{\partial^2 A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right|_{\hat{\bar{\theta}}=\bar{\theta}}, \tag{4.39}$$

where

$$\begin{aligned} \frac{\partial^2 A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} &= \left\{\mathbb{E}_{p_{s_1}}\left[l_i'''(\hat{\bar{\theta}})\right] - \mathbb{E}_{p_{s_0}}\left[l_i'''(\hat{\bar{\theta}})\right]\right\} + 2\sum_{v\in\mathcal{V}} l_{i_0}''(\hat{\bar{\theta}}) \left[\frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}} - \frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right] \\ &+ \sum_{v\in\mathcal{V}} l_{i_0}'(\hat{\bar{\theta}}) \left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} - \frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right]. \end{aligned} \tag{4.40}$$

From lemma (1), we also get:

$$\begin{aligned}\left.\frac{\partial^2 A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right|_{\hat{\bar{\theta}}=\bar{\theta}} &= 2\sum_{v\in\mathcal{V}} l''_{i_0}(\bar{\theta})\left[\frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}} - \frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right]_{\hat{\bar{\theta}}=\bar{\theta}} \\ &+ \sum_{v\in\mathcal{V}} l'_{i_0}(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} - \frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right]_{\hat{\bar{\theta}}=\bar{\theta}}.\end{aligned} \tag{4.41}$$

In order to derive an explicit formula for the third partial derivative of $\beta^*(\hat{\bar{\theta}})$ w.r.t $\hat{\bar{\theta}}_{i_0}$ at the true parameter $\bar{\theta}$ we need to compute analytically $\left.\frac{\partial^2 A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right|_{\hat{\bar{\theta}}=\bar{\theta}}$ and $\dot{s}_1^{(i_0)}(\bar{\theta})$. Their expressions are presented in the theorem 8 in Appendix A.4.

Part (v) in Theorem (8) especially highlights the importance of the proposed development on expressing Chernoff bound by using the concept of Boltzmann's distributions. We will see in the numerical part how these developments can be used to fulfill our proposed analysis.

And last but not least, in A.5 of Appendix we discuss about how to approximate $\log\beta(\hat{\bar{\theta}})$ in the case of multiple estimated parameters using Taylor expansion up to the third order. To do this, we have to rewrite the Taylor expansion (4.28), plus the appearance of the third order term, as follow:

$$\begin{aligned}\log\beta(\hat{\bar{\theta}}) &\cong \log\beta(\bar{\theta}) + \tfrac{N}{4}\sum_{i=1}^{m}\sum_{j=1}^{m}\left.\frac{\partial^2\log\beta(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i\partial\hat{\bar{\theta}}_j}\right|_{\hat{\bar{\theta}}=\bar{\theta}}(\hat{\bar{\theta}}_i-\bar{\theta}_i)(\hat{\bar{\theta}}_j-\bar{\theta}_j) \\ &+ \tfrac{N}{12}\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\left.\frac{\partial^3\log\beta(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}\right|_{\hat{\bar{\theta}}=\bar{\theta}}(\hat{\bar{\theta}}_i-\bar{\theta}_i)(\hat{\bar{\theta}}_j-\bar{\theta}_j)(\hat{\bar{\theta}}_k-\bar{\theta}_k)\end{aligned} \tag{4.42}$$

and this reduces to calculate $\left.\frac{\partial^3\log\beta(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}\right|_{\hat{\bar{\theta}}=\bar{\theta}}$. Remind that we let:

$$\beta^*(\hat{\bar{\theta}}) = \frac{2}{N}\log\beta(\hat{\bar{\theta}}),$$

so we just have to find to formula for $\left.\frac{\partial^3\beta^*(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}\right|_{\hat{\bar{\theta}}=\bar{\theta}}$ which is presented in Appendix A.5.

However, it should be emphasized here that until now we cannot derive the probability distribution for the expansion of $\log\beta(\hat{\bar{\theta}})$ up to the third order due to the fact that the cube of a normal distribution is indeterminate (see [20, 95]). It is reminded that the aim of this part is to find a better estimation of the behavior of $\log\beta(\hat{\bar{\theta}})$ w.r.t the variation $\rho(\hat{\bar{\theta}})$ of the estimated parameters.

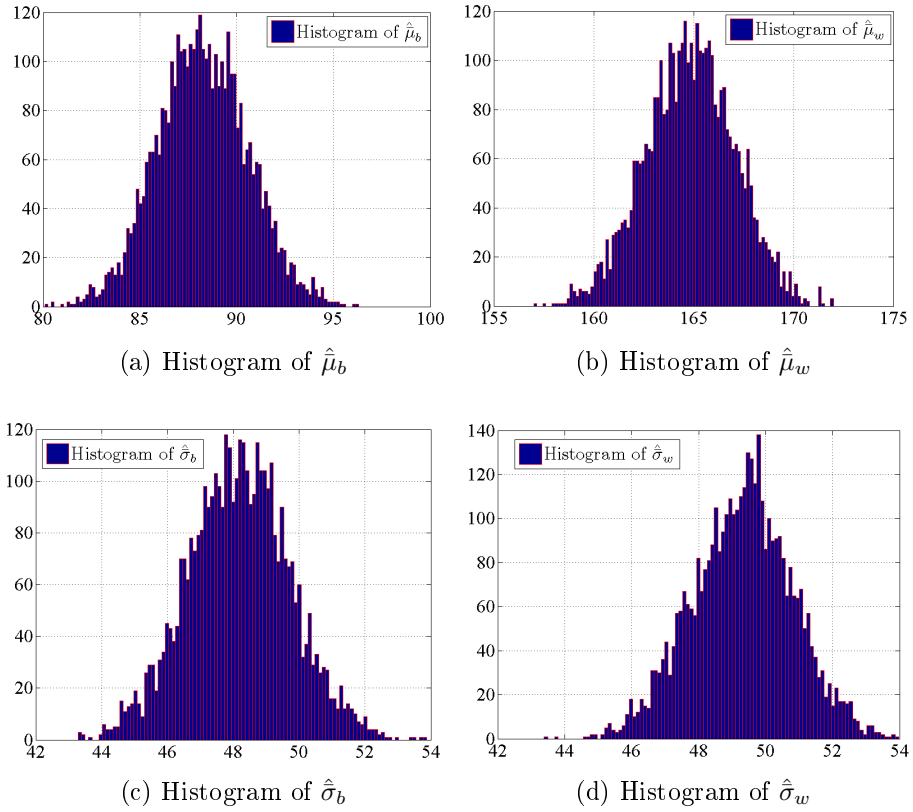
4.4 Numerical results

In order to perform our analysis numerically, we have to construct a MLE scheme for parameter estimation. It is known from the subsection (2.1.3) that the Expectation

Maximization (EM) algorithm is an iterative method for finding maximum likelihood. Without loss of generality we assume that $T_{Z|\hat{X}=0,\bar{\theta}}$ and $T_{Z|\hat{X}=1,\bar{\theta}}$ are modeled by truncated discrete normal distributions with $\bar{\theta} = (\bar{\mu}_b, \bar{\sigma}_b^2, \bar{\mu}_w, \bar{\sigma}_w^2)$ such that $T_{Z|\hat{X},\bar{\theta}}$ is a mixture of two truncated Gaussians (for instance, see (3.7) for $b = 2$). We then develop an EM algorithm for this particular mixture to estimate the set of unknown parameters.

4.4.1 EM algorithms on truncated data

First, in order to interpret why we have to develop and to adapt the EM algorithm to fulfill our analysis, let's see in Fig. 4.1 that if we use the classical EM algorithm for a mixture of two continous Gaussians the results are completely inaccurate even when the number of observation N_{obs} is large. For example, in Fig. 4.1, the estimated values $\hat{\bar{\mu}}_b$ of the true mean of black bits $\bar{\mu}_b = 50$ are mostly on the range $[80, 95]$ and the estimated values $\hat{\bar{\sigma}}_b$ of the true standard deviation $\bar{\sigma}_b = 42$ are mostly on the range $[44, 54]$. Similarly, $\hat{\bar{\mu}}_w \in [157, 172]$ while the true mean of white bits $\bar{\mu}_w = 150$ and $\hat{\bar{\sigma}}_w \in [44, 54]$ while the true standard deviation of white bits $\bar{\sigma}_w = 42$.



(a) Histogram of $\hat{\bar{\mu}}_b$

(b) Histogram of $\hat{\bar{\mu}}_w$

(c) Histogram of $\hat{\bar{\sigma}}_b$

(d) Histogram of $\hat{\bar{\sigma}}_w$

Figure 4.1: Histogram of four estimated parameters $\hat{\bar{\mu}}_b, \hat{\bar{\sigma}}_b, \hat{\bar{\mu}}_w, \hat{\bar{\sigma}}_w$ of true parameters $(50, 150, 42, 42)$ via classical EM algorithm for truncated data. Here, $N_{\text{obs}} = 10^4$, $N_{\text{iter}} = 5.10^3$, parameters of the main channel are $(55, 155, 40, 40)$.

In our setup, the observations, collected from the identified fake codes of opponent's channel, as supposed, are restricted to be integer (grey level) in the range $[0, 255]$. As introduced in (3.2), the probability density function of the opponent channel has the form:

$$\begin{aligned} P_{Z|X,\bar{\theta}}(Z = v \mid X = 0, \bar{\theta}) \\ &= (1 - P_{e,B})T_{Z|\hat{X},\bar{\theta}}(v \mid \hat{X} = 0, \bar{\theta}) \\ &+ P_{e,B}T_{Z|\hat{X},\bar{\theta}}(v \mid \hat{X} = 1, \bar{\theta}) \end{aligned}$$

or

$$\begin{aligned} P_{Z|X,\bar{\theta}}(Z = v \mid X = 1, \bar{\theta}) \\ &= (1 - P_{e,W})T_{Z|\hat{X},\bar{\theta}}(v \mid \hat{X} = 1, \bar{\theta}) \\ &+ P_{e,W}T_{Z|\hat{X},\bar{\theta}}(v \mid \hat{X} = 0, \bar{\theta}). \end{aligned}$$

Let call the mixing weights $\pi_{0,B} = 1 - P_{e,B}$, $\pi_{1,B} = P_{e,B}$ and $\pi_{0,W} = P_{e,W}$, $\pi_{1,W} = 1 - P_{e,W}$. The j-th $(j = 0, 1)$ component truncated density function $T_{Z|\hat{X},\bar{\theta}}(v \mid \hat{X} = j, \bar{\theta}) = \frac{f_j(v|\bar{\theta})}{\sum_{u=0}^{255} f_j(u|\bar{\theta})}$ where $f_j(v \mid \bar{\theta})$ are discrete normal distributions. Recall that $z_{j,n}$ be indicator random variables for the component membership. We can express the log-likelihood functions for complete data of size $N = N_b + N_w$ as

$$\mathcal{L}^{\mathcal{T}}(\bar{\theta}) = \mathcal{L}^{\mathcal{T}}(\bar{\theta} \mid B) + \mathcal{L}^{\mathcal{T}}(\bar{\theta} \mid W) \tag{4.43}$$

where

$$\begin{aligned} \mathcal{L}^{\mathcal{T}}(\bar{\theta} \mid B) &= \sum_{n=1}^{N_b} \sum_{j=0,1} \langle z_{j,n} \rangle^B \left[\log \pi_{j,B} + \log T_{Z|\hat{X},\bar{\theta}}(v_n \mid \hat{X} = j, \bar{\theta}) \right] \\ &= \sum_{n=1}^{N_b} \sum_{j=0,1} \langle z_{j,n} \rangle^B \left[\log \pi_{j,B} + \log f_j(v_n \mid \bar{\theta}) - \log \sum_{u=0}^{255} f_j(u \mid \bar{\theta}) \right] \end{aligned} \tag{4.44}$$

and

$$\begin{aligned} \mathcal{L}^{\mathcal{T}}(\bar{\theta} \mid W) &= \sum_{n=1}^{N_w} \sum_{j=0,1} \langle z_{j,n} \rangle^W \left[\log \pi_{j,W} + \log T_{Z|\hat{X},\bar{\theta}}(v_n \mid \hat{X} = j, \bar{\theta}) \right] \\ &= \sum_{n=1}^{N_w} \sum_{j=0,1} \langle z_{j,n} \rangle^W \left[\log \pi_{j,W} + \log f_j(v_n \mid \bar{\theta}) - \log \sum_{u=0}^{255} f_j(u \mid \bar{\theta}) \right] \end{aligned} \tag{4.45}$$

After some steps of mathematical computation we can have in E-step:

$$\langle z_{j,n} \rangle^B = \frac{\pi_{j,B} T_{Z|\hat{X},\bar{\theta}}(v_n \mid \hat{X} = j, \bar{\theta})}{\sum_{k=0,1} \pi_{k,B} T_{Z|\hat{X},\bar{\theta}}(v_n \mid \hat{X} = k, \bar{\theta})} \tag{4.46}$$

and

$$\langle z_{j,n} \rangle^W = \frac{\pi_{j,B} T_{Z|\hat{X},\bar{\theta}}(v_n \mid \hat{X} = j, \bar{\theta})}{\sum\limits_{k=0,1} \pi_{k,B} T_{Z|\hat{X},\bar{\theta}}(v_n \mid \hat{X} = k, \bar{\theta})} \tag{4.47}$$

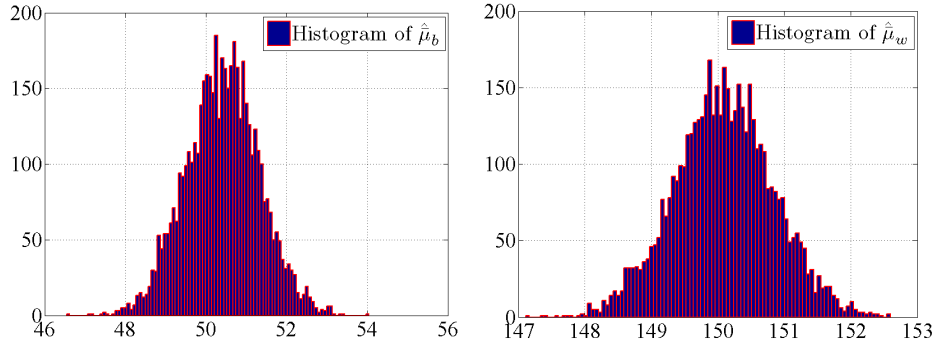
and in M-step the following update rules

$$\begin{aligned} \bar{\mu}_j &= \frac{\sum\limits_{n=1}^{N_b} \langle z_{j,n} \rangle^B v_n + \sum\limits_{n=1}^{N_w} \langle z_{j,n} \rangle^W v_n}{\sum\limits_{n=1}^{N_b} \langle z_{j,n} \rangle^B + \sum\limits_{n=1}^{N_w} \langle z_{j,n} \rangle^W} - M_j^1 \\ \bar{\sigma}_j^2 &= \frac{\sum\limits_{n=1}^{N_b} \langle z_{j,n} \rangle^B (v_n - \bar{\mu}_j)^2 + \sum\limits_{n=1}^{N_b} \langle z_{j,n} \rangle^W (v_n - \bar{\mu}_j)^2}{\sum\limits_{n=1}^{N_b} \langle z_{j,n} \rangle^B + \sum\limits_{n=1}^{N_w} \langle z_{j,n} \rangle^W} - M_j^2 \end{aligned} \tag{4.48}$$

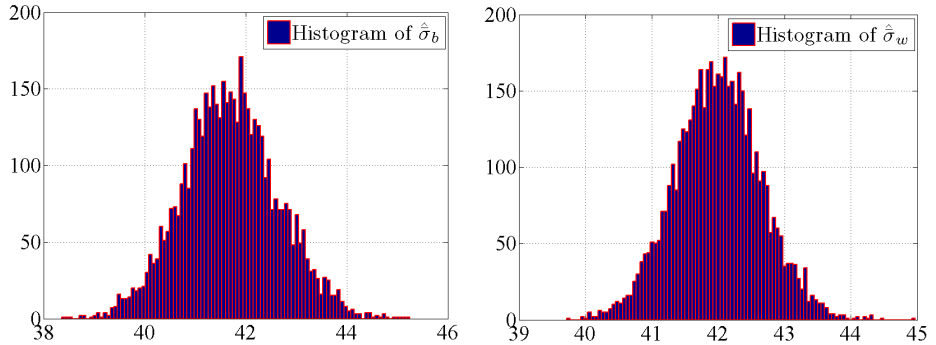
where

$$\begin{aligned} M_j^1 &= \frac{\sum_{u=0}^{255} (u - \bar{\mu}_j) f_j(u|\bar{\theta})}{\sum_{u=0}^{255} f_j(u|\bar{\theta})} \\ M_j^2 &= \frac{\sum_{u=0}^{255} \left[(u - \bar{\mu}_j)^2 - \bar{\sigma}_j^2 \right] f_j(u|\bar{\theta})}{\sum_{u=0}^{255} f_j(u|\bar{\theta})} \end{aligned} \tag{4.49}$$

In Fig. 4.2, we can see that the parameter estimations are accurate: the means of estimated parameters are close to the actual parameters and the standard deviations are not important compared with the range of possible values of the model parameters.



(a) Histogram of $\hat{\bar{\mu}}_b$ (b) Histogram of $\hat{\bar{\mu}}_w$



(c) Histogram of $\hat{\bar{\sigma}}_b$ (d) Histogram of $\hat{\bar{\sigma}}_w$

Figure 4.2: Histogram of four estimated parameters $\hat{\bar{\mu}}_b, \hat{\bar{\sigma}}_b, \hat{\bar{\mu}}_w, \hat{\bar{\sigma}}_w$ of true parameters $(50, 150, 42, 42)$ via modified EM algorithm for truncated data. Here, $N_{\text{obs}} = 10^4$, $N_{\text{iter}} = 5.10^3$, parameters of the main channel are $(55, 155, 40, 40)$.

4.4.2 Fisher information for mixture of truncated discrete normal distribution

In our analysis, in order to compute the mean of $\log \beta(\alpha, \hat{\bar{\theta}})$ and express the error spread region for the authentication performance, it is required to compute the covariance matrix $\Sigma_{\hat{\bar{\theta}}}$ associated with the ML estimators $\hat{\bar{\theta}}$. Using the Cramer-Rao lower bound, it is known that $\Sigma_{\hat{\bar{\theta}}}$ can be approximated by the inverse of the Fisher information $I_m(\bar{\theta})$ of a sequence of N $i.i.d$ random variables conditioned by the vector of parameter $\bar{\theta}$ of size m (see (2.37)). This matrix, in general, cannot be computed explicitly because $\bar{\theta}$ is unknown but it may be approximated by the observed Fisher information $J_m(\hat{\bar{\theta}})$ at the ML estimators $\hat{\bar{\theta}}$ (2.43) when N is sufficiently large. However, in order to conduct our analysis, we suppose that we know the true parameters $\bar{\theta}$ of the main channels. Therefore, we can numerically calculate the actual Fisher information matrices $I_m(\bar{\theta})$ and use it as an approximation for $\Sigma_{\hat{\bar{\theta}}}$.

In our setting, generally we consider $\bar{\theta} = (\bar{\mu}_b, \bar{\sigma}_b^2, \bar{\mu}_w, \bar{\sigma}_w^2)$ hence $m = 4$ and we need to express the formulas for each $I_4(\bar{\theta}_h, \bar{\theta}_k)$ $(h, k = 1, ..., 4)$:

$$I_4(\bar{\theta}_h, \bar{\theta}_k) = I_4^B(\bar{\theta}_h, \bar{\theta}_k) + I_4^W(\bar{\theta}_h, \bar{\theta}_k)$$

where $I_4^B(\bar{\theta}_h, \bar{\theta}_k)$ and $I_4^W(\bar{\theta}_h, \bar{\theta}_k)$ are respectively the (h,k)-entries of the Fisher information matrix $\mathbf{I}_4$ calculated by using N_b observed data from black bits and N_w observed data from white bits. Without loss of generality, we show here the formulas for $I_4^W(\bar{\theta}_h, \bar{\theta}_k)$ and obtain the similar ones for $I_4^B(\bar{\theta}_h, \bar{\theta}_k)$. If we take $g_{\bar{\theta}} \equiv P_{Z|X=1,\bar{\theta}}$ and we can easily find (for $i \in \{0, 1\}$):

$$\begin{aligned} \frac{\partial \log g_{\bar{\theta}}(v)}{\partial \mu_i} &= \frac{\pi_{i,W}}{B_i \sigma_i} \left[\left(\frac{v-\mu_i}{\sigma_i} \right) + \frac{A_i}{B_i} \right] \frac{f_i(v|\bar{\theta})}{g_{\bar{\theta}}(v)} \\ \frac{\partial \log g_{\bar{\theta}}(v)}{\partial \sigma_i^2} &= \frac{\pi_{i,W}}{B_i \sigma_i^2} \left[\frac{1}{2} \left(\frac{v-\mu_i}{\sigma_i} \right)^2 + \frac{C_i}{B_i} - \frac{1}{2} \right] \frac{f_i(v|\bar{\theta})}{g_{\bar{\theta}}(v)} \end{aligned} \tag{4.50}$$

where

$$\begin{aligned} A_i &= f_i(255 \mid \bar{\theta}) - f_i(0 \mid \bar{\theta}) \\ B_i &= \textstyle\sum_{u=1}^{255} f_i(u \mid \bar{\theta}) \\ C_i &= \left(\frac{255-\mu_i}{2\sigma_i^2} \right) f_i(255 \mid \bar{\theta}) + \frac{\mu_i}{2\sigma_i^2} f(0 \mid \bar{\theta}). \end{aligned} \tag{4.51}$$

Let's denote

$$\mathcal{M}_{i_1 j_1}^{(i,j)} = \sum_{u=0}^{u=255} \left(\frac{u-\mu_i}{\sigma_i} \right)^{i_1} \left(\frac{u-\mu_j}{\sigma_j} \right)^{j_1} \frac{f_i(u \mid \bar{\theta}) f_j(u \mid \bar{\theta})}{f(u \mid \bar{\theta})}, \qquad i_1, j_1 \in \{0, 1, 2\},$$

then we have the following formulas after several steps of calculation:

$$\begin{aligned} I_4^W(\mu_i, \mu_j) &= \frac{\pi_{i,W}\pi_{j,W}}{B_i B_j \sigma_i \sigma_j} \mathcal{M}_{11}^{(i,j)} + \frac{A_j \pi_{i,W}\pi_{j,W}}{B_i B_j^2 \sigma_i \sigma_j} \mathcal{M}_{10}^{(i,j)} \\ &+ \frac{A_i A_j \pi_{i,W}\pi_{j,W}}{B_i^2 B_j^2 \sigma_i \sigma_j} \mathcal{M}_{00}^{(i,j)} + \frac{A_i \pi_{i,W}\pi_{j,W}}{B_i^2 B_j \sigma_i \sigma_j} \mathcal{M}_{01}^{(i,j)} \end{aligned} \tag{4.52}$$

$$\begin{aligned} I_4^W(\mu_0, \sigma_i^2) &= \frac{\pi_{0,W}\pi_{i,W}}{2B_0 B_i \sigma_0 \sigma_i^2} \left[\mathcal{M}_{12}^{(0,i)} - \mathcal{M}_{10}^{(0,i)} \right] + \frac{C_i \pi_{0,W}\pi_{i,W}}{B_0 B_i^2 \sigma_0 \sigma_i^2} \mathcal{M}_{10}^{(0,i)} \\ &+ \frac{A_0 \pi_{0,W}\pi_{i,W}}{2B_0^2 B_i \sigma_0 \sigma_i^2} \left[\mathcal{M}_{02}^{(0,i)} - \mathcal{M}_{00}^{(0,i)} \right] + \frac{A_0 C_i \pi_{0,W}\pi_{i,W}}{B_0^2 B_i^2 \sigma_0 \sigma_i^2} \mathcal{M}_{00}^{(0,i)} \end{aligned} \tag{4.53}$$

$$\begin{aligned} I_4^W(\mu_1,\sigma_i^2) &= \frac{\pi_{1,W}\pi_{i,W}}{2B_1B_i\sigma_1\sigma_i^2}\left[\mathcal{M}_{21}^{(i,1)}-\mathcal{M}_{01}^{(i,1)}\right]+\frac{C_i\pi_{1,W}\pi_{i,W}}{B_1B_i^2\sigma_1\sigma_i^2}\mathcal{M}_{01}^{(i,1)} \\ &+ \frac{A_1\pi_{1,W}\pi_{i,W}}{2B_1^2B_i\sigma_1\sigma_i^2}\left[\mathcal{M}_{20}^{(i,1)}-\mathcal{M}_{00}^{(i,1)}\right]+\frac{A_1C_i\pi_{1,W}\pi_{i,W}}{B_1^2B_i^2\sigma_1\sigma_i^2}\mathcal{M}_{00}^{(i,1)} \end{aligned} \tag{4.54}$$

$$\begin{aligned} I_4^W(\sigma_i^2,\sigma_i^2) &= \frac{\pi_{i,W}^2}{4B_i^2\sigma_i^4}\left[\mathcal{M}_{22}^{(i,i)}-2\mathcal{M}_{11}^{(i,i)}+\mathcal{M}_{00}^{(i,i)}\right] \\ &+ \frac{C_i\pi_{i,W}^2}{B_i^3\sigma_i^4}\left[\mathcal{M}_{11}^{(i,i)}-\mathcal{M}_{00}^{(i,i)}\right]+\frac{C_i^2\pi_{i,W}^2}{B_i^4\sigma_i^4}\mathcal{M}_{00}^{(i,i)} \end{aligned} \tag{4.55}$$

$$\begin{aligned} I_4^W(\sigma_0^2,\sigma_1^2) &= \frac{\pi_{0,W}\pi_{1,W}}{4B_0B_1\sigma_0^2\sigma_1^2}\left[\mathcal{M}_{22}^{(0,1)}-\mathcal{M}_{20}^{(0,1)}-\mathcal{M}_{02}^{(0,1)}+\mathcal{M}_{00}^{(0,1)}\right] \\ &+ \frac{C_1\pi_{0,W}\pi_{1,W}}{2B_0B_1^2\sigma_0^2\sigma_1^2}\left[\mathcal{M}_{20}^{(0,1)}-\mathcal{M}_{00}^{(0,1)}\right]+\frac{C_0C_1\pi_{0,W}\pi_{1,W}}{B_0^2B_1^2\sigma_0^2\sigma_1^2}\mathcal{M}_{00}^{(0,1)} \\ &+ \frac{C_0\pi_{0,W}\pi_{1,W}}{2B_1B_0^2\sigma_1^2\sigma_0^2}\left[\mathcal{M}_{02}^{(0,1)}-\mathcal{M}_{00}^{(0,1)}\right]. \end{aligned} \tag{4.56}$$

The quadratic form of the error $\rho(\hat{\bar{\theta}})$ in (2.42) is then:

$$\rho(\hat{\bar{\theta}}) = \left(\hat{\bar{\theta}}-\bar{\theta}\right)^T \mathbf{I}_4(\bar{\theta})\left(\hat{\bar{\theta}}-\bar{\theta}\right) \overset{\text{asym}}{\sim} \chi_4^2 \tag{4.57}$$

which is bounded by two 4-dimensional ellipsoids:

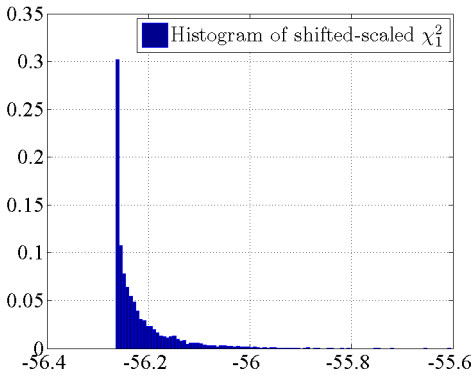
$$\mathcal{R} = \left\{\hat{\bar{\theta}} : \chi_{4,\gamma_1}^2 \leq \rho(\hat{\bar{\theta}}) \leq \chi_{4,\gamma_2}^2\right\}, \tag{4.58}$$

where χ_{4,γ_1}^2 and χ_{4,γ_2}^2 are critical values w.r.t γ_1 and γ_2, i.e., $\Pr[\rho(\hat{\bar{\theta}}) \leq \chi_{4,\gamma_1}^2] = \gamma_1$ and $\Pr[\rho(\hat{\bar{\theta}}) \leq \chi_{4,\gamma_2}^2] = \gamma_2$.

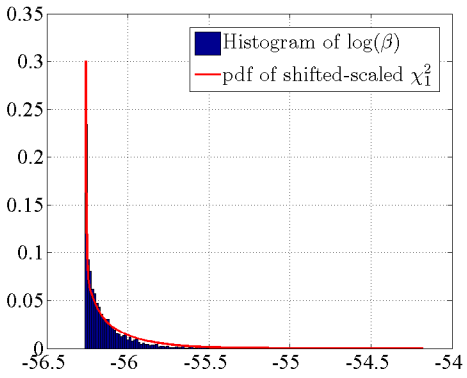
4.4.3 Impact of estimation on authentication performance

First, in case of one parameter, we can show the interplay between the error spread region ρ and the probability of type II error $\log\beta$. Precisely, based on the statistical quantities of ρ, we can derive the corresponding statistical quantities of $\log\beta$ and vice versa.

Figs. (4.3) and (4.4) show that in the case of one estimated parameter, $\log\beta(\alpha,\hat{\bar{\sigma}}_b)$ and $\log\beta(\alpha,\hat{\bar{\sigma}}_w)$ precisely follow a shifed-scaled chi-squared distributions of 1-degree of freedom (see Eq. 4.33), given that $\rho(\hat{\bar{\sigma}}_b)$ and $\rho(\hat{\bar{\sigma}}_w)$ are chi-squared distributions of 1-degree.



(a) Histogram of the second order expansion of $\log \beta(\alpha, \hat{\bar{\sigma}}_b)$ via Monte Carlo sampling.



(b) Histogram of $\log \beta(\alpha, \hat{\bar{\sigma}}_b)$, and the pdf curve via Monte Carlo sampling of the second order expansion.

Figure 4.3: Comparison of histograms of $\log \beta(\alpha, \hat{\bar{\sigma}}_b)$ and its second order expansion. Here, $\theta = (70, 30^2, 160, 30^2)$, $\bar{\theta} = (70, 35^2, 160, 35^2)$, $\bar{\sigma}_b$ is unknown. $N_{\text{obs}} = 8.10^3$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$, $\alpha = 10^{-16}$.

In Fig. 4.5, we suppose that only $\bar{\mu}_w$ is unknown and we run the EM algorithm N_{iter} times using each time N_{obs} observations and obtain a set of $\hat{\bar{\mu}}_w$. The scatter plot of Figure [4.5] represents the computed values of $\log \beta(\alpha, \hat{\bar{\mu}}_w)$ coming from the Asymptotic Expression and is compared with the analytical expression ((4.33)). Here $\rho(\hat{\bar{\mu}}_w) \sim \chi^2_1$, covariance matrix of the estimators becomes the variance $\text{Var}(\hat{\bar{\mu}}_w)$, hence $\text{Var}^{-1}(\hat{\bar{\mu}}_w) \simeq I_1(\bar{\mu}_w)$. The first derivative of log-likelihood ratio $l'_3(\bar{\theta})$ w.r.t $\hat{\bar{\mu}}_w$ at $\hat{\bar{\mu}}_w = \bar{\mu}_w$ is then:

$$\left.\frac{\partial \log P_{Z|X=0,\bar{\theta}}}{\partial \hat{\bar{\mu}}_w}\right|_{\hat{\bar{\mu}}_w = \bar{\mu}_w} = \frac{P_{e,B}}{P_{Z|X=0,\bar{\theta}}} \left[\frac{(v-\bar{\mu}_w)^2}{2\bar{\sigma}_w^4 S_1} - \frac{S_2}{S_1^2}\right] e^{-\frac{(v-\bar{\mu}_w)^2}{2\bar{\sigma}_w^2}} \tag{4.59}$$

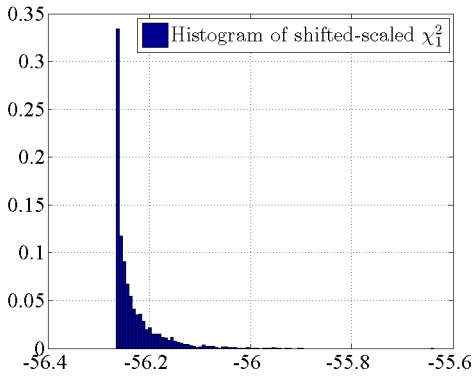
or

$$\left.\frac{\partial \log P_{Z|X=1,\bar{\theta}}}{\partial \hat{\bar{\mu}}_w}\right|_{\hat{\bar{\mu}}_w = \bar{\mu}_w} = \frac{(1-P_{e,W})}{P_{Z|X=1,\bar{\theta}}} \left[\frac{(v-\bar{\mu}_w)^2}{2\bar{\sigma}_w^4 S_1} - \frac{S_2}{S_1^2}\right] e^{-\frac{(v-\bar{\mu}_w)^2}{2\bar{\sigma}_w^2}} \tag{4.60}$$

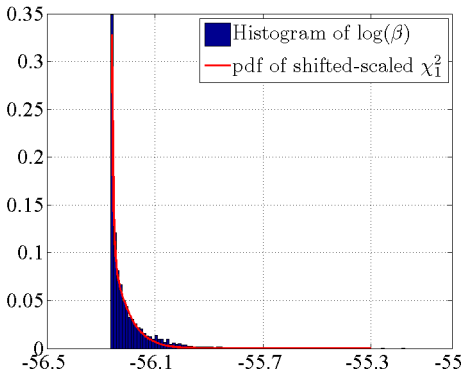
where

$$\begin{aligned} S_1 &= \sum_{u=0}^{255} e^{-\frac{(u-\bar{\mu}_w)^2}{2\bar{\sigma}_w^2}} \\ S_2 &= \sum_{u=0}^{255} \frac{(u-\bar{\mu}_w)^2}{2\sigma_w^4} e^{-\frac{(u-\bar{\mu}_w)^2}{2\bar{\sigma}_w^2}}. \end{aligned} \tag{4.61}$$

The slope of the analytical expression (see Eq. 4.31) of $\log \beta(\alpha, \hat{\bar{\mu}}_w)$ w.r.t $\rho(\hat{\bar{\mu}}_w)$ can be found as:



(a) Histogram of the second order expansion of $\log \beta(\alpha, \hat{\bar{\sigma}}_w)$ via Monte Carlo sampling.



(b) Histogram of $\log \beta(\alpha, \hat{\bar{\sigma}}_w)$, and the pdf curve via Monte Carlo sampling of the second order expansion.

Figure 4.4: Comparison of histograms of $\log \beta(\alpha, \hat{\bar{\sigma}}_w)$ and its second order expansion. Here, $\theta = (70, 30^2, 160, 30^2)$, $\bar{\theta} = (70, 35^2, 160, 35^2)$, $\bar{\sigma}_w$ is unknown. $N_{\text{obs}} = 8.10^3$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$, $\alpha = 10^{-16}$.

$$\gamma(\alpha, \bar{\mu}_w) = \frac{N}{4} \left[\frac{\partial^2 \beta^*(\hat{\bar{\theta}} \mid X = 0)}{\partial \hat{\bar{\mu}}_w^2} \bigg|_{\hat{\bar{\mu}}_w = \bar{\mu}_w} + \frac{\partial^2 \beta^*(\hat{\bar{\theta}} \mid X = 1)}{\partial \hat{\bar{\mu}}_w^2} \bigg|_{\hat{\bar{\mu}}_w = \bar{\mu}_w} \right] \times \text{Var}(\hat{\bar{\mu}}_w). \quad (4.62)$$

For additional comparisons, we represent the statistical linear regression estimated from the set of $\hat{\bar{\mu}}_w$ and expressed as $\log \beta(\alpha, \hat{\bar{\mu}}_w) = 0.1463\rho(\hat{\bar{\mu}}_w) - 44.4586$ with the goodness of fit coefficient 0.9994. We see that the slope of statistical linear regression is quite the same with $\gamma(\alpha, \bar{\mu}_w)$ and hence the analytical line and statistical line coincide.

In Fig. 4.6, we analyze the impact of the estimation error on the ROC curves. We select a 95% of error region for $\hat{\bar{\mu}}_w$, i.e., $\rho(\hat{\bar{\mu}}_w)$ is bounded by two critical levels $\chi^2_{1,0.025}$ and $\chi^2_{1,0.975}$ such that $\Pr[\rho(\hat{\bar{\mu}}_w) \leq \chi^2_{1,0.025}] = 0.025$ and $\Pr[\rho(\hat{\bar{\mu}}_w) \leq \chi^2_{1,0.975}] = 0.975$, and we thus obtain a corresponding 95% error region for $\log \beta(\alpha, \hat{\bar{\mu}}_w)$. We then derive two critical ROC curves $C^{0.025}_{\min}$ and $C^{0.975}_{\max}$ computed analytically from $\chi^2_{1,0.025}$ and $\chi^2_{1,0.975}$ and we choose the mean value for $\rho(\hat{\bar{\mu}}_w)$ to find the mean ROC curve C_{mean}. After that, we compare three analytical ROC curves $C^{0.025}_{\min}$, $C^{0.975}_{\max}$ and C_{mean} with the three ones ("Min ROC", "Max ROC" and "Mean ROC") computed from the dataset of $\rho(\hat{\bar{\mu}}_w)$ (see the legend in the Figs. 4.6, 4.8 and 4.13) and we observe that our approximation is accurate.

In Fig. 4.7, we execute the same analysis as in Fig. 4.5 but now when $\bar{\mu}_b$ is unknown. Similarly, we can find the first derivative of log-likelihood ratio $l'_1(\bar{\theta})$ w.r.t $\hat{\bar{\mu}}_b$ at $\hat{\bar{\mu}}_b = \bar{\mu}_b$ as follow:

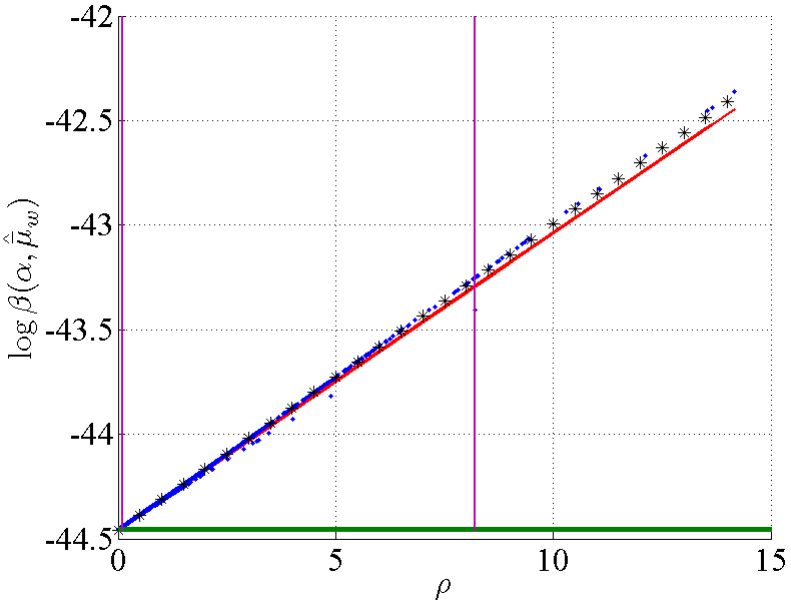


Figure 4.5: Comparison of the log-form of the probability of type II error $\beta(\alpha, \bar{\theta})$ (horizontal line) based on true opponent's parameters $\bar{\theta}$, the probability of type II error $\beta(\alpha, \hat{\bar{\theta}})$ (dots) based on estimated opponent's parameters $\hat{\bar{\theta}}$ with their statistical regression (stars) and the analytical line (straight line) for dataset $\theta = (80, 63^2, 170, 63^2)$, $\bar{\theta} = (85, 65^2, 160, 65^2)$, supposed that only one parameter $\bar{\mu}_w$ is unknown by the receiver, i.e, $\hat{\bar{\theta}} = (85, 65^2, \hat{\bar{\mu}}_w, 65^2)$. The vertical straight and dash line show the critical value $\chi^2_{1,0.025}$ and $\chi^2_{1,0.975}$. Here, $N_{\text{obs}} = 3.10^3$, $\alpha = 10^{-16}$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$, $\rho = \rho(\hat{\bar{\theta}}) = \rho(\hat{\bar{\mu}}_w)$ and $\gamma(\alpha, \bar{\mu}_w) = 0.14186$.

$$\left.\frac{\partial \log P_{Z|X=0,\bar{\theta}}}{\partial \hat{\bar{\mu}}_b}\right|_{\hat{\bar{\mu}}_b = \bar{\mu}_b} = \frac{(1 - P_{e,B})}{P_{Z|X=0,\bar{\theta}}} \left[\frac{(v - \bar{\mu}_b)^2}{2\bar{\sigma}_b^4 S_3} - \frac{S_4}{S_3^2}\right] e^{-\frac{(v-\bar{\mu}_b)^2}{2\bar{\sigma}_b^2}} \tag{4.63}$$

or

$$\left.\frac{\partial \log P_{Z|X=1,\bar{\theta}}}{\partial \hat{\bar{\mu}}_b}\right|_{\hat{\bar{\mu}}_b = \bar{\mu}_b} = \frac{P_{e,W}}{P_{Z|X=1,\bar{\theta}}} \left[\frac{(v - \bar{\mu}_b)^2}{2\bar{\sigma}_b^4 S_3} - \frac{S_4}{S_3^2}\right] e^{-\frac{(v-\bar{\mu}_b)^2}{2\bar{\sigma}_b^2}} \tag{4.64}$$

where

$$\begin{aligned} S_3 &= \sum_{u=0}^{255} e^{-\frac{(u-\bar{\mu}_b)^2}{2\bar{\sigma}_b^2}} \\ S_4 &= \sum_{u=0}^{255} \frac{(u-\bar{\mu}_b)^2}{2\sigma_b^4} e^{-\frac{(u-\bar{\mu}_b)^2}{2\bar{\sigma}_b^2}}. \end{aligned} \tag{4.65}$$

Moreover, the statistical linear regression can be found from the dataset of $\hat{\bar{\mu}}_b$ and expressed as $\log \beta(\alpha, \hat{\bar{\mu}}_b) = 0.1987\rho(\hat{\bar{\mu}}_b) - 40.10725$. We also analyze the impact of the

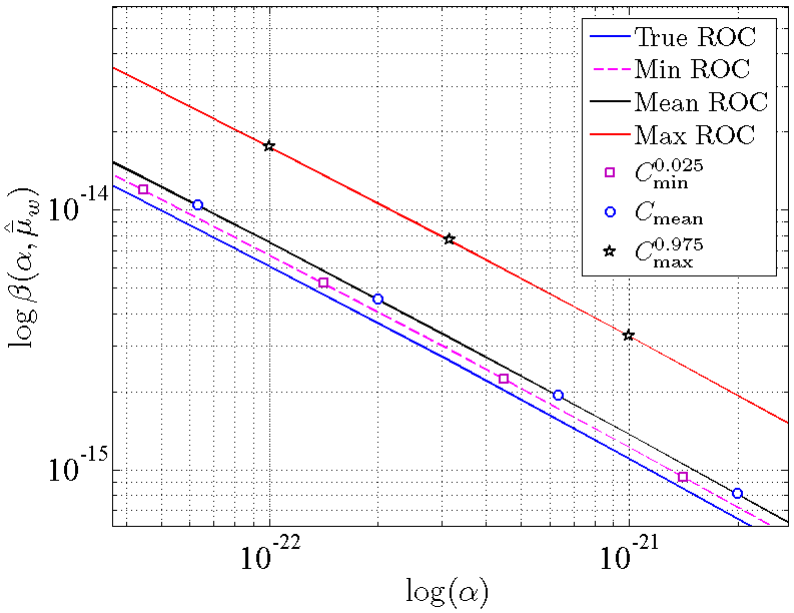


Figure 4.6: Comparison between three analytical ROC curves $C_{\min}^{0.025}$ (squares), $C_{\max}^{0.975}$ (stars) and C_{mean} (circles) with true min ROC curve (dash line), true max ROC curve (red line) and true mean ROC curve (black line) computed directly from $N_{\text{iter}} = 5000$ data of $\hat{\bar{\mu}}_w$.

estimation of $\bar{\mu}_b$ on the ROC curves in Fig. 4.8. Again, in this figure, we can see the accuracy of our analysis.

For the case of multiple estimated parameters, without loss of generality, we suppose that there are four unknown parameters, ie., $\bar{\theta} = (\bar{\mu}_b, \bar{\sigma}_b^2, \bar{\mu}_w, \bar{\sigma}_w^2)$. Here, we want to present the numerical results for the implementation of the CDF and the quantile of $\log \beta(\alpha, \hat{\bar{\theta}})$ in order to analyze the losses in authentication performance.

Generally, in the case of multiple parameters, we cannot theoretically relate the impact of variation of error estimation to the probability of type II error as in the case of one parameter estimation, albeit in Fig. 4.9 when the standard deviations of the opponent channels $\bar{\sigma}_b$, $\bar{\sigma}_w$ are small, we observe that $\log \beta(\alpha, \hat{\bar{\theta}})$ are quite linear w.r.t $\rho(\hat{\bar{\theta}})$.

To see more clearly the difference, Fig. 4.10 shows that the log of probability of type II error $\log \beta(\alpha, \hat{\bar{\theta}})$ is absolutely nonlinear w.r.t the variation $\rho(\hat{\bar{\theta}})$.

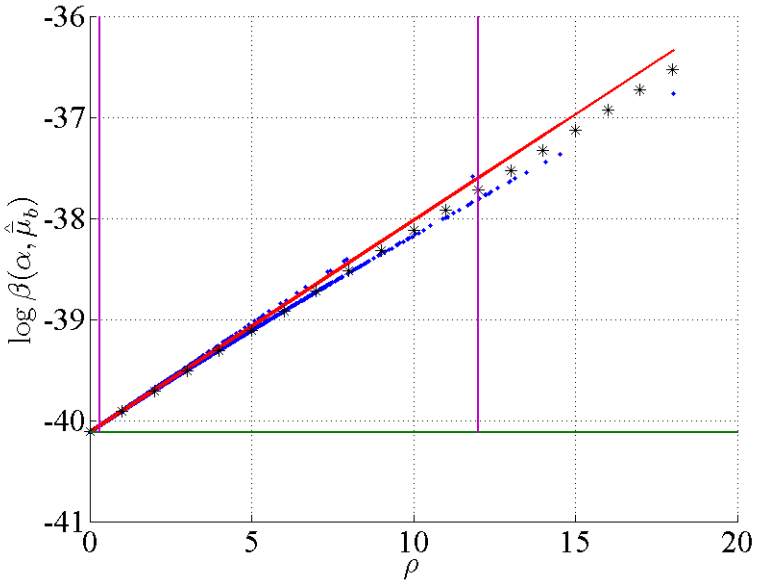
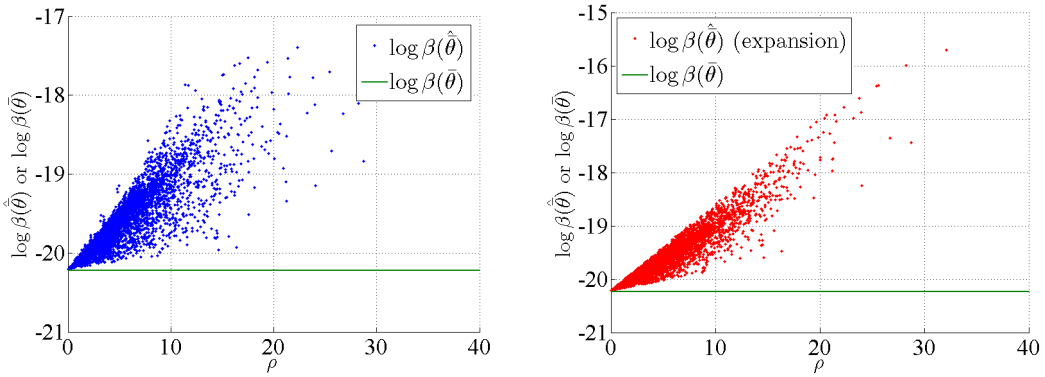


Figure 4.7: Comparison of the log-form of the probability of type II error $\beta(\alpha, \bar{\theta})$ (horizontal line) based on true opponent's parameters $\bar{\theta}$, the probability of type II error $\beta(\alpha, \hat{\bar{\theta}})$ (dots) based on estimated opponent's parameters $\hat{\bar{\theta}}$ with their statistical regression (plus) and the analytical line (straight line) for dataset $\theta = (80, 63^2, 170, 63^2)$, $\bar{\theta} = (85, 65^2, 160, 65^2)$, supposed that only one parameter $\bar{\mu}_b$ is unknown by the receiver, i.e, $\hat{\bar{\theta}} = (\hat{\bar{\mu}}_b, 65^2, 160, 65^2)$. The vertical straight and dash line show the critical value $\chi^2_{1,0.025}$ and $\chi^2_{1,0.975}$. Here, $N_{\text{obs}} = 2.10^3$, $\alpha = 10^{-18}$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$, $\rho = \rho(\hat{\bar{\theta}}) = \rho(\hat{\bar{\mu}}_b)$ and $\gamma(\alpha, \bar{\mu}_b) = 0.20969$.



(a) $\log \beta(\hat{\bar{\theta}})$ via Asymptotic Expression

(b) $\log \beta(\hat{\bar{\theta}})$ via Taylor expansion (4.37)

Figure 4.10: Comparison of the log-form of the probability of type II error $\beta(\alpha, \bar{\theta})$ (horizontal line) based on true opponent's parameters $\bar{\theta}$, the probability of type II error $\beta(\alpha, \hat{\bar{\theta}})$ (dots) based on estimated opponent's parameters $\hat{\bar{\theta}}$. Here we consider $\theta = (80, 63^2, 170, 63^2)$, $\bar{\theta} = (80, 60^2, 170, 60^2)$ and $\hat{\bar{\mu}}_b, \hat{\bar{\sigma}}^2_b, \hat{\bar{\mu}}_w, \hat{\bar{\sigma}}^2_w$ are unknown. $N_{\text{obs}} = 3.10^3$, $\alpha = 10^{-16}$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$.

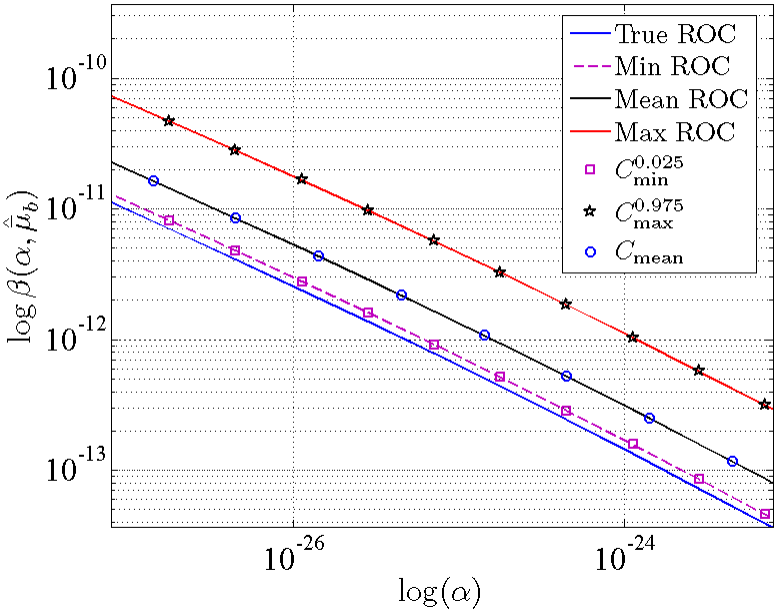
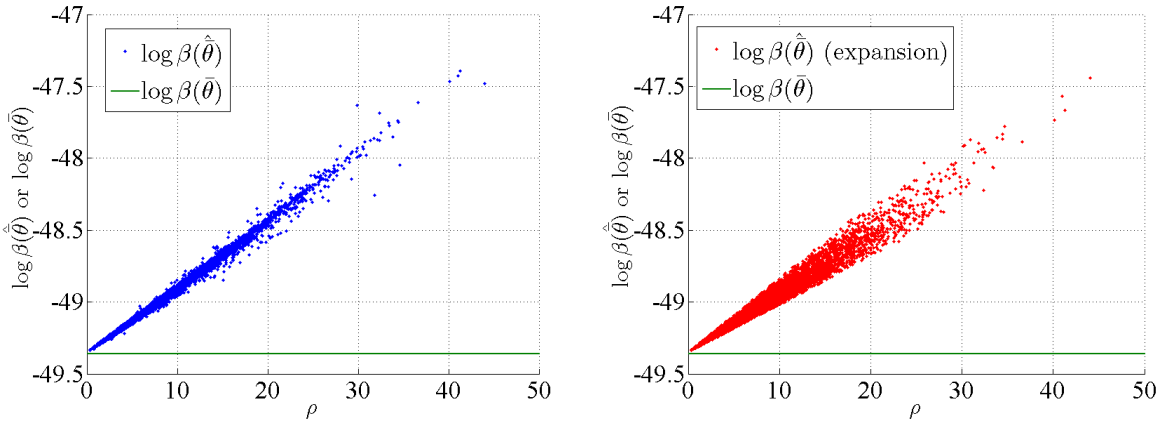


Figure 4.8: Comparison between three analytical ROC curves $C_{\min}^{0.025}$ (squares), $C_{\max}^{0.975}$ (stars) and C_{mean} (circles) with true min ROC curve (dash line), true max ROC curve (red line) and true mean ROC curve (black line) computed directly from $N_{\text{iter}} = 5000$ data of $\hat{\bar{\mu}}_b$.

Hopefully, we can directly derive the confidence region for $\log \beta(\alpha, \hat{\bar{\theta}})$ without using indirectly the role of the total variation $\rho(\hat{\bar{\theta}})$ using [92] and [58].

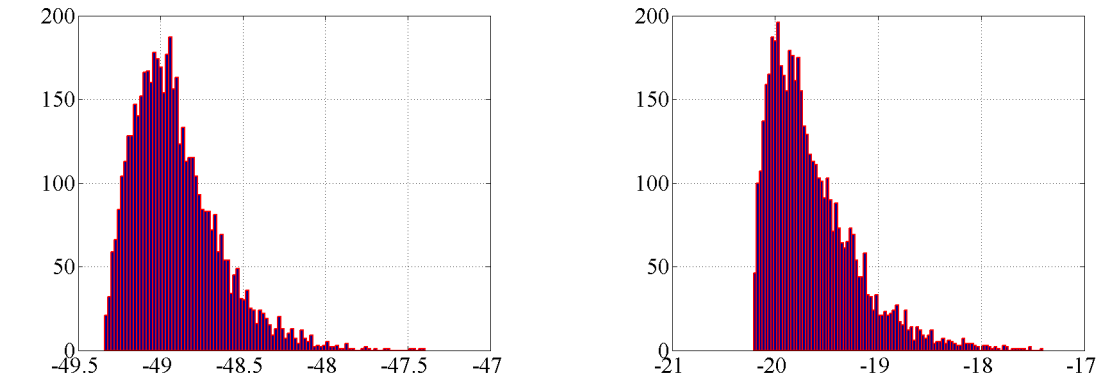
From these results, we can extend our analysis on the ROC curves for one estimated parameter to the analysis of ROC curves in general case of vector of estimated parameters. First, let us show the histogram for $\log \beta(\alpha, \hat{\bar{\theta}})$ in Fig. 4.11 to see how the distribution of $\log \beta(\alpha, \hat{\bar{\theta}})$ looks like.



(a) $\log \beta(\hat{\hat{\theta}})$ via Asymptotic Expression

(b) $\log \beta(\hat{\hat{\theta}})$ via Taylor expansion (4.37)

Figure 4.9: Comparison of the log-form of the probability of type II error $\beta(\alpha, \bar{\theta})$ (horizontal line) based on true opponent's parameters $\bar{\theta}$, the probability of type II error $\beta(\alpha, \hat{\hat{\theta}})$ (dots) based on estimated opponent's parameters $\hat{\hat{\theta}}$. Here we consider $\theta = (55, 40^2, 155, 40^2)$, $\bar{\theta} = (50, 42^2, 150, 42^2)$ and $\hat{\mu}_b, \hat{\sigma}_b^2, \hat{\mu}_w, \hat{\sigma}_w^2$ are unknown. $N_{\text{obs}} = 4.10^3$, $\alpha = 10^{-16}$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$.



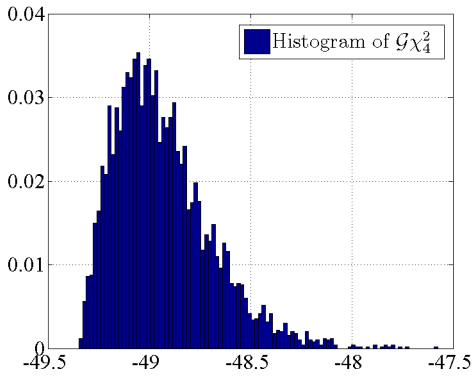
(a) $\theta = (55, 40^2, 155, 40^2)$; $\bar{\theta} = (50, 42^2, 150, 42^2)$

(b) $\theta = (80, 63^2, 170, 63^2)$; $\bar{\theta} = (80, 60^2, 170, 60^2)$

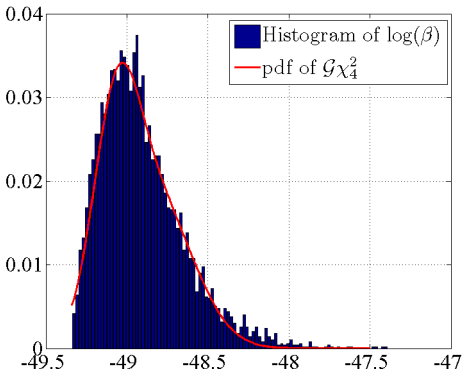
Figure 4.11: Histogram of $\log \beta(\alpha, \hat{\hat{\theta}})$ via Asymptotic Expression. Here, $N_{\text{iter}} = 5.10^3$ and a) $N_{\text{obs}} = 4.10^3$, b) $N_{\text{obs}} = 3.10^3$.

In Fig. 4.12, it can be shown that the distribution of $\log \beta(\alpha, \hat{\hat{\theta}})$ could be accurately approximated by a generalized chi-squared distribution $\mathcal{G}\chi_4^2$ (see Eq. 4.37).

Based on the algorithm proposed in [58], we find that the quantile for $\log \beta(\alpha, \hat{\hat{\theta}})$ in Fig. 4.9 at level 2.5% is $x_{0.025} = -49.2083$ and at 97.5% is $x_{0.975} = -48.35566$. We also obtain the average value $x_{\text{mean}} = -49.01286$ for $\log \beta(\alpha, \hat{\hat{\theta}})$ using Proposition 5. Let F_β



(a) Histogram of the second order expansion of $\log \beta(\alpha, \hat{\bar{\theta}})$ via Monte Carlo sampling (see (4.37)).



(b) Histogram of $\log \beta(\alpha, \hat{\bar{\theta}})$, and the pdf curve via Monte Carlo sampling of the second order expansion.

Figure 4.12: Comparison of histograms of $\log \beta(\alpha, \hat{\bar{\theta}})$ and its second order expansion. Here, $\theta = (55, 40^2, 155, 40^2)$, $\bar{\theta} = (50, 42^2, 150, 42^2)$, $\bar{\theta}$ is unknown. $N_{\text{obs}} = 8.10^3$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$, $\alpha = 10^{-16}$.

be the CDF for $\log \beta(\alpha, \hat{\bar{\theta}})$, we can get back the values for F_β at $x_{0.025}$ and at $x_{0.975}$, based on the algorithm proposed in [92], as:

$$F_\beta(x_{0.025}) = F_\beta(-49.2083) = 0.02432 \approx 0.025$$

and

$$F_\beta(x_{0.975}) = F_\beta(-48.35566) = 0.977808 \approx 0.975.$$

Similarly, in Fig. 4.10 the quantile for $\log \beta(\alpha, \hat{\bar{\theta}})$ at level 2.5% is $x_{0.025} = -19.9603$ and at 97.5% is $x_{0.975} = -18.8928$. The average value of $\log \beta(\alpha, \hat{\bar{\theta}})$ is $x_{\text{mean}} = -19.5283$. And based on [92], it yields to:

$$F_\beta(x_{0.025}) = F_\beta(-19.9603) = 0.02502 \approx 0.025$$

and

$$F_\beta(x_{0.975}) = F_\beta(-18.8928) = 0.975094 \approx 0.975.$$

As in the case of one estimated parameter, we now analyze the impact of the estimation of four parameters in authentication performance in Fig. 4.13based on the dataset used in Fig. 4.10. We also derive the analytical ROC curves $C_{\min}^{0.025}$, $C_{\max}^{0.975}$ and C_{mean} from using $x_{0.025}$, $x_{0.975}$ and x_{mean} based on the quadratic form (4.37) and the computation of the quantile uses a numerical method derived from [58]. We then

compare with the true ones attained from N_{iter} data used in Fig. 4.10, and once again we can see that our analysis is accurate.

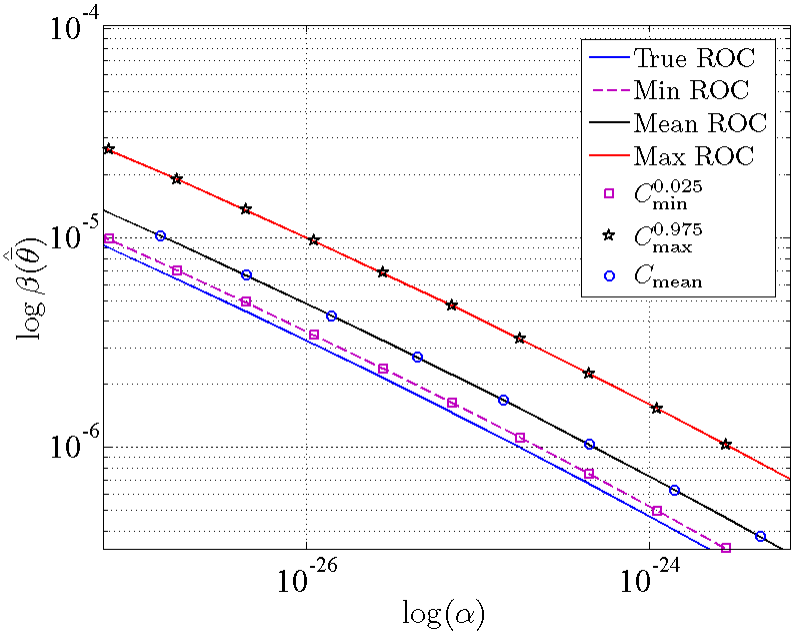


Figure 4.13: Comparison between three analytical ROC curves $C_{\text{min}}^{0.025}$ (squares), $C_{\text{max}}^{0.975}$ (stars) and C_{mean} (circles) with true min ROC curve (dash line), true max ROC curve (red line) and true mean ROC curve (black line) computed directly from $N_{\text{iter}} = 5000$ data used in Fig. 4.10 of four estimated parameters $\hat{\bar{\mu}}_b, \hat{\bar{\sigma}}_b^2, \hat{\bar{\mu}}_w, \hat{\bar{\sigma}}_w^2$.

4.4.4 A more accurate approximation for $\log \beta(\alpha, \hat{\bar{\theta}})$ using third order expansion

In this part, we show the numerical results for the expansion of $\log \beta(\alpha, \hat{\bar{\theta}})$ up to the third order and see how this development fits with the direct computation based on Asymptotic Expression.

In Fig. (4.14), we assume that there is only $\bar{\mu}_w$ in the set of parameters $\bar{\theta}$ needs to be estimated and we will show how a worse estimation for $\bar{\mu}_w$ impacts the authentication performance, and how our development in Theorem (8) can be used to approximate accurately the performance of $\log \beta(\alpha, \hat{\bar{\mu}}_w)$. In Fig. 4.14, the blue dots are the values of $\log \beta(\alpha, \hat{\bar{\mu}}_w)$ using directly Asymptotic Expression while the red dots are the ones using Taylor expansion up to third order of $\log \beta(\alpha, \hat{\bar{\mu}}_w)$ around the true value $\bar{\mu}_w$. We can see two set of dots are very close to each other and this shows that our development is reliable.

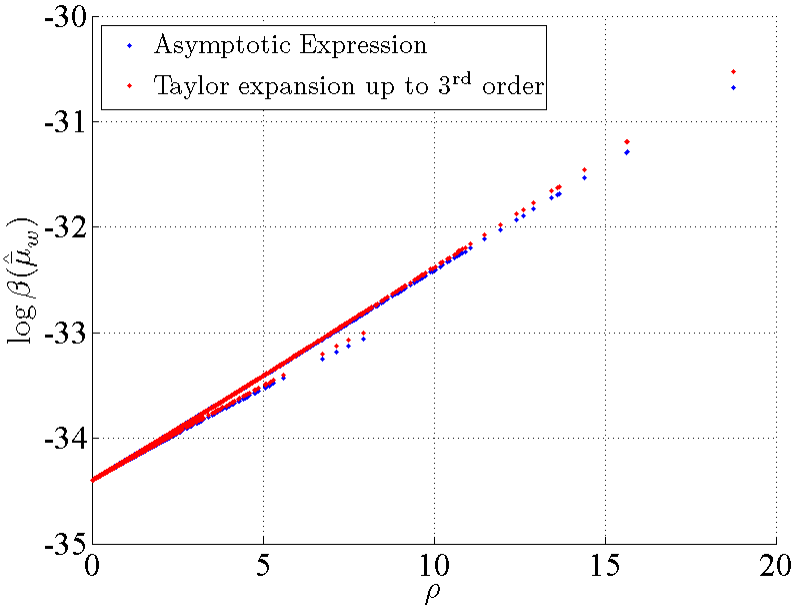


Figure 4.14: Scatter plot of the probability of type II error v.s the Taylor expansion up to the third derivative in case of one estimated parameter $\hat{\bar{\mu}}_w$. The parameters set of the main and opponent channels are dataset $\theta = (80, 63^2, 170, 63^2)$ and $\bar{\theta} = (85, 65^2, 160, 65^2)$. Here we use a fixed probability of type I error $\alpha = 10^{-18}$ and $N_{\text{obs}} = 2.10^3$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$.

A similar analysis is carried out in Fig. 4.15 in the case where $\bar{\sigma}_w$ is unknown.We can notice that the Taylor expansion up to the second order around the true parameter $\bar{\sigma}_w$ (the black line) is not accurate enough to approximate $\log \beta(\alpha, \hat{\bar{\sigma}}_w)$ (blue dots), while the one up to third order (red dots), once again, provides a rather accurate match between the Asymptotic expression and the approximation up to the third derivative.

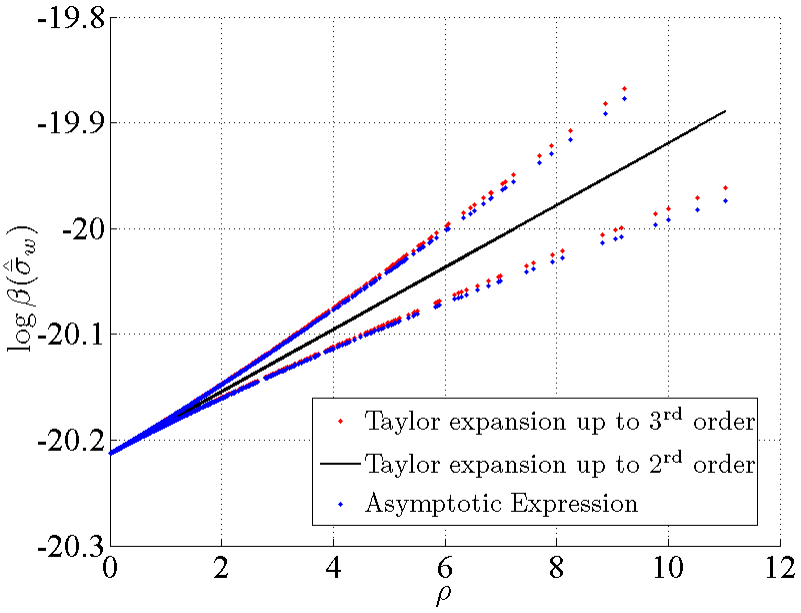


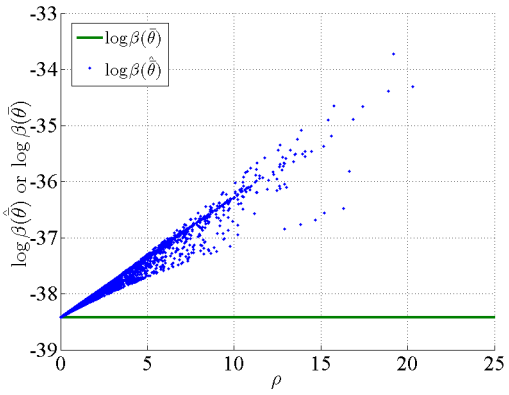
Figure 4.15: Scatter plot of the probability of type II error v.s the Taylor expansion up to the third derivative in case of one estimated parameter $\hat{\bar{\sigma}}_w$. The parameters set of the main and opponent channels are dataset $\theta = (80, 63^2, 170, 63^2)$ and $\bar{\theta} = (80, 60^2, 170, 60^2)$. Here we use a fixed probability of type I error $\alpha = 10^{-16}$ and $N_{\text{obs}} = 2.10^3$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$.

Continue analyzing the accuracy of third order expansion of $\log \beta(\alpha, \hat{\bar{\theta}})$ in case of multiple estimated parameters, without loss of generality, we consider the case when $\bar{\mu}_b, \bar{\mu}_w$ are unknown and need to be estimated in Fig. 4.16. In this case the Taylor expansion has the form:

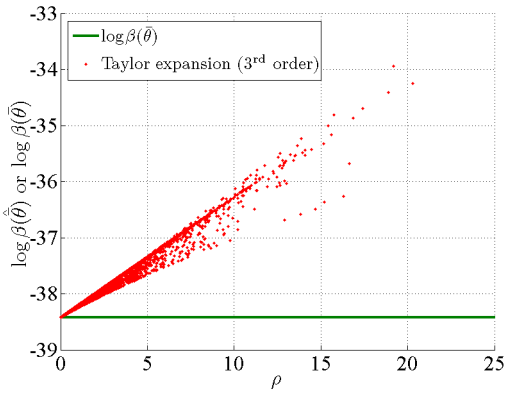
$$\begin{aligned}\log \beta(\hat{\bar{\mu}}_b, \hat{\bar{\mu}}_w) &= \frac{1}{2}\left[\frac{\partial^2\beta}{\partial\hat{\bar{\mu}}_b^2}(a)v_b^2 + 2\frac{\partial^2\beta}{\partial\hat{\bar{\mu}}_b\partial\hat{\bar{\mu}}_w}(a)v_b v_w + \frac{\partial^2\beta}{\partial\hat{\bar{\mu}}_w^2}(a)v_w^2\right] \\ &+ \frac{1}{6}\left[\frac{\partial^3\beta}{\partial\hat{\bar{\mu}}_b^3}(a)v_b^3 + 3\frac{\partial^3\beta}{\partial\hat{\bar{\mu}}_b^2\partial\hat{\bar{\mu}}_w}(a)v_b^2 v_w + 3\frac{\partial^3\beta}{\partial\hat{\bar{\mu}}_b\partial\hat{\bar{\mu}}_w^2}(a)v_b v_w^2 + \frac{\partial^3\beta}{\partial\hat{\bar{\mu}}_w^3}(a)v_w^3\right]\end{aligned} \tag{4.66}$$

where $a = (\bar{\mu}_b, \bar{\mu}_w)$, $v_b = (\hat{\bar{\mu}}_b - \bar{\mu}_b)$, $v_w = (\hat{\bar{\mu}}_w - \bar{\mu}_w)$.

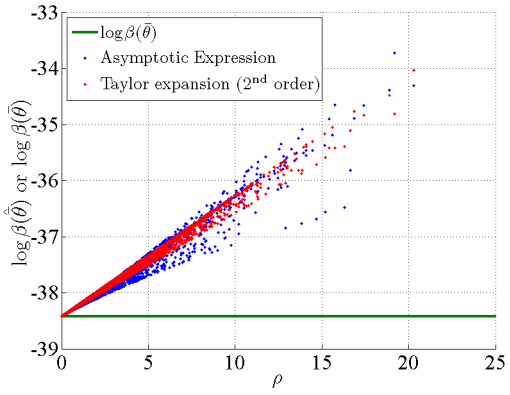
We can obviously see in Fig. 4.16 that the 3rd order expansion approximates the values of $\log \beta(\hat{\bar{\mu}}_b, \hat{\bar{\mu}}_w)$ better than 2nd order expansion.



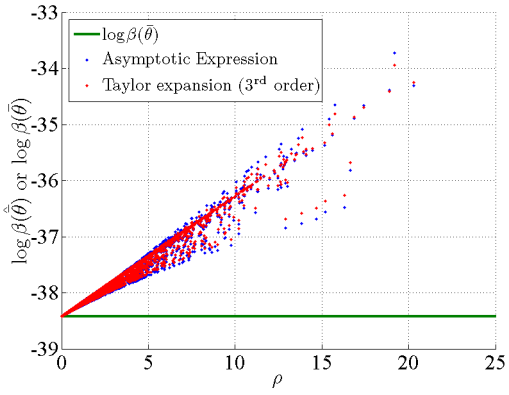
(a) $\log \beta(\alpha, \hat{\hat{\theta}})$ with estimated $\hat{\hat{\mu}}_b$, $\hat{\hat{\mu}}_w$



(b) 3^{rd} order expansion of $\log \beta(\alpha, \hat{\hat{\theta}})$



(c) $\log \beta(\alpha, \hat{\hat{\theta}})$ v.s 2^{nd} order expansion



(d) $\log \beta(\alpha, \hat{\hat{\theta}})$ v.s 3^{rd} order expansion

Figure 4.16: Scatter plot of the probability of type II error v.s the Taylor expansion up to the second and the third order in case of two estimated parameters $(\hat{\hat{\mu}}_b, \hat{\hat{\mu}}_w)$. The parameters set of the main and opponent channels are dataset $\theta = (80, 63^2, 170, 63^2)$ and $\bar{\theta} = (85, 65^2, 160, 65^2)$. Here we use a fixed probability of type I error $\alpha = 10^{-16}$ and $N_{\text{obs}} = 2.10^3$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$.

All figures (4.14), (4.15) and (4.16) show that the development of the third order expansion of $\log \beta(\alpha, \hat{\hat{\theta}})$ is meaningful and necessary to achieve a better approximation for the behavior of $\log \beta(\alpha, \hat{\hat{\theta}})$ w.r.t the estimation error.

4.4.5 Asymptote of authentication performance w.r.t the sample size

Last but not least, we want to show the asymptotic property of authentication performance w.r.t the number of observations. As we've already known that the larger the number of observation, the better the estimation and thus it is obviously to see that the

authentication performance will be better as well. Moreover, this consequently makes the confidence region of $\log \beta(\alpha, \hat{\bar{\theta}})$ more and more smaller. Here, we show this property numerically in Fig. 4.17. Note that Min ROC is denoted for the ROC curve computed at level 2.5%, Max ROC for the one computed at level 97.5% and Mean ROC for the average values computed from using (5) of $\log \beta(\alpha, \hat{\bar{\theta}})$.

To do this, first we run EM algorithm described in (4.4.1) for different numbers of observations to estimate four unknown parameters of the opponent channel. For each N_{obs}, we use the number of iteration $N_{\text{iter}} = 5.10^3$. In order to compute the critical points of $\log \beta(\alpha, \hat{\bar{\theta}})$ at levels 2.5% and 97.5%, we use the same way as we did in Fig. 4.13.

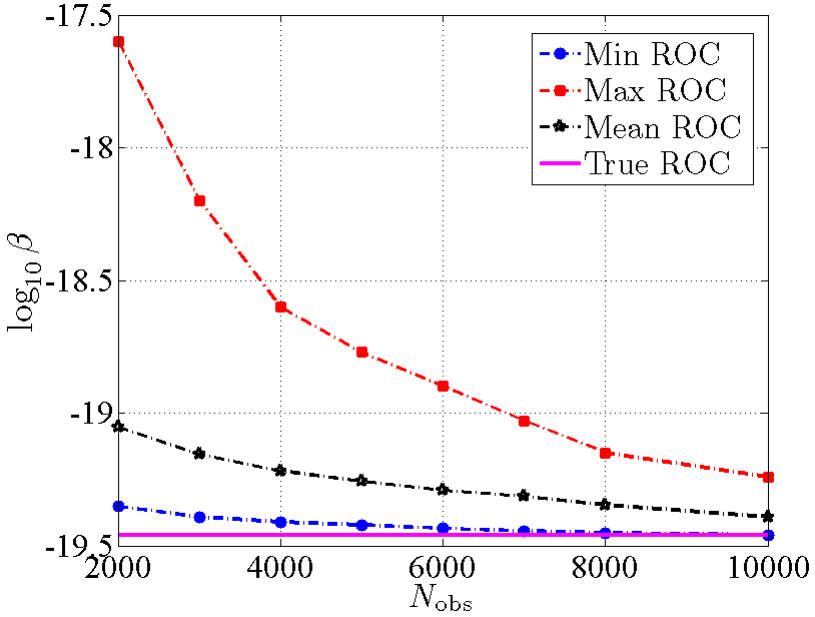


Figure 4.17: The asymptotic property of $\log_{10} \beta(\alpha, \hat{\bar{\theta}})$ w.r.t the number of observations N_{obs}. Here, $\theta = (70, 50^2, 160, 50^2)$ and $\bar{\theta} = (70, 52^2, 160, 52^2)$, $\alpha = 10^{-16}$, $N_{\text{iter}} = 5.10^3$, $N = 2.10^3$.

We can easily see in Fig. 4.17 when N_{obs} is increased, the authentication performance (here, without loss of generality we use $\log_{10} \beta(\alpha, \hat{\bar{\theta}})$ instead of $\log \beta(\alpha, \hat{\bar{\theta}})$) is more and more better. According to this figure, when $N_{\text{obs}} = 10^4$, the Min Roc is quite the same with the the actual ROC curve and the Max ROC curve is very close to it. While in the case of small N_{obs} (for example, $N_{\text{obs}} = 2.10^3$), the confidence region of $\log_{10} \beta(\alpha, \hat{\bar{\theta}})$ is really significant.

4.5 Conclusions of chapter 4

In this chapter we have proposed to analyze the impact of parameter estimation on the authentication performances. This was possible by relying on the asymptotic expression of the error probability presented in the previous chapter, and by computing a Taylor expansion linking the estimation error and the error probability.

The quadratic approximation derived from the expansion enables to compute the distributions of the error probabilities and to derive critical or average Receiver Operating Curves. Moreover, a higher order expansion is considered to approximate better the behavior of the authentication performance w.r.t the estimation error.

We also show practically the asymptotic property of the authentication w.r.t the number of observations used to estimate the opponent parameters. This suggests that if we get a large enough number of observations, the impact of estimation is not important anymore.

The next chapter explain how the accurate expression of the error probabilities can be used in order to optimize the print and scan channel.

Chapter 5

Optimization of the printing-scanning Channels

> "*(Game theory) is essentially a structural theory. It uncovers the logical structure of a great variety of conflict situations and describes this structure in mathematical terms. Sometimes the logical structure of a conflict situation admits rational decisions; sometimes it does not.*"
>
> Anatol Rapoport

In this chapter we propose to cast the problem of authentication of printed documents using binary graphical codes into an optimization game (called min-max game) between the legitimate source and the opponent, each player tries to select the best print and scan channel to minimize or maximize his authentication performance. This game is popularly used in many disciplines such as game theory, economics, engineering and technology, etc. This game is possible when considering accurate computations of

the type I and type II probability errors and by using stochastic processes to model the print and scan channel.

Moreover, we also propose a mathematical interpretation to fulfill the min-max game analytically based on the theory of nonlinear constrained optimization.

5.1 Passive and active opponents

We adopt in this chapter a methodology related to security by considering that the legitimate source or/and the adversary may try to modify their print and scan channel in order to maximize/minimize the authentication performances of the system. Practically this means that the channel can be chosen by using a given quality of paper, an ink of appropriate density and/or by adopting a given resolution. For example if the legitimate source wants to decrease the noise variance, he can choose to use oversampling to replicate the dots, on the contrary if the legitimate source wants to increase the noise variance, he can use a paper of lesser quality. It is important to recall that because the opponent will have to print a binary version of its observation, and because a printing device at this very high resolution can only print binary images, the opponent will in any case have to print with decoding errors after estimation $\hat{X}^N$.

From a mathematical point of view, this game can be considered as an optimization problem where the main goal of the designer of the authentication system (or the sender) is the achievement of the optimal ROC curves, i.e., for a given probability of type I error α, he tries to find a channel that minimizes the probability of type II error β.

We analyze three practical scenarios that are described below:

– The legitimate source and the opponent have identical printing and scanning devices (by devices we mean printer, ink, paper, scanner), practically this means that they use exactly the same setup. Therefore, we can suppose that the set of possible channels used by the opponent is the same than the one used by the legitimate source. In this case the legitimate source will try to look for the channel $\mathcal{C}_{\text{p}}$ such that for a given probability of type I error α, the legitimate party will have a probability of type II error β^* which is the smallest value among all the possible probabilities of type II β involved in the set of channels $\mathcal{C}$, i.e.,

$$\beta^* \equiv \beta(\mathcal{C}_{\text{p}}) = \min_{\mathcal{C}} \beta(\alpha). \tag{5.1}$$

In this case, opponent is defined to be passive and has no strategy but duplicating the graphical code. We can refer this game as a passive game.

– The opponent can modify its printing channel $\mathcal{C}_o$, practically it means that he can modify one or several parameters of his printing setup. As a matter of fact, we can assume that he changes the variance of its noise given that it will be the most efficient way for him to confuse the receiver. The opponent thus tries

to maximize the probability of type II error by choosing his adequate printing channel, whereas the legitimate source will adopt a printing channel $\mathcal{C}_l$ which minimizes the probability of type II error. We end up with so-called a min-max game (or minimax game) in game theory, where the optimal β^* (obtained by a couple of channels $\mathcal{C}_{\mathrm{a}}^l$, $\mathcal{C}_{\mathrm{a}}^o$) is not only the largest value among all the possible probabilities of type II β involved in the set of channels $\mathcal{C}_0$, but also the smallest value among all the possible probabilities of type II β involved in the set of channels $\mathcal{C}_l$, i.e.,

$$\beta^* \equiv \beta(\mathcal{C}_{\mathrm{a}}^l, \mathcal{C}_{\mathrm{a}}^o) = \min_{\mathcal{C}_l} \max_{\mathcal{C}_o} \beta(\alpha). \tag{5.2}$$

Because for this scenario we assume that the receiver has a perfect knowledge of the opponent channel, we will denote it as *an active opponent facing an informed receiver* since the opponent tries to adapt his strategy by selecting exactly $\mathcal{C}_{\mathrm{a}}^o$ in order to degrade the authentication performance. This game can be seen as a deterministic minimax game.

– The last scenario is more general in the sense that we assume here that the receiver will have to estimate the opponent channel. Because of the Kerckhoffs' principle we assume that the opponent knows also this fact and that he will try to maximize a statistic of the type II probability β, such as for example $\mathbb{E}[\beta(\hat{\bar{\theta}})]$, to conduct the game. i.e.,

$$\beta^* = \min_{\mathcal{C}_l} \max_{\mathcal{C}_o} \mathbb{E}\left[\beta(\hat{\bar{\theta}})\right]. \tag{5.3}$$

Here, the scenario presents *an active opponent facing a non-informed receiver*, the term "non-informed" coming from the fact that the receiver needs to estimate the opponent parameters beforehand in oder to compute the hypothesis test. We will use the results obtained from the preposition (5) in the previous chapter in order to conduct this analysis.This game can be seen as a random minimax game..

In order to study the this game rigorously, we would like to develop the problems (5.1) and (5.2) more explicitly in the next section.

5.2 Min-max games as optimization problems

If the problems formulated by equations (5.1), (5.2) and (5.3) can be easily intuitively solved when the opponent vector parameter $\bar{\theta}$ is one dimensional, it is necessary to provide a formulation to find a solution in the N-dimensional case mathematically. We consequently aim to provide the mathematical expressions of the passive game (5.1) and the active deterministic game (5.2). Because the active random game (5.3) is more complicated, we postpone its mathematical analysis for future researches.

We can consider these games as constrained optimization problems. The constraint comes from the fact that in the entire thesis, we suppose that the probability of type I error α is fixed. Taking into account this fact, here we let $\alpha = \alpha_0$ with α_0 is a constant and find the optimization for the probability of type II error β.

We know that we can solve a constrained optimization problem by means of Lagrange multiplier method (see [21], [56]) and for any function $f(x) > 0$, we can look for the optimum using $\log f(x)$ instead of $f(x)$. Therefore, our problem turns to find the optimal $\log \beta$ given fixed $\log \alpha = \log \alpha_0$. Remind that $\log \alpha$ and $\log \beta$ can be approximated from (4.8) and (4.9) and hence we can rewrite:

– For the passive game,

$$\begin{aligned} &\underset{\mathcal{C}}{\text{minimize}} \; -N_c D_{KL}(p_{s_0} \parallel p_1) \\ &\text{subject to} \; -N_c D_{KL}(p_{s_0} \parallel p_0) = \log \alpha_0 \end{aligned} \tag{5.4}$$

– For the active deterministic game,

$$\begin{aligned} &\min_{\mathcal{C}_l} \max_{\mathcal{C}_o} \; -N_c D_{KL}(p_{s_0} \parallel p_1) \\ &\text{subject to} \; -N_c D_{KL}(p_{s_0} \parallel p_0) = \log \alpha_0 \end{aligned} \tag{5.5}$$

where $N_c = \frac{N}{2}$.

First, we start to conduct the constrained optimization problem for the case of passive game. Suppose that θ_0 is such parameter governing the channel $\mathcal{C}$ in the problem (5.4), we consider the Lagrange multiplier function of (5.4) as:

$$F(s_0, \theta_0, \lambda) = -N_c D_{KL}(p_{s_0} \parallel p_1) - \lambda \left[-N_c D_{KL}(p_{s_0} \parallel p_0) - \log \alpha_0 \right]. \tag{5.6}$$

If we put

$$\begin{aligned} f(s_0, \theta_0) &= -N_c D_{KL}(p_{s_0} \parallel p_0) - \log \alpha_0 \\ g(s_0, \theta_0) &= s_0 \mathbb{E}(l_0') + \mathbb{E}\left[\tfrac{\partial \log p_0}{\partial \theta_0} \right] \\ &= s_0 \mathbb{E}\left[\tfrac{\partial \log p_1}{\partial \theta_0} \right] + (1 - s_0) \mathbb{E}\left[\tfrac{\partial \log p_0}{\partial \theta_0} \right] \end{aligned} \tag{5.7}$$

where $l_0' = \frac{\partial l(\theta_0)}{\partial \theta_0}$ and the expectation is taken w.r.t p_{s_0}, then in order to solve (5.4), it necessarily leads to solve the system of nonlinear equations below (see detail computation in Appendix A.6):

$$\begin{cases} f(s_0, \theta_0) = 0 \\ g(s_0, \theta_0) = 0 \end{cases} \tag{5.8}$$

For the deterministic active game, the scenario is more complicated and we cannot take advantage of the Lagrange multiplier. However, because the legitimate channel $\mathcal{C}_l$

and the opponent channel $\mathcal{C}_o$ are independent, we can solve this problem by splitting $\mathcal{C}_l$ into M discreet setups $(\mathcal{C}_l^1, \mathcal{C}_l^2, ..., \mathcal{C}_l^M)$ to perform the minimization and we can use the result of the maximization can be solved by Lagrangian formulation. Therefore, the game (5.5) can be rewritten as:

$$\begin{aligned} \min_{(\mathcal{C}_l^1, \mathcal{C}_l^2, ..., \mathcal{C}_l^M)} \max_{\mathcal{C}_o} \quad & -N_c D_{KL}(p_{s_0} \parallel p_1) \\ \text{subject to} \quad & -N_c D_{KL}(p_{s_0} \parallel p_0) = \log \alpha_0. \end{aligned} \tag{5.9}$$

and the pseudo algorithm can be written as:

Algorithme 5.1 Optimization algorithm for the deterministic active game.

For each legitimate channel $\mathcal{C}_l^i$ do:\\
 Compute $\max_{\mathcal{C}_o} \; -N_c D_{KL}(p_{s_0} \parallel p_1)$ subject to $-N_c D_{KL}(p_{s_0} \parallel p_0) = \log \alpha_0$
 end For
Choose the minimum as the solution

If we assume that θ_1 is the parameter governing the opponent channel $\mathcal{C}_o$we will use the Lagrange multiplier method for each $\mathcal{C}_l^i$. This leads to the fact that we need to solve M corresponding systems of nonlinear equations:

$$\begin{cases} -N_c D_{KL}(p_{s_0} \parallel p_0) - \log \alpha_0 &= 0 \\ \mathbb{E}\left[\frac{\partial \log p_1}{\partial \theta_1}\right] &= 0 \end{cases} \tag{5.10}$$

Numerical methods

Because it is not possible to obtain close form solution of these sets of equations, it is required to use numerical optimizers such as the Newton-Raphson method which is based on the computation of the Jacobian matrix J (see Appendix A.6) and then find J^{-1} to approximate the actual solution through an iterative procedure.

However, sometimes it is difficult to compute J^{-1} because J is close to singular and overflows the number range. We can overcome this issue by using quasi Newton algorithm (see [36, 77]) which approximate J^{-1} through an iterative formula instead of calculating it directly. We can also use function "fsolve" in Matlab toolbox (see [25, 29]) to find the solution for the system of nonlinear equations (5.8) or (5.10).

Another approach for solving passive game is applying a certain class of algorithms which is best known for solving nonlinear constrained optimization problems (see in [85, 99, 75]). Taking into account this advantage, we can practically use "fmincon" and "fminimax" functions from Matlab optimization toolbox in order to find the solution for the games (5.1) and (5.2) (see [25, 29]).

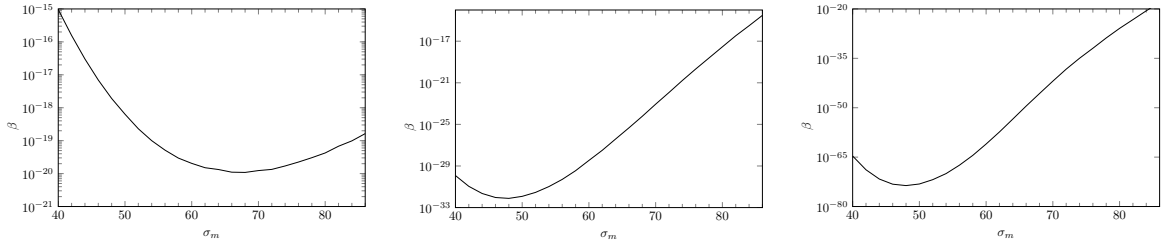
5.3 Numerical results

For the sake of simplicity, in this part, we conduct this analysis for the generalized Gaussian model and Lognormal model, where we assume respectively that the means μ_b, μ_w and the modes M_b, M_w (see Eqs. 3.7 and 3.9) are constant for the main and the opponent channels (which implies that the scanning process has the same calibration for the two types of images). We also assume that the main channel and the opponent channel variances σ^2 and $\bar{\sigma}^2$ are identical for black and white dots.

5.3.1 Passive game

Here the opponent has to undergo a channel identical to the main channel, the only parameter of the optimization problem 5.1 is consequently $\sigma = \sigma_m$. If the generalized Gaussian channels are used, Fig. 5.1 presents the evolution of $\beta(\alpha)$ w.r.t. σ_m for $\alpha = 10^{-6}$ and $\mu_b = 50$, $\mu_w = 150$. For each channel configuration, we can find an optimal configuration, this configuration offers a smaller probability of error for $b = 6$ than for $b = 2$ or $b = 1$. It is not surprising to notice that in each case, β is important whenever σ_m is very small (i.e. when the print and scan noise is very small hence the estimation of the original code is easy) or very large (i.e. when the print and scan noise is so important that the original and forgery become equally noisy).

It is not surprising to notice that in each case $\beta(\alpha)$ is important whenever σ_m is very small, i.e. when the print and scan noise is negligible hence the estimation of the original code by the opponent is easy; or very large; i.e. when the print and scan noise is so important that the original and forgery become equally noisy. The legitimate source will consequently avoid a channel that generates noise of very small or very large variance.



$b = 1$, $\beta^* \approx 10^{-20}$ $b = 2$, $\beta^* \approx 10^{-32}$ $b = 6$, $\beta^* \approx 10^{-74}$

Figure 5.1: Evolution of $\beta(\alpha)$ w.r.t. σ_m ($\alpha = 10^{-6}$) in case of generalized Gaussian distribution. Main and opponent channels are identical, $\mu_b = 50$, $\mu_w = 150$.

Similarly, Fig. 5.2 shows the evolution of the probability of type II error w.r.t the standard deviation in case of using Lognormal channels for different modes.

Fig. 5.3 shows the result of the optimization for different setups in the case of truncated Gaussian distributions when the variance is optimized. We can notice the

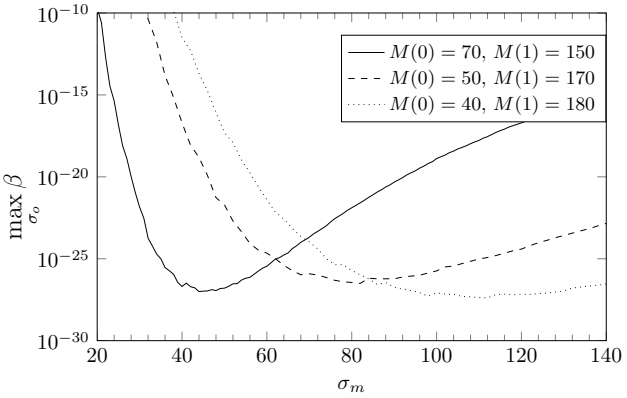
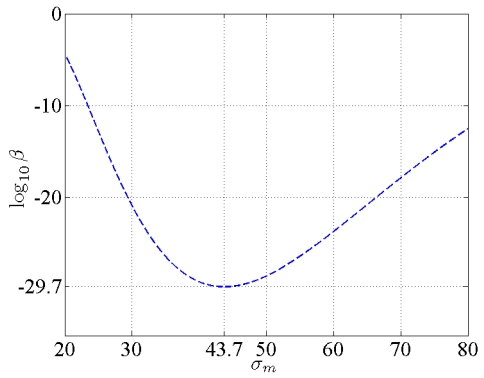
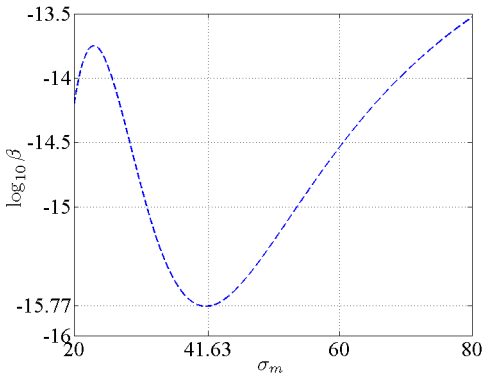


Figure 5.2: Evolution of $\beta(\alpha)$ w.r.t. σ_m ($\alpha = 10^{-6}$) in case of Lognormal distribution for different modes.

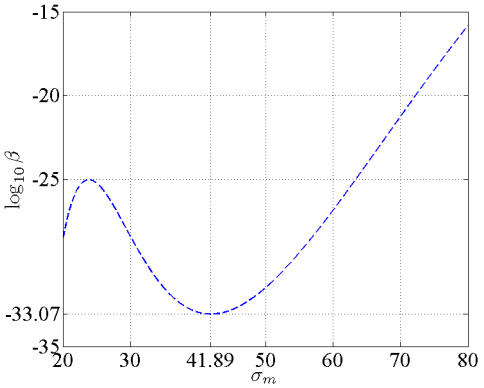
existence of a minimum in each case, however the variations of β w.r.t. σ_m can be rather small, as for the setup proposed in Figure 5.3b.



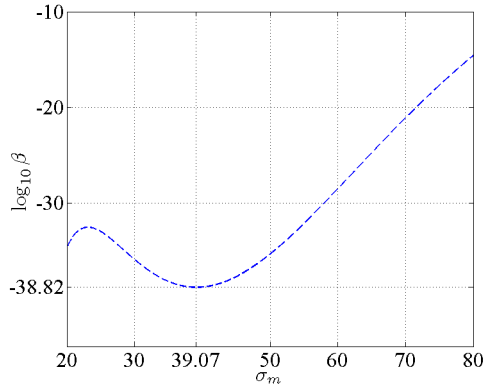
(a) $\mu_b = \bar{\mu}_b = 70$, $\mu_w = \bar{\mu}_w = 160$, $\sigma_m \approx 43.7$, $\beta^* \approx 10^{-29.7}$.



(b) $\mu_b = \bar{\mu}_b = 50$, $\mu_w = \bar{\mu}_w = 150$, $\sigma_m \approx 41.63$, $\beta^* \approx 10^{-15.77}$.



(c) $\mu_b = \bar{\mu}_b = 50$, $\mu_w = 150$, $\bar{\mu}_w = 160$, $\sigma_m \approx 41.89$, $\beta^* \approx 10^{-33.07}$.



(d) $\mu_b = 70$, $\bar{\mu}_b = 80$, $\mu_w = 158$, $\bar{\mu}_w = 160$, $\sigma_m \approx 39.07$, $\beta^* \approx 10^{-38.82}$.

Figure 5.3: Evolution of $\beta(\alpha)$ w.r.t. σ_m ($\alpha = 10^{-6}$) in case of Gaussian distributions for different setups of main and opponent channels.

5.3.2 Active deterministic game

First, the opponent is supposed to be able to use the generalized Gaussian channel with different variance $\bar{\sigma}^2 = \sigma_o^2$ than the main channel $\sigma^2 = \sigma_m^2$ and we try to solve the game defined in (5.2). Fig. 5.4 shows the evolutions of $\beta(\alpha)$ w.r.t σ_o for different σ_m. We can see that in each case it is in the opponent interest to optimize his channel, i.e., he tries to select such channels that maximize the possibility of getting type II error of the receiver.

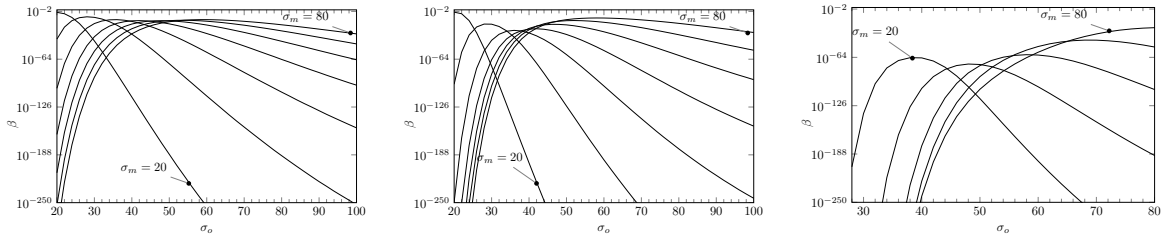
Note that even if we assume that the opponent print and scan channel is perfect ($\hat{X}^N = Z^N$), because the input of the printer has to be binary and because the opponent will make decoding errors by estimating the original code, the copied printed code will be necessarily different from the original printed code (see Figure 3.1), which implies a

perfect discrimination between the two hypotheses.

Fig. 5.5 shows the evolution of best opponent strategy $\max_{\sigma_o} \beta(\alpha)$ w.r.t σ_m. This figure reflects the purpose of the receiver when he wants to minimize the possibility of getting type II error considering the best setup of the opponent. By comparing it with Fig. 5.1, we can see that the opponent's probability of type II error can be multiplied by one or several orders of magnitude ($\times 10^7$ for $b = 1$, $\times 10^5$ for $b = 2$ and $\times 10$ for $b = 6$).

The active scenario offers a saddle point satisfying (5.2) either for generalized Gaussian or Lognormal distribution. This means that even if the opponent owns ideally perfect print and scan devices ($\sigma_0 \to 0$, $o^N = \hat{x}^N$), it is not to his advantage to use it since the authentication is still efficient due to the decoding errors he will create by generating the binary code $\hat{X}^N$.

Another general remark is to notice that the optimal opponent parameters are very close to the optimal parameters of the passive scenario, which means that the opponent has little room to maneuver when choosing his best attack (see Fig. 5.1 and Fig. 5.5, 5.6) and nearly no room when the noise is close to uniform (for example, $b = 6$).



$b = 1$ $b = 2$ $b = 6$

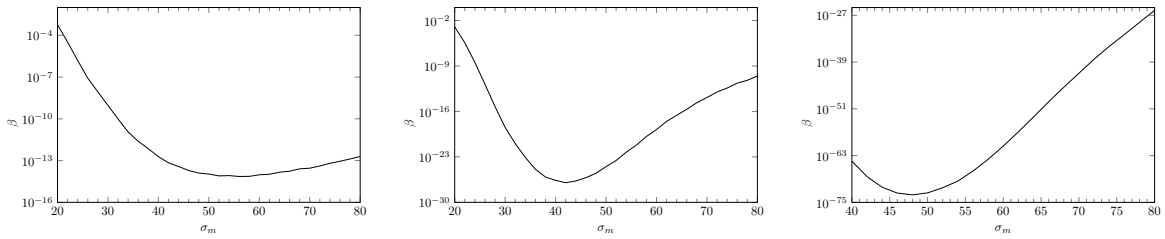
Figure 5.4: Evolution of the probability of type II error $\beta(\alpha)$ w.r.t σ_o for different σ_m in case of generalized Gaussian distribution. The plots arriving from left to right show σ_m varying from 20 to 80 with an increment of 10. Here, $\mu_b = 50$, $\mu_w = 150$, $\alpha = 10^{-6}$.

For the generalized Gaussian distribution, it is important to notice that for distributions of same variance, dense distributions yields to better authentication performance than sparse distributions for both scenarios (see Fig. 5.1 and Fig. 5.5). Unsurprisingly, this is due to the fact that a distribution close to uniform tend to create a bigger overlap between the two decision regions than a sparse distribution that will generate codes mainly lying in the original one.

For the Lognormal distribution we can easily see that the authentication performances are almost similar for different values of modes, both for a passive and an active opponent (see Fig. 5.2 and Fig. 5.6). However, the larger the difference, the larger the optimal standard deviation, which means that it is in the designer strategy to force the opponent to generate decoding errors in this case.

We have shown practically that for both the generalized Gaussian and Lognormal distributions the game can be tractable, and that it is in the interest of the legitimate source to adopt a channel which is close to the uniform distribution.

Figs. 5.7 and 5.8 shows the result of the optimization for different setups in the case of truncated Gaussian distributions when the variance is optimized. It is interesting to notice that in some cases (the setups considered in Fig. 5.8) the local minimum occurs for the minimal standard deviation of the selected range. This happens particularly for the setup where the parameters μ of the legitimate channel and the opponent channel are different, which means that for such setups, this difference between other characteristics of the channel is enough to enable accurate authentication. In this case the optimal strategy is to adopt the print-and-scan system presenting the best fidelity for the legitimate printer, because the opponent differs w.r.t to parameters that he cannot control.



$b = 1$, $\beta^* \approx 10^{-14}$ $b = 2$, $\beta^* \approx 10^{-27}$ $b = 6$, $\beta^* \approx 10^{-73}$

Figure 5.5: Evolution of best opponent strategy $\max_{\sigma_o} \beta$ w.r.t σ_m in case of generalized Gaussian distribution. Here, $\mu_b = 50$, $\mu_w = 150$, $\alpha = 10^{-6}$.

For the Lognormal distribution, we can notice in Fig. 5.6 that the opponent's probability of type II error stays the same when the distribution is close to uniform.

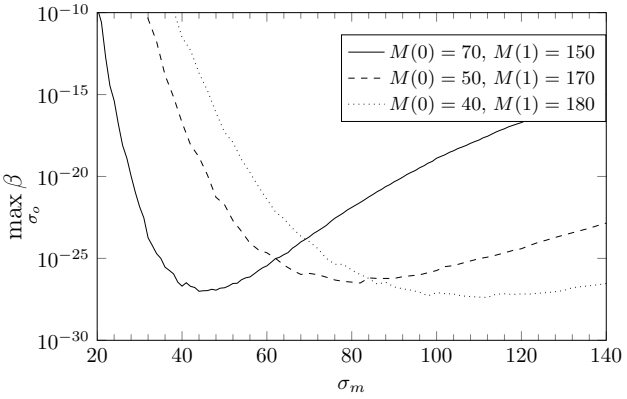
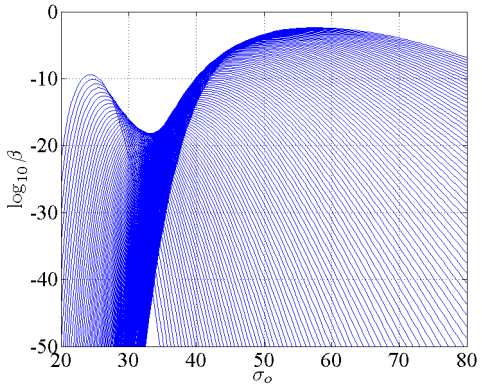
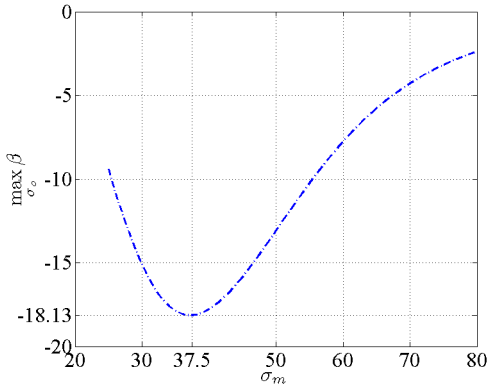


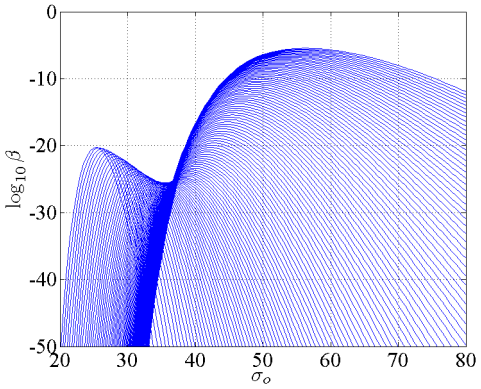
Figure 5.6: Evolution of best opponent strategy $\max_{\sigma_o} \beta$ w.r.t σ_m ($\alpha = 10^{-6}$) in case of Lognormal distribution for different modes.



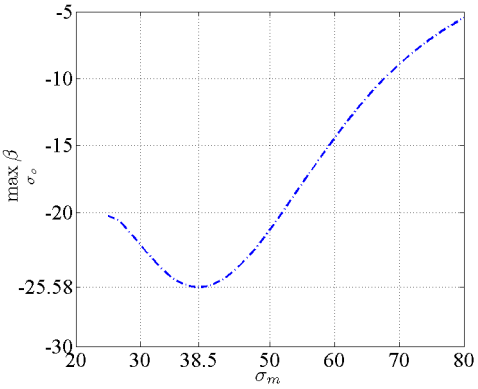
(a) Evolution $\beta(\alpha)$ w.r.t σ_o for different σ_m. Here, $\mu_b = \bar{\mu}_b = 60$, $\mu_w = \bar{\mu}_w = 150$.



(b) Evolution of best opponent strategy $\max_{\sigma_o}\beta$ w.r.t σ_m. Here, $\mu_b = \bar{\mu}_b = 60$, $\mu_w = \bar{\mu}_w = 150$.

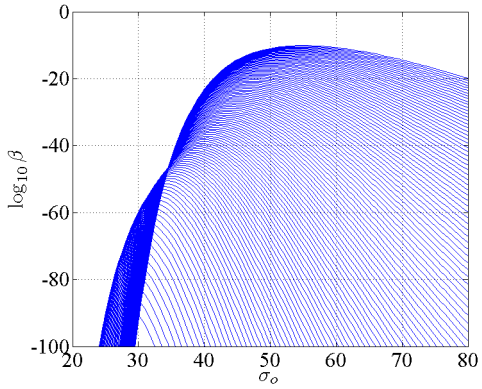


(c) Evolution $\beta(\alpha)$ w.r.t σ_o for different σ_m. Here, $\mu_b = 65$, $\mu_w = 170$, $\bar{\mu}_b = 70$, $\bar{\mu}_w = 160$.

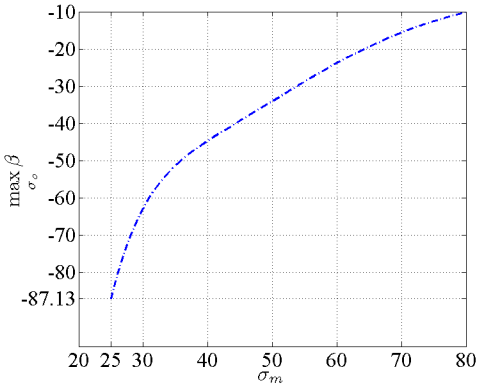


(d) Evolution of best opponent strategy $\max_{\sigma_o}\beta$ w.r.t σ_m. Here, $\mu_b = 65$, $\mu_w = 170$, $\bar{\mu}_b = 70$, $\bar{\mu}_w = 160$.

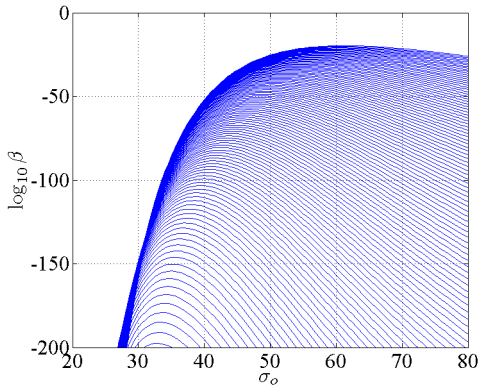
Figure 5.7: Left: Evolution $\beta(\alpha)$ w.r.t σ_o for different σ_m, right: Evolution of the best opponent strategy $\max_{\sigma_o}\beta$. $\alpha = 10^{-6}$.



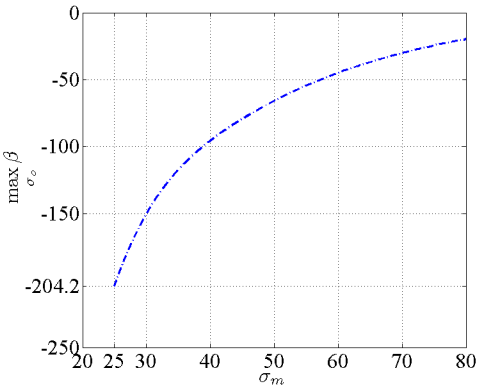
(a) Evolution $\beta(\alpha)$ w.r.t σ_o for different σ_m. Here, $\mu_b = 50$, $\mu_w = 170$, $\bar{\mu}_b = 70$, $\bar{\mu}_w = 160$.



(b) Evolution of best opponent strategy $\max_{\sigma_o}\beta$ w.r.t σ_m. Here, $\mu_b = 50$, $\mu_w = 170$, $\bar{\mu}_b = 70$, $\bar{\mu}_w = 160$.



(c) Evolution $\beta(\alpha)$ w.r.t σ_o for different σ_m. Here, $\mu_b = 50$, $\mu_w = 170$, $\bar{\mu}_b = 80$, $\bar{\mu}_w = 155$.



(d) Evolution of best opponent strategy $\max_{\sigma_o}\beta$ w.r.t σ_m. Here, $\mu_b = 50$, $\mu_w = 170$, $\bar{\mu}_b = 80$, $\bar{\mu}_w = 155$.

Figure 5.8: Left: Evolution $\beta(\alpha)$ w.r.t σ_o for different σ_m, right: Evolution of the best opponent strategy $\max_{\sigma_o}\beta$. $\alpha = 10^{-6}$.

5.3.3 Active random game

In Fig. 5.9, we see the difference between the evolutions of the active deterministic game and the active random game for a setup where the standard deviation σ_m is tuned. It is not surprising to notice that the larger σ_m, the more important the difference between the deterministic and the random game since it is well known that large variances are more difficult to be estimated than small ones.

The difference between these two games is however relatively small, and the optimal parameter is only slightly modified ($\sigma_m \approx 36.5$ for the random game vs $\sigma_m \approx 37$ for the deterministic game).

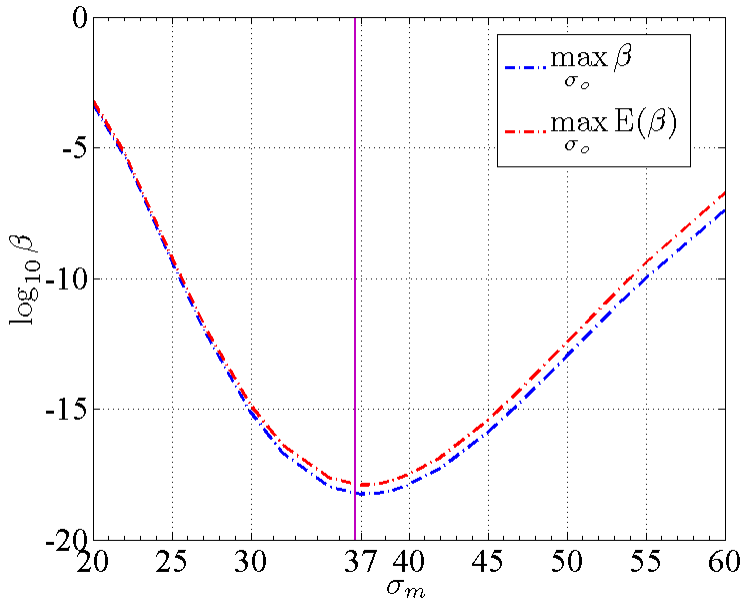


Figure 5.9: Evolution of best opponent strategy $\max_{\sigma_o}\beta$ and $\max_{\sigma_o}\mathbb{E}(\beta)$ w.r.t σ_m in case of Gaussian distribution. Here, $\mu_b = \bar{\mu}_b = 70$, $\mu_w = \bar{\mu}_w = 160$, $\alpha = 10^{-12}$. The horizontal line ($\sigma_m \approx 36.5$) presents the minimum of $\max_{\sigma_o}\mathbb{E}(\beta)$. It is very close to the minimum of $\max_{\sigma_o}\beta$ ($\sigma_m \approx 37$).

We analyze another numerical result in Fig. 5.10 for the case where there is no saddle point and the means of black and white dots are close to each other. Here, the optimal points of the active deterministic game and the active random game are exactly the same.

Similarly to the former example, the difference between the deterministic and the random game is significant for only large σ_m.

In this case, we can notice that the probabilities of type II error of both games gradually goes to 1 when σ_m increases. This fact is not surprising since the means of black and white dots are close to each other, the distributions of black and white dots tend to be very similar for large variances.

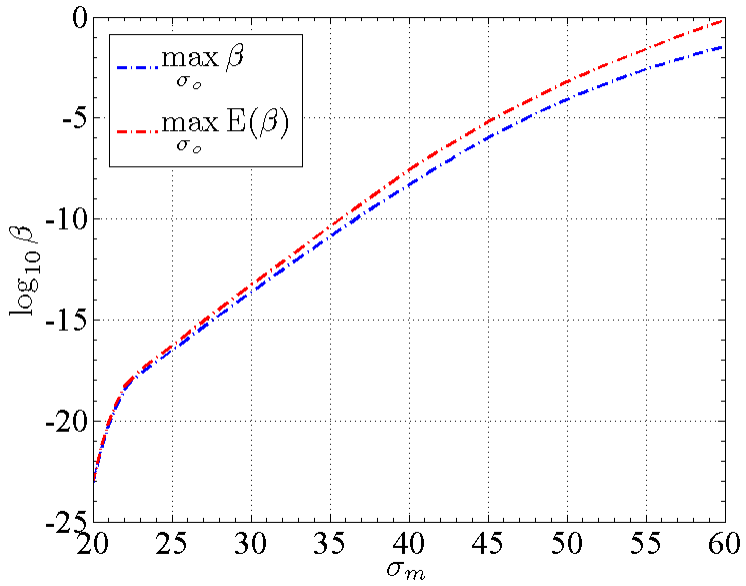


Figure 5.10: Evolution of best opponent strategy $\max_{\sigma_o}\beta$ and $\max_{\sigma_o}\mathbb{E}(\beta)$ w.r.t σ_m in case of Gaussian distribution. Here, $\mu_b = \bar{\mu}_b = 100$, $\mu_w = \bar{\mu}_w = 140$, $\alpha = 10^{-12}$. The minimum of $\max_{\sigma_o}\mathbb{E}(\beta)$ and the minimum of $\max_{\sigma_o}\beta$ are the same.

5.4 Conclusions of Chapter 5

In this chapter, we have modeled the optimization of the print and scan channels as a min-max game between the legitimate receiver and the opponent.

Our first conclusions are the facts that (i) the authentication performance is better for dense noises than for sparse noises for both scenarios, and (ii) for both families of distribution, the opponent optimal parameters are close to the legitimate source parameters, and (iii) the legitimate source can find a configuration which maximizes the authentication performance.

Moreover, the results obtained within the two previous chapters lead us to perform optimization in the more general context that includes the problem of parameter estimation and it's impact on this security game.

Chapter 6

Conclusions and perspectives

"In the end, it's not the years in your life that count. It's the life in your years. "

Abraham Lincoln

We summarize here all the achievements of this thesis and we provide general conclusions and comments regarding the results we have achieved.

Last but not least, we would like also present our current works, several ideas in preparation and sketch out the directions for the future researches.

6.1 Conclusions

This thesis has presented a framework for authentication using graphical codes that is based on statistical hypothesis testing. The main advantages of such a methodology are:

– the possibility to rely on the Neyman-Pearson lemma and to derive optimal tests (i.e. test minimizing a type II error for a given type I error) whenever the model of the legitimate printer and the model of the opponent are known. This strategy has the advantage to be very flexible since it can be used for various distributions and it helps us to compare the performances of different printing models, as proposed in the chapter 5.

- the computation of accurate approximations of the error probabilities based on the asymptotic expression that we have derived in chapter 3. Theses approximations are based on the Chernoff's bound and are useful whenever the error probabilities are very small (e.g. $< 10^{-3}$).

- the fact that, using Boltzmann's distributions, it is possible to approximate the distribution of the error probabilities when the opponent channel has to be estimated beforehand. This analysis derived in chapter 4 enables to compute the average performance of the authentication system or the bounds of a confidence interval, but also to analyze the evolution of theses statistics w.r.t. the number of observed codes.

- the possible optimization of the channel of a legitimate printing facing a passive or an active adversary. In the case where the opponent can tune his channel and the receiver has to estimate the opponent channel, we deal with a random game with an active adversary that is solvable (at least numerically) by using the results of chapter 3 and chapter 4. This analysis enables to point out the categories of distributions that are compatible with our authentication problem.

6.2 Perspectives

In this section, we would like to indicate several objective limitation of our analysis and possible approach that can be used to overcome these drawbacks.

Throughout the thesis, we propose and analyze effectively the classical binary testing for the sequences of $i.i.d$ discrete random variables and its application to the analysis of authentication performance. However, until now we have not come up yet to the case in which graphical codes are composed from such a memory source sequences. One of the possible problem that can be considered is to generate the graphical codes from using a Markov chain. In [74], an asymptotic approximation for the probabilities of type I and type II error is proposed based on the ideas of large deviation principle. In Fig. 6.1, we can implement to obtain the ROC curve of a test between two transition matrices. However, the values of $\log \alpha$ and $\log \beta$ are still very important in comparison with the $i.i.d$ cases. This may be due to the fact that the configurations for the matrices are realistic. Therefore, the study for this problem will be taken into account in an upcoming research.

Another perspective stems from the fact that, due to the assumption of the print and scan process, we have only dealt with the case of discrete distributions for the main and the opponent channels. In order to extend our analysis in this thesis to a larger class of forensic or signal processing problems, it is required to consider a large family of continuous distributions for H_0 and H_1. It leads to look for the approximations schemes for the formulas (4.8) and (4.9). In the literature, there are several authors that have

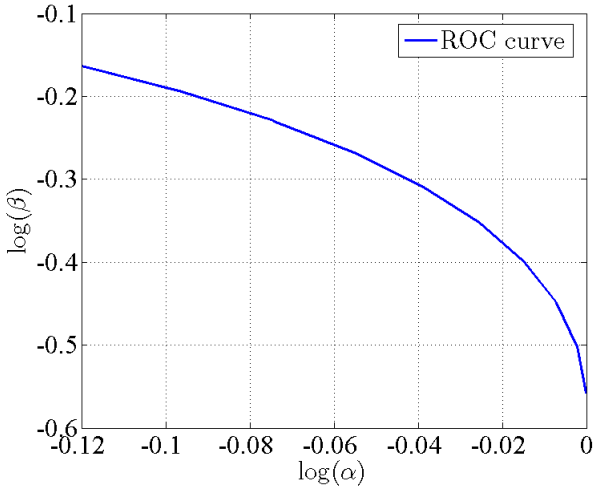


Figure 6.1: ROC curve computed from the test between two finite Markov chains. The size of transition matrices are 500×500. Each row of these matrices is generated by uniform distributions in $[a, b]$ $(a, b > 0)$.

worked on this problem, for example [52, 80]. Based on their works, we believe that we could apply our analysis to many other applications.

In the case when the distributions of the opponent channels are completely unknown by the receiver, our problem somehow is related to the problem of universal hypothesis testing which has already discussed in literature, for example [100, 44]. However, these study are still very deep in theory and have not been used for applications. We hope that we could bring these study to applications in the future research.

As a final remark, a more general perspective is to test the presented methodology on real print-and-scanned graphical codes. This is a complicated problem since in order to derive robust tests, we will need to select more realistic print and scan models (that probably are not i.i.d) and to infer their parameters, but also to test them on real acquisitions that need to be resynchronized and calibrated.

Chapitre 7

Résumé en Français

7.1 Contexte de la thèse

7.1.1 Les enjeux de l'authentification et l'authentification par codes graphiques

Les travaux de cette thèse analysent un système qui vise à endiguer la contrefaçon de documents et d'emballages via l'utilisation d'un code graphique authentifiant ; ce système cherche ainsi à faire face au problème majeur de la falsification.

La motivation pratique est d'importance : à titre d'exemple ce problème est extrêmement préoccupant lorsqu'il s'agit de contrefaçons d'emballages de médicaments, et donc à fortiori de médicaments. En 2005 l'Organisation Mondiale de la Santé a rapporté que 25% des médicaments vendus dans le tiers monde étaient des contrefaçons [4]. Depuis, ce phénomène n'a fait que s'amplifier, notamment via la commande de médicaments par Internet qui rend les possibilités de contrefaçon encore plus importantes [3].

Il existe plusieurs stratégies pour authentifier des contenants ou contenus qu'ils soient numériques ou physiques, nous pouvons par exemple citer :

- l'utilisation de la cryptographie asymétrique, qui permet d'authentifier l'émetteur d'un contenu numérique [35, 41],
- l'utilisation de données biométriques pour authentifier des documents liées à une personne, comme par exemple les passeports.
- l'utilisation de techniques de tatouage pour authentifier par exemple des images numériques [55]. Dans ce cas particulier il s'agit plus précisément d'un contrôle d'intégrité, c'est à dire à une vérification que le contenu n'a pas été modifié.
- enfin, l'utilisation de fonction physiques non clônables (PUF) qui proviennent souvent de circuits électroniques, et qui sont utilisées pour authentifier des matériels [94].

Le procédé que nous étudions dans cette thèse est semblable aux PUF car il repose sur l'utilisation d'un procédé non clônable puisque relevant d'un processus stochastique. Il s'agit ici d'un processus d'impression-acquisition d'un code graphique 2D imprimé à une définition très élevée (autour de 2400 points par pouce). Il permet alors d'authentifier n'importe quel support papier comme l'étiquette d'une bouteille de vin, une boite de médicaments, une facture ou un billet de banque.

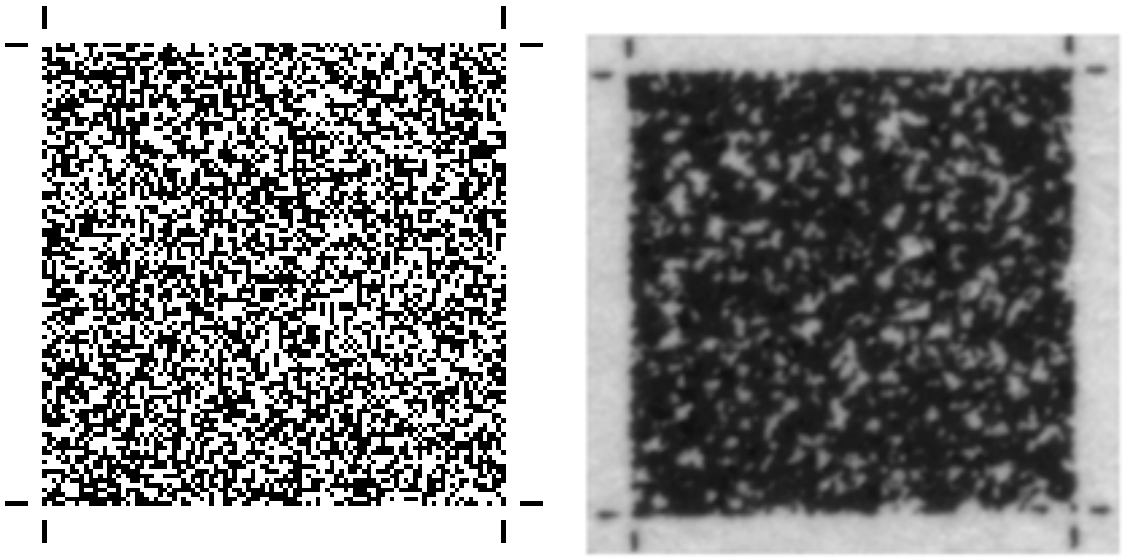
7.1.2 Système étudié

Cette thèse s'inscrit dans le cadre du projet ANR Estampille [7] et le système étudié est proche du système d'authentification développé par l'entreprise Advanced Track & Trace. Le principe de fonctionnement de ce schéma d'authentification est illustré sur la figure 7.2. Il peut se décomposer en 3 étapes :

- Etape 1 : Un code graphique, représenté par une matrice 2D binaire est généré à partir d'une clé secrète, puis est imprimé sur le contenu à protéger. Le processus d'impression peut être considéré comme non inversible, c'est à dire qu'il est impossible de reconstituer le code graphique binaire original à partir d'un code imprimé puis scanné. Cette propriété est essentielle et permet l'authentification : elle est due au fait que le code une fois imprimé subit un processus stochastique de part les propriétés aléatoires des particules composant l'encre à l'échelle microscopique (chaque point du code mesurant autour de $10\mu m$), mais aussi des procédés d'impressions, des positionnements aléatoires des fibres de papiers (voir figure 7.3), et de l'acquisition par un capteur numérique (voir figure 7.1).

- Etape 2 : Le contrefacteur (aussi appelé adversaire) de son coté va également chercher à dupliquer le code graphique. Pour cela il doit dans un premier temps générer une version binaire à partir de l'observation du code imprimé. C'est étape est obligatoire car l'imprimante utilisant sa résolution native, chaque élément du code binaire étant représenté par un point d'impression. Dans cette configuration l'imprimante ne peut imprimer qu'à partir d'une version binaire du code, chaque bit indiquant la présence ou l'absence d'un point d'impression. Puisque le procédé d'impression n'est pas inversible, le code binaire généré sera différent du code original. Cette étape se termine par une nouvelle impression du code graphique binaire généré par le contrefacteur, qui subit comme pour l'étape 1 un processus d'impression stochastique.

- Etape 3 : Le receveur (c'est à dire la personne qui authentifie les contenus), compare le code original, qu'il génère à nouveau au moyen de la clé secrète, au code observé provenant d'un code original ou d'un code contrefait. Comme nous le verrons par la suite, un test d'hypothèse peut lui permettre de décider si le code observé est authentique ou s'il provient d'une copie. Dans notre étude ce test est basé sur la connaissance à priori des modèles statistiques liées aux processus d'acquisition et de copie du code (voir section (3.1)).



(a) (b)

Figure 7.1: Effet du processus d'impression acquisition : (a) Code graphique avant impression. (b) Code graphique après impression (les segments autour des coins sont utilisés pour des besoins de synchronisation).

7.1.3 Liens avec les travaux existants

D'autres systèmes d'authentification sont semblables, le plus proche étant celui proposé par Picard *et al.* [82, 83] . Ce système diffère essentiellement par le test d'hypothèse qui est basé dans ce cas ci sur le comptage du nombre d'erreurs des codes acquis puis binarisés. Nous montrons dans la section 3.3 que cette stratégie n'est pas optimale au sens de l'authentification.

La sécurité de ce système a été étudiée par Baras et Cayre [18] dans le cadre d'une attaque par lots (attaque où l'adversaire essaye d'estimer le code graphique original à partir d'un lot de code graphiques imprimés). Les auteurs montrent que, même si le système est sensible à ce type d'attaques et que les performances d'authentification sont moindres, la présence d'éléments déterministes mais non inversibles dans le système d'impression permet de garantir la sécurité de ce système. Une autre étude menée par Diong et al. [37] a cherché à inférer la fonction permettant à partir d'éléments du code imprimé de retrouver des éléments du code original. Cette étude a montré qu'il est possible de diminuer l'erreur d'estimation par rapport à l'utilisation d'une binarisation optimale, sans pour autant garantir un taux d'erreur nul. Le caractère non-inversible du procédé d'impression-acquisition n'a donc pas pu être remis en cause.

Il existe également d'autres systèmes d'authentification similaires, mais reposant sur des supports différents. Dans [91] les auteurs proposent un procédé d'authentification utilisant la gravure laser des métaux comme procédé non inversible. Dans [51] le procédé d'authentification repose sur l'enregistrement des structures aléatoires des fibres de

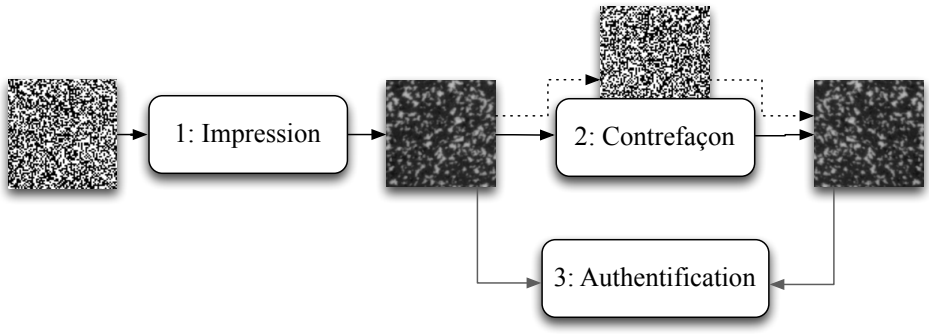


Figure 7.2: Description of the works of Estampille project

papier. Il est a noté toutefois que ce procédé nécessite l'appel à une base de données contenant l'ensemble des structures à authentifier alors que le procédé étudié dans cette thèse ne requière que la connaissance de la clé servant à générer le code original pour effectuer l'authentification.

Comme nous le verrons dans les sections 7.2 et 7.3, cette thèse repose sur l'utilisation d'une part de la théorie des tests d'hypothèses statistiques via l'utilisation du rapport de vraisemblance et d'autre de la théorie de l'estimation via l'utilisation de méthodes d'estimations classiques comme le maximum de vraisemblance ou l'algorithme "Expectation Maximization".

Notre méthodologie se rapproche de celle utilisée par [102] en stéganalyse et [96] en identification de modèles de capteurs, à savoir l'utilisation de tests d'hypothèses statistiques et de modélisation de procédés pour détecter dans le premier cas la présence d'information cachée et dans le second identifier un modèle d'appareil photo numérique.

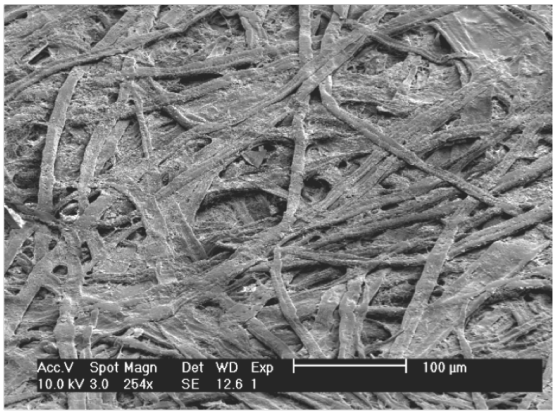


Figure 7.3: Caractère aléatoire des fibres de papier vues sous microscope. [6, 28].

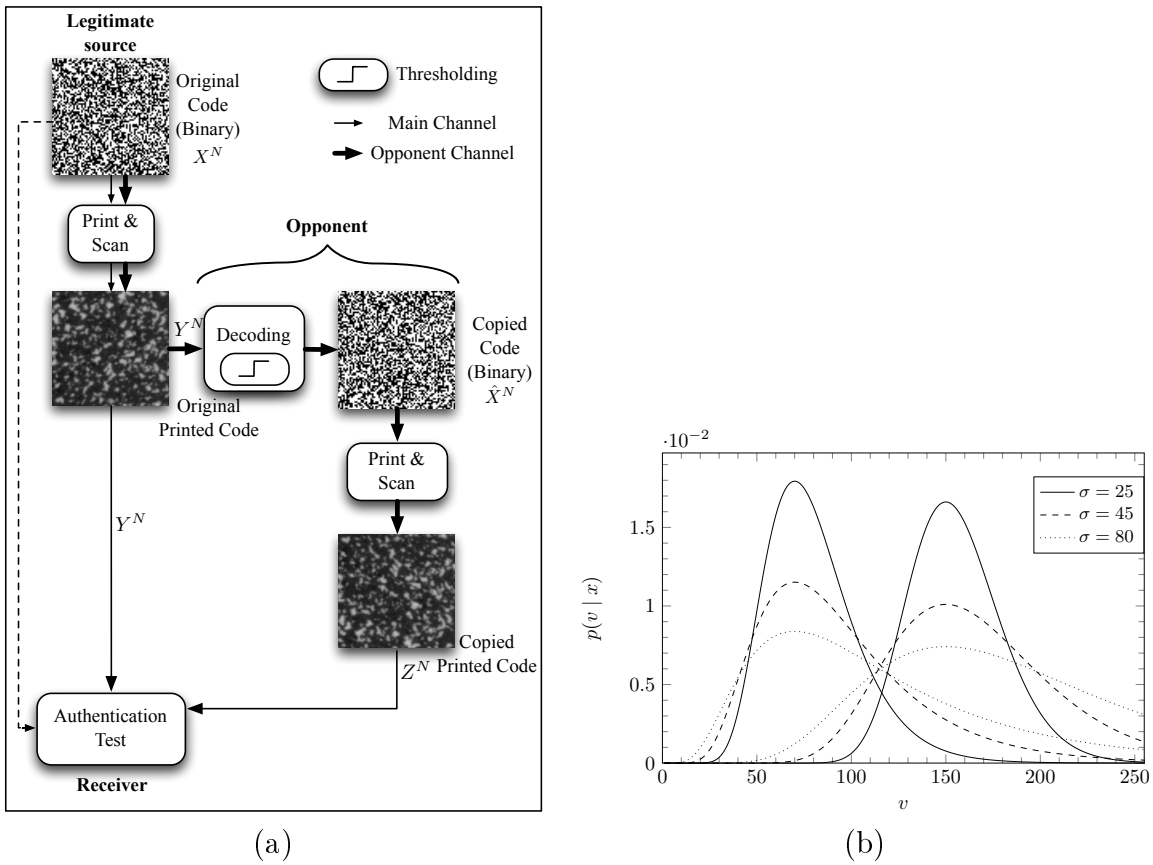


Figure 7.4: (a) : Authentification en utilisant des codes graphiques. (b) Modèles d'impression-acquisition pour les points d'encres (sur la gauche) et les parties blanches du papier pour des distributions Lognormales.

7.2 Authentification et tests d'hypothèses

7.2.1 Principes du système d'authentification

Dans un premier temps, nous sommes partis de l'hypothèse que le receveur a en sa connaissance :

- d'une part le modèle statistique d'impression-acquisition des codes imprimés originaux $P_{Y^N|X^N}\left(v^N \middle| x^N\right)$ où X^N est le vecteur aléatoire représentant le code original et Y^N le vecteur aléatoire représentant le code original imprimé,

- d'autre part le modèle statistique d'impression-acquisition des codes contrefaits $P_{Z^N|X^N}\left(v^N \middle| x^N\right)$ où Z^N est le vecteur aléatoire représentant le code contrefait.

Ces deux hypothèses se traduisent pratiquement par le fait que si les codes sont imprimés dans des conditions constantes (même papier, même encre, même imprimante,

...) à la fois chez l'imprimeur légitime et le contrefacteur, il sera possible d'estimer précisément les modèles d'impression-acquisition. Nous supposons que ces deux modèles d'impression-acquisition sont i.i.d., et notre méthodologie peut s'appliquer pour différentes distributions comme par exemples des modèles Gaussien ou Lognormaux (voir Figure 7.4 (b)). Le modèle du contrefacteur pour une valeur de code x donnée est calculé à partir d'un mélange de deux distributions, l'une pour l'impression-acquisition d'un point noir, l'autre pour l'impression-acquisition d'une zone restée blanche et les paramètres de ce mélange sont déterminés à partir de l'erreur commise par le contrefacteur lors de la binarisation du code original.

Soit H_0 l'hypothèse traduisant le fait que l'observation du code reçu o^N est un code original et soit H_1 l'hypothèse traduisant le fait que l'observation du code reçu o^N est un code contrefait. Dans ces conditions, le receveur peut utiliser la stratégie de Neyman-Pearson qui consiste à calculer le rapport de vraisemblance :

$$L = \log \frac{P_{Z^N|X^N}\left(o^N \middle| x^N, H_1\right)}{P_{Y^N|X^N}\left(o^N \middle| x^N, H_0\right)}, \tag{7.1}$$

et à décider H_0 ou H_1 en comparant ce rapport à un seuil λ garantissant une probabilité de non détection minimale pour une probabilité de fausse alarme inférieure à un niveau α :

$$L \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \tag{7.2}$$

Nous avons dans un premier temps comparé deux types d'observations, le premier suppose que le receveur binarise le code observé avant de calculer son test d'hypothèse, alors que le second type suppose que c'est l'image scannée en niveau de gris qui est directement utilisée comme observation o^N. Nous avons montré que la stratégie consistant à utiliser un code binaire n'est pas optimale dans le sens où pour une probabilité de fausse alarme donnée, la probabilité de non-détection d'un code contrefait est plus importante qu'avec l'utilisation d'un code scanné en niveau de gris. Par contre, d'un point de vue pratique cette stratégie peut comporter plusieurs avantages puisqu'elle ne nécessite pas la connaissance du canal d'impression du contrefacteur et qu'elle se traduit par un comptage du nombre d'erreurs entre le code observé et le code original.

7.2.2 Calcul précis des probabilités d'erreur

La probabilité de fausse alarme α, c'est à dire la probabilité de détecter un code original comme faux, s'exprime comme :

$$\alpha = \Pr\left(L \geq \lambda \mid H_0\right). \tag{7.3}$$

Classiquement, cette probabilité est calculée en invoquant le théorème central limite qui approxime la distribution de la variable aléatoire L par une distribution Gaussienne. Il

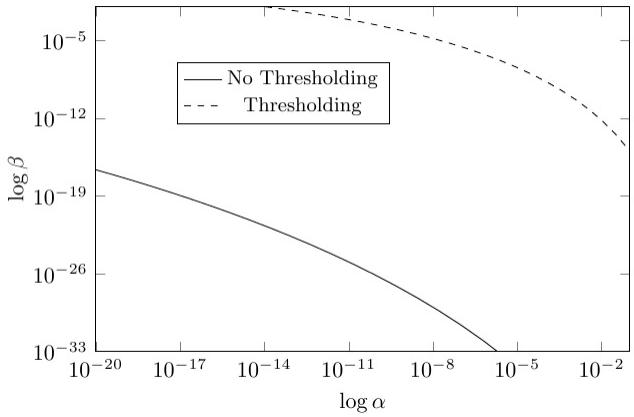


Figure 7.5: Comparaison des courbes ROC avec et sans binarisation. Ici l'utilisation directe de codes scannés en niveau de gris permet d'obtenir des performances en authentification bien supérieures. Modèle Gaussien, $N = 2.10^3$, $\sigma_b = \sigma_w = 52$.

s'avère cependant que pour une valeur très faible de cette probabilité, cette approximation n'est pas réaliste. Ce problème est identique pour la probabilité de non-détection β.

Nous avons ici utilisé la borne de Chernoff qui propose un encadrement supérieur de α ou de β à partir de la fonction génératrice des moments $g_L(s \mid H_0)$ de la variable L, ainsi :

$$\alpha \leq e^{-s\lambda} g_L(s \mid H_0) \text{ pour tout } s > 0, \tag{7.4}$$

avec $g_L(s \mid H_0) = \mathbb{E}_{P_L(L|H_0)}\left[e^{sL}\right]$, la borne la plus précise étant obtenue pour la valeur s_0 minimisant $e^{-s\lambda} g_L(s \mid H_0)$ sous H_0. En introduisant la fonction génératrice des moments semi-invariante $\mu(s\,;\, H_0) = \ln g_L(s \mid H_0)$ et en invoquant le théorème de Cramer [33, 48] pour N suffisamment grand, nous obtenons l'expression asymptotique suivante :

$$\alpha \underset{N\to\infty}{\to} \frac{1}{s_0\sqrt{N\pi\mu''(s_0)}} \exp\left\{\frac{N}{2}\left[\mu(s_0) - s_0\mu'(s_0)\right]\right\}. \tag{7.5}$$

Il est également possible d'obtenir une formule similaire pour la probabilité de non détection β.

7.2.3 Résultats obtenus

La figure 7.6 présente une comparaison entre les courbes ROC obtenues via l'expression asymptotique et via l'approximation Gaussienne et ce pour différents paramètres de la distribution Gaussienne généralisée. Nous pouvons constater que dans certains cas, notamment pour des distributions approchant la loi uniforme, que cette différence est conséquente. La précision de l'expression asymptotique est également corroborée par des simulations de Monté-Carlo qui utilisent un échantillonnage d'importance.

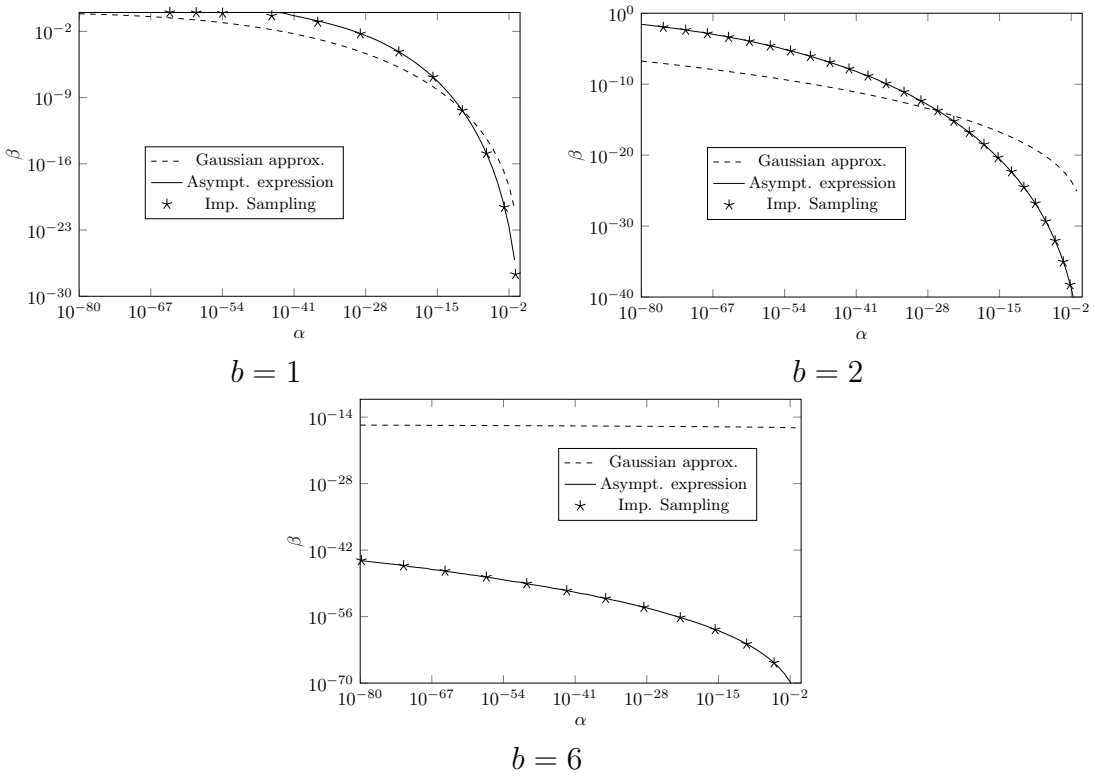


Figure 7.6: Comparaison entre l'approximation Gaussienne, l'expression asymptotique et les simulations de Monte-Carlo via échantillonage d'importance dans le cas de distributions Gaussienne généralisées $b = 1$, $b = 2$ and $b = 6$. Les canaux d'impression-acquisition pour l'imprimeur légitime et le contrefacteur sont identiques, $\mu_b = 50$, $\mu_w = 150$, $\sigma_b = 40$, $\sigma_w = 40$.

7.3 Impact de l'estimation du canal du contrefacteur sur les performances d'authentification

7.3.1 Cadre et objectifs de l'étude

Dans la section précédente nous faisions l'hypothèse que les systèmes d'impression-acquisition de l'imprimeur légitime et du contrefacteur étaient tous les deux connus. Nous prenons maintenant le cas plus réaliste où le modèle de l'imprimeur légitime reste connu, mais où celui du contrefacteur reste à estimer. Dans ce scénario, le receveur cherche, à partir d'un ensemble de N_{obs} codes contrefaits qu'il a en sa possession, à estimer les paramètres du modèle d'impression-acquisition du contrefacteur.

Puisque les paramètres d'estimation sont entachés d'une erreur, nous cherchons ici à quantifier l'impact de cette erreur d'estimation sur les performances globales du système d'authentification, l'objectif étant pour une probabilité de fausse alarme donnée, de calculer les performances moyennes en terme de probabilité de non-détection du sys-

tème, ou encore les performances minimales ou maximales pour un taux de confiance donné.

7.3.2 Relation entre l'erreur d'estimation et la probabilité de fausse détection

Après avoir constater une relation quasi-linéaire entre le carré de l'erreur d'estimation $||\hat{\theta} - \theta||^2$ et le logarithme de la probabilité de non-détection $\log \beta$, nous avons cherché à l'expliquer en analysant le développement limité de $\log \beta$ pour des erreurs d'estimation faibles. Pour cela nous avons calculé la dérivée première, seconde et troisième intervenant dans le développement. De part l'optimalité du test pour une erreur d'estimation nulle, il est facile de montrer que la dérivée première est nulle. Les variations par rapport à l'erreur quadratique et celle du troisième ordre nécessitent l'utilisation de la distribution de Boltzmann $p_{s_0}(\hat{\theta})$ au point s_0 (voir équation (4.21)), et en introduisant le rapport de vraisemblance $l(\hat{\theta}) = \log p_1(v \mid \hat{\theta})/p_0(v \mid \hat{\theta})$ nous obtenons la dépendance quadratique suivante :

$$\log \beta(\hat{\theta}) \approx \log \beta(\theta) + \frac{N}{4}(\hat{\theta} - \theta)^T \mathcal{H}(\theta)(\hat{\theta} - \theta) \tag{7.6}$$

où $\mathcal{H}(\theta)$ est la matrice Hessienne de la fonction $\log \beta(\theta)$ explicitée par :

$$\begin{aligned}
\mathcal{H}_{i,i}(\theta) &= s_1(\theta)\left[\tfrac{\mathrm{cov}^2(l,l'_i)}{\mathrm{Var}(l)} - \mathrm{Var}(l'_i)\right], \\
\mathcal{H}_{i,k}(\theta) &= s_1(\theta)\left[\tfrac{\mathrm{cov}(l,l'_i)\mathrm{cov}(l,l'_k)}{\mathrm{Var}(l)} - \mathrm{cov}(l'_i,l'_k)\right], \\
\mathrm{cov}(l,l'_i) &= \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})l'_i(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right]\mathbb{E}_{p_{s_0}}\left[l'_i(\bar{\theta})\right], \\
\mathrm{cov}(l'_i,l'_k) &= \mathbb{E}_{p_{s_0}}\left[l'_i(\theta)l'_k(\theta)\right] - \mathbb{E}_{p_{s_0}}\left[l'_i(\theta)\right]\mathbb{E}_{p_{s_0}}\left[l'_k(\theta)\right], \\
\mathrm{Var}(l) &= \mathbb{E}_{p_{s_0}}\left[l^2(\theta)\right] - \mathbb{E}^2_{p_{s_0}}\left[l(\theta)\right], \\
\mathrm{Var}(l'_i) &= \mathbb{E}_{p_{s_0}}\left[\left(l'_i(\bar{\theta})\right)^2\right] - \mathbb{E}^2_{p_{s_0}}\left[l'_i(\bar{\theta})\right].
\end{aligned} \tag{7.7}$$

Une étude plus précise de la dérivée troisième montre d'une part que celle-ci est dans la plupart des cas négligeable, d'autre part qu'elle permet d'expliquer en grande partie des différence entre l'approximation quadratique et les mesures pratiques.

7.3.3 Modélisation de la distribution des probabilités de fausse détection

Une fois l'approximation quadratique établie, il est possible de modéliser la distribution des $\log \beta$ en partant du principe qu'un estimateur utilisant une estimation par

maximum de vraisemblance ou bien par algorithme EM génère une erreur d'estimation dont les marginales sont Gaussiennes. Après normalisation, il en ressort donc que le terme quadratique de 7.6 se comporte comme une loi χ^2 et donc que $\log \beta$ peut être s'approximer par une loi χ^2 généralisée, comme illustrée sur la Figure 7.7.

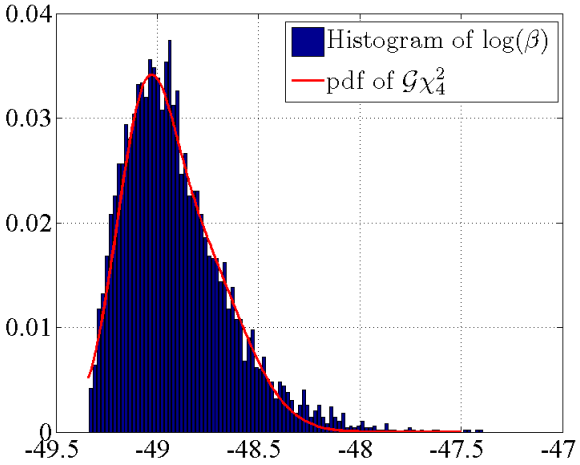


Figure 7.7: Histogramme et densité de probabilité de $\log \beta(\alpha, \theta)$ pour 4 paramètres..

7.3.4 Résultats obtenus

Nous sommes à présent capables de calculer des statistiques de $\log \beta$ pour un α donné, comme notamment sa valeur moyenne, ou encore ses valeurs minimales ou maximales à un seuil de confiance de 95%, et d'analyser l'évolution décroissante de ses statistiques en fonction du nombre de codes graphiques observés. A partir de là, il est possible de quantifier précisément les performances du système d'authentification dans un scénario pratique donné.

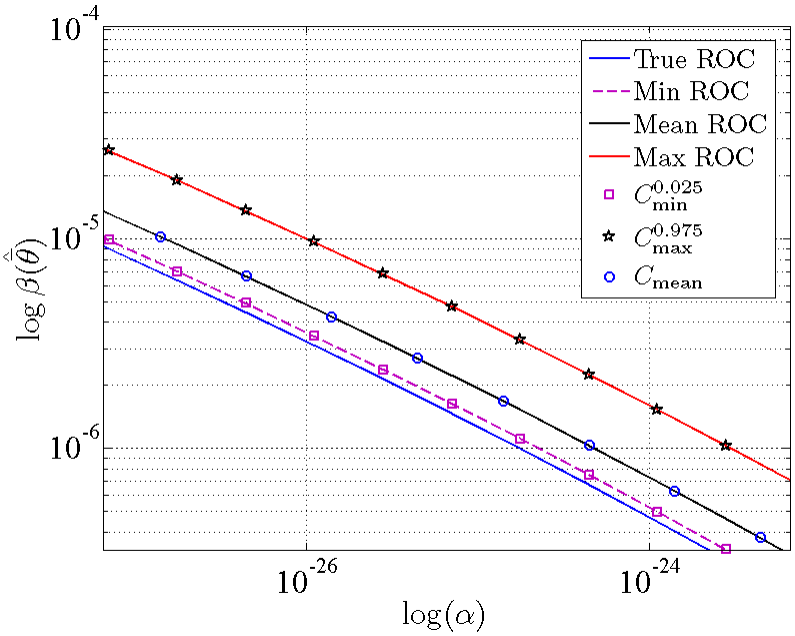


Figure 7.8: Courbes ROC obtenues après l'estimation de quatre paramètres du canal du contrefacteur via l'algorithme EM (modélisation Gaussienne) $\hat{\bar{\mu}}_b, \hat{\bar{\sigma}}_b^2, \hat{\bar{\mu}}_w, \hat{\bar{\sigma}}_w^2$ pour différentes statistiques (moyenne, minimum, maximum pour un seuil de confiance de 95%) en utilisant les formules litérales en via simulations. Comparaison avec la connaissance des paramètres du canal (courbe "True").

7.4 Optimisation du canal d'impression

7.4.1 Scénarios envisagés

Nous cherchons ici à optimiser le système d'authentification via la spécification de ses procédés d'impression-acquisition. Nous analysons trois scénarios pratiques :

1. A partir d'un modèle d'impression-acquisition donné, nous cherchons dans un premier temps à trouver les paramètres du modèle qui permettront de minimiser la probabilité de non-détection β du système d'authentification. Cette optimisation revient en pratique à sélectionner le type d'imprimante, d'encre, ou de papier qui permettront d'obtenir de bonnes performances. Dans ce cas ci, nous faisons l'hypothèse que l'adversaire est passif et qu'il se contentera d'utiliser le même système d'impression-acquisition que l'imprimeur légitime.

2. Le deuxième scénario correspond à un scénario de sécurité à proprement dit puisqu'ici nous prenons en compte un adversaire cherchant à modifier son modèle d'impression-acquisition afin de détériorer les performances du système d'authentification. L'objectif ici est d'envisager une attaque au pire des cas en cherchant le

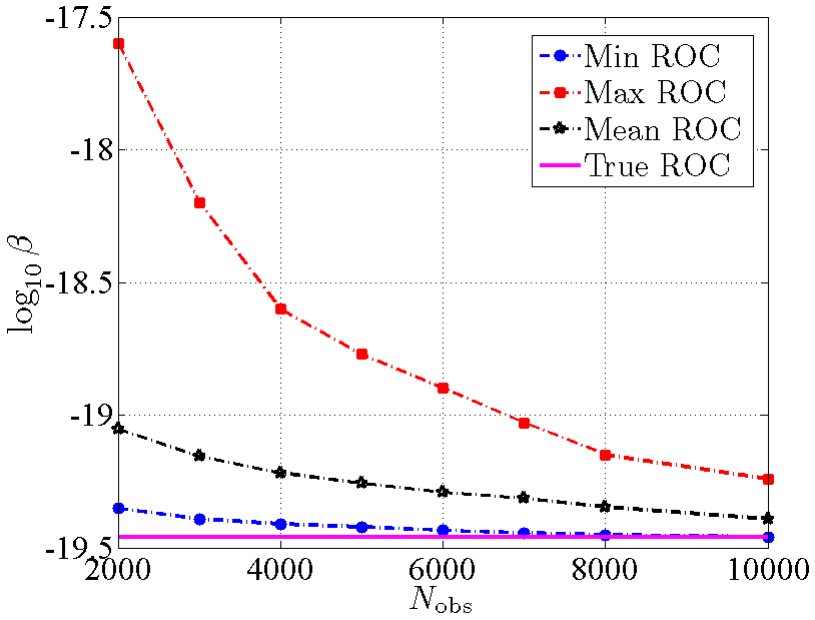


Figure 7.9: Evolution de $\log_{10}\beta(\alpha,\hat{\hat{\theta}})$ en fonction du nombre d'observations N_{obs}, $\alpha = 10^{-16}$.

modèle d'impression-acquisition de l'imprimeur légitime qui permettra d'obtenir les meilleurs performances de détection une fois que l'adversaire aura sélectionné son modèle le plus néfaste. Dans ce scénario l'adversaire est actif puisqu'il est capable de modifier son canal d'impression-acquisition et nous partons du principe que le receveur connait le canal du contrefacteur.

3. Le troisième scénario est similaire au second dans le sens où encore une fois l'adversaire est actif, mais nous faisons l'hypothèse ici que le receveur doit estimer le canal du contrefacteur avant de procéder à l'authentification.

7.4.2 Formalisation des problèmes

Le premier problème peut être formalisé par la recherche au sein d'une famille paramétrique donnée de canaux d'impression-acquisition $\mathcal{C}$, les paramètres du canal minimisant la probabilité de non détection β, nous cherchons donc la probabilité β^* telle que :

$$\beta^* = \min_{\mathcal{C}} \beta(\alpha). \tag{7.8}$$

Dans le second cas, l'optimisation consiste à résoudre un jeu min max pour deux familles de canaux, l'un appelé $\mathcal{C}_l$ pour l'imprimeur légitime, l'autre appelé $\mathcal{C}_o$ pour le contrefacteur. Dans le cas où le receveur connait le canal du contrefacteur, nous cherchons donc la probabilité β^* telle que :

$$\beta^* = \min_{\mathcal{C}_l} \max_{\mathcal{C}_o} \beta(\alpha). \tag{7.9}$$

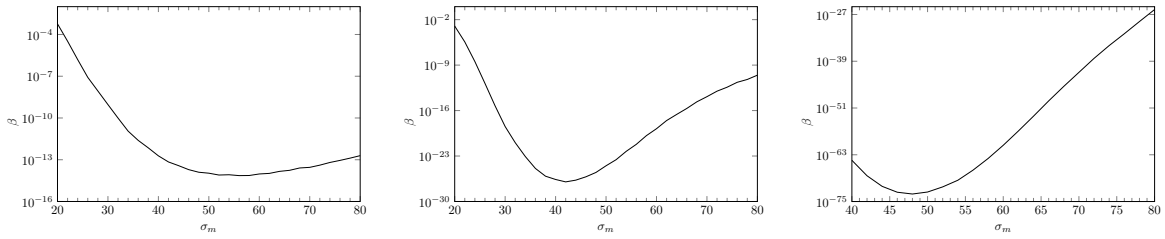
Dans le cas où le receveur doit estimer le canal du contrefacteur, nous utilisons les résultats sur les performances après estimation du canal pour, par exemple, optimiser les performances moyennes $\mathbb{E}[\beta(\alpha)]$, nous cherchons donc la probabilité β^* telle que :

$$\beta^* = \min_{\mathcal{C}_l} \max_{\mathcal{C}_o} \mathbb{E}[\beta(\alpha)]. \tag{7.10}$$

7.4.3 Résultats obtenus

La figure 7.10 présente un exemple de résultats obtenus dans le scénario qui considère un contrefacteur actif. Nous voyons que pour chacun de ces exemples (ce n'est cependant pas vrai dans tous les cas), la stratégie optimale pour l'imprimeur certifiée est d'éviter un procédé d'impression-acquisition peu bruité qui favoriserai une estimation facile du code original par le contrefacteur, mais d'éviter également un procédé trop dégradé pour lequel le bruit important empêcherait la distinction entre code originaux et codes contrefaits. Ces résultats montrent également l'intérêt d'utiliser un canal proche de la loi uniforme, c'est à dire paramètre b grand qui amène un β faible, par rapport à un canal proche d'une loi parcimonieuse, c'est à dire un$b$ faible qui amène un β grand.

La figure 7.11 illustre la différence entre l'optimisation effectuée sans et avec estimation du canal. Nous notons que dans cet exemple la différence de performance croit en fonction du paramètre σ_m mais que le résultat de l'optimisation est peu différent d'un scénario à l'autre.



$b = 1$, $\beta^* \approx 10^{-14}$ $b = 2$, $\beta^* \approx 10^{-27}$ $b = 6$, $\beta^* \approx 10^{-73}$

Figure 7.10: Evolution de la meilleure stratégie du contrefacteur $\max_{\sigma_o} \beta$ en fonction de l'écart type σ_m d'une distribution Gaussienne généralisée pour différents paramètres b de cette distribution. $\mu_b = 50$, $\mu_w = 150$, $\alpha = 10^{-6}$.

7.5 Conclusions

Cette thèse nous a permis d'appréhender des problèmes théoriques d'importance en authentification, et qui plus est pour une méthode d'authentification utilisant des codes

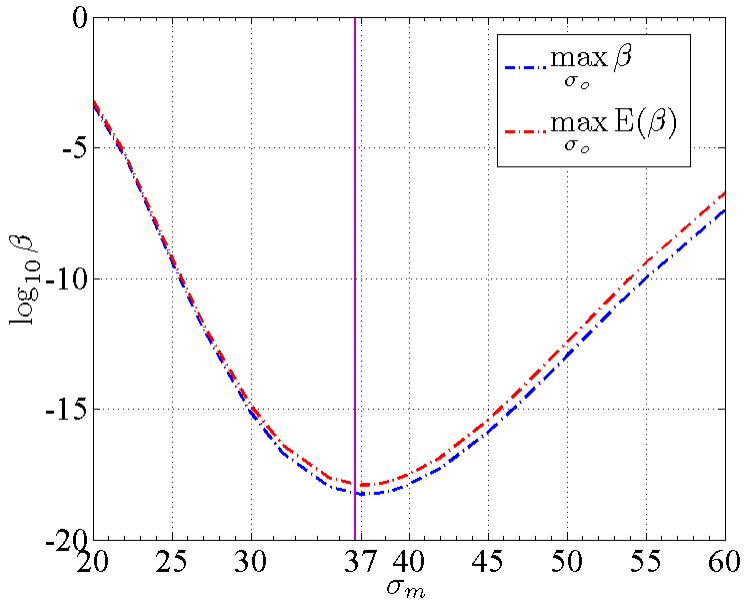


Figure 7.11: Comparaisons entre les optimisations $\max_{\sigma_o}\beta$ et $\max_{\sigma_o}\mathbb{E}(\beta)$ en fonction de σ_m pour une distribution Gaussienne. Ici, $\mu_b = \bar{\mu}_b = 70$, $\mu_w = \bar{\mu}_w = 160$, $\alpha = 10^{-12}$.

graphiques. En utilisant la stratégie du test de Neyman-Pearson comme test d'authentification, nous avons dans un premier temps chercher à mesurer précisément les probabilités de fausse alarme et de non détection de ce test, et ce pour des probabilités faibles. Dans un second temps nous avons voulu quantifier l'impacte d'une étape préalable d'estimation des paramètres d'estimation sur les performances de ce test, et pour cela nous avons cherché à estimer la distribution des probabilité de non-détection pour une probabilité de fausse alarme donnée. Notre dernière contribution a cherché à trouver les paramètres des systèmes d'impression-acquisition qui permettent de maximiser les performances du système d'authentification et ceci pour trois scénarios distincts.

Nous avons comme perspectives d'utiliser ces différentes méthodologies sur des codes graphiques imprimés et non pas générés à partir de modèles statistiques.

Bibliography

[1] According to counterfeiting intelligence bureau. http://www.iccwbo.org/products-and-services/fighting-commercial-crime/counterfeiting-intelligence-bureau/.

[2] Counterfeiting statistics. http://www.iacc.org/counterfeiting-statistics.

[3] Pharmaceutical crime. http://www.interpol.int/Crime-areas/Pharmaceutical-crime/Pharmaceutical-crime.

[4] Global congress addresses international counterfeits threat immediate action required to combat threat to finance/health. http://www.wcoomd.org/en/media/newsroom/2005/november, 2005.

[5] Counterfeiting and piracy endangers global economic recovery, say global congress leaders. http://www.wipo.int/pressroom/en/articles/2009/article_0054.html, 2009.

[6] Fingerprinting blank paper using commodity scanners. https://freedom-to-tinker.com/blog/felten/fingerprinting-blank-paper-using-commodity-scanners/, 2010.

[7] Secure printings using graphical codes. http://www.agence-nationale-recherche.fr/en/anr-funded-project/?tx_lwmsuivibilan_pi2[CODE]=ANR-10-CORD-0019, 2010.

[8] China piracy costs almost million jobs: Us study. http://phys.org/news/2011-05-china-piracy-million-jobs.html, 2011.

[9] Counterfeit economy's impact on u.s firms and government. http://ticker.baruchconnect.com/article/counterfeit-economys-impact-on-u-s-firms-and-government/, 2012.

[10] Counterfeit goods becoming more dangerous. http://money.cnn.com/2012/09/27/news/economy/counterfeit-goods, 2012.

[11] Counterfeiting. http://tommytoy.typepad.com/tommy-toy-pbt-consultin/counterfeiting/, 2012.

[12] A mind-blowing number of counterfeit goods come from china. `http://www.businessinsider.com/most-counterfeit-goods-are-from-china-2013-6`, 2013.

[13] Counterfeit consumer goods. `http://en.wikipedia.org/wiki/Counterfeit_consumer_goods`, 2014.

[14] Paul D Allison. Missing data: Quantitative applications in the social sciences. *British Journal of Mathematical and Statistical Psychology*, 55(1):193–196, 2002.

[15] Philippe Antoine. Tomo3d project. 2011.

[16] Henry S Baird. The state of the art of document image degradation modelling. In *Digital Document Processing*, pages 261–279. Springer, 2007.

[17] C. Baras and F. Cayre. Towards a realistic channel model for security analysis of authentication using graphical codes. In *Information Forensics and Security (WIFS), 2013 IEEE International Workshop on*, pages 115–119. IEEE, 2013.

[18] Cléo Baras and François Cayre. 2d bar-codes for authentication: A security approach. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, pages 1760–1766. IEEE, 2012.

[19] Fokko Beekhof, Sviatoslav Voloshynovskiy, and Farzad Farhadzadeh. Content authentication and identification under informed attacks. In *WIFS*, pages 133–138, 2012.

[20] Christian Berg. The cube of a normal distribution is indeterminate. *The Annals of Probability*, pages 910–913, 1988.

[21] Dimitri P Bertsekas. Constrained optimization and Lagrange multiplier methods. *Computer Science and Applied Mathematics, Boston: Academic Press, 1982*, 1, 1982.

[22] Jeff A Bilmes et al. A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. *International Computer Science Institute*, 4(510):126, 1998.

[23] Richard E Blahut. Hypothesis testing and information theory. *Information Theory, IEEE Transactions on*, 20(4):405–417, 1974.

[24] R Darrell Bock and Murray Aitkin. Marginal maximum likelihood estimation of item parameters: Application of an em algorithm. *Psychometrika*, 46(4):443–459, 1981.

[25] Mary Ann Branch and Andrew Grace. *MATLAB: optimization toolbox: user's guide version 1.5*. The MathWorks, 1996.

[26] James A Bucklew. Large deviation techniques in decision, simulation, and estimation. *Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1990*, 1, 1990.

[27] Gilles Celeux and Gérard Govaert. A classification em algorithm for clustering and two stochastic versions. *Computational statistics & Data analysis*, 14(3):315–332, 1992.

[28] William Clarkson, Tim Weyrich, Adam Finkelstein, Nadia Heninger, J Alex Halderman, and Edward W Felten. Fingerprinting blank paper using commodity scanners. In *Security and Privacy, 2009 30th IEEE Symposium on*, pages 301–314. IEEE, 2009.

[29] Thomas Coleman, Mary Ann Branch, and Andrew Grace. *Optimization Toolbox for Use with MATLAB: User's Guide, Version 2.* Math Works, Incorporated, 1999.

[30] Thomas M Cover and Joy A Thomas. Elements of information theory 2nd edition. 2006.

[31] Anand G Dabak and Don H Johnson. Relations between Kullback-Leibler distance and Fisher information. 2002.

[32] A De Maio and S Iommelli. Coincidence of the rao test, wald test, and glrt in partially homogeneous environment. *Signal Processing Letters, IEEE*, 15:385–388, 2008.

[33] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 2. Springer, 1998.

[34] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.

[35] Dorothy E Denning. Digital signatures with rsa and other public-key cryptosystems. *Communications of the ACM*, 27(4):388–392, 1984.

[36] John E Dennis, Jr and Jorge J Moré. Quasi-Newton methods, motivation and theory. *SIAM review*, 19(1):46–89, 1977.

[37] Mouhamadou L Diong, Patrick Bas, Chloé Pelle, and Wadih Sawaya. Document authentication using 2D codes: Maximizing the decoding performance using statistical inference. In *Communications and Multimedia Security*, pages 39–54. Springer, 2012.

[38] Ahmet Emir Dirik and Bertrand Haas. Copy detection pattern-based document protection for variable media. *Image Processing, IET*, 6(8):1102–1113, 2012.

[39] Gunther Eberhard. Method for authentication between two electronic devices, December 5 1995. US Patent 5,473,689.

[40] Bradley Efron and David V Hinkley. Assessing the accuracy of the maximum likelihood estimator: Observed versus expected fisher information. *Biometrika*, 65(3):457–483, 1978.

[41] Taher ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In *Advances in Cryptology*, pages 10–18. Springer, 1985.

[42] Patrick Emmel and Roger David Hersch. Modeling ink spreading for color prediction. *Journal of Imaging Science and Technology*, 46(3):237–246, 2002.

[43] Jianqing Fan, Chunming Zhang, and Jian Zhang. Generalized likelihood ratio statistics and wilks phenomenon. *Annals of statistics*, pages 153–193, 2001.

[44] Meir Feder and Neri Merhav. Universal composite hypothesis testing: A competitive minimax approach. *Information Theory, IEEE Transactions on*, 48(6):1504–1517, 2002.

[45] Meir Feder and Ehud Weinstein. Parameter estimation of superimposed signals using the em algorithm. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(4):477–489, 1988.

[46] David F Ferraiolo and D Richard Kuhn. Role-based access controls. *arXiv preprint arXiv:0903.2171*, 2009.

[47] Deborah Frincke. Developing secure objects. In *Proceedings of the 19th National Information Systems Security Conference, Baltimore Maryland*, pages 410–419, 1996.

[48] Robert G Gallager. *Information theory and reliable communication*, volume 2. Springer, 1968.

[49] Matthew D Gaubatz, Steven J Simske, and Shawn Gibson. Distortion metrics for predicting authentication functionality of printed security deterrents. In *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pages 1489–1492. IEEE, 2009.

[50] Zoubin Ghahramani, Geoffrey E Hinton, et al. The em algorithm for mixtures of factor analyzers. Technical report, Technical Report CRG-TR-96-1, University of Toronto, 1996.

[51] Tobias Haist and Hans J Tiziani. Optical detection of random features for high security applications. *Optics communications*, 147(1):173–179, 1998.

[52] John R Hershey and Peder A Olsen. Approximating the kullback leibler divergence between gaussian mixture models. In *ICASSP (4)*, pages 317–320, 2007.

[53] Anh Thu Phan Ho, Bao An Mai Hoang, Wadih Sawaya, and Patrick Bas. Document authentication using graphical codes: Reliable performance analysis and channel optimization. *EURASIP Journal on Information Security*, 2014(1):9, 2014.

[54] Anh Thu Phan Ho, Bao An Hoang Mai, Wadih Sawaya, Patrick Bas, et al. Authentication using graphical codes: Optimisation of the print and scan channels. *EUSIPCO 2014*, 2014.

[55] Anthony TS Ho, Xunzhan Zhu, and Yong Liang Guan. Image content authentication using pinned sine transform. *EURASIP Journal on Advances in Signal Processing*, 2004(14):2174–2184, 1900.

[56] Kazufumi Ito and Karl Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15. SIAM, 2008.

[57] Anil K Jain, M Narasimha Murty, and Patrick J Flynn. Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3):264–323, 1999.

[58] Stefan Jaschke, Claudia Klüppelberg, and Alexander Lindner. Asymptotic behavior of tails and quantiles of quadratic forms of Gaussian vectors. *Journal of multivariate analysis*, 88(2):252–273, 2004.

[59] DR Jensen and Herbert Solomon. Approximations to joint distributions of definite quadratic forms. *Journal of the American Statistical Association*, 89(426):480–486, 1994.

[60] Samuel Karlin and Herman Rubin. The theory of decision procedures for distributions with monotone likelihood ratio. *The Annals of Mathematical Statistics*, pages 272–299, 1956.

[61] Helmut Kipphan. *Handbook of print media: technologies and production methods*. Springer, 2001.

[62] David A Kodde and Franz Đ¡ Palm. Notes and comments wald criteria for jointly testing equality and inequality. *Econometrica*, 54(5):1243–1248, 1986.

[63] Erich L Lehmann and Joseph P Romano. *Testing statistical hypotheses*. springer, 2006.

[64] Ching-Yung Lin and Shih-Fu Chang. Distortion modeling and invariant extraction for digital image print-and-scan process. In *Int. Symp. Multimedia Information Processing*, 1999.

[65] Thomas A Louis. Finding the observed information matrix when using the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 226–233, 1982.

[66] Bao An Hoang Mai, Wadih Sawaya, Patrick Bas, et al. Image model and printed document authentication: A theoretical analysis. In *IEEE International Conference on Image Processing*, 2014.

[67] Arak M Mathai and Serge B Provost. Quadratic forms in random variables, vol. 126 of statistics: Textbooks and monographs, 1992.

[68] Ueli M Maurer. Authentication theory and hypothesis testing. *Information Theory, IEEE Transactions on*, 46(4):1350–1356, 2000.

[69] Geoffrey McLachlan and Thriyambakam Krishnan. *The EM algorithm and extensions*, volume 382. John Wiley & Sons, 2007.

[70] Geoffrey McLachlan and David Peel. *Finite mixture models*. John Wiley & Sons, 2004.

[71] Geoffrey J McLachlan and Kaye E Basford. Mixture models. inference and applications to clustering. *Statistics: Textbooks and Monographs, New York: Dekker, 1988*, 1, 1988.

[72] Thomas B Moeslund and Erik Granum. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, 81(3):231–268, 2001.

[73] Bengt Muthén and Kerby Shedden. Finite mixture modeling with mixture outcomes using the em algorithm. *Biometrics*, 55(2):463–469, 1999.

[74] S Natarajan. Large deviations, hypotheses testing, and source coding for finite markov chains. *Information Theory, IEEE Transactions on*, 31(3):360–365, 1985.

[75] Arkadi S Nemirovski and Michael J Todd. Interior-point methods for optimization. *Acta Numerica*, 17:191–234, 2008.

[76] Q.-T. Nguyen, Yves Delignon, Lionel Chagas, and François Septier. Printer technology authentication from micrometric scan of a single printed dot. In *IS&T/SPIE Electronic Imaging*, pages 1–7, 2014.

[77] Thong Q. Nguyen, Y. Delignon, L. Chagas, and F. Septier. Printer identification from micro-metric scale printing. In *Proceedings of ICASSP*, pages 6277–6280, 2014.

[78] Kamal Nigam, Andrew Kachites McCallum, Sebastian Thrun, and Tom Mitchell. Text classification from labeled and unlabeled documents using em. *Machine learning*, 39(2-3):103–134, 2000.

[79] Margaret Norris and Elisa H Barney Smith. Printer modeling for document imaging. In *CISST*, pages 14–20, 2004.

[80] Fernando Pérez-Cruz. Kullback-leibler divergence estimation of continuous distributions. In *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, pages 1666–1670. IEEE, 2008.

[81] Anh Thu Phan Ho, Bao An Mai Hoang, Wadih Sawaya, and Patrick Bas. Document authentication using graphical codes: impacts of the channel model. In *Proceedings of the first ACM workshop on Information hiding and multimedia security*, pages 87–94. ACM, 2013.

[82] J. Picard, C. Vielhauer, and N. Thorwirth. Towards fraud-proof id documents using multiple data hiding technologies and biometrics. *SPIE Proceedings–Electronic Imaging, Security and Watermarking of Multimedia Contents VI*, pages 123–234, 2004.

[83] J Picard and J Zhao. Improved techniques for detecting, analyzing, and using visible authentication patterns, july 28 2005. *WO Patent WO/2005/067,586*.

[84] Justin Picard. Digital authentication with copy-detection patterns. In *Electronic Imaging 2004*, pages 176–183. International Society for Optics and Photonics, 2004.

[85] Florian A Potra and Stephen J Wright. Interior-point methods. *Journal of Computational and Applied Mathematics*, 124(1):281–302, 2000.

[86] SO Rice. Distribution of quadratic forms in normal random variables-evaluation by numerical integration. *SIAM Journal on Scientific and Statistical Computing*, 1(4):438–448, 1980.

[87] Ravi S Sandhu and Pierangela Samarati. Access control: principle and practice. *Communications Magazine, IEEE*, 32(9):40–48, 1994.

[88] Louis L Scharf. *Statistical signal processing*, volume 98. Addison-Wesley Reading, MA, 1991.

[89] FW Scholz. Maximum likelihood estimation. *Encyclopedia of statistical sciences*, 1985.

[90] Shayle R Searle. *Linear models*. John Wiley & Sons, 2012.

[91] Saloomeh Shariati, François-Xavier Standaert, Laurent Jacques, Benoit Macq, Mohamed Amin Salhi, and Philippe Antoine. Random profiles of laser marks. In *WIC Symposium on Information Theory in the Benelux*, pages 27–34, 2010.

[92] J Sheil and I O'Muircheartaigh. Algorithm as 106: The distribution of non-negative quadratic forms in normal variables. *Applied Statistics*, pages 92–98, 1977.

[93] Madhusudan Singh, Hanna M Haverinen, Parul Dhagat, and Ghassan E Jabbour. Inkjet printing process and its applications. *Advanced materials*, 22(6):673–685, 2010.

[94] G Edward Suh and Srinivas Devadas. Physical unclonable functions for device authentication and secret key generation. In *Proceedings of the 44th annual Design Automation Conference*, pages 9–14. ACM, 2007.

[95] M Luisa Targhetta. On a family of indeterminate distributions. *Journal of Mathematical Analysis and Applications*, 147(2):477–479, 1990.

[96] T THAI, Rémi Cogranne, and Florent Retraint. Camera model identification based on the heteroscedastic noise model. 2014.

[97] Akaraphunt Vongkunghae, Jang Yi, and Richard B Wells. A printer model using signal processing techniques. *Image Processing, IEEE Transactions on*, 12(7):776–783, 2003.

[98] Samuel S Wilks et al. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9(1):60–62, 1938.

[99] Margaret Wright. The interior-point revolution in optimization: History, recent developments, and lasting consequences. *Bulletin of the American mathematical society*, 42(1):39–56, 2005.

[100] Ofer Zeitouni and Michael Gutman. On universal hypotheses testing via large deviations. *Information Theory, IEEE Transactions on*, 37(2):285–290, 1991.

[101] Ofer Zeitouni, Jacob Ziv, and Neri Merhav. When is the generalized likelihood ratio test optimal? *Information Theory, IEEE Transactions on*, 38(5):1597–1602, 1992.

[102] Cathel Zitzmann, Rémi Cogranne, Florent Retraint, Igor Nikiforov, Lionel Fillatre, and Philippe Cornu. Statistical decision methods in hidden information detection. In *Information Hiding*, pages 163–177. Springer, 2011.

Index

Appendix A

Materials, Proofs and Extensions

"Appendix usually means "small outgrowth from large intestine," but in this case it means "additional information accompanying main text." Or are those really the same things? Think carefully before you insult this book. "

Pseudonymous Bosch

A.1 Boltzmann's distributions and probabilities of error

Proposition 6. *We always have for any $s \in (0, 1)$:*

$$\mu(s) + (1 - s)\mu'(s) \quad = \quad -D_{KL}(p_s||p_1), \tag{A.1}$$

Similarly we can obtain :

$$\mu(s) - s\mu'(s) \;=\; -D_{KL}(p_s||p_0). \tag{A.2}$$

Proof. These equations, (A.2) and (A.1), are very important in our analysis, and they are true also $\forall s$:

$$\begin{aligned}
D_{KL}(p_s||p_0) &= \int p_s(v) \log \frac{p_s(v)}{p_0(v)} dv, \\
&= \int \frac{p_0^{1-s}(v) p_1^s(v)}{N_s} \log \frac{\frac{p_0^{1-s}(v) p_1^s(v)}{N_s}}{p_0(v)} dv, \\
&= \frac{\int p_0(v) \left(\frac{p_1(v)}{p_0(v)}\right)^s \log \left(\frac{p_1(v)}{p_0(v)}\right)^s dv}{N_s} - \log N_s, \\
&= \frac{s\mathbb{E}_{p_0}\left[l(V) \exp(sl(V)\right]}{E_{p_0}\left[\exp(sl(V)\right]} - \log E_{p_0}\left[\exp(sl(V)\right], \\
&= s\mu'(s) - \mu(s).
\end{aligned}$$

where $l(v) = \log\left(\frac{p_1(v)}{p_0(v)}\right)$. □

Proposition 7. *Let X_t be a random variable with density function p_t, the mean and variance of X_t satisfy*

$$\mathbb{E}_{p_t}(X_t) = \mu'(t) \tag{A.3}$$

and

$$\mathrm{Var}_{p_t}(X_t) = \mu''(t) \tag{A.4}$$

Proof. Taking directly the first and second derivatives of $\mu(t)$, and using the definition of mean and variance of X_t w.r.t p_t. □

A.2 Proof of lemma 2

Proof. We choose the same threshold for the LLR test so we can rewrite this condition mathematically as:

$$\mathbb{E}_{p_{s_1}}\left[l(\hat{\bar{\theta}})\right] = \mathbb{E}_{p_{s_0}}\left[l(\hat{\bar{\theta}})\right] \tag{A.5}$$

Taking derivative on both sides of (A.5) yields:

$$\sum_{v\in\mathcal{V}} l_i'(\hat{\bar{\theta}})p_{s_1}(\hat{\bar{\theta}}) + \sum_{v\in\mathcal{V}} l(\hat{\bar{\theta}})\frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i} = \sum_{v\in\mathcal{V}} l_i'(\hat{\bar{\theta}})p_{s_0}(\hat{\bar{\theta}}) + \sum_{v\in\mathcal{V}} l(\hat{\bar{\theta}})\frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i}. \tag{A.6}$$

At $\hat{\bar{\theta}} = \bar{\theta}$, using lemma (1), Eq. A.6 becomes

$$\sum_{v\in\mathcal{V}} l(\bar{\theta})\left.\frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i}\right|_{\hat{\bar{\theta}}=\bar{\theta}} = \sum_{v\in\mathcal{V}} l(\bar{\theta})\left.\frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i}\right|_{\hat{\bar{\theta}}=\bar{\theta}}. \tag{A.7}$$

Because it is easy to see

$$\frac{\partial p_t(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i} = \left[\frac{\partial K(t,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i} - \sum_{v\in\mathcal{V}}\frac{\partial K(t,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i}p_t(\hat{\bar{\theta}})\right]p_t(\hat{\bar{\theta}})$$

where

$$K(t,\hat{\bar{\theta}}) = t(\hat{\bar{\theta}})l(\hat{\bar{\theta}})$$

So (A.7) is equivalent to

$$\begin{aligned} &\mathbb{E}_{p_{s_1}}\left[l(\bar{\theta})\left.\tfrac{\partial K(s_1,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}}\right|_{\hat{\bar{\theta}}=\bar{\theta}}\right] - \mathbb{E}_{p_{s_1}}\left[l(\bar{\theta})\right]\mathbb{E}_{p_{s_1}}\left[\left.\tfrac{\partial K(s_1,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}}\right|_{\hat{\bar{\theta}}=\bar{\theta}}\right] \\ = \;&\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\left.\tfrac{\partial K(s_0,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}}\right|_{\hat{\bar{\theta}}=\bar{\theta}}\right] - \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right]\mathbb{E}_{p_{s_0}}\left[\left.\tfrac{\partial K(s_0,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}}\right|_{\hat{\bar{\theta}}=\bar{\theta}}\right]. \end{aligned} \tag{A.8}$$

Since:

$$\frac{\partial K(s_1,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}} = t'(\hat{\bar{\theta}})l(\hat{\bar{\theta}}) + t(\hat{\bar{\theta}})l'(\hat{\bar{\theta}}) \tag{A.9}$$

and using (4.23), it can be shown that (A.8) is equivalent to

$$\begin{aligned} &\left[\dot{s}_1^{(i)}(\bar{\theta}) - \dot{s}_0^{(i)}(\bar{\theta})\right]\left\{\mathbb{E}_{p_{s_0}}\left[l^2(\bar{\theta})\right] - \mathbb{E}^2_{p_{s_0}}\left[l(\bar{\theta})\right]\right\} \\ &- \left\{\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})l_i'(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right]\mathbb{E}_{p_{s_0}}\left[l_i'(\bar{\theta})\right]\right\} \;=\; 0 \end{aligned}$$

with $\dot{s}_j^{(i)}(\bar{\theta}) = \left.\frac{\partial s_j(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_i}\right|_{\hat{\bar{\theta}}=\bar{\theta}}$ $(j=0,1)$, and hence

$$\dot{s}_1^{(i)}(\bar{\theta}) - \dot{s}_0^{(i)}(\bar{\theta}) = \frac{\mathrm{cov}(l,l_i')}{\mathrm{Var}(l)}. \tag{A.10}$$

Let

$$A(\hat{\bar{\theta}}) = \mathbb{E}_{p_{s_1}}\left[l_i'(\hat{\bar{\theta}})\right] - \mathbb{E}_{p_{s_0}}\left[l_i'(\hat{\bar{\theta}})\right], \tag{A.11}$$

from (4.22), we have

$$\frac{\partial^2 \beta^*(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i^2} = \frac{\partial s_1(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} A(\hat{\bar{\theta}}) + s_1(\hat{\bar{\theta}}) \frac{\partial A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} \tag{A.12}$$

in which

$$\frac{\partial A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} = \mathbb{E}_{p_{s_1}} \left[l_i''(\hat{\bar{\theta}}) \right] - \mathbb{E}_{p_{s_0}} \left[l_i''(\hat{\bar{\theta}}) \right] + \sum_{v \in \mathcal{V}} l_i'(\hat{\bar{\theta}}) \frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} - \sum_{v \in \mathcal{V}} l_i'(\hat{\bar{\theta}}) \frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i}.$$

Then at $\hat{\bar{\theta}} = \bar{\theta}$,

$$\begin{aligned} \left. \frac{\partial A(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i} \right|_{\hat{\bar{\theta}} = \bar{\theta}} &= \left[\dot{s}_1^{(i)}(\bar{\theta}) - \dot{s}_0^{(i)}(\bar{\theta}) \right] \left\{ \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) l_i'(\bar{\theta}) \right] - \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[l_i'(\bar{\theta}) \right] \right\} \\ &- \left\{ \mathbb{E}_{p_{s_0}} \left[\left(l_i'(\bar{\theta}) \right)^2 \right] - \mathbb{E}_{p_{s_0}}^2 \left[l_i'(\bar{\theta}) \right] \right\}. \end{aligned} \tag{A.13}$$

Finally, it is trivial to finish the proof and obtain (4.26). □

A.3 Proof of proposition 5

Below we give the analytical computation stated in preposition 5 for the mean and variance of the quadratic form (4.28):

Proof. Let $Y = \hat{\bar{\theta}} - \bar{\theta}$, we have $\mathbb{E}(Y) = 0$ and $\mathbb{E}(YY^T) = \Sigma_{\hat{\bar{\theta}}}$. Since $Y^T \mathcal{H}^* Y$ is a scalar, it is equal to its trace, hence:

$$\begin{aligned} \mathbb{E}\left[Y^T \mathcal{H}^* Y \right] &= \mathbb{E}\left[\mathrm{tr}\left(Y^T \mathcal{H}^* Y \right) \right] \\ &= \mathbb{E}\left[\mathrm{tr}\left(\mathcal{H}^* Y Y^T \right) \right] \\ &= \mathrm{tr}\left[\mathbb{E}\left(\mathcal{H}^* Y Y^T \right) \right] \\ &= \mathrm{tr}\left(\mathcal{H}^* \Sigma_{\hat{\bar{\theta}}} \right). \end{aligned}$$

Apply theorem 1 in page 55 of the book [90], we then have:

$$\mathrm{Var}\left[Y^T \mathcal{H}^* Y \right] = 2\mathrm{tr}\left[\left(\mathcal{H}^* \Sigma_{\hat{\bar{\theta}}} \right)^2 \right].$$

So the proposition is proven. □

A.4 The third order expansion of $\log \beta(\alpha, \hat{\bar{\theta}})$ - one parameter

Let first denote generally:

$$\begin{aligned}
\mathrm{cov}(l_{i_0}^{(m)}, l_{i_0}^{(n)}) &= \mathbb{E}_{p_{s_0}}\left[l_{i_0}^{(m)}(\bar{\theta}) l_{i_0}^{(n)}(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l_{i_0}^{(m)}(\bar{\theta})\right]\mathbb{E}_{p_{s_0}}\left[l_{i_0}^{(n)}(\bar{\theta})\right] \quad \text{for } m \neq n \\
\mathrm{Var}(l_{i_0}^{(m)}) &= \mathbb{E}_{p_{s_0}}\left[\left(l_{i_0}^{(m)}(\bar{\theta})\right)^2\right] - \mathbb{E}^2_{p_{s_0}}\left[l_{i_0}^{(m)}(\bar{\theta})\right]
\end{aligned}$$

where $l_{i_0}^{(m)} = \left.\frac{\partial^m l(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_i^m}\right|_{\hat{\bar{\theta}}=\bar{\theta}}$ and if m is equal $1, 2$ or 3 then $l_{i_0}^{(m)}$ becomes l'_{i_0}, l''_{i_0} or l'''_{i_0} respectively.

In order to achieve a better approximation for the authentication performance we need to come up the following theorem:

Theorem 8. *At the point $\hat{\bar{\theta}} = \bar{\theta}$, these below equalities are always true:*

$$\begin{aligned}
(i) \quad & \sum_{v\in\mathcal{V}} l''_{i_0}(\bar{\theta})\left[\frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}} - \frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right]_{\hat{\bar{\theta}}=\bar{\theta}} \\
&= \left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\mathrm{cov}(l, l''_{i_0}) - \mathrm{cov}(l''_{i_0}, l'_{i_0}),
\end{aligned}$$

$$\begin{aligned}
(ii) \quad & \sum_{v\in\mathcal{V}} l'_{i_0}(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} - \frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right]_{\hat{\bar{\theta}}=\bar{\theta}} \\
&= \left(\ddot{s}_1^{(i_0)}(\bar{\theta}) - \ddot{s}_0^{(i_0)}(\bar{\theta})\right)\mathrm{cov}(l, l'_{i_0}) \\
&+ 2\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\mathrm{Var}(l'_{i_0}) - \mathrm{cov}(l''_{i_0}, l'_{i_0}) \\
&+ \left[\left(\dot{s}_1^{(i_0)}(\bar{\theta})\right)^2 - \left(\dot{s}_0^{(i_0)}(\bar{\theta})\right)^2\right]\left[\mathrm{cov}(l'_{i_0}, l^2) - 2\mathbb{E}(l)\mathrm{cov}(l, l'_{i_0})\right] \\
&+ 2s_0(\bar{\theta})\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\left[\mathrm{cov}(l, (l'_{i_0})^2) - 2\mathbb{E}(l'_{i_0})\mathrm{cov}(l, l'_{i_0})\right] \\
&- 2\dot{s}_1^{(i_0)}(\bar{\theta})\left[\mathrm{cov}(l, (l'_{i_0})^2) - 2\mathbb{E}(l'_{i_0})\mathrm{cov}(l, l'_{i_0})\right] \\
&- \left(s_1(\bar{\theta}) + s_0(\bar{\theta})\right)\left[\mathrm{cov}(l'_{i_0}, (l'_{i_0})^2) - 2\mathbb{E}(l'_{i_0})\mathrm{Var}(l'_{i_0})\right]
\end{aligned}$$

$$
\begin{aligned}
(iii) \qquad & \sum_{v\in\mathcal{V}} l(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}^2}-\frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}^2}\right]_{\hat{\bar{\theta}}=\bar{\theta}} \\
= \;& \left(\ddot{s}_1^{(i_0)}(\bar{\theta})-\ddot{s}_0^{(i_0)}(\bar{\theta})\right)\mathrm{Var}(l) \\
+ \;& 2\left(\dot{s}_1^{(i_0)}(\bar{\theta})-\dot{s}_0^{(i_0)}(\bar{\theta})\right)\mathrm{cov}(l,l'_{i_0})-\mathrm{cov}(l''_{i_0},l) \\
+ \;& \left[\left(\dot{s}_1^{(i_0)}(\bar{\theta})\right)^2-\left(\dot{s}_0^{(i_0)}(\bar{\theta})\right)^2\right]\left[\mathrm{cov}(l,l^2)-2\mathbb{E}(l)\mathrm{Var}(l)\right] \\
+ \;& 2s_0(\bar{\theta})\left(\dot{s}_1^{(i_0)}(\bar{\theta})-\dot{s}_0^{(i_0)}(\bar{\theta})\right)\left[\mathrm{cov}(l'_{i_0},l^2)-2\mathbb{E}(l)\mathrm{cov}(l,l'_{i_0})\right] \\
- \;& 2\dot{s}_1^{(i_0)}(\bar{\theta})\left[\mathrm{cov}(l'_{i_0},l^2)-2\mathbb{E}(l)\mathrm{cov}(l,l'_{i_0})\right] \\
- \;& \left(s_1(\bar{\theta})+s_0(\bar{\theta})\right)\left[\mathrm{cov}(l,(l'_{i_0})^2)-2\mathbb{E}(l'_{i_0})\mathrm{cov}(l,l'_{i_0})\right]
\end{aligned}
$$

$$
(iv) \qquad 2\,\frac{\partial A(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}\bigg|_{\hat{\bar{\theta}}=\bar{\theta}}+\sum_{v\in\mathcal{V}} l(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}^2}-\frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}^2}\right]_{\hat{\bar{\theta}}=\bar{\theta}}=0
$$

$$
(v) \qquad \dot{s}_0^{(i_0)}(\bar{\theta})=-s_0(\bar{\theta})\tfrac{\mathrm{cov}(l,l'_{i_0})}{\mathrm{Var}(l)}
$$

$$
(vi) \qquad \left[\dot{s}_1^{(i_0)}(\bar{\theta})\right]^2-\left[\dot{s}_0^{(i_0)}(\bar{\theta})\right]^2=2\dot{s}_0^{(i_0)}(\bar{\theta})\tfrac{\mathrm{cov}(l,l'_{i_0})}{\mathrm{Var}(l)}+\left[\tfrac{\mathrm{cov}(l,l'_{i_0})}{\mathrm{Var}(l)}\right]^2
$$

Proof. We respectively prove (i) to (vi) as follow
(i) We follow the same way as in the proof in lemma 2.

(ii) Remind that:

$$
K(t,\hat{\bar{\theta}})=t(\hat{\bar{\theta}})l(\hat{\bar{\theta}})
$$

and

$$
\frac{\partial p_t(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}=\left[\frac{\partial K(t,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}-\sum_{v\in\mathcal{V}}\frac{\partial K(t,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}p_t(\hat{\bar{\theta}})\right]p_t(\hat{\bar{\theta}}),
$$

we then have

$$
\begin{aligned}
\tfrac{\partial K(t,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}} &= t'_{i_0}(\hat{\bar{\theta}})l(\hat{\bar{\theta}})+t(\hat{\bar{\theta}})l'_{i_0}(\hat{\bar{\theta}}) \\
\tfrac{\partial^2 K(t,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}^2} &= t''_{i_0}(\hat{\bar{\theta}})l(\hat{\bar{\theta}})+2t'_{i_0}(\hat{\bar{\theta}})l'_{i_0}(\hat{\bar{\theta}})+t(\hat{\bar{\theta}})l''_{i_0}(\hat{\bar{\theta}}) \qquad (A.14) \\
\left[\tfrac{\partial K(t,\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}}\right]^2 &= \left[t'_{i_0}(\hat{\bar{\theta}})\right]^2 l^2(\hat{\bar{\theta}})+t^2(\hat{\bar{\theta}})\left[l'_{i_0}(\hat{\bar{\theta}})\right]^2+2l(\hat{\bar{\theta}})l'_{i_0}(\hat{\bar{\theta}})t(\hat{\bar{\theta}})t'_{i_0}(\hat{\bar{\theta}})
\end{aligned}
$$

where $t'_{i_0}(\hat{\bar{\theta}})=\frac{\partial t(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}$, $t''_{i_0}(\hat{\bar{\theta}})=\frac{\partial^2 t(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}^2}$ and $l'_{i_0}(\hat{\bar{\theta}})=\frac{\partial l(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}$, $l''_{i_0}(\hat{\bar{\theta}})=\frac{\partial^2 l(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}^2}$. Moreover, we obtain the general formula for the second derivative of $p_t(\hat{\bar{\theta}})$ w.r.t $\hat{\bar{\theta}}_{i_0}$ as:

$$\begin{aligned}\frac{\partial^2 p_t(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} &= p_t(\hat{\bar{\theta}})\left\{\frac{\partial^2 K(t,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} - \sum_{v\in\mathcal{V}} \frac{\partial^2 K(t,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} p_t(\hat{\bar{\theta}})\right. \\ &+ \left[\frac{\partial K(t,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right]^2 - \sum_{v\in\mathcal{V}} \left[\frac{\partial K(t,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right]^2 p_t(\hat{\bar{\theta}}) \\ &- \left. 2\left[\sum_{v\in\mathcal{V}} \frac{\partial K(t,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}} p_t(\hat{\bar{\theta}})\right]\left[\frac{\partial K(t,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}} - \sum_{v\in\mathcal{V}} \frac{\partial K(t,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}} p_t(\hat{\bar{\theta}})\right]\right\}. \end{aligned} \tag{A.15}$$

We have denoted:

$$\begin{aligned} p_{s_0}(\hat{\bar{\theta}}) &= \frac{e^{s_0(\hat{\bar{\theta}})l(\hat{\bar{\theta}})} p_0(\hat{\bar{\theta}})}{\mathbb{E}_{p_0}\left[e^{s_0(\hat{\bar{\theta}})l(\hat{\bar{\theta}})}\right]} \\ p_{s_1}(\hat{\bar{\theta}}) &= \frac{e^{s_1(\hat{\bar{\theta}})l(\hat{\bar{\theta}})} p_1(\hat{\bar{\theta}})}{\mathbb{E}_{p_1}\left[e^{s_1(\hat{\bar{\theta}})l(\hat{\bar{\theta}})}\right]}, \end{aligned} \tag{A.16}$$

hence

$$\sum_{v\in\mathcal{V}} l'_{i_0}(\hat{\bar{\theta}})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} - \frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right] = B_1(\hat{\bar{\theta}}) - B_3(\hat{\bar{\theta}}) + B_2(\hat{\bar{\theta}}) - B_4(\hat{\bar{\theta}}) + 2\left[B_6(\hat{\bar{\theta}}) - B_5(\hat{\bar{\theta}})\right] \tag{A.17}$$

where

$$B_1(\hat{\bar{\theta}}) = \mathbb{E}_{p_{s_1}}\left[l'_{i_0}(\hat{\bar{\theta}})\frac{\partial^2 K(s_1(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right] - \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\hat{\bar{\theta}})\frac{\partial^2 K(s_0(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right] \tag{A.18}$$

$$B_2(\hat{\bar{\theta}}) = \mathbb{E}_{p_{s_1}}\left[l'_{i_0}(\hat{\bar{\theta}})\left(\frac{\partial K(s_1(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right)^2\right] - \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\hat{\bar{\theta}})\left(\frac{\partial K(s_0(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right)^2\right] \tag{A.19}$$

$$B_3(\hat{\bar{\theta}}) = \mathbb{E}_{p_{s_1}}\left[l'_{i_0}(\hat{\bar{\theta}})\right]\mathbb{E}_{p_{s_1}}\left[\frac{\partial^2 K(s_1(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right] - \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\hat{\bar{\theta}})\right]\mathbb{E}_{p_{s_0}}\left[\frac{\partial^2 K(s_0(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2}\right] \tag{A.20}$$

$$B_4(\hat{\bar{\theta}}) = \mathbb{E}_{p_{s_1}}\left[l'_{i_0}(\hat{\bar{\theta}})\right]\mathbb{E}_{p_{s_1}}\left[\left(\frac{\partial K(s_1(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right)^2\right] - \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\hat{\bar{\theta}})\right]\mathbb{E}_{p_{s_0}}\left[\left(\frac{\partial K(s_0(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right)^2\right] \tag{A.21}$$

$$\begin{aligned} B_5(\hat{\bar{\theta}}) &= \mathbb{E}_{p_{s_1}}\left[\frac{\partial K(s_1(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right]\mathbb{E}_{p_{s_1}}\left[l'_{i_0}(\hat{\bar{\theta}})\frac{\partial K(s_1(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right] \\ &- \mathbb{E}_{p_{s_0}}\left[\frac{\partial K(s_0(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right]\mathbb{E}_{p_{s_1}}\left[l'_{i_0}(\hat{\bar{\theta}})\frac{\partial K(s_0(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}\right] \end{aligned} \tag{A.22}$$

$$B_6(\hat{\bar{\theta}}) = \mathbb{E}_{p_{s_1}}\left[l'_{i_0}(\hat{\bar{\theta}})\right]\left\{\mathbb{E}_{p_{s_1}}\left[\frac{\partial K(s_1(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}\right]\right\}^2 - \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\hat{\bar{\theta}})\right]\left\{\mathbb{E}_{p_{s_0}}\left[\frac{\partial K(s_0(\hat{\bar{\theta}}),\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}\right]\right\}^2. \tag{A.23}$$

At $\hat{\bar{\theta}} = \bar{\theta}$, using (4.23) and (A.14) we have:

$$\begin{aligned} B_1(\bar{\theta}) &= \left(\ddot{s}_1^{(i_0)}(\bar{\theta}) - \ddot{s}_0^{(i_0)}(\bar{\theta})\right)\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})l'_{i_0}(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})l''_{i_0}(\bar{\theta})\right] \\ &+ 2\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\mathbb{E}_{p_{s_0}}\left[\left(l'_{i_0}(\bar{\theta})\right)^2\right] \end{aligned} \tag{A.24}$$

$$\begin{aligned} B_2(\bar{\theta}) &= \left(\left[\dot{s}_1^{(i_0)}(\bar{\theta})\right]^2 - \left[\dot{s}_0^{(i_0)}(\bar{\theta})\right]^2\right)\mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})l^2(\bar{\theta})\right] \\ &+ 2s_0(\bar{\theta})\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\left(l'_{i_0}(\bar{\theta})\right)^2\right] \\ &- \left(s_1^{(i_0)}(\bar{\theta}) + s_0^{(i_0)}(\bar{\theta})\right)\mathbb{E}_{p_{s_0}}\left[\left(l'_{i_0}(\bar{\theta})\right)^3\right] - 2\dot{s}_1^{(i_0)}(\bar{\theta})\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\left(l'_{i_0}(\bar{\theta})\right)^2\right] \end{aligned} \tag{A.25}$$

$$\begin{aligned} B_3(\bar{\theta}) &= \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right]\left[\left(\ddot{s}_1^{(i_0)}(\bar{\theta}) - \ddot{s}_0^{(i_0)}(\bar{\theta})\right)\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l''_{i_0}(\bar{\theta})\right]\right. \\ &+ \left. 2\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right]\right] \end{aligned} \tag{A.26}$$

$$\begin{aligned} B_4(\bar{\theta}) &= \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right]\left\{\left(\left[\dot{s}_1^{(i_0)}(\bar{\theta})\right]^2 - \left[\dot{s}_0^{(i_0)}(\bar{\theta})\right]^2\right)\mathbb{E}_{p_{s_0}}\left[l^2(\bar{\theta})\right]\right. \\ &+ 2s_0(\bar{\theta})\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})l'_{i_0}(\bar{\theta})\right] \\ &- \left.\left(s_1^{(i_0)}(\bar{\theta}) + s_0^{(i_0)}(\bar{\theta})\right)\mathbb{E}_{p_{s_0}}\left[\left(l'_{i_0}(\bar{\theta})\right)^2\right] - 2\dot{s}_1^{(i_0)}(\bar{\theta})\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})l'_{i_0}(\bar{\theta})\right]\right\} \end{aligned} \tag{A.27}$$

$$
\begin{aligned}
B_5(\bar{\theta}) &= \left(\left[\dot{s}_1^{(i_0)}(\bar{\theta})\right]^2 - \left[\dot{s}_0^{(i_0)}(\bar{\theta})\right]^2\right) \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right] \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta}) l'_{i_0}(\bar{\theta})\right] \\
&+ s_0(\bar{\theta})\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right) \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right] \mathbb{E}_{p_{s_0}}\left[\left(l'_{i_0}(\bar{\theta})\right)^2\right] \\
&+ s_0(\bar{\theta})\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right) \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right] \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta}) l'_{i_0}(\bar{\theta})\right] \\
&- \left(s_1^{(i_0)}(\bar{\theta}) + s_0^{(i_0)}(\bar{\theta})\right) \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right] \mathbb{E}_{p_{s_0}}\left[\left(l'_{i_0}(\bar{\theta})\right)^2\right] \\
&- \dot{s}_1^{(i_0)}(\bar{\theta})\left\{\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right] \mathbb{E}_{p_{s_0}}\left[\left(l'_{i_0}(\bar{\theta})\right)^2\right] + \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right] \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta}) l'_{i_0}(\bar{\theta})\right]\right\}
\end{aligned} \tag{A.28}
$$

$$
\begin{aligned}
B_6(\bar{\theta}) &= \left(\left[\dot{s}_1^{(i_0)}(\bar{\theta})\right]^2 - \left[\dot{s}_0^{(i_0)}(\bar{\theta})\right]^2\right) \left\{\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right]\right\}^2 \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right] \\
&+ 2 s_0(\bar{\theta})\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right) \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right] \left\{\mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right]\right\}^2 \\
&- \left(s_1^{(i_0)}(\bar{\theta}) + s_0^{(i_0)}(\bar{\theta})\right) \left\{\mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right]\right\}^3 \\
&- 2\dot{s}_1^{(i_0)}(\bar{\theta}) \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right] \left\{\mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right]\right\}^2.
\end{aligned} \tag{A.29}
$$

We then get

$$
\begin{aligned}
B_1(\bar{\theta}) - B_3(\bar{\theta}) &= \left(\ddot{s}_1^{(i_0)}(\bar{\theta}) - \ddot{s}_0^{(i_0)}(\bar{\theta})\right) \left\{\mathbb{E}_{p_{s_0}}\left[l(\bar{\theta}) l'_{i_0}(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l(\bar{\theta})\right] \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right]\right\} \\
&+ 2\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right) \left\{\mathbb{E}_{p_{s_0}}\left[\left(l'_{i_0}(\bar{\theta})\right)^2\right] - \left(\mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right]\right)^2\right\} \\
&- \left\{\mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta}) l''_{i_0}(\bar{\theta})\right] - \mathbb{E}_{p_{s_0}}\left[l'_{i_0}(\bar{\theta})\right] \mathbb{E}_{p_{s_0}}\left[l''_{i_0}(\bar{\theta})\right]\right\}
\end{aligned} \tag{A.30}
$$

$$
\begin{aligned}
&B_2(\bar{\theta}) - B_4(\bar{\theta}) = \\
&\left(\left[\dot{s}_1^{(i_0)}(\bar{\theta}) \right]^2 - \left[\dot{s}_0^{(i_0)}(\bar{\theta}) \right]^2 \right) \left\{ \mathbb{E}_{p_{s_0}} \left[l^2(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right] - \mathbb{E}_{p_{s_0}} \left[l^2(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \right\} + \\
&2 s_0(\bar{\theta}) \left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta}) \right) \left\{ \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \left(l'_{i_0}(\bar{\theta}) \right)^2 \right] - \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right] \right\} - \\
&\left(s_1^{(i_0)}(\bar{\theta}) + s_0^{(i_0)}(\bar{\theta}) \right) \left\{ \mathbb{E}_{p_{s_0}} \left[\left(l'_{i_0}(\bar{\theta}) \right)^3 \right] - \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[\left(l'_{i_0}(\bar{\theta}) \right)^2 \right] \right\} - \\
&2 \dot{s}_1^{(i_0)}(\bar{\theta}) \left\{ \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \left(l'_{i_0}(\bar{\theta}) \right)^2 \right] - \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right] \right\}
\end{aligned}
\tag{A.31}
$$

$$
\begin{aligned}
&2 \left[B_6(\bar{\theta}) - B_5(\bar{\theta}) \right] = \\
&2 \left(\left[\dot{s}_1^{(i_0)}(\bar{\theta}) \right]^2 - \left[\dot{s}_0^{(i_0)}(\bar{\theta}) \right]^2 \right) \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \left\{ \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] - \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right] \right\} + \\
&2 s_0(\bar{\theta}) \left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta}) \right) \left\{ 2 \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \left\{ \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \right\}^2 \right. \\
&\left. - \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[\left(l'_{i_0}(\bar{\theta}) \right)^2 \right] - \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right] \right\} - \\
&2 \left(s_1^{(i_0)}(\bar{\theta}) + s_0^{(i_0)}(\bar{\theta}) \right) \left[\left\{ \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \right\}^3 - \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[\left(l'_{i_0}(\bar{\theta}) \right)^2 \right] \right] - \\
&2 \dot{s}_1^{(i_0)}(\bar{\theta}) \left\{ 2 \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \left\{ \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \right\}^2 \right. \\
&\left. - \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[\left(l'_{i_0}(\bar{\theta}) \right)^2 \right] - \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right] \right\}.
\end{aligned}
\tag{A.32}
$$

Because we get

$$
\begin{aligned}
&\sum_{v \in \mathcal{V}} l'_{i_0}(\bar{\theta}) \left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} - \frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}^2} \right]_{\hat{\bar{\theta}} = \bar{\theta}} \\
= \;& B_1(\bar{\theta}) - B_3(\bar{\theta}) + B_2(\bar{\theta}) - B_4(\bar{\theta}) + 2 \left[B_6(\bar{\theta}) - B_5(\bar{\theta}) \right]
\end{aligned}
$$

hence it yields

$$
\begin{aligned}
&\sum_{v\in\mathcal{V}} l'_{i_0}(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}^2_{i_0}} - \frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}^2_{i_0}}\right]_{\hat{\bar{\theta}}=\bar{\theta}} \\
= \;& B_1(\bar{\theta}) - B_3(\bar{\theta}) + B_2(\bar{\theta}) - B_4(\bar{\theta}) + 2\left[B_6(\bar{\theta}) - B_5(\bar{\theta})\right] \\
= \;& \left(\ddot{s}_1^{(i_0)}(\bar{\theta}) - \ddot{s}_0^{(i_0)}(\bar{\theta})\right)\mathrm{cov}(l, l'_{i_0}) \\
+ \;& 2\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\mathrm{Var}(l'_{i_0}) - \mathrm{cov}(l''_{i_0}, l'_{i_0}) \\
+ \;& \left[\left(\dot{s}_1^{(i_0)}(\bar{\theta})\right)^2 - \left(\dot{s}_0^{(i_0)}(\bar{\theta})\right)^2\right]\left[\mathrm{cov}(l'_{i_0}, l^2) - 2\mathbb{E}(l)\mathrm{cov}(l, l'_{i_0})\right] \\
+ \;& 2s_0(\bar{\theta})\left(\dot{s}_1^{(i_0)}(\bar{\theta}) - \dot{s}_0^{(i_0)}(\bar{\theta})\right)\left[\mathrm{cov}(l, (l'_{i_0})^2) - 2\mathbb{E}(l'_{i_0})\mathrm{cov}(l, l'_{i_0})\right] \\
- \;& 2\dot{s}_1^{(i_0)}(\bar{\theta})\left[\mathrm{cov}(l, (l'_{i_0})^2) - 2\mathbb{E}(l'_{i_0})\mathrm{cov}(l, l'_{i_0})\right] \\
- \;& \left(s_1(\bar{\theta}) + s_0(\bar{\theta})\right)\left[\mathrm{cov}(l'_{i_0}, (l'_{i_0})^2) - 2\mathbb{E}(l'_{i_0})\mathrm{Var}(l'_{i_0})\right].
\end{aligned}
\tag{A.33}
$$

So, (ii) is proven.

(iii) Similar to (ii).

(iv) Taking the second derivative from both sides of (A.5) yields:

$$
\begin{aligned}
&\sum_{v\in\mathcal{V}} l''_{i_0}(\hat{\bar{\theta}})p_{s_1}(\hat{\bar{\theta}}) + 2\sum_{v\in\mathcal{V}} l'_{i_0}(\hat{\bar{\theta}})\frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}} + \sum_{v\in\mathcal{V}} l(\hat{\bar{\theta}})\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}^2_{i_0}} \\
= \;& \sum_{v\in\mathcal{V}} l''_{i_0}(\hat{\bar{\theta}})p_{s_0}(\hat{\bar{\theta}}) + 2\sum_{v\in\mathcal{V}} l'_{i_0}(\hat{\bar{\theta}})\frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}} + \sum_{v\in\mathcal{V}} l(\hat{\bar{\theta}})\frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}^2_{i_0}}.
\end{aligned}
\tag{A.34}
$$

At $\hat{\bar{\theta}} = \bar{\theta}$, from (A.11), the above equation is equivalent to:

$$
2\,\frac{\partial A(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}}\bigg|_{\hat{\bar{\theta}}=\bar{\theta}} + \sum_{v\in\mathcal{V}} l(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}^2_{i_0}} - \frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}^2_{i_0}}\right]_{\hat{\bar{\theta}}=\bar{\theta}} = 0. \tag{A.35}
$$

Fom (iii) and (iv), we can calculate exactly $\ddot{s}_1^{(i_0)}(\bar{\theta}) - \ddot{s}_0^{(i_0)}(\bar{\theta})$.

(v) In our analysis, we suppose that $\alpha^*(\hat{\bar{\theta}}) = \frac{2}{N}\log\alpha(\hat{\bar{\theta}})$ is fixed, hence $\log\alpha^*(\hat{\bar{\theta}})$ is fixed and so

$$
\frac{\partial\log\alpha^*(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_{i_0}} = 0. \tag{A.36}
$$

From using (4.8), we can express:

$$\log \alpha^*(\hat{\bar{\theta}}) = -D_{KL}(p_{s_0}||p_0). \tag{A.37}$$

It is able to see:

$$D_{KL}(p_{s_0}||p_0) = s_0(\hat{\bar{\theta}}) \sum_{v \in \mathcal{V}} l(\hat{\bar{\theta}}) p_{s_0}(\hat{\bar{\theta}}) - \log \sum_{v \in \mathcal{V}} e^{K(s_0, \hat{\bar{\theta}})} p_0(\hat{\bar{\theta}}) \tag{A.38}$$

$$\frac{\partial D_{KL}(p_{s_0}||p_0)}{\partial \hat{\bar{\theta}}} = s_0(\hat{\bar{\theta}}) \sum_{v \in \mathcal{V}} l(\hat{\bar{\theta}}) \frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}}. \tag{A.39}$$

We know that at $\hat{\bar{\theta}} = \bar{\theta}$,

$$\begin{aligned} \sum_{v \in \mathcal{V}} l(\bar{\theta}) \frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_{i_0}} \Big|_{\hat{\bar{\theta}} = \bar{\theta}} &= \dot{s}_0^{(i_0)}(\bar{\theta}) \mathbb{E}_{p_{s_0}} \left[l^2(\bar{\theta}) \right] + s_0(\bar{\theta}) \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right] \\ &- \dot{s}_0^{(i_0)}(\bar{\theta}) \left\{ \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \right\}^2 + s_0(\bar{\theta}) \mathbb{E}_{p_{s_0}} \left[l(\bar{\theta}) \right] \mathbb{E}_{p_{s_0}} \left[l'_{i_0}(\bar{\theta}) \right]. \end{aligned} \tag{A.40}$$

Therefore,

$$\frac{\partial D_{KL}(p_{s_0}||p_0)}{\partial \hat{\bar{\theta}}} \Big|_{\hat{\bar{\theta}} = \bar{\theta}} = s_0(\bar{\theta}) \dot{s}_0^{(i_0)}(\bar{\theta}) \text{Var} \left[l(\bar{\theta}) \right] + s_0^2(\bar{\theta}) \text{cov} \left[l(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right]. \tag{A.41}$$

From (A.36), we have:

$$s_0(\bar{\theta}) \dot{s}_0^{(i_0)}(\bar{\theta}) \text{Var} \left[l(\bar{\theta}) \right] + s_0^2(\bar{\theta}) \text{cov} \left[l(\bar{\theta}) l'_{i_0}(\bar{\theta}) \right] = 0. \tag{A.42}$$

Consequently,

$$\dot{s}_0^{(i_0)}(\bar{\theta}) = -s_0(\bar{\theta}) \frac{\text{cov}(l, l'_{i_0})}{\text{Var}(l)}, \tag{A.43}$$

and (v) is proven.

(vi) From (A.10), we get:

$$\left[\dot{s}_1^{(i_0)}(\bar{\theta}) \right]^2 = \left[\dot{s}_0^{(i_0)}(\bar{\theta}) \right]^2 + \frac{\text{cov}^2(l, l'_{i_0})}{\text{Var}^2(l)} + 2 \dot{s}_0^{(i_0)}(\bar{\theta}) \frac{\text{cov}(l, l'_{i_0})}{\text{Var}(l)} \tag{A.44}$$

and so

$$\left[\dot{s}_1^{(i_0)}(\bar{\theta}) \right]^2 - \left[\dot{s}_0^{(i_0)}(\bar{\theta}) \right]^2 = 2 \dot{s}_0^{(i_0)}(\bar{\theta}) \frac{\text{cov}(l, l'_{i_0})}{\text{Var}(l)} + \left[\frac{\text{cov}(l, l'_{i_0})}{\text{Var}(l)} \right]^2. \tag{A.45}$$

From (v) and (vi), we can rewrite:

$$\left[\dot{s}_1^{(i_0)}(\bar{\theta})\right]^2 - \left[\dot{s}_0^{(i_0)}(\bar{\theta})\right]^2 = \left[\frac{\text{cov}(l, l'_{i_0})}{\text{Var}(l)}\right]^2 \left[1 - 2s_0(\bar{\theta})\right]. \tag{A.46}$$

Finally, using $(i)-(vi)$ in the above theorem, we can easily obtain the explicit formula for $\frac{\partial^3 \beta^*(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_{i_0}^3}\Big|_{\hat{\tilde{\theta}}=\bar{\theta}}$ and hence for $\frac{\partial^3 \log \beta(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_{i_0}^3}\Big|_{\hat{\tilde{\theta}}=\bar{\theta}}$. □

A.5 The third order expansion of $\log \beta(\alpha, \hat{\tilde{\theta}})$ - multiple parameters

In order to compute $\frac{\partial^3 \beta^*(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_i \partial \hat{\tilde{\theta}}_j \partial \hat{\tilde{\theta}}_k}\Big|_{\hat{\tilde{\theta}}=\bar{\theta}}$, we follow the same steps as in the case of computing $\frac{\partial^3 \beta^*(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_{i_0}^3}\Big|_{\hat{\tilde{\theta}}=\bar{\theta}}$ with some modification. First, generally we have:

$$\begin{aligned}
&\frac{\partial^3 \beta^*(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_i \partial \hat{\tilde{\theta}}_j \partial \hat{\tilde{\theta}}_k} \\
= \;& \dot{s}_1^{(j)}(\hat{\tilde{\theta}}) \left\{ \mathbb{E}_{p_{s_1}} \left[l''_{ik}(\hat{\tilde{\theta}}) \right] - \mathbb{E}_{p_{s_0}} \left[l''_{ik}(\hat{\tilde{\theta}}) \right] + \sum_{v \in \mathcal{V}} l'_i(\hat{\tilde{\theta}}) \left[\frac{\partial p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_k} - \frac{\partial p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_k} \right] \right\} \\
+ \;& \dot{s}_1^{(k)}(\hat{\tilde{\theta}}) \left\{ \mathbb{E}_{p_{s_1}} \left[l''_{ij}(\hat{\tilde{\theta}}) \right] - \mathbb{E}_{p_{s_0}} \left[l''_{ij}(\hat{\tilde{\theta}}) \right] + \sum_{v \in \mathcal{V}} l'_i(\hat{\tilde{\theta}}) \left[\frac{\partial p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j} - \frac{\partial p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j} \right] \right\} \\
+ \;& s_1(\hat{\tilde{\theta}}) \left\{ \mathbb{E}_{p_{s_1}} \left[l'''_{ijk}(\hat{\tilde{\theta}}) \right] - \mathbb{E}_{p_{s_0}} \left[l'''_{ijk}(\hat{\tilde{\theta}}) \right] + \sum_{v \in \mathcal{V}} l''_{ij}(\hat{\tilde{\theta}}) \left[\frac{\partial p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_k} - \frac{\partial p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_k} \right] \right\} \\
+ \;& s_1(\hat{\tilde{\theta}}) \left\{ \sum_{v \in \mathcal{V}} l''_{ik}(\hat{\tilde{\theta}}) \left[\frac{\partial p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j} - \frac{\partial p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j} \right] + \sum_{v \in \mathcal{V}} l'_i(\hat{\tilde{\theta}}) \left[\frac{\partial^2 p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j \partial \hat{\tilde{\theta}}_k} - \frac{\partial^2 p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j \partial \hat{\tilde{\theta}}_k} \right] \right\} \\
+ \;& \ddot{s}_1^{(ij)}(\hat{\tilde{\theta}}) \left\{ \mathbb{E}_{p_{s_1}} \left[l'_i(\hat{\tilde{\theta}}) \right] - \mathbb{E}_{p_{s_0}} \left[l'_i(\hat{\tilde{\theta}}) \right] \right\}.
\end{aligned} \tag{A.47}$$

Thus at $\hat{\tilde{\theta}} = \bar{\theta}$, using lemma 1 we similarly get:

$$\begin{aligned}
\frac{\partial^3 \beta^*(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_i \partial \hat{\tilde{\theta}}_j \partial \hat{\tilde{\theta}}_k}\Big|_{\hat{\tilde{\theta}}=\bar{\theta}} = \;& \dot{s}_1^{(j)}(\bar{\theta}) \sum_{v \in \mathcal{V}} l'_i(\bar{\theta}) \left[\frac{\partial p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_k} - \frac{\partial p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_k} \right]_{\hat{\tilde{\theta}}=\bar{\theta}} \\
+ \;& \dot{s}_1^{(k)}(\bar{\theta}) \sum_{v \in \mathcal{V}} l'_i(\bar{\theta}) \left[\frac{\partial p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j} - \frac{\partial p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j} \right]_{\hat{\tilde{\theta}}=\bar{\theta}} \\
+ \;& s_1(\bar{\theta}) \sum_{v \in \mathcal{V}} l''_{ij}(\bar{\theta}) \left[\frac{\partial p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_k} - \frac{\partial p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_k} \right]_{\hat{\tilde{\theta}}=\bar{\theta}} \\
+ \;& s_1(\bar{\theta}) \sum_{v \in \mathcal{V}} l''_{ik}(\bar{\theta}) \left[\frac{\partial p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j} - \frac{\partial p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j} \right]_{\hat{\tilde{\theta}}=\bar{\theta}} \\
+ \;& s_1(\bar{\theta}) \sum_{v \in \mathcal{V}} l'_i(\bar{\theta}) \left[\frac{\partial^2 p_{s_1}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j \partial \hat{\tilde{\theta}}_k} - \frac{\partial^2 p_{s_0}(\hat{\tilde{\theta}})}{\partial \hat{\tilde{\theta}}_j \partial \hat{\tilde{\theta}}_k} \right]_{\hat{\tilde{\theta}}=\bar{\theta}}
\end{aligned} \tag{A.48}$$

where for $x = 0, 1$

$$\frac{\partial p_{s_x}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_u} = \left[\frac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_u} - \sum_{v\in\mathcal{V}} \frac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_u} p_{s_x}(\hat{\bar{\theta}})\right] p_{s_x}(\hat{\bar{\theta}}) \qquad u = j,k$$

and

$$\begin{aligned}
\tfrac{\partial^2 p_{s_x}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j \partial \hat{\bar{\theta}}_k} &= p_{s_x}(\hat{\bar{\theta}}) \left\{ \tfrac{\partial^2 K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j \partial \hat{\bar{\theta}}_k} - \sum_{v\in\mathcal{V}} \tfrac{\partial^2 K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j \partial \hat{\bar{\theta}}_k} p_{s_x}(\hat{\bar{\theta}}) \right. \\
&+ \tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j} \tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k} - \sum_{v\in\mathcal{V}} \left[\tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j} \tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k}\right] p_{s_x}(\hat{\bar{\theta}}) \\
&- \left[\sum_{v\in\mathcal{V}} \tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j} p_{s_x}(\hat{\bar{\theta}})\right] \left[\tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k} - \sum_{v\in\mathcal{V}} \tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k} p_{s_x}(\hat{\bar{\theta}})\right] \\
&- \left. \left[\sum_{v\in\mathcal{V}} \tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k} p_{s_x}(\hat{\bar{\theta}})\right] \left[\tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j} - \sum_{v\in\mathcal{V}} \tfrac{\partial K(s_x,\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j} p_{s_x}(\hat{\bar{\theta}})\right] \right\}.
\end{aligned}$$

To obtain the explicit formula for (A.48), we need the following equations (for the sake of simplicity, we drop out the index of parameter $\bar{\theta}$):

$$\begin{aligned}
\sum_{v\in\mathcal{V}} l_i'(\bar{\theta}) \left[\tfrac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j} - \tfrac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j}\right]_{\hat{\bar{\theta}}=\bar{\theta}} &= \left[\dot{s}_1^{(j)} - \dot{s}_0^{(j)}\right] \mathrm{cov}(l, l_i') - \mathrm{cov}(l_i', l_j') \\
\sum_{v\in\mathcal{V}} l_j'(\bar{\theta}) \left[\tfrac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k} - \tfrac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k}\right]_{\hat{\bar{\theta}}=\bar{\theta}} &= \left[\dot{s}_1^{(k)} - \dot{s}_0^{(k)}\right] \mathrm{cov}(l, l_j') - \mathrm{cov}(l_j', l_k') \qquad \text{(A.49)} \\
\sum_{v\in\mathcal{V}} l_k'(\bar{\theta}) \left[\tfrac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j} - \tfrac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j}\right]_{\hat{\bar{\theta}}=\bar{\theta}} &= \left[\dot{s}_1^{(j)} - \dot{s}_0^{(j)}\right] \mathrm{cov}(l, l_k') - \mathrm{cov}(l_k', l_j')
\end{aligned}$$

and

$$\begin{aligned}
\sum_{v\in\mathcal{V}} l_{ij}''(\bar{\theta}) \left[\tfrac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k} - \tfrac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_k}\right]_{\hat{\bar{\theta}}=\bar{\theta}} &= \left[\dot{s}_1^{(k)} - \dot{s}_0^{(k)}\right] \mathrm{cov}(l, l_{ij}'') - \mathrm{cov}(l_k', l_{ij}'') \\
\sum_{v\in\mathcal{V}} l_{ik}''(\bar{\theta}) \left[\tfrac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j} - \tfrac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial \hat{\bar{\theta}}_j}\right]_{\hat{\bar{\theta}}=\bar{\theta}} &= \left[\dot{s}_1^{(j)} - \dot{s}_0^{(j)}\right] \mathrm{cov}(l, l_{ik}'') - \mathrm{cov}(l_j', l_{ik}'')
\end{aligned} \qquad \text{(A.50)}$$

where

$$\left[\dot{s}_1^{(u)} - \dot{s}_0^{(u)}\right] = \frac{\mathrm{cov}(l, l_u')}{\mathrm{Var}(l)}, \qquad u = j,k \qquad \text{(A.51)}$$

$$
\begin{aligned}
&\sum_{v\in\mathcal{V}} l_i'(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}-\frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}\right]_{\hat{\bar{\theta}}=\bar{\theta}}\\
=&\left[\ddot{s}_1^{(jk)}-\ddot{s}_0^{(jk)}\right]\mathrm{cov}(l,l_i')-\mathrm{cov}(l_i',l_{jk}'')\\
+&\left[\dot{s}_1^{(j)}-\dot{s}_0^{(j)}\right]\mathrm{cov}(l_i',l_k')+\left[\dot{s}_1^{(k)}-\dot{s}_0^{(k)}\right]\mathrm{cov}(l_i',l_j')\\
+&\left[s_1\dot{s}_1^{(j)}-s_0\dot{s}_0^{(j)}\right]\left[\mathrm{cov}(l_i',ll_k')-\mathbb{E}(l)\mathrm{cov}(l_i',l_k')-\mathbb{E}(l_k')\mathrm{cov}(l,l_i')\right]\\
+&\left[s_1\dot{s}_1^{(k)}-s_0\dot{s}_0^{(k)}\right]\left[\mathrm{cov}(l_i',ll_j')-\mathbb{E}(l)\mathrm{cov}(l_i',l_j')-\mathbb{E}(l_j')\mathrm{cov}(l,l_i')\right]\\
+&\left[\dot{s}_1^{(j)}\dot{s}_1^{(k)}-\dot{s}_0^{(j)}\dot{s}_0^{(k)}\right]\left[\mathrm{cov}(l^2,l_i')-2\mathbb{E}(l)\mathrm{cov}(l,l_i')\right]\\
-&\left[s_1+s_0\right]\left[\mathrm{cov}(l_i',l_j'l_k')-\mathbb{E}(l_j')\mathrm{cov}(l_i',l_k')-\mathbb{E}(l_k')\mathrm{cov}(l_i',l_j')\right]
\end{aligned}
\tag{A.52}
$$

where $\ddot{s}_1^{(jk)}-\ddot{s}_0^{(jk)}$ is computed by using:

$$
\begin{aligned}
0=&\sum_{v\in\mathcal{V}} l_j'(\bar{\theta})\left[\frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_k}-\frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_k}\right]_{\hat{\bar{\theta}}=\bar{\theta}}+\sum_{v\in\mathcal{V}} l_k'(\bar{\theta})\left[\frac{\partial p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_j}-\frac{\partial p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_j}\right]_{\hat{\bar{\theta}}=\bar{\theta}}\\
+&\sum_{v\in\mathcal{V}} l(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}-\frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}\right]_{\hat{\bar{\theta}}=\bar{\theta}}.
\end{aligned}
\tag{A.53}
$$

The condition (A.53) comes from the fact that we use the same threshold for the hypothesis testing. Note that similar to (A.52), we also have:

$$
\begin{aligned}
&\sum_{v\in\mathcal{V}} l(\bar{\theta})\left[\frac{\partial^2 p_{s_1}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}-\frac{\partial^2 p_{s_0}(\hat{\bar{\theta}})}{\partial\hat{\bar{\theta}}_j\partial\hat{\bar{\theta}}_k}\right]_{\hat{\bar{\theta}}=\bar{\theta}}\\
=&\left[\ddot{s}_1^{(jk)}-\ddot{s}_0^{(jk)}\right]\mathrm{Var}(l)-\mathrm{cov}(l,l_{jk}'')\\
+&\left[\dot{s}_1^{(j)}-\dot{s}_0^{(j)}\right]\mathrm{cov}(l,l_k')+\left[\dot{s}_1^{(k)}-\dot{s}_0^{(k)}\right]\mathrm{cov}(l,l_j')\\
+&\left[s_1\dot{s}_1^{(j)}-s_0\dot{s}_0^{(j)}\right]\left[\mathrm{cov}(l^2,l_k')-\mathbb{E}(l)\mathrm{cov}(l,l_k')-\mathbb{E}(l_k')\mathrm{Var}(l)\right]\\
+&\left[s_1\dot{s}_1^{(k)}-s_0\dot{s}_0^{(k)}\right]\left[\mathrm{cov}(l^2,l_j')-\mathbb{E}(l)\mathrm{cov}(l,l_j')-\mathbb{E}(l_j')\mathrm{Var}(l)\right]\\
+&\left[\dot{s}_1^{(j)}\dot{s}_1^{(k)}-\dot{s}_0^{(j)}\dot{s}_0^{(k)}\right]\left[\mathrm{cov}(l^2,l)-2\mathbb{E}(l)\mathrm{Var}(l,)\right]\\
-&\left[s_1+s_0\right]\left[\mathrm{cov}(l,l_j'l_k')-\mathbb{E}(l_j')\mathrm{cov}(l,l_k')-\mathbb{E}(l_k')\mathrm{cov}(l,l_j')\right]
\end{aligned}
\tag{A.54}
$$

All $\mathbb{E}$, Var and cov are taken w.r.t the Boltzmann's distribution $p_{s_0}(\hat{\bar{\theta}})$.

A.6 Constrained optimization using Lagrange multiplier method

We remind that the Lagrange multiplier function of the passive game (5.4) is defined as:

$$\begin{aligned} F(s_0, \theta_0, \lambda) &= \log\beta - \lambda\left[\log\alpha - \log\alpha_0\right] \\ &= -N_c D_{KL}(p_{s_0} \parallel p_1) - \lambda\left[-N_c D_{KL}(p_{s_0} \parallel p_0) - \log\alpha_0\right]. \end{aligned} \tag{A.55}$$

In order to solve the problem (5.4), we need to solve a system of non-linear equations below:

$$\begin{cases} \frac{\partial F(s_0,\theta_0,\lambda)}{\partial s_0} &= 0 \\ \frac{\partial F(s_0,\theta_0,\lambda)}{\partial \theta_0} &= 0 \\ \frac{\partial F(s_0,\theta_0,\lambda)}{\partial \lambda} &= 0 \end{cases}. \tag{A.56}$$

First we have find the explicit formulas for the partial derivatives of $F(s_0, \theta_0, \lambda)$ w.r.t s_0, θ_0 and λ respectively. Remember that $\log\alpha$ and $\log\beta$ can be performed as (see (4.7)):

$$\begin{aligned} \log\alpha &= N_c\left[\mu(s_0) - s_0\mu'(s_0)\right] \\ \log\beta &= N_c\left[\mu(s_0) + (1-s_0)\mu'(s_0)\right] \end{aligned} \qquad 0 < s_0 < 1, \tag{A.57}$$

so

$$\begin{aligned} \frac{\partial\log\alpha}{\partial s_0} &= -N_c s_0\mu''(s_0) \\ \frac{\partial\log\beta}{\partial s_0} &= N_c(1-s_0)\mu''(s_0) \end{aligned}. \tag{A.58}$$

Using (A.4), we then have:

$$\begin{aligned} \frac{\partial\log\alpha}{\partial s_0} &= -N_c s_0 F(s_0) \\ &= -N_c s_0\left\{\mathbb{E}\left[l^2(\theta_0)\right] - \left(\mathbb{E}\left[l(\theta_0)\right]\right)^2\right\} \\ &= -N_c s_0 \mathrm{Var}\left[l(\theta_0)\right] \end{aligned} \tag{A.59}$$

and similarly,

$$\frac{\partial\log\beta}{\partial s_0} = N_c(1-s_0)\ \mathrm{Var}\left[l(\theta_0)\right]. \tag{A.60}$$

Thus

$$\begin{aligned} \frac{\partial F(s_0,\theta_0,\lambda)}{\partial s_0} &= \frac{\partial \log \beta}{\partial s_0} - \lambda \frac{\partial \log \alpha}{\partial s_0} \\ &= N_c \left[1 - s_0 + \lambda s_0\right] \mathrm{Var}\left[l(\theta_0)\right]. \end{aligned} \tag{A.61}$$

From $\frac{\partial F(s_0,\theta_0,\lambda)}{\partial s_0} = 0$, we derive:

$$\lambda = \frac{s_0 - 1}{s_0}. \tag{A.62}$$

Next, we try to find $\frac{\partial F(s_0,\theta_0,\lambda)}{\partial \theta_0}$ by first noting that:

$$\begin{aligned} \frac{1}{N_c}\frac{\partial \log \alpha}{\partial \theta_0} &= -s_0 \sum_{v\in\mathcal{V}} l'(\theta_0) p_{s_0}(\theta_0) - s_0 \sum_{v\in\mathcal{V}} l(\theta_0) \frac{\partial p_{s_0}(\theta_0)}{\partial \theta_0} \\ &+ \sum_{v\in\mathcal{V}} s_0 l'(\theta_0) p_{s_0}(\theta_0) + \sum_{v\in\mathcal{V}} \frac{\partial \log p_0}{\partial \theta_0} p_{s_0}(\theta_0) \\ &= -s_0 \sum_{v\in\mathcal{V}} s_0 l(\theta_0) l'(\theta_0) p_{s_0}(\theta_0) + \sum_{v\in\mathcal{V}} \frac{\partial \log p_0}{\partial \theta_0} p_{s_0}(\theta_0) \\ &+ s_0 \left(\sum_{v\in\mathcal{V}} l(\theta_0) p_{s_0}(\theta_0)\right)\left(\sum_{v\in\mathcal{V}} s_0 l'(\theta_0) p_{s_0}(\theta_0)\right) \\ &- s_0 \sum_{v\in\mathcal{V}} l(\theta_0) \frac{\partial \log p_0}{\partial \theta_0} p_{s_0}(\theta_0) + s_0 \left(\sum_{v\in\mathcal{V}} l(\theta_0) p_{s_0}(\theta_0)\right)\left(\sum_{v\in\mathcal{V}} \frac{\partial \log p_0}{\partial \theta_0} p_{s_0}(\theta_0)\right), \end{aligned} \tag{A.63}$$

hence (for the sake of simplicity we drop out the index of parameter θ_0)

$$\begin{aligned} \frac{\partial \log \alpha}{\partial \theta_0} &= N_c s_0 \left\{-s_0 \mathbb{E}\left[l l'\right] + s_0 \mathbb{E}\left[l\right]\mathbb{E}\left[l'\right] - \mathbb{E}\left[l \frac{\partial \log p_0}{\partial \theta_0}\right] + \mathbb{E}\left[l\right]\mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right]\right\} \\ &+ N_c \mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right] \\ &= N_c \left\{-s_0^2 \mathrm{cov}(l, l') - s_0 \mathrm{cov}\left(l, \frac{\partial \log p_0}{\partial \theta_0}\right) + \mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right]\right\}. \end{aligned} \tag{A.64}$$

Similarly,

$$\frac{\partial \log \beta}{\partial \theta_0} = N_c \left\{\mathbb{E}\left[l'\right] - (s_0 - 1) s_0 \mathrm{cov}(l, l') - (s_0 - 1)\mathrm{cov}\left(l, \frac{\partial \log p_0}{\partial \theta_0}\right) + \mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right]\right\}. \tag{A.65}$$

Therefore,

$$\begin{aligned}\frac{\partial F(s_0,\theta_0,\lambda)}{\partial \theta_0} &= -N_c\{-\mathbb{E}[l'] + (s_0-1)s_0\text{cov}(l,l') \\ &+ (s_0-1)\text{cov}\left(l, \frac{\partial \log p_0}{\partial \theta_0}\right) - \mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right] \\ &+ -\lambda s_0^2\text{cov}(l,l') - \lambda s_0 \text{cov}\left(l, \frac{\partial \log p_0}{\partial \theta_0}\right) + \lambda \mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right]\Big\}. \end{aligned} \tag{A.66}$$

From (A.62) we easily get:

$$\frac{\partial F(s_0,\theta_0,\lambda)}{\partial \theta_0} = N_c\left\{\mathbb{E}[l'] + \frac{1}{s_0}\mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right]\right\}. \tag{A.67}$$

Finally,

$$\frac{\partial F(s_0,\theta_0,\lambda)}{\partial \lambda} = N_c D_{KL}(p_{s_0} \parallel p_0) + \log \alpha_0 . \tag{A.68}$$

From (A.61), (A.67), (A.68) and (A.56), it reduces to solve the following system of non-linear equations:

$$\begin{cases} -N_c D_{KL}(p_{s_0} \parallel p_0) - \log \alpha_0 = 0 \\ s_0\mathbb{E}[l'] + \mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right] = 0 \end{cases}. \tag{A.69}$$

In order to solve (A.69), we first try to use Newton-Raphson method, hence we need to compute the Jacobian matrix. Let's call:

$$\begin{aligned} f(s_0,\theta_0) &= -N_c D_{KL}(p_{s_0} \parallel p_0) - \log \alpha_0 \\ g(s_0,\theta_0) &= s_0\mathbb{E}[l'] + \mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right]. \end{aligned} \tag{A.70}$$

Similar to (A.59), we get:

$$\frac{\partial f(s_0,\theta_0)}{\partial s_0} = -N_c s_0 \text{Var}[l(\theta_0)], \tag{A.71}$$

and we can easily obtain:

$$\frac{\partial g(s_0,\theta_0)}{\partial s_0} = \mathbb{E}[l'] + s_0\text{cov}(l,l') + \text{cov}\left(l, \frac{\partial \log p_0}{\partial \theta_0}\right). \tag{A.72}$$

Now we have only compute $\frac{\partial f(s_0,\theta_0)}{\partial \theta_0}$ and $\frac{\partial g(s_0,\theta_0)}{\partial \theta_0}$. From (A.64), we have:

$$\frac{\partial f(s_0,\theta_0)}{\partial \theta_0} = N_c\left\{-s_0^2\text{cov}(l,l') - s_0\text{cov}\left(l, \frac{\partial \log p_0}{\partial \theta_0}\right) + \mathbb{E}\left[\frac{\partial \log p_0}{\partial \theta_0}\right]\right\}. \tag{A.73}$$

Note that:

$$\begin{aligned}\frac{\partial g(s_0,\theta_0)}{\partial \theta_0} &= s_0\left\{\sum_{v\in\mathcal{V}} l''(\theta_0)p_{s_0}(\theta_0)+\sum_{v\in\mathcal{V}} l'(\theta_0)\frac{\partial p_{s_0}(\theta_0)}{\partial\theta_0}\right\}\\ &+ \sum_{v\in\mathcal{V}}\frac{\partial^2\log p_0}{\partial\theta_0^2}p_{s_0}(\theta_0)+\sum_{v\in\mathcal{V}}\frac{\partial\log p_0}{\partial\theta_0}\frac{\partial p_{s_0}(\theta_0)}{\partial\theta_0}\end{aligned} \tag{A.74}$$

where we have

$$\begin{aligned}l'(\theta_0)\frac{\partial p_{s_0}(\theta_0)}{\partial\theta_0} &= s_0 l'^2 p_{s_0}-l'p_{s_0}\mathbb{E}\left[s_0 l'\right]\\ &+ l'\frac{\partial\log p_0}{\partial\theta_0}p_{s_0}-l'p_{s_0}\mathbb{E}\left[\frac{\partial\log p_0}{\partial\theta_0}\right]\end{aligned} \tag{A.75}$$

hence

$$\sum_{v\in\mathcal{V}} l'(\theta_0)\frac{\partial p_{s_0}(\theta_0)}{\partial\theta_0}=s_0\mathrm{Var}\left[l'\right]+\mathrm{cov}\left(l',\frac{\partial\log p_0}{\partial\theta_0}\right). \tag{A.76}$$

And similarly,

$$\sum_{v\in\mathcal{V}}\frac{\partial\log p_0}{\partial\theta_0}\frac{\partial p_{s_0}(\theta_0)}{\partial\theta_0}=s_0\mathrm{cov}\left(l',\frac{\partial\log p_0}{\partial\theta_0}\right)+\mathrm{Var}\left[\frac{\partial\log p_0}{\partial\theta_0}\right]. \tag{A.77}$$

From (A.74), (A.76) and (A.77), we obtain:

$$\begin{aligned}\frac{\partial g(s_0,\theta_0)}{\partial\theta_0} &= s_0\left\{\mathbb{E}\left[l''\right]+s_0\mathrm{Var}\left[l'\right]+\mathrm{cov}\left(l',\frac{\partial\log p_0}{\partial\theta_0}\right)\right\}\\ &+ \left\{\mathbb{E}\left[\frac{\partial^2\log p_0}{\partial\theta_0^2}\right]+s_0\mathrm{cov}\left(l',\frac{\partial\log p_0}{\partial\theta_0}\right)+\mathrm{Var}\left[\frac{\partial\log p_0}{\partial\theta_0}\right]\right\}.\end{aligned} \tag{A.78}$$

From (A.71), (A.73), (A.72) and (A.78) we obtain the explicit formula of the Jacobian matrix of the problem (A.70) as:

$$J=\begin{bmatrix}\frac{\partial f(s_0,\theta_0)}{\partial s_0} & \frac{\partial f(s_0,\theta_0)}{\partial\theta_0}\\ \frac{\partial g(s_0,\theta_0)}{\partial s_0} & \frac{\partial g(s_0,\theta_0)}{\partial\theta_0}\end{bmatrix}. \tag{A.79}$$

Analyse de performance d'un système d'authentification utilisant des codes graphiques

Résumé:
Nous étudions dans cette thèse l'influence d'un système d'authentification utilisant des codes graphiques 2D modifiés lors de l'impression par un procédé physique non-clônable. Un tel procédé part du principe qu'à très haute résolution le système d'impression acquisition peut être modélisé comme un processus stochastique, de part le caractère aléatoire de la disposition des fibres de papiers, de mélange des particules d'encre, de l'adressabilité de l'imprimante ou encore du bruit d'acquisition. Nous considérons un scénario où l'adversaire pourra estimer le code original et essaiera de le reproduire en utilisant son propre système d'impression. La première solution que nous proposons pour arriver à l'authentification est d'utiliser un test d'hypothèse à partir des modèles à priori connus et sans mémoire des canaux d'impression-acquisition de l'imprimeur légitime et du contrefacteur. Dans ce contexte nous proposons une approximation fiable des probabilités d'erreur via l'utilisation de bornes exponentiels et du principe des grandes déviations. Dans un second temps, nous analysons un scénario plus réaliste qui prends en compte une estimation a priori du canal du contrefacteur et nous mesurons l'impact de cette étape sur les performances du système d'authentification. Nous montrons qu'il est possible de calculer la distribution des probabilité de non-détection et d'en extraire par exemple ses performances moyennes. La dernière partie de cette thèse propose d'optimiser, au travers d'un jeu minmax, le canal de l'imprimeur légitime afin de maximiser ses performances d'authentification tout en envisageant une attaque au pire des cas de la part du contrefacteur.

Mots-clefs : authentification, codes graphiques, tests d'hypothèses, probabilités d'erreurs, théorie de l'estimation

Performance Analysis of an Authentication Method relying on Graphical Codes

Abstract:
We study in this thesis the impact of an authentication system based on 2D graphical codes that are corrupted by a physically unclonable noise such as the one emitted by a printing process. The core of such a system is that a printing process at very high resolution can be seen as a stochastic process and hence produces noise, this is due to the nature of different elements such as the randomness of paper fibers, the physical properties of the ink drop, the dot addressability of the printer, etc. We consider a scenario where the opponent may estimate the original graphical code and tries to reproduce the forged one using his printing process in order to fool the receiver. Our first solution to perform authentication is to use hypothesis testing on the observed memoryless sequences of a printed graphical code considering the assumption that we are able to perfectly model the printing process. The proposed approach arises from error exponent using exponential bounds as a direct application of the large deviation principle. Moreover, when looking for a more practical scenario, we take into account the estimation of the printing process used to generate the graphical code of the opponent, and we see how it impacts the performance of the authentication system. We show that it is both possible to compute the distribution of the probability of non-detection and to compute the average performance of the authentication system when the opponent channel has to be estimated. The last part of this thesis addresses the optimization problem of the printing channel controlled by the legitimate manufacturer in order to maximize his ability to detect a forged graphical code within a min-max game.

Key words : authentication, graphical codes, hypothesis testing, error probabilities, estimation theory