Available online at www.sciencedirect.com

ScienceDirect

Signal Processing 93 (2013) 1724–1738

Signal Processing

Application of Hypothesis Testing Theory for Optimal Detection of LSB Matching Data Hiding[☆]

Rémi Cogranne*, and Florent Retraint**

ICD - LM2S - Université de Technologie de Troyes - UMR STMR CNRS
12, rue Marie Curie - B.P. 2060 - 10010 Troyes cedex - France

Abstract

This paper addresses the problem of detecting the presence of data hidden in digital media by the Least Significant Bit (LSB) matching scheme. In a theoretical context of known digital medium parameters, two important results are presented. First, the use of hypothesis testing theory allows the design of the Most Powerful (MP) test. Second, a study of the MP test provides the opportunity to analytically calculate its statistical properties in order to warrant a given probability of false-alarm. In practice when detecting LSB matching, the unknown medium parameters have to be estimated. Based on a local model of medium content, two different estimations which lead to two different tests are present. A numerical comparison with state-of-the-art detectors shows the good performance of the proposed tests and highlights the relevance of the proposed methodology.

Keywords: Hypothesis testing theory, Information hiding, Optimal detection, Nuisance parameters, Steganography and steganalysis, Information forensics.

1. Introduction and Contributions.

Steganography concerns the reliable transmission of a secret message buried in a host digital medium, such as digital image or audio file. This data hiding technique has been mainly used in information security applications and has receive an increasing interest in the past decade. While a cyphered messages can easily be detected the detection of data hidden within innocuous-looking digital media remains a difficult problem. More generally, the goal of steganalysis is to obtain any information about the potential steganographic system from an unknown medium.
The "prisoners problem" [39], illustrates a typical scenario of steganography and steganalysis. Alice and Bob, two prisoners, communicate by imperceptibly embedding a secret binary message $\boldsymbol{M}$ into a cover-object $\boldsymbol{C}$ to obtain an innocuous looking stego-object $\boldsymbol{S}$. Then, Alice send the stego-object $\boldsymbol{S}$ to Bob through a public channel. Wendy, the warden, examines all their communications in order to detect whether the inspected object $\mathbf{Z}$, contains a secret message $\boldsymbol{M}$ or not.

1.1. State of the Art

Many steganographic tools are nowadays easily available on the Internet making steganography within the reach of anyone, for legitimate or malicious usage. It is thus crucial for security forces to be able to reliably detect steganographic content

[*] Corresponding author.

[**] With the financial support from the Prevention of and Fight against Crime Programme of the European Union European Commission - Directorate-General Home Affairs (2centre.eu project). Research partially funded by Troyes University of Technology (UTT) strategic program COLUMBO.

Email address: `remi.cogranne@utt.fr` (Rémi Cogranne)
URL: `http://lm2s.utt.fr/fr/membres/cogrannr.html` (Rémi Cogranne)

among a (possibly very large) set of media files. In this operational context, the detection of a rather simple but most commonly found stegosystem is more important than the detection of a very complex but rarely encountered stegosystem. The vast majority of downloadable steganographic tools insert the secret information in the LSB plane [5, 6]. More precisely, two embedding functions have been widely studied: LSB replacement and LSB matching, also known as LSB ±1. The LSB replacement method consists in substituting cover medium LSB by bits of secret message. The LSB matching scheme have been proposed as an improvement of LSB replacement; when the hidden bit to be inserted does not match cover medium sample LSB, it is proposed to randomly increment or decrement cover sample value (see [2, 11, 15] and the references therein). While substantial progress has recently been made in the detection of LSB replacement, the steganalysis of LSB matching remains a much harder problem [3, 45]. Therefore, the detection of steganographic algorithms based on LSB matching embedding remains a live research topic.

It can be noted that many methods have been proposed to improve LSB replacement and/or LSB matching schemes. On the one hand, by using coding theory it has been proposed to improved embedding efficiency. Roughly speaking, the idea is to gather several samples and, using coding theory, to embed more than one bits of hidden data for each modification of cover medium samples [34, 36, 40, 41]. On the other hands, focusing on image steganography, it has been proposed to choose the pixels in which it is more difficult to detect hidden bits. It has thus been proposed to locate edges within an image on the underlying assumption that textured areas are difficult to model which, consequently, makes the hidden data detection more difficult [28, 30, 43, 44]. Similarly, recently proposed algorithms select pixels location by minimizing a distortion function, see [25] and the references therein.

1.2. Contributions of the Paper

The recently proposed steganalyzers dedicated to LSB matching can be roughly divided into two categories [3, 31, 35]. On the one hand, most of the latest detectors are based on supervised machine learning methods and use targeted [4, 29, 46] or universal features set [17, 32, 47]. As in all applications of machine learning, a difficult problem is to choose an appropriate feature set. Moreover, the problem of measuring classification error probabilities remains open in the framework of statistical learning [37]. On the other hand, it has been observed in [18] that LSB matching acts as a low-pass filter on the Histogram Characteristic Function (HCF) of a digital image. This pioneering work lead to an entire family of histogram-based detectors [22, 45]. While histogram based detectors has been shown to be very efficient, they have been designed with a limited exploitation of cover medium model and hypothesis testing. Hence, their statistical properties are evaluated through numerical simulations but remain analytically unknown.

In the operational context described above, the proposed steganalyzer must be immediately applicable without any training or tuning phase. For this reason, the use of a machine learning based detector is hardly possible. Moreover, the most important challenge for the steganalyst is to provide detection algorithms with an analytical expression for the false-alarm and missed-detection probabilities without which the "uncertainty" of the result can not be "measured." The prior art LSB matching steganalyzers are certainly very interesting and efficient, but the *ad hoc* design of these algorithms does not permit to calculate the detection errors probability. In addition, only a few theoretical results exist because steganalysis has seldom been thoroughly studied using hypothesis testing theory.

Alternatively, the first step in the direction of hypothesis testing has been made in [12, 8] for LSB replacement to design a statistical test with known statistical properties. In the present paper, the statistical methodology proposed in the case of LSB replacement [7, 8], based on a model of non identically distributed Gaussian samples, is extended for the LSB matching. It should be highlighted that for this extension is not immediate because by changing the embedding model the hypothesis testing problem is modified. Moreover, for the case of LSB matching, the design of an optimal test has never been studied even when all the signal parameters are known. Therefore, the goal of this paper is threefold:

1. Define the most powerful (MP) test in the theoretical case when the cover medium parameters are known, namely the expectation and noise variance of each sample.
2. Analytically calculate the statistical performance of the MP test in terms of the false-alarm and missed-detection probabilities. This result particularly allows us to calculate the decision threshold which warrants a prescribed false-alarm probability, this is the well known Constant False Alarm Rate (CFAR) detection [38]. Moreover, the statistical performance of the MP test also provides an upper bound, which remains unknown, on the detection power of any detection scheme.
3. Design a practical efficient implementation of this test based on a simple local estimation of expectation and variance of each sample.

Compared to the previous published works [9, 10] this paper proposes two main innovations: 1) it especially focuses on low signal-to-noise ratio for which the hidden data detection is the hardest, and 2) this allows us to study the statistical properties of proposed test and provide an upper bound on the performance one can expect for any detector which aims at detecting LSB matching data hiding.

1.3. Organization of the Paper

The paper is organized as follows. The problem of LSB matching steganalysis is casted within the framework of hypothesis testing in Section 2. Following the Neyman-Pearson approach, the MP Likelihood Ratio Test (LRT) is presented in Section 3 and its statistical performance is calculated in Section 4. Finally, the proposed practical implementation of the Generalized LRT (GLRT) is presented in Section 5. To show the relevance of the proposed approach, numerical results on large natural databases of natural images and uncompressed sound are shown in Section 6. Section 7 concludes the paper.

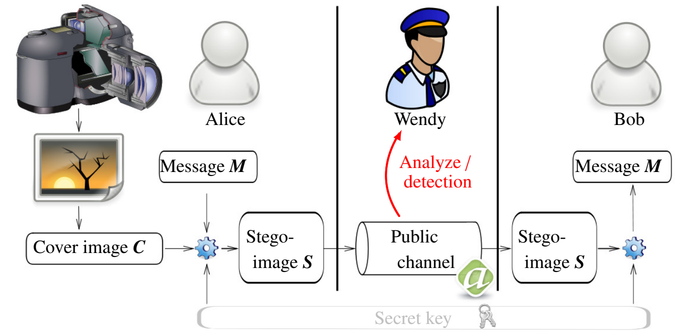


Figure 1: Graphical representation of both the problem of steganalysis and the goal of steganography to estalish a secure communication.

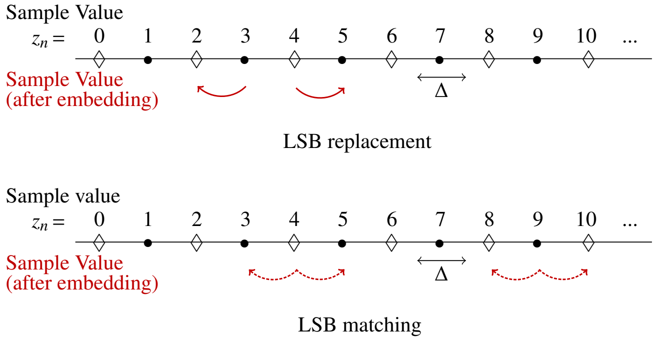

Sample Value
$z_n =$ 0 1 2 3 4 5 6 7 8 9 10 ...
Sample Value (after embedding)
Δ
LSB replacement
Sample value
$z_n =$ 0 1 2 3 4 5 6 7 8 9 10 ...
Sample Value (after embedding)
Δ
LSB matching

Figure 2: Comparison between the LSB replacement (on the top) and the LSB matching (on the bottom) embedding scheme. With the LSB replacement scheme even values can only be incremented and odd values can only decremented. On the contrary, the LSB replacement can increment or decrement every sample value.

2. Hidden Information Detection Problem Statement.

2.1. Statistical model of media

This article studies uncompressed digital medium and, without loss of generality, focuses on natural images, *i.e.* recorded with some imaging device. Hence, the column vector $\boldsymbol{C} = (c_1, \dots, c_N)^T$ represents medium, of $N = N_x{\times}N_y$ pixels for a grayscale image. The set of quantized levels is denoted $\mathcal{Z} = \{0; \dots ; 2^B - 1\}$ as samples values are usually unsigned integers encoded with B bits. Each cover sample c_n results from the quantization:

$$c_n = Q_\Delta(y_n), \qquad (1)$$

where $y_n \in \mathbb{R}^+$ denotes the analogical sample value recorded by the acquisition device and Q_Δ represents the uniform quantization with a step Δ defines as:

$$Q_\Delta(x) = k \Leftrightarrow x \in [\Delta(k - 1/2) \,;\, \Delta(k + 1/2))\,.$$

Seeking simplicity, it is assumed in this paper that the saturation effect is absent, *i.e.* the probability of exceeding the quantizer boundaries $-\Delta/2$ and $\Delta(2^B - 1 + 1/2)$ is negligible. Indeed, taking into account this phenomenon is possible at the cost of much more complicated notations.
The recorded sample value can be decomposed as [14, 7]:

$$y_n = \mu_n + \xi_n, \qquad (2)$$

where μ_n is a deterministic parameter corresponding to the mathematical expectation of y_n. On the opposite, ξ_n is a random variable representing all the noises corrupting the cover medium during acquisition. For most of the digital media, ξ_n is accurately modeled as a realization of a zero-mean Gaussian random variable $\Xi_n \sim \mathcal{N}(0, \sigma_n^2)$ with variance σ_n^2. It is important to note that the analogical samples are statistically independent [14, 19] and that the variance σ_n^2 varies from sample to sample, for instance because of photo-counting shot noise of digital images. It thus follows from Equations (1) and (2) that c_n follows a distribution denoted P_{θ_n} which entirely characterized by the parameter $\theta_n = (\mu_n \sigma_n, \Delta)^T$, where $\mathbf{A}^T$ represents the transpose of $\mathbf{A}$. For the sake of definition let θ_n belongs to a compact set $\Theta_n \subset \mathbb{R}^3$ and let define $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$, $\boldsymbol{\theta} \in \Theta = \Theta_1 \times \dots \times \Theta_N$ represents the whole medium parameter.
The distribution P_{θ_n} is defined by its probability mass function (pmf) $P_{\theta_n} = (p_{\theta_n}[0], \dots, p_{\theta_n}[2^B - 1])$ where:

$$\begin{aligned} \forall k \in \mathcal{Z}\,,\ p_{\theta_n}[k] &= \Phi\left(\frac{\Delta(k + 1/2) - \mu_n}{\sigma_n}\right) - \Phi\left(\frac{\Delta(k - 1/2) - \mu_n}{\sigma_n}\right), \\ &= \frac{1}{\sigma_n} \int_{\Delta(k-1/2)}^{\Delta(k+1/2)} \phi\left(\frac{u - \mu_n}{\sigma_n}\right) du. \qquad (3) \end{aligned}$$

In this paper, ϕ denotes the standard Gaussian probability distribution function (pdf) $\phi(u) = \frac{1}{\sqrt{2\pi}} \exp(u^2/2)$ and Φ represents the standard Gaussian cumulative distribution function (cdf) defined by $\Phi(x) = \int_{-\infty}^{x} \phi(u) du$.
In virtue of the mean value theorem, the probability $p_{\theta_n}[k]$ defined in Equation (3) can be written as:

$$p_{\theta_n}[k] = \frac{1}{\sigma_n} \int_{\Delta(k-1/2)}^{\Delta(k+1/2)} \phi\left(\frac{u - \mu_n}{\sigma_n}\right) du = \frac{\Delta}{\sigma_n} \phi\left(\frac{\Delta k - \mu_n}{\sigma_n}\right) + \epsilon, \quad (4)$$

where the (small) corrective term ϵ can be written, using the well known Taylor expansion of the function ϕ [42, p.931] as:

$$\epsilon = 0 + o\left(\frac{\Delta^2}{\sigma_n^2}\right) \qquad (5)$$

and the notation $y = o(x)$ means that y/x tends to 0 as x tends to 0, see details in [8, 48]. This paper focus on the case $\Delta \ll \sigma_n$ in which the problem of hidden information detection is the hardest; in practice this situation is especially satisfies for raw digital image or uncompressed audio files usually, encoded using $B = 12, 14$ or even 16 bits.

To model statistically stego-samples from (3) - (4), the two following assumptions are usually adopted [12, 16] :

- The message is previously compressed and/or cyphered prior to its insertion, hence, each hidden bits of message $\boldsymbol{M} = (m_1, \dots, m_L)^T$ is drawn from a binomial distribution $\mathcal{B}(1, 1/2)$, *i.e.* m_l is either 0 or 1 with the same probability.

- The samples used to carry hidden bits are chosen pseudo-randomly using a secret key, hence, each cover sample c_n are used with the same probability.

Let the embedding rate $R \in (0,1]$, $R = {}^{L}\!/_{N}$ be the number of hidden bits per cover sample and let $\boldsymbol{S} = \{s_1, \dots, s_N\}$ be the values of stego-samples, *i.e.* after insertion of hidden information. This impact of information hiding is captured by denoting

$$\forall n \in \{0, \dots, N\}, \left\{ \begin{array}{c} \mathbb{P}[s_n = c_n] = (1{-}R), \\ \mathbb{P}[s_n = c_n + \mathrm{ins}(m_l, c_n)] = R, \end{array} \right. \tag{6}$$

where $\mathrm{ins}(m_l, c_n)$ represents the value added to c_n to insert the hidden bit m_l.

In the case of LSB replacement, the insertion of each hidden bit is done by setting the LSB of c_n to the value of m_l ; hence the function ins can be written $\mathrm{ins}(m_l, c_n) = (-1)^{c_n}$ when the LSB of c_n differs from m_l.

The particularity of LSB matching lies in its insertion function $\mathrm{ins} : \{0;1\} \times \mathcal{Z} \mapsto \{-1;0;1\}$ which allows to change not only the LSB of c_n. Whenever the LSB of c_n is equal to m_l, *i.e.* when $\mathrm{lsb}(c_n) = c_n \mathrm{mod} 2 = m_l$, there is no need to change c_n, hence $\mathrm{ins}(m_l, c_n) = 0$. On the contrary, whenever $\mathrm{lsb}(c_n) \neq m_l$, the insertion must change the LSB of c_n, which is done by adding or subtracting 1 with the same probabilities:

$$\left\{ \begin{array}{lcl} \mathbb{P}[\mathrm{ins}(m_l, c_n) = 1 \,|\, \mathrm{lsb}(c_n) \neq m_l] & = & 1/2 \\ \mathbb{P}[\mathrm{ins}(m_l, c_n) = -1 \,|\, \mathrm{lsb}(c_n) \neq m_l] & = & 1/2. \end{array} \right. \tag{7}$$

Since each hidden bit m_l follows the binomial distribution $\mathcal{B}(1, 1/2)$, a straightforward calculation finally shows that $\mathbb{P}[\mathrm{lsb}(c_n) = m_l] = \mathbb{P}[\mathrm{lsb}(c_n) \neq m_l] = 1/2$. Hence, as described in [18, 45, 4], it follows from (6)–(7) that for all $n \in \{1, \dots, N\}$, the pmf of the stego-sample s_n after embedding at rate R with LSB matching is given by $Q^R_{\theta_n} = \left(q^R_{\theta_n}[0], \dots, q^R_{\theta_n}[2^b - 1]\right)$ with $\forall k \in \mathcal{Z}$:

$$q^R_{\theta_n}[k] = \frac{R}{4}\left(p_{\theta_n}[k{-}1] + p_{\theta_n}[k{+}1]\right) + \left(1{-}\frac{R}{2}\right) p_{\theta_n}[k]. \tag{8}$$

2.2. Hypothesis Testing Problem Statement

When analyzing an unknown medium $\mathbf{Z}$ the general goal of LSB matching steganalysis is to decide between the two following hypotheses:

$$\left\{ \begin{array}{l} \mathcal{H}_0 = \left\{ z_n \sim P_{\theta_n}, \forall n \in \{1, \dots, N\}, \forall \boldsymbol{\theta} \in \Theta \right\} \\ \mathcal{H}_1 = \left\{ z_n \sim Q^R_{\theta_n}, \forall n \in \{1, \dots, N\}, \forall \boldsymbol{\theta} \in \Theta, \forall R \in (0,1] \right\}. \end{array} \right. \tag{9}$$

The goal is to find a test $\delta : \mathcal{Z}^N \mapsto \{\mathcal{H}_0, \mathcal{H}_1\}$, such that hypothesis $\mathcal{H}_i$ is accepted if $\delta(\mathbf{Z}) = \mathcal{H}_i$ (see [26] for details about statistical hypothesis testing). As explained in the introduction, in an operational forensics context the most important challenge is first, to warrant a prescribed (very low) false-alarm probability and second, to maximize the detection power defined by:

$$\beta_\delta = \mathbb{P}_1[\delta(\mathbf{Z}) = \mathcal{H}_1],$$

where $\mathbb{P}_i(\cdot)$ stands for the probability under hypotheses $\mathcal{H}_i$, $i = \{0;1\}$. Therefore, let $\mathcal{K}_\alpha$ be the class of tests with an upper-bounded false-alarm probability α_0 defined by

$$\mathcal{K}_\alpha = \left\{ \delta : \sup_{\theta \in \Theta} \mathbb{P}_0[\delta(\mathbf{Z}) = \mathcal{H}_1] \leq \alpha_0 \right\}. \tag{10}$$

The statement of hidden information as formulated in Equation (9) in the framework of hypothesis testing highlights two major difficulties which should be consider differently. First, because the embedding R is unknown the hypothesis $\mathcal{H}_1$ is composite and the ultimate goal is to find a Uniformly Most Powerful (UMP) test that maximizes the power whatever R might be. Unfortunately, the tested hypotheses (9) do not have a monotonic likelihood ratio, hence a UMP test scarcely exists, see [26, theorem 3.4.1] for details. Second, only the medium $\mathbf{Z}$ is available for analysis and the parameters $\theta_1, \dots, \theta_N$ act as nuisance parameters which prevents detection of hidden information; in addition, an accurate estimation of each sample parameter $\theta_n = (\mu_n, \sigma_n, \Delta)^T$ is well-known to be difficult and open problem of signal and image processing.

The main goals of this paper are, first, to design a test whose statistical performance is theoretically established and, second, to propose a practical implementation of this test to analyze an unknown medium. In other words, the problem of finding a most powerful (MP) test for any embedding rate $R \in [0;1]$ lays outside the scope of this paper. The reader interested in this theoretical aspect might referrer to [8, 48].

As a consequence, the problem of testing composite hypothesis is simplified in Section 3 by designing the Likelihood Ratio Test (LRT) is the case when all the samples are modified by ±1. Then the statistical performance of this test are established in the general case of $R \in (0,1]$ in Section 4. Finally, to propose a practical implementation when no information on analyzed medium is available, the problem of dealing with nuisance parameters $\theta_1, \dots, \theta_N$ is addressed using a linear parametric model of samples in Section 5.

3. Optimal Likelihood Ratio Test for Simple Hypotheses.

Let us start with the simplest case, when the embedding rate R and, for all n, the parameters θ_n are known. In this case, the hypothesis testing problem (9) is reduced to a test between two simple hypotheses.

In virtue of the Neyman-Pearson lemma, see [26, Theorem 3.2.1], the most powerful (MP) test over the class $\mathcal{K}_{\alpha_0}$ (10) is the LRT given by the following decision rule:

$$\delta_R(\mathbf{Z}) = \left\{ \begin{array}{lccc} \mathcal{H}_0 & \text{if} & \Lambda_R(\mathbf{Z}) & \leq & \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{if} & \Lambda_R(\mathbf{Z}) & > & \tau_{\alpha_0}, \end{array} \right. \tag{11}$$

where τ_{α_0} is the solution of $\mathbb{P}_0[\delta(\mathbf{Z}) > \tau_{\alpha_0}] = \alpha_0$, to insure that $\delta_R \in \mathcal{K}_{\alpha_0}$, and the likelihood ratio (LR) $\Lambda_R(\mathbf{Z})$ is given, from the statistical independence between samples, by:

$$\Lambda_R(\mathbf{Z}) = \sum_{n=1}^{N} \Lambda_R(z_n) = \sum_{n=1}^{N} \log\left(\frac{q^R_{\theta_n}[z_n]}{p_{\theta_n}[z_n]} \right). \tag{12}$$

It worth noting that the optimality of the LRT holds for any decision problem but requires a knowledge of the statistical distribution of observations under both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. Hence, the previous woks on LSB replacement [7, 8, 48] can no longer be used for LSB replacement because the statistical definition of alternative hypothesis $\mathcal{H}_1$ is different and the LR thus changes.

Using the definition of samples distribution under alternative hypothesis $\mathcal{H}_1$ (8), the LR can be written:

$$\Lambda_R(\mathbf{Z}) = \sum_{n=1}^{N} \log\left(\frac{R}{4} \frac{p_{\theta_n}[z_n-1] + p_{\theta_n}[z_n+1]}{p_{\theta_n}[z_n]} + \left(1 - \frac{R}{2}\right)\right). \quad (13)$$

It can be noted that $\Lambda_R(z_n)$ depends on the sample values z_n through the quantity:

$$\Lambda_2(z_n) = \log\left(\frac{1}{2} \frac{p_{\theta_n}[z_n-1] + p_{\theta_n}[z_n+1]}{p_{\theta_n}[z_n]} \right), \quad (14)$$

which corresponds to the the likelihood ratio for the conceptual case of $R = 2$.

In other words, Equation (14) corresponds to the likelihood ratio for testing the two following hypotheses: $\mathcal{H}_0 : \{\,\mathbf{Z} \text{ is a cover medium}\,\}$ vs $\mathcal{H}_1 : \{\,\text{each sample of } \mathbf{Z} \text{ is modified by } \pm 1\,\}$. Indeed, considering the case $R = 2$ permits us to clarify the present methodology, which is then extended to the more general case of $R \in (0, 1]$ in Section 4.2.

The exact expression for the LR $\Lambda_2(z_n)$ is complicated due to the corrective terms ϵ defined in (4). However, as established in Equation (5), the calculation shows that these corrective terms are negligible when $\Delta \ll \sigma_n$ which is the case considered in this paper. Therefore, it is proposed to neglect ϵ in order to obtain a simplified expression for the LR $\Lambda_2(z_n)$. From Equation (4), it is shown in the Appendix A.1 that this approximation permits us to write:

$$\begin{aligned} \frac{p_{\theta_n}[z_n-1]}{p_{\theta_n}[z_n]} &= \exp\left(-\frac{\Delta^2}{2\sigma_n^2}\right)\exp\left(\frac{-\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right), \\ \frac{p_{\theta_n}[z_n+1]}{p_{\theta_n}[z_n]} &= \exp\left(-\frac{\Delta^2}{2\sigma_n^2}\right)\exp\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right). \end{aligned} \quad (15)$$

By using the results from Equation (15), the LR $\Lambda_2(z_n)$ can be written as:

$$\Lambda_2(z_n) = \log\left[\exp\left(\frac{-\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right) + \exp\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right)\right] - \log(2) - \frac{\Delta^2}{2\sigma_n^2} \quad (16)$$

Again, these calculations are detailed in the Appendix A.1.

3.1. Statistical Analysis of the LR $\Lambda_2(z_n)$.

In order to establish the performance of the LR test δ_R (11) in the case $R = 2$, a thorough analysis of the statistics $\Lambda_2(z_n)$ is necessary. From Equation (16), it is obvious that such statistical study is not straightforward. Indeed, the LR $\Lambda_2(z_n)$ only depends on the observation z_n through the quantity

$$\log\left[\cosh\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right)\right]$$

whose statistical distribution is difficult to establish.

For the sake of clarity, it is proposed in the present paper to define the variable $v(z_n;\theta_n)$, abbreviated as v_n, as follows:

$$\begin{aligned} v_n &\overset{\text{def.}}{=} \log\left[\exp\left(\frac{-\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right) + \exp\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right)\right] - \log(2) \\ &= \Lambda_2(z_n) + \frac{\Delta^2}{2\sigma_n^2}. \end{aligned} \quad (17)$$

The main interest of the random variable Υ_n, whose v_n is a realization, is that it only differs from the LR $\Lambda_2(z_n)$ by a constant when inspected medium parameters are known. Moreover, the following Theorem 1 provide an accurate approximation of the statistical distribution of v_n.

Theorem 1. *Let v_n, defined in Equation (17), be distributed according to a distribution P_{Υ_n} whose pdf is denoted f_{Υ_n} and let P_{Γ_n} denotes the Gamma distribution with shape parameter $k = 1/2$ and scale parameter $\theta = {}^{\Delta^2}\!/_{\sigma_n^2}$ whose pdf is denoted f_{γ_n}. Assuming that $\theta_n = (\Delta, \mu_n, \sigma_n)^T$ are known and that z_n is described by the statistical model given in Equation (4), it asymptotically holds that as ${}^{\Delta}\!/_{\sigma_n}$ tends to 0 then v_n tends to be distributed as a Gamma distribution in the following senses:*

$$\lim_{\frac{\Delta}{\sigma_n}\to 0} D_{KL}(P_{\Upsilon_n}, P_{\Gamma}) = \int_{\mathbb{R}^+} f_{\Upsilon_n}(x)\log\left(\frac{f_{\Upsilon_n}(x)}{f_{\Gamma_n}(x)}\right) = 0, \quad (18)$$

$$\textit{and} \ \sup_{x\in\mathbb{R}^+} \left| f_{\Upsilon_n}(x) - f_{\Gamma_n}(x) \right| \to 0 \ \textit{as} \ \frac{\Delta}{\sigma_n} \to 0 \quad (19)$$

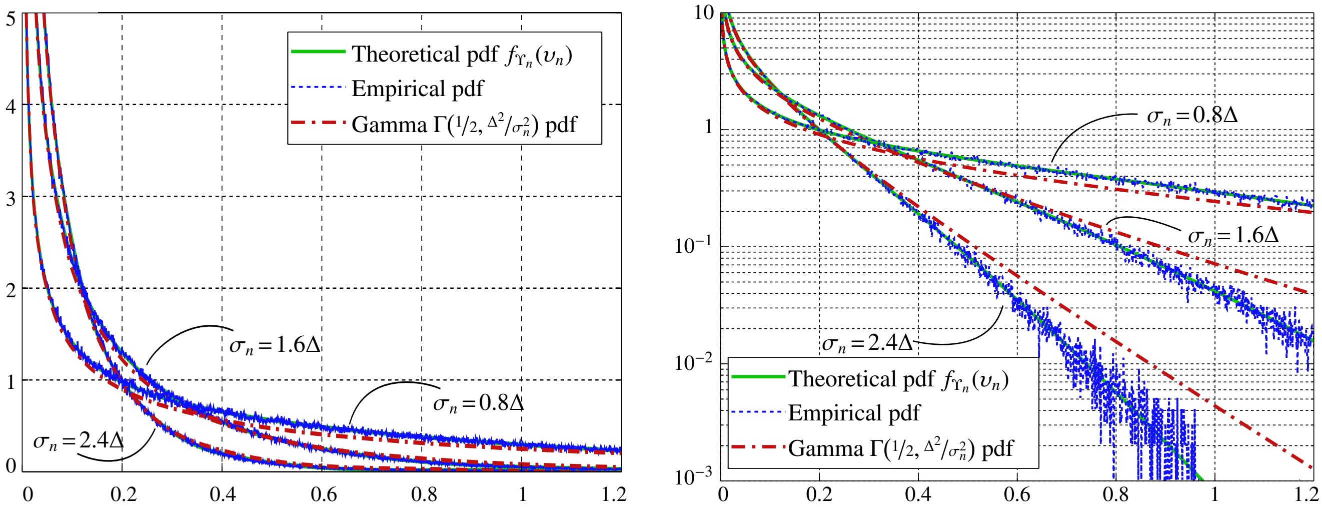
Proof. The proof of Theorem 1 is given in Appendix A □

The Figures 3 and 4 provide a graphical representation of this convergence. First, Figure 3 shows a comparison between the exact distribution of random variable v_n and the Gamma distribution $\Gamma\left({}^1\!/_2; {}^{\Delta^2}\!/_{\sigma_n^2}\right)$ for few values of noise standard deviation $\sigma_n = \{0.8, 1.6, 2.4\}$. Similarly, Figure 4 shows a comparison between the two first moments of v_n and the two first moments of a $\Gamma\left({}^1\!/_2; {}^{\Delta^2}\!/_{\sigma_n^2}\right)$ distribution. These figures obviously show the accuracy of theorem 1. This theorem is of crucial importance to study the statistical properties of the LR $\Lambda_2(z_n)$ and, thus, to establish the performance of proposed LR test.

4. Statistical Performance of the LR test.

4.1. Case of simple hypotheses, when $R = 2$.

In this section it is first proposed to study the statistical performance for the case of simple hypotheses, when $R = 2$. The results are then extended to the general case of $R \in (0; 1]$ in Section 4.2. To calculate easily the statistical performance of the LR test δ_R (11), the asymptotic approach is of crucial interest. Indeed, even though Theorem 1 establishes that v_n tends to be distributed as a Gamma distribution, it is not easy to explicit the distribution of the sum $\sum_{n=1}^{N} v_n$, see [33] for a detailed study. Moreover, from a practical point of view, the assumption that N grows to infinity is relevant due to the very large number of samples in digital media.



(a) Comparison between exact, empirical and approximated pdf of random variable υ_n, linear scale.

(b) Comparison between exact, empirical and approximated pdf of random variable υ_n, logarithmic scale.

Figure 3: Graphical comparison between the exact, the empirically obtained and the approximated pdf of random variable υ_n. The presented results correspond to $\sigma_n = \{0.8, 1.6, 2.4\}, \Delta = 1$. Empirical results are obtained from a Monte-Carlo simulation with 10^8 realizations.

Hence, to study the asymptotic distribution of the LR $\Lambda_2(\mathbf{Z})$, under both hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, it is proposed to apply the well known Lindeberg's central limit theorem (CLT) [26, theorem 11.2.5] from which it follows that under $\mathcal{H}_i$, $i = \{0, 1\}$:

$$\frac{\sum_{n=1}^{N} \Lambda_2(z_n) - \mathbb{E}_i\left[\Lambda_2(z_n)\right]}{\sqrt{\sum_{n=1}^{N} \mathbb{V}ar_i\left[\Lambda_1(z_n)\right]}} \rightsquigarrow \mathcal{N}(0,1), \qquad (20)$$

where for $i = \{0; 1\}$, the notations $\mathbb{E}_i[\cdot]$ and $\mathbb{V}ar_i[\cdot]$ respectively denote the mathematical expectation and the variance under $\mathcal{H}_i$ and $\rightsquigarrow$ represents the convergence in distribution as N tends to infinity.
It should be noted that the application of Lindeberg's CLT (20) required that the well known Lindeberg's condition [26, Eq. (11.11)] is satisfied. Here, Lindeberg's condition can easily be verified by using the Lyapounov's condition [26, Eq. (11.12)] which implies Lindeberg's condition. For the sake of clarity, let the mean expectation and the mean variance of $\Lambda_2(z_n)$ under hypotheses $\mathcal{H}_i$, $i = \{0, 1\}$ be defined as follows:

$$m_i = \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}_i\Big[\Lambda_2(z_n)\Big] \quad \text{and} \quad s_i^2 = \frac{1}{N}\sum_{n=1}^{N} \mathbb{V}ar_i\Big[\Lambda_2(z_n)\Big]. \quad (21)$$

In the present paper, it is aimed to design a reliable algorithm for LSB matching detection. Hence, it is very important that the decision threshold of the proposed test does not depend on any inspected medium parameter. To this end, let the test $\widetilde{\delta}_2$ associated with the "normalized" LR $\widetilde{\Lambda}_2(\mathbf{Z})$ be defined as follows:

$$\widetilde{\delta}_2 = \begin{cases} \mathcal{H}_0 & \text{if} \quad \widetilde{\Lambda}_2(\mathbf{Z}) \le \widetilde{\tau}_{\alpha_0}, \\ \mathcal{H}_1 & \text{if} \quad \widetilde{\Lambda}_2(\mathbf{Z}) > \widetilde{\tau}_{\alpha_0}. \end{cases} \qquad (22)$$

$$\text{where} \;\; \widetilde{\Lambda}_2(\mathbf{Z}) \overset{\text{def.}}{=} \frac{\sum\limits_{n=1}^{N} \Lambda_2(z_n) - \mathbb{E}_0\left[\Lambda_2(z_n)\right]}{\sqrt{\sum_{n=1}^{N} \mathbb{V}ar_0\left[\Lambda_1(z_n)\right]}} \qquad (23)$$
$$= \frac{1}{s_0\sqrt{N}}\left(\sum_{n=1}^{N} \Lambda_2(z_n) - N m_0\right).$$

As previously discussed, the calculation of the expectation and the variance of the LR $\Lambda_2(z_n)$ is not straightforward. However, Theorem 1 states that as Δ/σ_n tends to 0, the random variable υ_n tends to be distributed as a Gamma random variable. In such a situation, the moments of υ_n can be thus be approximated by the moments of a Gamma random variable $\Gamma(1/2, \Delta^2/\sigma_n^2)$ whose expectation is $\Delta^2/2\sigma_n^2$ and whose variance is $\Delta^4/2\sigma_n^4$. Hence, it follows from Theorem 1 and from the relation (17) between υ_n and $\Lambda_2(z_n)$ that:

$$\mathbb{E}_0[\Lambda_2(z_n)] = \mathbb{E}_0[\upsilon_n] - \frac{\Delta^2}{2\sigma_n^2} \to 0 \text{ as } \frac{\Delta}{\sigma_n} \to 0 \qquad (24)$$

$$\mathbb{V}ar_0[\Lambda_2(z_n)] = \mathbb{V}ar_0[\upsilon_n] \to \frac{\Delta^4}{2\sigma_n^4} \text{ as } \frac{\Delta}{\sigma_n} \to 0 \qquad (25)$$

The comparison between the empirical two first moments of υ_n, obtained from a Monte-Carlo simulation, and their asymptotic approximation are illustrated in Figure 4.

Consequently, using the moments of the LR Λ_2 given in Equations (24) and (25), the quantities m_0 and s_0^2 (21) are given, as ${}^{\Delta}/_{\sigma_n}$ tends to 0, by:

$$m_0 = 0 \text{ and } s_0^2 = \frac{\Delta^4}{2\bar{\sigma}^4} \text{ with } \frac{1}{\bar{\sigma}^4} = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{\sigma_n^4}. \tag{26}$$

It follows from (26) that the "normalized" LR $\widetilde{\Lambda}_2(\mathbf{Z})$ defined in Equation (23) is given by:

$$\widetilde{\Lambda}_2(\mathbf{Z}) = \frac{\sqrt{2}\bar{\sigma}^2}{\sqrt{N}\Delta^2}\sum_{n=1}^{N}\Lambda_2(z_n), \tag{27}$$

It Finally follows from Lindeberg CLT (20), that under hypothesis $\mathcal{H}_0$ the "normalized" LR $\widetilde{\Lambda}_2$ (27) satisfies:

$$\widetilde{\Lambda}_2(\mathbf{Z}) \rightsquigarrow \mathcal{N}(0,1) \quad \text{under} \quad \mathcal{H}_0, \tag{28}$$

asymptotically as ${}^{\Delta}/_{\sigma_n}$ tends to 0. Therefore, a short algebra establishes the following theorem.

Theorem 2. *For any given probability of false alarm $\alpha_0 \in (0,1]$, assuming that the parameter $\boldsymbol{\theta}$ is known, the decision threshold $\widetilde{\tau}_{\alpha_0}$ given by:*

$$\widetilde{\tau}_{\alpha_0} = \Phi^{-1}(1-\alpha_0) \tag{29}$$

where $\Phi^{-1}(\cdot)$ is the Gaussian inverse cumulative distribution, asymptotically warrants that the test $\widetilde{\delta}_2$ (22) is in $\mathcal{K}_{\alpha_0}$.

Proof. Using the result (28), it asymptotically holds that for any $\widetilde{\tau}_{\alpha_0} \in \mathbb{R}$:

$$\alpha_0(\widetilde{\delta}_2) = \mathbb{P}_0\left[\widetilde{\Lambda}_2(\mathbf{Z}) > \widetilde{\tau}_{\alpha_0}\right] = 1 - \Phi\left(\widetilde{\tau}_{\alpha_0}\right).$$

Hence, because Φ is strictly increasing, one has:

$$(1-\alpha_0(\widetilde{\delta}_2)) = \Phi(\widetilde{\tau}_{\alpha_0}) \Leftrightarrow \widetilde{\tau}_{\alpha_0} = \Phi^{-1}\left(1-\alpha_0(\delta_2)\right), \tag{30}$$

which proves Theorem 2. □

The main conclusion of Theorem 2 is that the decision threshold $\widetilde{\tau}_{\alpha_0}$ depends neither on the embedding rate R nor the medium parameters θ_n. Hence, by using the "normalized" LR $\widetilde{\Lambda}_2(\mathbf{Z})$, the same threshold allows to warrant a prescribed false-alarm probability α_0 whatever the analyzed medium and the embedding rate are.

The Lindeberg CLT (20) also implies that to asymptotically calculate the detection power of LR test $\widetilde{\delta}_2$ (22), one only needs to calculate the first moments of $\widetilde{\Lambda}_2(\mathbf{Z})$ under hypothesis $\mathcal{H}_1$. Indeed, it holds from Linderberg CLT that under hypothesis $\mathcal{H}_1$ the "normalized" LR $\widetilde{\Lambda}_2$ (23) - (27) satisfies:

$$\widetilde{\Lambda}_2(\mathbf{Z}) \rightsquigarrow \mathcal{N}\left(\frac{\sqrt{N}(m_1-m_0)}{s_0}, \frac{s_1^2}{s_0^2}\right) \sim \mathcal{N}\left(m_1\sqrt{N}\frac{2\bar{\sigma}^2}{\Delta^2}, s_1^2\frac{2\bar{\sigma}^4}{\Delta^4}\right) \tag{31}$$

It should be noted that under hypothesis $\mathcal{H}_0$ the expectation and the variance of the LR $\Lambda_2(z_n)$ could have been established using Theorem 1. On the opposite, the expectation and the variance of the LR $\Lambda_2(z_n)$ under hypothesis $\mathcal{H}_1$ can not explicitly given in a closed form. Hence, it is proposed, for clarity, to keep the notations m_1 and s_1^2 to represent, respectively, the mean expectation and the mean variance of the LR $\Lambda_2(z_n)$ under hypothesis $\mathcal{H}_1$ which are given, as Δ/σ_n tends to 0, by:

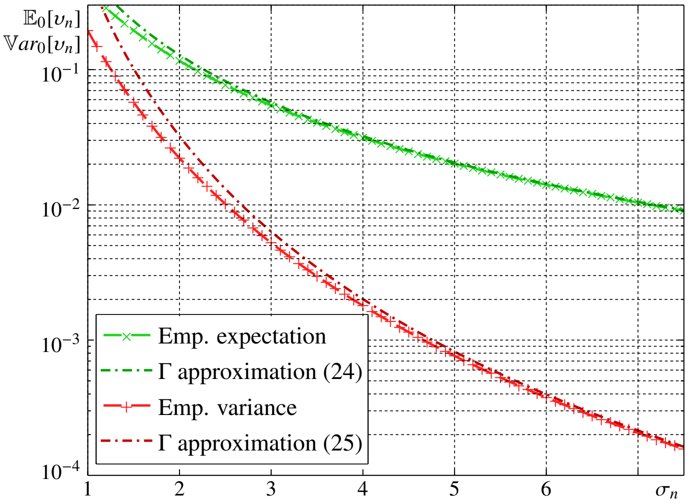


Figure 4: Graphical comparison between the proposed approximation υ_n two first moments and the empirically obtained moments. The presented results were obtained with $\sigma_n \in [\Delta, 7.5\Delta]$ and a Monte-Carlo simulation with 10^6 realizations.

$$m_1[\Lambda_2(z_n)] \overset{\text{def.}}{=} \frac{1}{N}\sum_{n=1}^{N}\int_{\mathbb{R}}\frac{\Delta}{2\sigma_n}\left(\phi\left(\frac{x+\Delta-\mu_n}{\sigma_n}\right)+\phi\left(\frac{x-\Delta-\mu_n}{\sigma_n}\right)\right)$$
$$\upsilon_n(x;\theta_n)\,dx - \frac{1}{2\sigma_n^2}, \tag{32}$$
$$s_1^2[\Lambda_2(z_n)] \overset{\text{def.}}{=} \frac{1}{N}\sum_{n=1}^{N}\int_{\mathbb{R}}\frac{\Delta}{2\sigma_n}\left(\phi\left(\frac{x+\Delta-\mu_n}{\sigma_n}\right)+\phi\left(\frac{x-\Delta-\mu_n}{\sigma_n}\right)\right)$$
$$\upsilon_n(x;\theta_n)^2 - \mathbb{E}_1[\Lambda_2(z_n)]^2\,dx. \tag{33}$$

Even though, the quantities m_1 and s_1^2 as given in (32)–(33) have a rather complicated expression, their numerical calculation is straightforward as long as the parameters θ_n are known.

From the asymptotic distributions (28) - (31) of the LR $\widetilde{\Lambda}_2(\mathbf{Z})$, under hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively, the detection power of the LR test $\widetilde{\delta}_2$ (22) is given by the following theorem.

Theorem 3. *For any $\alpha_0 \in (0,1]$, assuming that the parameter $\boldsymbol{\theta}$ is known, the power function $\widetilde{\beta}_{\delta_2}$ associated with the test $\widetilde{\delta}_2$ (22) is asymptotically given, as $N \to \infty$, by:*

$$\widetilde{\beta}_{\delta_2} = 1 - \Phi\left(\frac{s_0}{s_1}\Phi^{-1}(1-\alpha_0) + \frac{\sqrt{N}(m_0-m_1)}{s_1}\right),$$
$$= 1 - \Phi\left(\frac{1}{s_1}\left[\frac{\Delta^2}{\sqrt{2}\bar{\sigma}^2}\Phi^{-1}(1-\alpha_0) - \sqrt{N}m_1\right]\right). \tag{34}$$

Proof. It follows from (31) that for any decision threshold $\widetilde{\tau}_{\alpha_0} \in \mathbb{R}$ the power of the test $\widetilde{\delta}_2$ (22) is given by:

$$\widetilde{\beta}_{\delta_2} = \mathbb{P}_1\left[\widetilde{\Lambda}_2(\mathbf{Z}) > \widetilde{\tau}_{\alpha_0}\right] = 1 - \Phi\left(\frac{1}{s_1}\frac{\Delta^2}{\sqrt{2}\bar{\sigma}^2}\left(\widetilde{\tau}_{\alpha_0} - \sqrt{N}m_1\frac{\sqrt{2}\bar{\sigma}^2}{\Delta^2}\right)\right).$$

By substituting $\widetilde{\tau}_{\alpha_0}$ by the value given in Theorem 2, a short algebra leads to the relation (34). □

Even tough the expression of LRT power function $\widetilde{\beta}_{\delta_2}$ (34) depends on m_1 and s_1 which are not given in simple forms (32) - (33), two important remarks can be formulated. First, the m_1 and s_1 can easily be numerically evaluated in order to compute the detection power according to parameter $\boldsymbol{\theta}$. Moreover, the most important is that the decision threshold which allows to warrant a prescribe false-probability can be easily calculated from Theorem 2.

4.2. General case of $R \in (0, 1]$.

The case for which the embedding rate R can take any value in (0; 1] is treated in a similar manner as the case $R = 2$. The problem of designing an optimal test for hidden information detection has been shown to be particularly difficult in [48] for the LSB replacement algorithm. A thorough design a MP test uniformly with respect to the embedding rate lies outside of the scope of this paper which mainly studies the MP test for $R = 2$ and its practical implementation. Hence, it is proposed to use the test $\widetilde{\delta}_2$ (22) whatever the embedding rate R might be. Once again, the asymptotic distribution (28) - (31) are used to solve the decision problem (9).

The alternative hypothesis $\mathcal{H}_R$, that $\mathbf{Z}$ contains a stego-medium with embedding rate $R \in]0; 1]$, can be considered as a combination of stego and cover samples. Hence, the use of the law of total expectation and the law of total variance is relevant to calculate the two first moments of the LR $\widetilde{\Lambda}_2(\mathbf{Z})$. Using the moments given in (24)–(33), for the case $R = 2$, a short calculation gives:

$$m_R = \frac{R}{2} m_1, \tag{35}$$

$$s_R^2 = \frac{R}{2} s_1^2 + \frac{R}{2}\left(1 - \frac{R}{2}\right) m_1^2 + \left(1 - \frac{R}{2}\right)\left(\frac{\Delta^4}{2\bar{\sigma}^4}\right). \tag{36}$$

In other words, by using the test $\widetilde{\delta}_2$ (22) for any $R \in]0; 1]$ only the detection power is impacted. Indeed, the null hypothesis does not change, hence, the asymptotic distribution (28) of the LR $\widetilde{\Lambda}_2(\mathbf{Z})$ under $\mathcal{H}_0$ as well as the decision threshold $\widetilde{\tau}_{\alpha_0}$ (29) remain the same. This point is highlighted in the following theorem.

Theorem 4. *For any $\alpha_0 \in (0, 1]$, assuming that the parameter $\boldsymbol{\theta}$ is known, the power function $\widetilde{\beta}_{\delta_R}$ associated with the test $\widetilde{\delta}_2$ (22) is asymptotically given for any $R \in]0; 1]$ by:*

$$\widetilde{\beta}_{\delta_R} = 1 - \Phi\left(\frac{s_0}{s_R}\Phi^{-1}(1-\alpha_0) + \frac{R}{2}\frac{\sqrt{N}(m_0 - m_1)}{s_R}\right),$$

$$= 1 - \Phi\left(\frac{1}{s_R}\left[\frac{\Delta^2}{2\bar{\sigma}^2}\Phi^{-1}(1-\alpha_0) - \frac{R}{2}\sqrt{N} m_1\right]\right). \tag{37}$$

It should be reminded that the proposed LR test $\widetilde{\delta}_2$ (22) is based on the case of simple hypotheses, when $R = 2$. Figure 5 emphasizes the relevance of the proposed approach which extends the application of $\widetilde{\delta}_2$ to the case of composite hypotheses, when $R \in (0; 1[$, with the main goal of designing a test with known statistical properties. Figure 5 presents a comparison between the power function of the proposed test and the power function of the "clairvoyant" detector that knows R and, hence, can apply the most powerful test δ_R (11). The numerical comparison present in Figure 5 shows that the loss of the power is negligible. This particularly highlights that Theorem 4 provides an accurate upper bound on the performance one can expect for any detector which aims at detecting LSB matching data hiding.

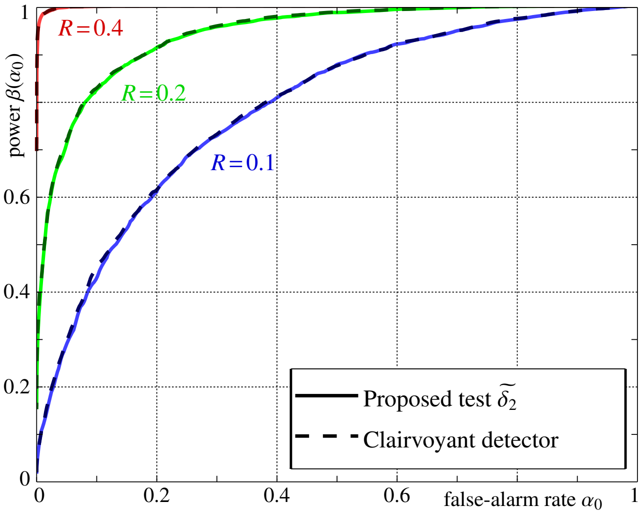


Figure 5: Numerical comparison between Proposed LR test $\widetilde{\delta}_2$ (22), and the clairvoyant detector which knows the embedding rate $R = 0.1$ ans, thus, uses the LR test design for this rate. Results were obtained from a Monte-Carlo simulation with 5.10^4 realizations using *Lena* image cropped to 128×128 samples and addition of a Gaussian white noise with $\sigma = 2$.

Finally, Equations (34) and (37) clearly shows that, as in many application of CLT, the power of the proposed test grows with the square root of sample number. Hence, Theorems 4 complies with the square root law of steganographic capacity [23]. In facts, from (37), a short algebra immediately permits us to establish that:

$$\lim_{\sqrt{N}/L \to 0} \widetilde{\beta}_{\delta_R} = 1 \quad \text{and} \quad \lim_{\sqrt{N}/L \to \infty} \widetilde{\beta}_{\delta_R} = \alpha_0. \tag{38}$$

5. Practical Design of LR test: Dealing with Nuisance Parameters.

In practice, the application of the test $\widetilde{\delta}_2$ (22) is compromised because neither the expectation μ_n nor the variance σ_n^2 of samples are known. In such a situation, an usual solution consist in replacing the unknown values by their Maximum Likelihood Estimation (MLE), denoted $\widehat{\mu}_n$ and $\widehat{\sigma}_n^2$, respectively, to design a Generalized Likelihood Ratio Test (GLRT).

However, accurate estimation of the parameters μ_n and σ_n is a difficult problem but necessary to obtain a high detection performance. In the following Sections 5.1 and 5.2 two different solutions are presented. The underlying idea is that most of digital media acquired with a recording device can be represented as signals whose properties vary smoothly from samples to samples. For instance, aa image is blurred by the optical system, hence, pixels expectation vary smoothly along each column (or row).

5.1. Polynomial Linear Parametric Model of Medium.

Following the general methodology [21], it is first proposed in this paper to use a local model of cover medium. The inspected medium $\mathbf{Z}$ is thus considered as a set of K non-overlapping signals of L samples, with $N \approx KL$. Similarly to the scalar case (1) - (2), let define for all $k \in \{1, \dots, K\}$ the k-th vector $\mathbf{z}_k = (\mathbf{z}_{k,1}, \dots, \mathbf{z}_{k,L})^T$ as:

$$\mathbf{z}_k = Q_\Delta(\mathbf{y}_k) \, , \, \mathbf{y}_k = \boldsymbol{\mu}_k + \boldsymbol{\xi}_k \sim \mathcal{N}(\boldsymbol{\mu}_k, \sigma_k^2 \mathbf{I}_L), \qquad (39)$$

where the operation of uniform quantization Q_Δ is applied on each sample individually, σ_k^2 is the samples variance assumed constant on each segment and $\mathbf{I}_L$ is the identity matrix of size $L \times L$.

The literature proposes a wide range of mathematical model to locally approximate vectors $\boldsymbol{\mu}_k = (\mu_{k,1}, \dots, \mu_{k,L})^T$ of expectations. In the present paper, it is proposed to use the following linear parametric model [8, 9]:

$$\boldsymbol{\mu}_k = \mathbf{H}\mathbf{x}_k, \qquad (40)$$

where $\mathbf{H}$ is a known full rank matrix of size $L \times p$, with $p < L$, and $\mathbf{x}_k \in \mathbb{R}^p$ is the nuisance parameter describing expectation of signal $\mathbf{z}_k$.

The hypothesis testing theory is relatively well developed for models such as (40). Indeed, such a model permits to reject the nuisance parameters μ_k by using the theory of invariance. In practice, the nuisance parameter rejection is usually done by using matrix $\mathbf{P}_H^\perp = \mathbf{I}_L - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$ because $\mathbf{P}_H^\perp \mathbf{H}\mathbf{x} = \mathbf{0}_L$, with $\mathbf{0}_L$ the null of vector of $\mathbb{R}^L$. Note that theory of invariance is used here but formally holds for non-quantized observations [26, chap. 6].

For the numerical simulation presented in this paper, a polynomial of degree $p-1$ was used; matrix $\mathbf{H}$ is thus given as:

$$\mathbf{H}^T = \begin{pmatrix} 1 & 1/L & \dots & (1/L)^{p-2} & (1/L)^{p-1} \\ 1 & 2/L & \dots & (2/L)^{p-2} & (2/L)^{p-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} \qquad (41)$$

From model (39) - (40), a short algebra shows that the maximum likelihood estimators (MLE) are given by:

$$\widehat{\boldsymbol{\theta}}_k = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T \Delta\mathbf{z}_k \text{ and } \widehat{\sigma}_k^2 = \frac{\|\mathbf{P}_H^\perp \Delta\mathbf{z}_k\|_2^2}{L-p}. \qquad (42)$$

One can note that the residuals can be written $\Delta\mathbf{z}_k - \widehat{\boldsymbol{\theta}}_k = \mathbf{P}_H^\perp \Delta\mathbf{z}_k$. Theory of invariance thus permits to reject nuisance parameters and to write from (16) the proposed GLR calculated on k-th segment as:

$$\Lambda^{\mathrm{glr}}(\mathbf{z}_k) = \sum_{l=1}^{L} \log \left[\exp\left(\frac{\Delta^2 \mathbf{P}_H^\perp \mathbf{z}_k}{\widehat{\sigma}_k^2} \right) + \exp\left(\frac{-\Delta^2 \mathbf{P}_H^\perp \mathbf{z}_k}{\widehat{\sigma}_k^2} \right) \right] - \log(2) - \frac{\Delta^2}{2\widehat{\sigma}_k^2}$$

Hence, the proposed GLRT is finally given as:

$$\delta^{\mathrm{glr}}(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{if} \quad \Lambda^{\mathrm{glr}}(\mathbf{Z}) \leq \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{if} \quad \Lambda^{\mathrm{glr}}(\mathbf{Z}) > \tau_{\alpha_0}, \end{cases} \qquad (43)$$

with:

$$\Lambda^{\mathrm{glr}}(\mathbf{Z}) = \frac{\sqrt{2}\,\widehat{\sigma}^2}{\sqrt{K(L-p)}\Delta^2} \sum_{k=1}^{K} \Lambda^{\mathrm{glr}}(\mathbf{z}_k). \qquad (44)$$

It should be noted that, for the sake of clarity, the proposed linear parametric model is presented in one-dimension. However, the numerical results presented in Section 6 for digital images also includes the application of GLR (43) using a two-dimensional polynomial model. In this case, the segments in one-dimension are replaced by non-overlapping blocks of L^2 pixels and, of course, the matrix $\mathbf{H}$ (41) is designed consequently.

5.2. Autoregressive Model of Medium.

Another approach to deal with the nuisance parameter $\boldsymbol{\theta}$ is possible by using a local autoregressive (AR) model of inspected medium. Such model have been especially studied in time series analysis [13] and image processing [27]. Such approach have been used for LSB replacement detection and lead to the well-known Weighted Stego-image steganalysis (WS), initially proposed in [16]. The authors propose to locally estimate the parameter μ_n by filtering the inspected image so that $\widehat{\mu}_n$ correspond to the mean of the four surrounding pixels. Similarly, the local variance of the four surrounding pixels is used to estimate σ_n^2. The WS method has been studied thoroughly in [24]; the authors especially enhanced the estimation of pixel expectations by testing different local filters.

In the present paper, it is proposed to use the WS filtering method to estimate the parameters μ_n and σ_n^2. Note that those estimations does not correspond to a rigorous statistical estimation but rather to a simple ad-hoc procedure. Following the WS method, it is proposed for digital images to estimate each $\mu_{m,n}$ as follows [10]:

$$\begin{aligned} \widehat{\mu}_{m,n} = \; & \frac{\Delta}{2}(z_{m,n-1} + z_{m,n+1} + z_{m-1,n} + z_{m+1,n}) \\ & - \frac{\Delta}{4}(z_{m-1,n-1} + z_{m-1,n+1} + z_{m+1,n-1} + z_{m+1,n+1}). \end{aligned} \qquad (45)$$

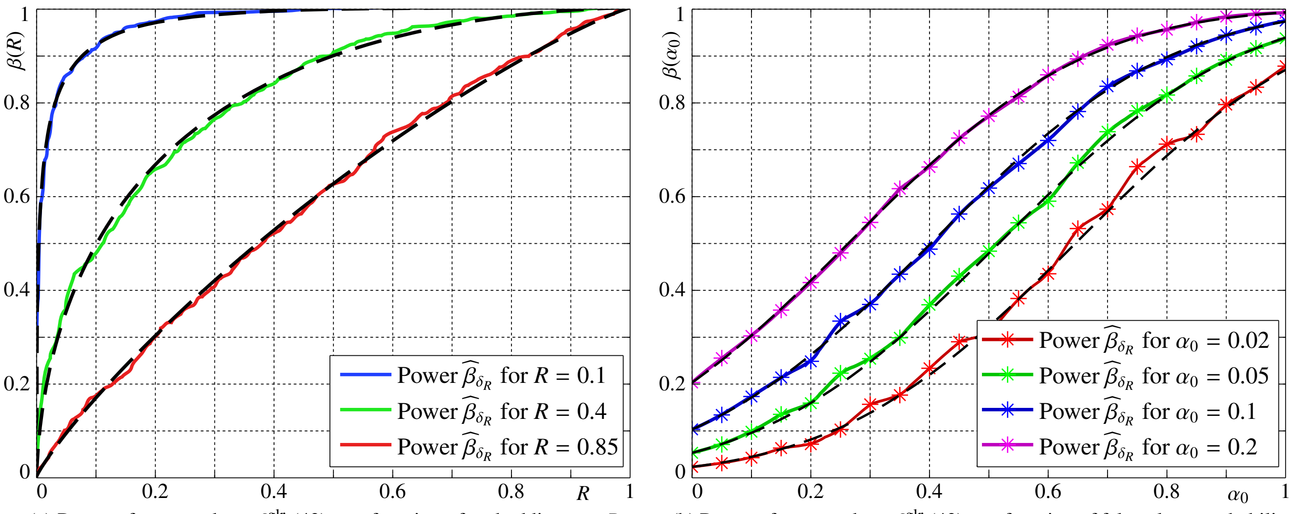
As previously described, the local variance of the four surrounding pixels is used to estimate pixels noise variance; this can formally be denoted:

$$\begin{aligned} \widehat{\sigma}_{m,n}^2 = \frac{1}{3} \Big[& (\Delta z_{m,n-1} - \widehat{\mu}_{m,n})^2 + (\Delta z_{m,n+1} - \widehat{\mu}_{m,n})^2 \\ & + (\Delta z_{m-1,n} - \widehat{\mu}_{m,n})^2 + (\Delta z_{m+1,n} - \widehat{\mu}_{m,n})^2 \Big]. \end{aligned} \qquad (46)$$

With the estimations $\widehat{\mu}_{m,n}$ and $\widehat{\sigma}_{m,n}^2$ defined in Equations (45) - (46), the proposed test based on AR model is given by:

$$\delta^{\mathrm{ar}}(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{if} \quad \Lambda^{\mathrm{ar}}(\mathbf{Z}) \leq \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{if} \quad \Lambda^{\mathrm{ar}}(\mathbf{Z}) > \tau_{\alpha_0}, \end{cases} \qquad (47)$$

$$\text{with } \Lambda^{\mathrm{ar}}(\mathbf{Z}) = \frac{\sqrt{2}\,\widehat{\sigma}^2}{\sqrt{MN}\Delta^2} \sum_{m=1}^{M} \sum_{n=1}^{N} \Lambda^{\mathrm{glr}}(z_{m,n}) \qquad (48)$$



(a) Power of proposed test δ^{glr} (43) as a function of embedding rate R.

(b) Power of proposed test δ^{glr} (43) as a function of false-alarm probability α_0.

Figure 6: Comparisons on simulated data of empirical performance of proposed test and theoretically established detection power.

and:

$$\Lambda^{\text{ar}}(z_{m,n}) = \log\left[\exp\left(\frac{\Delta(\widehat{\mu}_{m,n}-\Delta z_{m,n})}{\widehat{\sigma}^2_{m,n}}\right)+\exp\left(\frac{-\Delta(\widehat{\mu}_{m,n}-\Delta z_{m,n})}{\widehat{\sigma}^2_{m,n}}\right)\right] - \log(2) - \frac{\Delta^2}{2\widehat{\sigma}^2_{m,n}}.$$

Finally, it should be noted that, the direct use of the estimated variance $\widehat{\sigma}^2_n$ may lead to numerical instability due to the terms $\widehat{\sigma}^{-2}$. For instance, on areas which are over or under-exposed, the estimated variance might be zero which obviously causes a computational problem. To tackle this problem, it is proposed in the present paper to add a (small) constant $w = 0.2$ to the estimated variance. This technique is similar to the weighted of variance estimation proposed in the WS algorithm [10, 16, 24].

6. Numerical Simulations

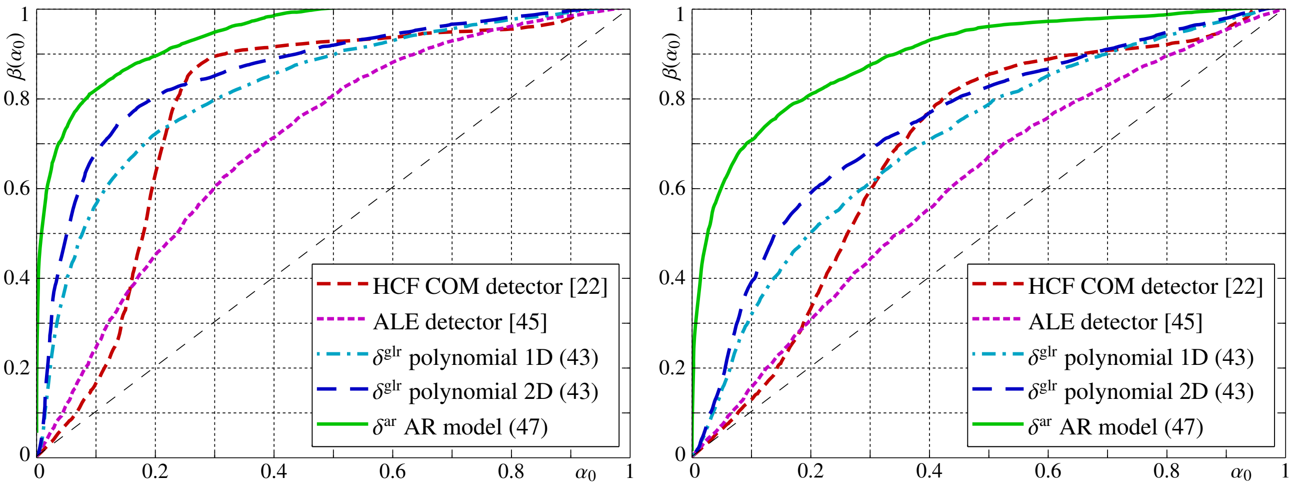
One of the main motivations for this paper is to show that the hypothesis testing theory can be applied in practice to design a reliable LSB matching detector.

As previously discussed, the reliability of the proposed tests heavily depends of the possibility to theoretically predict the parameters of proposed test in practice. To verify that the proposed test performs as established by Theorems 2 - 4 a numerical simulation was performed on simulated data. The Monte -Carlo simulation was conducted by generated 128^2 i.i.d samples which follows a Gaussian distribution with $\sigma_n = 3.78$ and $\Delta = 1$. Those set of samples were analyzed 10^4 times before and after have been subjected to random hidden message embedding at various rate R. The results are presented in Figure 6 in to different manners. First, Figure 6a offers a comparison between theoretical and empirical detection power as a function of prescribed false-alarm probability α_0 for three different embedding rates $R = \{0.1, 0.4, 0.85\}$ which respectively correspond to low, medium and high embedding rates which enhanced readability of results. Figure 6a highlights that the expected detection power fits almost perfectly the observed results. Similarly, Figure 6b presents a comparison between theoretical and empirical detection power as a function of embedding rate R for four different prescribed false-alarm probabilities $\alpha_0 = \{0.02, 0.05, 0.1, 0.2\}$ above which a warranted false-alarm probability is not meaningful while preserving readability of Figure 6b. The results presented in Figures 6 highlight the relevance of proposed approach which allows us to accurately establish the detection power of proposed test.

Another main motivation of the present paper is to design an efficient statistical test for LSB matching detection. To highlight the efficiency of proposed tests, a numerical comparison with state-of-the-art detectors on large digital media databases is required. The potential competitors for LSB matching detection are not as numerous as for LSB replacement. As briefly described in the introduction, the operational context selected in this paper eliminates all prior-art detectors based on machine learning. Almost every other detectors found in the literature are based on the histogram of inspected medium. For the present comparison, two recent histogram-based detectors, namely ALE [45] and the adjacency HCF COM [22] detector, were used due to their high detection performance.

Two different implementations of proposed GLR test Λ^{glr} (44) are presented in Figures 7 and 8. The test referred as "polynomial 1D" exploits the one-dimensional polynomial model of medium described in Section 5.1 with a segment size $L = 16$ and a degree $p - 1 = 6$. Alternatively, the test called "polynomial 2D" uses a two-dimension model of images with block size $L = 4$ and order 2.

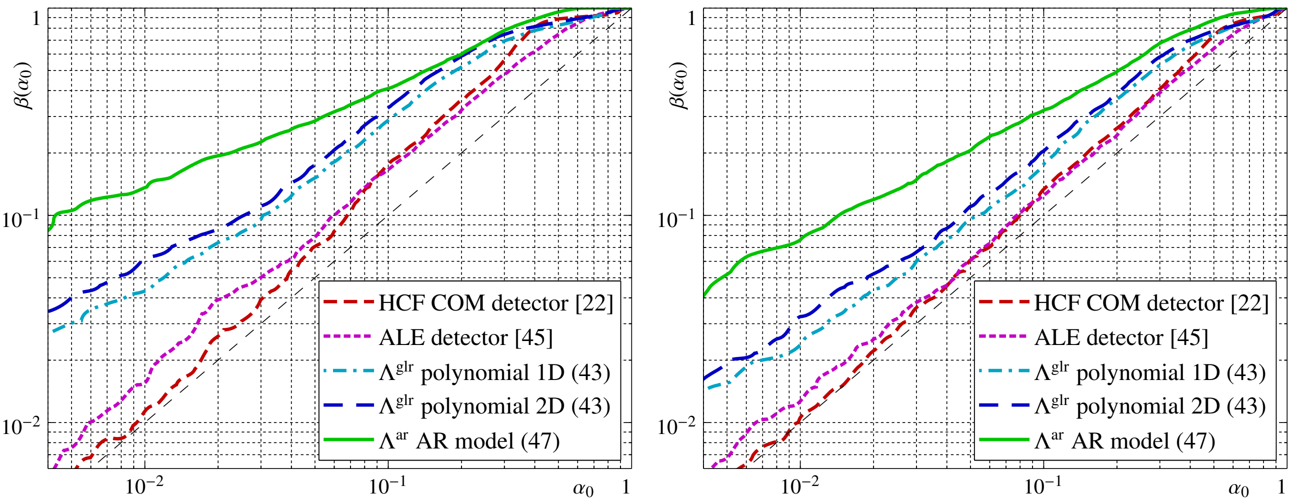
Figure 7 shows the results obtained with the 10 000 uncompressed grayscale images of size 512 × 512 from the second



(a) ROC curves presenting the detection performance of compared algorithms for $R = 1$.

(b) ROC curves presenting the detection performance of compared algorithms for $R = 0.5$.

Figure 7: Numerical comparisons of detectors performance using the 10 000 images from BOWS database.



(a) ROC curves presenting the detection performance of compared algorithms for $R = 1$.

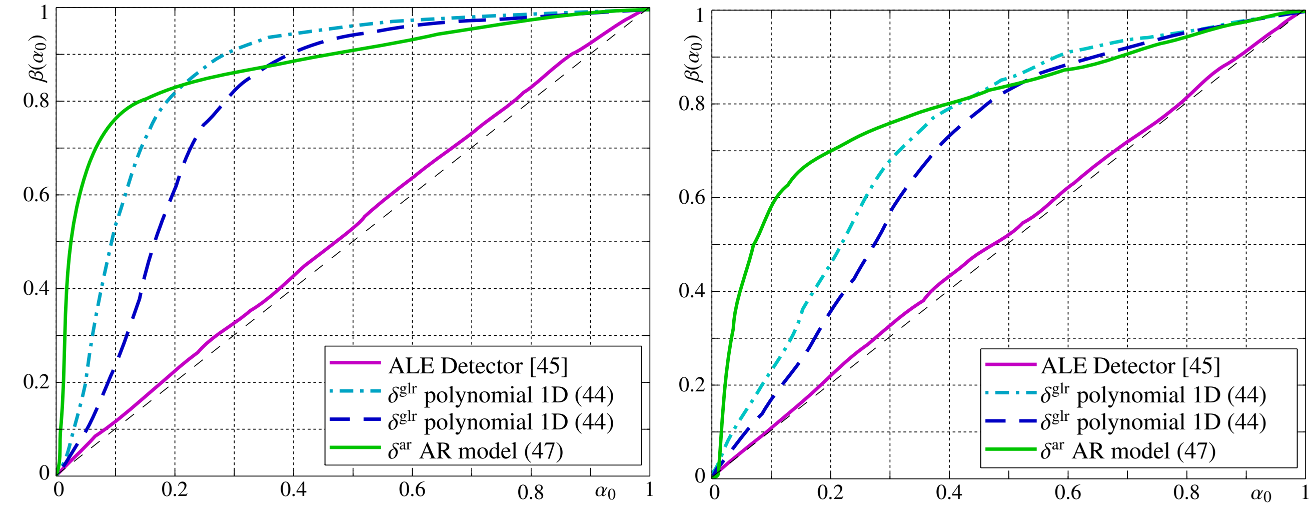
(b) ROC curves presenting the detection performance of compared algorithms for $R = 0.5$.

Figure 8: Numerical comparisons of detectors performance using the 10 000 images from BOSS database [1].

edition of Break Our Watermarking System (BOWS) contest[1]. The embedding rate was $R = 1$ in Figure 7a and $R = 0.5$ in Figure 7b. Both figures show that the proposed tests achieve a better detection power for any prescribed false-alarm probability. In addition, it should be highlight that the proposed test based on a AutoRegressive (AR) image model has a higher detection power that the proposed tests based on a polynomial image model. Similarly, the 2D polynomial image model seems to achieve a higher detection power that the 1D polynomial. Finally, the ALE and the HCF COM detector has very poor detection power for low false-alarm probabilities (typically $\alpha_0 \in [0, 0.2]$); however, the HCF COM detector achieve a rather good detection power for much higher false-alarm probability (typically $\alpha_0 \in [0.3, 0.4]$). The comprehension of this phenomenon is out the scope of present paper. Nevertheless, it should be noted that the design of a test which can only detect hidden information accurately with such a high false-alarm probability is hardly usable in a practical forensics application due to the high number of medium that must be inspected.

Similarly, the results presented in Figure 8 offers a numerical comparison of detection algorithm performance on a large image database. For a meaningful comparison, another image database, namely BOSSbase [1], was used. This database is made of 10 000 grayscale images all of size 512 × 512 pixels. The embedding rate was $R = 1$ in Figure 8a and $R = 0.5$ in Figure 8b. However, it should be noted that to highlight the efficiency of proposed test, the results are presented using a logarithmic scale; with this representation, Figures 8b and 8a em-

[1] Internet website of BOWS Contest: `http://bows2.ec-lille.fr/`



(a) ROC curves presenting the detection performance of compared algorithms for $R = 0.5$.

(b) ROC curves presenting the detection performance of compared algorithms for $R = 0.25$.

Figure 9: Extension of the proposed approach for different sort of digital media: Performance of the proposed test using the uncompressed music files database.

phasize the fact that proposed test has a much higher detection power than the competitors for low false-alarm probabilities.
For a complete comparison on algorithms efficiency, the mean computation time obtained on the BOSSbase [1] image is presented in Table 1. It can be noted that the computation time for ALE and HCF COM detectors is less than for the proposed tests. Indeed, those two competitors rely on image histogram whose calculation is much simple that the estimation of expectation and noise variance of each pixel. However, it can be noted in Table 1 that the proposed test using a one-dimensional polynomial model need a similar computation time compared to the HCF COM detector while the two-dimensional polynomial model requires twice more time. Obviously the proposed test using an AR model is the slower algorithm but its computation time remain reasonable compare to the other algorithm.

Finally, to show the potential application to other sorts of digital media, it is proposed to perform an experimentation using a database of uncompressed sounds. Those sounds were downloaded from the Internet website `http://archive.org/`. Among the wide range of audio files proposed on this website, a set of music files recorded using the Free Lossless Audio Codec (FLAC file format) was downloaded and then cropped to $1024^2 = 1048576$ non-overlapping samples to create a database of 11 200 sound files. It should be noted that the HCF COM detector cannot be used on such file as it relies on a two-dimension Discrete Fourier Transform (DFT) of pixels co-occurrence values; similarly, the proposed test using a 2D polynomial model of medium could not be applied and the test based on AR medium model was adapted by using a one-dimension convolution kernel, see Section 5.2.
The results presented in Figure 9a are obtained with an embedding rate $R = 0.5$ and the results presented in Figure 9b with $R = 0.25$. Once again, those results show the relevance of proposed approach which permits to design a test with high detection power. More important, it emphasis that contrary the prior-art, the proposed approach is sufficiently general to allow a straightforward extension to inspect other sort of digital media. On the opposite, the detector proposed in [22] cannot be used while the detector proposed in [45] performs poorly.

Algorithm	Mean computation time (ms)
δ^{glr} polynomial 1D (43)	36.80
δ^{glr} polynomial 2D (43)	75.10
δ^{ar} AR model (47)	129.67
HCF COM detector [22]	19.92
ALE Detector [45]	25.69

Table 1: Mean computation time on BOSSbase [1] images for the different LSB replacement detector compared in Section 6. The algorithms were running using *Matlab* software on a *Intel®Core™2 Duo CPU E8400, 3.00GHz.*

7. Conclusion and future works.

The first step to fill the gap between hypothesis testing theory and steganalysis was recently proposed in [12, 7, 48]. This paper extends this first step to the case of LSB matching. By casting the problem of LSB matching steganalysis in the framework of hypothesis testing theory, the most powerful likelihood ratio test is designed. Then, a thorough statistical study permits the analytical calculation of its performance in terms of the false-alarm probability and detection power. To apply this test in practice, unknown medium parameters have to be estimated. Based on two simple estimations of unknown parameters, two practical tests are proposed.
The relevance of the proposed approach is emphasized through numerical experimentation. Compared to two leading

histogram-based detectors, the proposed practical tests achieve a better detection power. In addition, numerical results show that the proposed approach is sufficiently general to be applied on digital images and audio signals contrary to the prior-art detectors which mainly focuses on digital images.

Appendix A. Exact and Approximated Distributions of the LR $\Lambda_2(z_n)$

The goal of this Appendix A is to provide a demonstration of Theorem 1 in two steps: first, Appendix A.1 provide an exact expression of random variable υ_n distribution. By using this exact expression, the approximation as a Gamma distribution proposed in Theorem 1 is demonstrated in Appendix A.2.

Appendix A.1. Exact distribution of LR $\Lambda_2(z_n)$

Let us first recall that υ_n is defined from Equation 17 as:

$$\upsilon_n = \log\left[\exp\left(\frac{-\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right) + \exp\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right)\right] - \log(2),$$

$$= \log\left[\frac{\exp\left(\frac{-\Delta(\mu_n-\Delta z_n)}{\sigma_n^2}\right) + \exp\left(\frac{\Delta(\mu_n-\Delta z_n)}{\sigma_n^2}\right)}{2}\right].$$

with $y_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$ and $z_n = Q_\Delta(y_n)$.
It is first propose to provide an exact expression of random variable Υ_n cdf, then this expression is used to obtain an expression of the pdf. By definition the cdf is defined as:

$$\forall x \in \mathbb{R}, \, F_{\Upsilon_n}(x) = \mathbb{P}\left[\upsilon_n \leq x\right].$$

Due to the strict monotonicity of function $\exp(x)$, this probability can be written as:

$$F_{\Upsilon_n}(x) = \mathbb{P}\left[\log\left[\frac{\exp\left(\frac{-\Delta(\mu_n-\Delta z_n)}{\sigma_n^2}\right) + \exp\left(\frac{\Delta(\mu_n-\Delta z_n)}{\sigma_n^2}\right)}{2}\right] \leq x\right],$$

$$= \mathbb{P}\left[\frac{\exp\left(\frac{\Delta(\mu_n-\Delta z_n)}{\sigma_n^2}\right) + \exp\left(\frac{\Delta(\mu_n-\Delta z_n)}{\sigma_n^2}\right)^{-1}}{2} \leq \exp(x)\right]. \quad \text{(A.1)}$$

In addition, it is straightforward to verify that the quadratic equation of the form:

$$\frac{z + z^{-1}}{2} \leq x \Leftrightarrow z^2 - 2xz + 1 \leq 0,$$

admits, with respect to z, the following set of solutions:

$$x - (x^2 - 1)^{\frac{1}{2}} \leq z \leq x + (x^2 - 1)^{\frac{1}{2}}. \quad \text{(A.2)}$$

Therefore, let us define for clarity:

$$\begin{aligned} x_0^\star &= \exp(x) - \sqrt{\exp(x)^2 - 1}, \\ \text{and} \quad x_1^\star &= \exp(x) + \sqrt{\exp(x)^2 - 1}. \end{aligned} \quad \text{(A.3)}$$

Putting the solutions (A.3) in Equation (A.1) lead to following expression of υ_n cdf:

$$F_{\Upsilon_n}(x) = \mathbb{P}\left[\exp\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right) \in \left(x_0^\star, x_1^\star\right)\right]$$

$$= \mathbb{P}\left[\exp\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right) \leq x_1^\star\right] - \mathbb{P}\left[\exp\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right) \leq x_0^\star\right]. \quad \text{(A.4)}$$

From the properties of Gaussian random variable one has:

$$\mathbb{P}\left[\exp\left(\frac{\Delta(\mu_n - \Delta z_n)}{\sigma_n^2}\right) \leq x\right] = \Phi\left(\frac{\sigma_n}{\Delta}\log(x)\right), \quad \text{(A.5)}$$

which corresponds to a special case of log-normal random variable cdf [20, Chap. 14].
Finally, putting results (A.5) into Equation (A.4) permits to obtain the exact expression of υ_n cdf:

$$F_{\Upsilon_n}(x) = \Phi\left(\frac{\sigma_n}{\Delta}\log(x_1^\star)\right) - \Phi\left(\frac{\sigma_n}{\Delta}\log(x_0^\star)\right), \quad \text{(A.6)}$$

$$= \Phi\left(\frac{\sigma_n}{\Delta}\log\left[\exp(x) - \sqrt{\exp(x)^2 - 1}\right]\right)$$

$$- \Phi\left(\frac{\sigma_n}{\Delta}\log\left[\exp(x) - \sqrt{\exp(x)^2 - 1}\right]\right).$$

To obtain an exact expression of the pdf $f_{\Upsilon_n}(x)$ it suffice to differentiate $F_{\Upsilon_n}(x)$ with respect to x. By using Equation (A.6) a direct calculation permit to have:

$$f_{\Upsilon_n}(x) = \frac{\sigma_n}{\Delta\sqrt{2\pi}\sqrt{\exp(2x) - 1}}\exp(x)\left[\exp\left(-\frac{\sigma_n^2}{2\Delta^2}\log(x_1^\star)^2\right)\right.$$

$$\left. + \exp\left(-\frac{\sigma_n^2}{2\Delta^2}\log(x_0^\star)^2\right)\right]. \quad \text{(A.7)}$$

Appendix A.2. Proof of Theorem 1: Asymptotic LR $\Lambda_2(z_n)$ distribution

Let us denote f_Γ the pdf of a Gamma distribution with shape parameter $k = 1/2$ and scale parameter $\theta = \Delta^2/\sigma_n^2$ whose expression is:

$$f_\Gamma(x) = \frac{\theta^{-k}}{\Gamma(k)}x^{k-1}\exp\left(-\frac{x}{\theta}\right) = \frac{\sigma_n}{\Delta\sqrt{\pi x}}\exp\left(-\frac{\sigma_n^2}{\Delta^2}x\right)$$

From Equation (A.7), it is immediate to verify that:

$$\frac{f_{\Upsilon_n}(x)}{f_\Gamma(x)} = \sqrt{\frac{x}{2}}\frac{\exp\left(x(1 + \sigma_n^2/\Delta^2)\right)}{\sqrt{\exp(2x) - 1}}\left[\exp\left(-\frac{\sigma_n^2}{2\Delta^2}\log(x_1^\star)^2\right)\right.$$

$$\left. + \exp\left(-\frac{\sigma_n^2}{2\Delta^2}\log(x_0^\star)^2\right)\right].$$

Finally, a Taylor expansion shows that:

$$\frac{f_{\Upsilon_n}(x)}{f_\Gamma(x)} = 1 + \frac{x}{2} + \frac{x^2}{24}\left(1 - 8\frac{\sigma_n^2}{\Delta^2}\right) + o(x^2).$$

This first results can be interpreted as follows: it is obvious that for any $x \in \mathbb{R}$, the probability $\mathbb{P}[\upsilon \geq x] \sim 1 - \Phi(\sigma x)$ tends to zero as σ_n tends to infinity. Hence, the expression of the Taylor expansion clearly shows that as σ_n becomes arbitrarily large, the difference between the function $f_{\Upsilon_n}(x)$ and $f_\Gamma(x)$ becomes negligible in probability. This "intuition" is confirmed by the following calculations which shows that the Kullback-Leibler divergence between these two distributions tends to zero as σ_n grows.

First, let us denote that from Equation (A.7), a direct calculation permits to have:

$$\log\left(\frac{f_{\Upsilon_n}(x)}{f_\Gamma(x)}\right) = \frac{\log(\frac{x}{2})}{2} - \frac{\log(2x-1)}{2} + x\left(1+\frac{\sigma_n^2}{\Delta^2}\right)$$
$$-\log\left[\exp\left(-\frac{\sigma_n^2}{2\Delta^2}\log(x_1^\star)^2\right)+\exp\left(-\frac{\sigma_n^2}{2\Delta^2}\log(x_0^\star)^2\right)\right] \qquad \text{(A.8)}$$

From the Equations (A.7) and (A.8) the exact expression of Kullback-Leibler divergence it is obviously difficult to formally calculate. However, it is easy to show that a coarse upper bound of f_{Υ_n} is given by:

$$f_{\Upsilon_n}(x) \leq \frac{\sigma_n}{\Delta\sqrt{2\pi}}\left[\exp\left(-\frac{\sigma_n^2}{\Delta^2}x\right)\left(1+2^{-\sigma_n^2/\Delta^2}\right)\right] \qquad \text{(A.9)}$$

From Equation (A.8) and the bound provided in Equation (A.9) one finally has:

$$f_{\Upsilon_n}(x)\log\left(\frac{f_{\Upsilon_n}(x)}{f_\Gamma(x)}\right) \leq x\exp\left(-\frac{\sigma_n^2}{\Delta^2}x\right)$$
$$\Leftrightarrow \int_0^\infty f_{\Upsilon_n}(u)\log\left(\frac{f_{\Upsilon_n}(u)}{f_\Gamma(u)}\right)du \leq \int_0^\infty x\frac{\Delta^2}{\sigma_n^2}\exp\left(-\frac{\sigma_n^2}{\Delta^2}x\right)dx$$
$$\leq \frac{\Delta^4}{\sigma_n^4} \qquad \text{(A.10)}$$

Though the bound proposed in Equation (A.9) is rather large or coarse, it suffice to show that the Kullback-Leibler divergence is upper bounded by Δ^4/σ_n^4 and hence rapidly becomes negligible as $\sigma_n >> \Delta$. Note that a more precise bound could potentially be found because, as shown in Figure 4, even for not large σ_n^2/Δ the proposed approximation remain accurate.

Finally, an alternative demonstration is possible because a taught calculation shows that:

$$x_M = \left\{\arg\max_{x\in\mathbb{R}^+} |f_{\Upsilon_n}(x) - f_\Gamma(x)|^2\right\} \to 0 \text{ as } \frac{\Delta}{\sigma} \to \infty.$$

Hence, from the Taylor expansion of $\frac{f_{\Upsilon_n}(x)}{f_\Gamma(x)}$ one get that:

$$\lim_{\Delta/\sigma_n\to 0}\max_{x\in\mathbb{R}^+}\left|f_{\Upsilon_n}(x) - f_\Gamma(x)\right|_2^2 = 0.$$

[1] P. Bas, T. Filler, T. Pevný, Break our steganographic system — the ins and outs of organizing boss, in: Information Hiding, 13th International Workshop, LNCS vol.6958, Springer, 2011, pp. 59–70.

[2] R. Böhme, Advanced Statistical Steganalysis, Springer, 1st edition, 2010.

[3] G. Cancelli, G. Doerr, M. Barni, I. Cox, A comparative study of ±1 steganalyzers, in: IEEE Workshop on Multimedia Signal Processing, 2008, pp. 791 –796.

[4] G. Cancelli, G. Doerr, I. Cox, M. Barni, Detection of ±1 LSB steganography based on the amplitude of histogram local extrema, in: IEEE International Conference on Image Processing, 2008, pp. 1288 –1291.

[5] C.K. Chan, L. Cheng, Hiding data in images by simple LSB substitution, Pattern Recognition Elsevier 37 (2004) 469 – 474.

[6] A. Cheddad, J. Condell, K. Curran, P.M. Kevitt, Digital image steganography: Survey and analysis of current methods, Signal Processing Elsevier 90 (2010) 727 – 752.

[7] R. Cogranne, C. Zitzmann, L. Fillatre, I. Nikiforov, F. Retraint, P. Cornu, A cover image model for reliable steganalysis, in: Information Hiding, 13th International Workshop, LNCS vol.6958, Springer, 2011, pp. 178 – 192.

[8] R. Cogranne, C. Zitzmann, L. Fillatre, F. Retraint, I. Nikiforov, P. Cornu, Statistical decision by using quantized observations, in: IEEE International Symposium on Information Theory, pp. 1135 – 1139.

[9] R. Cogranne, C. Zitzmann, I. Nikiforov, F. Retraint, L. Fillatre, P. Cornu, Statistical Detection of LSB Matching in the Presence of Nuisance Parameters, in: IEEE Workshop on Statistical Signal Processing *(to be published)* .

[10] R. Cogranne, C. Zitzmann, F. Retraint, I. Nikiforov, L. Fillatre, P. Cornu, Statistical Detection of LSB Matching Using Hypothesis Testing Theory, in: Information Hiding *(to be published)*, LNCS, Springer, 2012.

[11] I. Cox, M. Miller, J. Bloom, J. Fridrich, T. Kalker, Digital Watermarking and Steganography, Morgan Kaufmann, 2nd edition edition, 2007.

[12] O. Dabeer, K. Sullivan, U. Madhow, S. Chandrasekaran, B. Manjunath, Detection of hiding in the least significant bit, IEEE Transactions on Signal Processing 52 (2004) 3046 – 3058.

[13] R. Dubes, A. Jain, Random field models in image analysis, Journal of applied statistics 16 (1989) 131–164.

[14] A. Foi, M. Trimeche, V. Katkovnik, K. Egiazarian, Practical poissonian-gaussian noise modeling and fitting for single-image raw-data, IEEE Transactions on Image Processing, 17 (2008) 1737–1754.

[15] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications, Cambridge University Press, 1st edition edition, 2009.

[16] J. Fridrich, M. Goljan, On estimation of secret message length in LSB steganography in spatial domain, in: in Proc. of the SPIE, volume 5306, pp. 23–34.

[17] M. Goljan, J. Fridrich, T. Holotyak, New blind steganalysis and its implications, in: in Proc. of the SPIE, volume 5306, pp. 1–13.

[18] J. Harmsen, W. Pearlman, Higher-order statistical steganalysis of palette images, in: in Proc. of the SPIE, volume 5020.

[19] G. Healey, R. Kondepudy, Radiometric CCD camera calibration and noise estimation, IEEE Trans. Pattern Anal. Mach. Intell. 16 (1994) 267–276.

[20] N.L. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, volume I, Wiley & Sons, 2nd edition edition, 1994.

[21] V. Katkovnik, K. Egiazarian, J. Astola, Local Approximation Techniques in Signal and Image Processing, SPIE Press, Monograph, 2006.

[22] A. Ker, Steganalysis of LSB matching in grayscale images, Signal Processing Letters, IEEE 12 (2005) 441 – 444.

[23] A.D. Ker, A capacity result for batch steganography, Signal Processing Letters, IEEE 14 (2007) 525–528.

[24] A.D. Ker, R. Böhme, Revisiting weighted stego-image steganalysis, in: in Proc. of the SPIE, volume 6819, pp. 501–517.

[25] S. Kouider, M. Chaumont, W. Puech, Technical Points about Adaptive Steganography by Oracle (ASO), in: *to be published in*EUSIPCO'2012, the 20th European Signal Processing Conference, 2012.

[26] E. Lehmann, J. Romano, Testing Statistical Hypotheses, Second Edition, Springer, 3rd edition, 2005.

[27] W. Li, A. McLeod, Distribution of the residual autocorrelations in multivariate arma time series models, Journal of the Royal Statistical Society. Series B (Methodological) (1981) 231–239.

[28] Y.C. Li, C.M. Yeh, C.C. Chang, Data hiding based on the similarity between neighboring pixels with reversibility, Digital Signal Processing Elsevier 20 (2010) 1116 – 1128.

[29] Q. Liu, A.H. Sung, Z. Chen, J. Xu, Feature mining and pattern classification for steganalysis of LSB matching steganography in grayscale images, Pattern Recognition Elsevier 41 (2008) 56 – 66.

[30] W. Luo, F. Huang, J. Huang, Edge adaptive image steganography based on LSB matching revisited, IEEE Transactions on Information Forensics and Security 5 (2010) 201 –214.

[31] X.Y. Luo, D.S. Wang, P. Wang, F.L. Liu, A review on blind detection for image steganography, Signal Processing Elsevier 88 (2008) 2138 – 2157.

[32] S. Lyu, H. Farid, Steganalysis using higher-order image statistics, IEEE Transactions on Information Forensics and Security 1 (2006) 111 – 119.

[33] P. Moschopoulos, The distribution of the sum of independent gamma random variables, Annals of the Institute of Statistical Mathematics 37 (1985) 541–544.

[34] C. Munuera, Steganography and error-correcting codes, Signal Processing Elsevier 87 (2007) 1528 – 1533.

[35] A. Nissar, A. Mir, Classification of steganalysis techniques: A study, Digital Signal Processing Elsevier 20 (2010) 1758 – 1770.

[36] H. Rifà -Pous, J. Rifà, Product perfect codes and steganography, Digital Signal Processing Elsevier 19 (2009) 764 – 769.

[37] C. Scott, Performance measures for Neyman-Pearson classification, IEEE Trans. Inform. Theory 53 (2007) 2852–2863.

[38] S. Kraut, L.L. Scharf, L.T.. McWhorter, Adaptive subspace detectors, IEEE Transactions on Signal Processing 49 (2001) 1 – 16.

[39] G. Simmons, The prisoners problem and the subliminal channel, CRYPTO (1983) 51–67.

[40] P. Tsai, Y.C. Hu, H.L. Yeh, Reversible image hiding scheme using predictive coding and histogram shifting, Signal Processing Elsevier 89 (2009) 1129 – 1143.

[41] R.Z. Wang, C.F. Lin, J.C. Lin, Image hiding by optimal LSB substitution and genetic algorithm, Pattern Recognition Elsevier 34 (2001) 671 – 683.

[42] E.W. Weisstein, CRC concise encyclopedia of mathematics, Chapman & Hall/CRC, 2003.

[43] D.C. Wu, W.H. Tsai, A steganographic method for images by pixel-value differencing, Pattern Recognition Letters Elsevier 24 (2003) 1613 – 1626.

[44] C.H. Yang, C.Y. Weng, S.J. Wang, H.M. Sun, Adaptive data hiding in edge areas of images with spatial LSB domain systems, IEEE Transactions on Information Forensics and Security 3 (2008) 488 –497.

[45] J. Zhang, I. Cox, G. Doerr, Steganalysis for LSB matching in images with high-frequency noise, in: IEEE Workshop on Multimedia Signal Processing, 2007, pp. 385 –388.

[46] T. Zhang, W. Li, Y. Zhang, E. Zheng, X. Ping, Steganalysis of LSB matching based on statistical modeling of pixel difference distributions, Information Sciences Elsevier 180 (2010) 4685 – 4694.

[47] T. Zhang, X. Ping, A new approach to reliable detection of LSB steganography in natural images, Signal Processing 83 (2003) 2085 – 2093.

[48] C. Zitzmann, R. Cogranne, F. Retraint, I. Nikiforov, L. Fillatre, P. Cornu, Statistical decision methods in hidden information detection, in: Information Hiding, 13th International Workshop, LNCS vol.6958, Springer, 2011, pp. 163 – 177.