IEEE COPYRIGHT AND CONSENT FORM

To ensure uniformity of treatment among all contributors, other forms may not be substituted for this form, nor may any wording of the form be changed. This form is intended for original material submitted to the IEEE and must accompany any such material in order to be published by the IEEE.Please read the form carefully and keep a copy for your files.

Author Online Use

6. Personal Servers. Authors and/or their employers shall have the right to post the accepted version of IEEE-copyrighted articles on their own personal servers or the servers of their institutions or employers without permission from IEEE, provided that the posted version includes a prominently displayed IEEE copyright notice and, when published, a full citation to the original IEEE publication, including a link to the article abstract in IEEE Xplore.Authors shall not post the final, published versions of their papers.

7. Classroom or Internal Training Use. An author is expressly permitted to post any portion of the accepted version of his/her own IEEE-copyrighted articles on the authors personal web site or the servers of the authors institution or company in connection with the authors teaching, training, or work responsibilities, provided that the appropriate copyright, credit, and reuse notices appear prominently with the posted material. Examples of permitted uses are lecture materials, course packs, e-reserves, conference presentations, or in-house training courses.

STATISTICAL DETECTION OF LSB MATCHING IN THE PRESENCE OF NUISANCE PARAMETERS

Rémi Cogranne, Cathel Zitzmann, Florent Retraint, Igor Nikiforov, Lionel Fillatre and Philippe Cornu

ICD - LM2S - Université de Technologie de Troyes - UMR STMR CNRS
12, rue Marie Curie - B.P. 2060 - 10010 Troyes cedex - France
E-mail : name.surname@utt.fr

ABSTRACT

This paper investigates the reliable detection of information embedded with the least significant bits (LSB) matching scheme. It is aimed to design a test with analytically predictable error probabilities. To this end, the problem of hidden information detection is cast in the framework of hypothesis testing theory. In order to deal with nuisance parameters a rejection approach is used with a statistical model of medium. The use of a linear parametric model permits to analytically express statistical performance of proposed test. Numerical simulations and comparisons with state-of-the-art detectors highlight the relevance of proposed approach.

Index Terms— Information forensics; Detection theory; LSB steganalysis; Nuisance parameters; parametric model.

1. INTRODUCTION AND CONTRIBUTIONS

Steganography and steganalysis have received an increasing interest in the past decade. Therefore, it has become a crucial and useful challenge to reliably detect media which contain hidden information. In such an operational context, the detection of simple but commonly found steganographic scheme is very important. The vast majority of downloadable steganographic tools insert the hidden information in the LSB plane. Hence, many different methods have been proposed to detect information hidden in the LSB of digital media, see [1] and therein references. Among the two schemes which embeds hidden data in the LSB plane, considerable progress have been made in the detection of steganographic algorithm based on LSB replacement. However, detection of information hidden using the LSB matching scheme remains a challenging problem [2].

The detectors dedicated to LSB matching steganography can be roughly divided into two categories. Most of the latest detector are based on supervised machine learning methods using targeted [2] or universal features [3]. On the opposite, it has been observed that LSB matching acts as a low-pass filter on medium histogram ; this finding lead to the family of histogram based detectors [4, 5].

In the operational context described above, proposed steganalyzer must be immediately applicable without any training or tuning phase. For this reason, the use of machine learning based detector is to proscribe. Moreover, the most important challenge is to propose a detection algorithm with analytically predictable error probabilities which remains an open problem of machine learning. The histogram-based detectors are very interesting, but these *ad hoc* algorithms have been designed with a limited exploitation of statistical cover model and hypothesis testing theory. Hence, their statistical performance can only be approximated by simulation on large databases and only few theoretical results exist.

An alternative approach is to design a test with known theoretical properties by using decision theory with a model of cover media [6]. To apply this approach in practice, the main difficulty is to deal with content of media (samples expectation) which acts as nuisance parameters which prevents detection. The present paper investigates the design of a statistical test when cover parameters are unknown. The original contribution is twofold: 1) the proposed approach explicitly takes into account the nuisance parameters by using a linear parametric model of medium, and 2) the statistical performances of presented test are asymptotically given explicitly. Numerical simulation and comparison with state-of-the-art detectors show the relevance of proposed methodology.

The paper is organized as follows. The problem of LSB matching steganalysis is stated in section 2. Section 3 presents the linear parametric model of cover medium. Based on this model, the Generalized LRT is presented in section 4 and its statistical performance are established. Finally, section 5 presents numerical simulations and section 6 concludes the paper.

2. LIKELIHOOD RATIO TEST (LRT) FOR TWO SIMPLE HYPOTHESIS.

Let $\mathbf{c}=\{c_n\}_{n=1}^{N}$ be a vector representing a digital cover medium of N samples. For most of uncompressed digital

This work is partially supported by French National Agency (ANR) through ANR-CSOSG Program (Project ANR-07-SECU-004).

media, each sample value $c_n \in \mathcal{Z} = \{0, \ldots, 2^b - 1\}$ results from the quantization :

$$c_n = Q(y_n), \qquad (1)$$

where $y_n \in \mathbb{R}^+$ denotes the raw sample intensity recorded by the acquisition device and Q represents the uniform quantization with unit step (integer part of y_n).

It is usually assumed that each sample value can be modeled as an independent Gaussian random variables [6–8]:

$$y_n = \theta_n + \xi_n \sim \mathcal{N}(\theta_n, \sigma_n^2), \qquad (2)$$

where θ_n is a deterministic parameter representing image content, and ξ_n is a zero-mean Gaussian noise. For the shake of definiteness, the expected vector $\boldsymbol{\theta} = \{\theta_n\}_{n=1}^N$ is assumed to belong to a compact set $\boldsymbol{\Omega}$.

From equations (1) and (2) the probability mass function (pmf) of c_n can be written $P_{\theta_n} = (p_{\theta_n}[0], \ldots, p_{\theta_n}[2^b - 1])$, where for all $k \in \mathcal{Z}$

$$p_{\theta_n}[k] = \frac{1}{\sigma_n} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \phi\left(\frac{u - \theta_n}{\sigma_n}\right) du \approx \frac{1}{\sigma_n} \phi\left(\frac{k - \theta_n}{\sigma_n}\right). \quad (3)$$

In this paper, ϕ denoted the standard Gaussian probability density function (pdf) $\phi(u) = (2\pi)^{-1/2} \exp(u^2/2)$ and Φ denotes the standard Gaussian cumulative distribution function (cdf) $\Phi(x) = \int_{-\infty}^{u} \phi(u) du$.

The LSB matching scheme consists in randomly incrementing or decrementing each sample value whose LSB differs from the bit to be inserted. After embedding the message $\mathbf{m} = (m_1, \ldots, m_M)^T$ at rate $R = M/N$ a short calculation shows that the pmf of c_n can be written by $Q_{R,\theta_n} = \{q_{R,\theta_n}[0], \cdots, q_{R,\theta_n}[2^b - 1])\}$, where for all $k \in \mathcal{Z}$ [2,4]

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

Let $\mathbf{z} = \{z_n\}_{n=1}^N$ denotes an inspected signal, which is either a cover or a steganographic medium. The hypothesis testing problem of steganalysis consists in choosing between:

$$\begin{cases} \mathcal{H}_0 = \{z_n \sim P_{\theta_n}, \boldsymbol{\theta} \in \boldsymbol{\Omega}, \forall n = 1, \ldots, N\} \\ \mathcal{H}_1 = \{z_n \sim Q_{R,\theta_n}, \boldsymbol{\theta} \in \boldsymbol{\Omega}, 0 < R \leq 1, \forall n = 1, \ldots, N\} \end{cases} \quad (5)$$

where $\boldsymbol{\theta} \in \boldsymbol{\Omega}$ is a nuisance parameter.

The goal is to find a test $\delta : \mathcal{Z}^N \mapsto \{\mathcal{H}_0; \mathcal{H}_1\}$ which accepts hypothesis $\mathcal{H}_i$ if $\delta(\mathbf{z}) = \mathcal{H}_i$ (see more details in [9]). Let

$$\mathcal{K}_{\alpha_0} = \left\{ \delta : \sup_{\boldsymbol{\theta} \in \Omega} \mathbb{P}_{\boldsymbol{\theta}}(\delta(\mathbf{z}) = \mathcal{H}_1) \leq \alpha_0 \right\}$$

be the class of tests with an upper-bounded false alarm probability α_0. Here $\mathbb{P}_{\boldsymbol{\theta}}(A)$ stands for the probability of the event A when z_n is generated by P_{θ_n} for all n. The power function $\beta_{R,\boldsymbol{\theta}}$ is the probability of hidden bits detection:

$$\beta_{R,\boldsymbol{\theta}} = \mathbb{P}_{R,\boldsymbol{\theta}}(\delta(\mathbf{z}) = \mathcal{H}_1).$$

where $\mathbb{P}_{R,\boldsymbol{\theta}}(A)$ stands for the probability of the event A when z_n is generated by Q_{R,θ_n} for all n.

When parameters R, θ_n and σ_n are known for all $n \in \{1, \ldots, N\}$, the problem (5) is to choose between two simple hypothesis. In such a situation, it follows from Neyman-Pearson lemma [9, theorem 3.2.1], that the Likelihood ratio test (LRT) is the most powerful (MP) test over the class $\mathcal{K}_{\alpha_0}$. In [7] it is shown that the likelihood (LR) ratio is given by:

$$\Lambda^{np}(z_n) = \log\left[\exp\left(\frac{z_n - \theta_n}{\sigma_n^2}\right) + \exp\left(\frac{\theta_n - z_n}{\sigma_n^2}\right)\right]. \quad (6)$$

and that for any known embedding rate R and for any given probability of false-alarm α_0, the power of the LRT is:

$$\beta_{R,\boldsymbol{\theta}}^{np} = 1 - \Phi\left(\frac{s_0}{s_R} \Phi^{-1}(1 - \alpha_0) - \frac{R\sqrt{N}(m_1 - m_0)}{s_R}\right). \quad (7)$$

with m_i and s_i^2 the mean expectation and the mean variance of the LR $\Lambda^{np}(z_n)$, see details in the companion paper [7].

However, in a practical situation, neither the embedding rate R nor the cover medium parameters θ_n and σ_n are known. In this situation, two difficulties occur. On the one hand, the content of cover medium, represented by $\boldsymbol{\theta} \in \Omega$ acts as a nuisance parameter because it has no interest to solve the decision problem (5) while it defines the pmf under both hypothesis. It is therefore necessary to design a statistical test which eliminates $\boldsymbol{\theta}$ by explicitly considering this nuisance parameter. On the other hand, due to quantization, the existence of a UMP test, which maximize the power $\beta_{R,\boldsymbol{\theta}}$ uniformly with respect to the rate R, is compromised. This problem is outside the scope of this paper which mainly focus on the presence of nuisance parameters $\boldsymbol{\theta}$. More details about the impact of quantization on statistical performance of LRT can be found in [6].

3. LINEAR PARAMETRIC MODEL OF COVER MEDIUM

Most of digital media acquired with a recording device can be represented as signals whose properties vary smoothly from samples to samples. For instance, an image is blurred by the optical system, hence, expectation of each column (or row) is smooth or low frequency. Following the general methodology [10], it is 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 $N \approx KL$ [6]. Similarly to the scalar case (1) - (2), let define for all $k \in \{1, \ldots, K\}$ the k-th vector $\mathbf{z}_k = (\mathbf{z}_{k,1}, \ldots, \mathbf{z}_{k,L})^T$ as:

$$\mathbf{z}_k = Q(\mathbf{y}_k), \; \mathbf{y}_k = \boldsymbol{\theta}_k + \boldsymbol{\xi}_k \sim \mathcal{N}(\boldsymbol{\theta}_k, \sigma_k^2 \mathbf{I}_L) \qquad (8)$$

where the operation of uniform quantization Q is applied on each samples 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 $\boldsymbol{\theta}_k = (\boldsymbol{\theta}_{k,1}, \ldots, \boldsymbol{\theta}_{k,L})^T$ of expectations. In the present paper, it is proposed to use the following linear parametric model:

$$\boldsymbol{\theta}_k = \mathbf{H}\mathbf{x}_k \tag{9}$$

where $\mathbf{H}$ is a known full rank matrix of size $L \times p$, $p < L$ and $\mathbf{x}_k \in \mathbb{R}^p$ is the nuisance parameter describing expectation of signal $\mathbf{z}_k$. The p columns of matrix $\mathbf{H}$ span the subspace Ω with, using the Cartesian product, $\boldsymbol{\Omega} = \Omega^K$.
The hypothesis testing theory is relatively well developed for models such as (9). Indeed, such a model permits to reject the nuisance parameters θ_k by using the theory of invariance. In practice the rejection 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 $\mathbf{P}_H^\perp \mathbf{H} = \mathbf{0}$. Note that theory of invariance is used here but formally holds for non-quantized observations [9, 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:

$$\mathbf{H}^T = \begin{pmatrix} 1 & 1 & \ldots & 1 & 1 \\ 1 & 2 & \ldots & 2^{p-2} & 2^{p-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & L & \ldots & L^{p-2} & L^{p-1} \end{pmatrix} \tag{10}$$

A more accurate model could be used for complex medium such as images, but model (10) is simple and highlights strengths and limits or proposed methodology, see section 5.

Let us denote $\mathbf{P}_{\boldsymbol{\theta}_k} = P_{\boldsymbol{\theta}_{k,1}} \times \ldots \times P_{\boldsymbol{\theta}_{k,L}}$ and $\mathbf{Q}_{R,\boldsymbol{\theta}_k} = Q_{R,\boldsymbol{\theta}_{k,1}} \times \ldots \times Q_{R,\boldsymbol{\theta}_{k,L}}$ the pmf of vector $\mathbf{z}_k$ under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively. Clearly, by using the model (8)-(9), definition of hypothesis in the problem (5) becomes $\mathcal{H}_0 = \{\mathbf{z}_m \sim P_{\boldsymbol{\theta}_k}, \boldsymbol{\theta} \in \boldsymbol{\Omega}\}$ vs $\mathcal{H}_1 = \{\mathbf{z}_m \sim Q_{R,\boldsymbol{\theta}_k}, \boldsymbol{\theta} \in \boldsymbol{\Omega}\}$.

4. PROPOSED TEST FOR DEALING WITH NUISANCE PARAMETERS

One of the main goal of this paper is to design a statistical test whose performance can be analytically calculated. In this section, it is proposed to use the LR (6) with estimations $\widehat{\boldsymbol{\theta}}_k$ and $\widehat{\sigma}_k$. From model (8) - (9), a short algebra shows that the maximum likelihood estimators (MLE) are given by:

$$\widehat{\boldsymbol{\theta}}_k = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{z}_k \;\text{ and }\; \widehat{\sigma}_k^2 = \frac{\|\mathbf{P}_H^\perp \mathbf{z}_k\|_2^2}{L-p}. \tag{11}$$

One can note that the residuals can be written $\mathbf{z}_k - \widehat{\boldsymbol{\theta}}_k = \mathbf{P}_H^\perp \mathbf{z}_k$. Theory of invariance thus permits to reject nuisance parameters and write from (6) the "approximated" GLR (due to quantization) as:

$$\widehat{\Lambda}(\mathbf{z}_k) = \sum_{k=1}^{K} \sum_{l=1}^{L} \log\left[\exp\left(\frac{\mathbf{z}_{k,l} - \widehat{\boldsymbol{\theta}}_{k,l}}{\widehat{\sigma}_k^2}\right) + \exp\left(\frac{-\mathbf{P}_H^\perp \mathbf{z}_k}{\widehat{\sigma}_k^2}\right)\right]$$

Let define the proposed "approximated" GLR test as follows:

$$\widehat{\delta} = \begin{cases} \mathcal{H}_0 \text{ if } \widehat{\Lambda}(\mathbf{Z}) = \sum_{k=1}^K \widehat{\Lambda}(\mathbf{z}_k) \leq \tau_{\alpha_0}, \\ \mathcal{H}_1 \text{ if } \widehat{\Lambda}(\mathbf{Z}) = \sum_{k=1}^K \widehat{\Lambda}(\mathbf{z}_k) > \tau_{\alpha_0}. \end{cases} \tag{12}$$

The following theorem provides an asymptotic expression of "approximated" GLR test $\widehat{\delta}$ parameters.

Theorem 1. *Assuming that $\boldsymbol{\theta} \in \boldsymbol{\Omega}$, the decision threshold:*

$$\widehat{\tau}_{\alpha_0} = s_0 \Phi^{-1}(1-\alpha) + \sqrt{K(L-p)}m_0 \tag{13}$$

*asymptotically, as $K \to \infty$, warrants that $\widehat{\delta} \in \mathcal{K}_{\alpha_0}$.
Choosing the decision threshold $\widehat{\tau}_{\alpha_0}$ given in (13) the power function $\widehat{\beta}_{R,\boldsymbol{\theta}}$ associated with the test $\widehat{\delta}$ is given by*

$$\widehat{\beta}_{R,\boldsymbol{\theta}} = 1 - \Phi\left(\frac{s_0}{s_R}\Phi^{-1}(1-\alpha_0) - R\sqrt{K(L-p)}\left(\frac{m_1 - m_0}{s_R}\right)\right) \tag{14}$$

with m_0, m_1, s_0 and s_R the two firsts moments of LR $\Lambda^{np}(\mathbf{Z})$ defined in the companion paper [7, eqs. (9) and (10)].

Proof of theorem 1 is omitted due to the lack of space.
The comparison between power functions of LRT $\beta_{R,\boldsymbol{\theta}}^{np}$ (6) and "approximated" GLR (due to quantization) $\widehat{\beta}_{R,\boldsymbol{\theta}}$ (14) shows that the loss of power is due to reduction of "free parameters" from $N \approx K \cdot L$ to $K(L-p)$

5. NUMERICAL RESULTS AND COMPARISON

This paper aims to design a test with analytically predictable performance. Hence, figure 1 presents a comparison between power of LR and GLR tests equations (7) and (14). Results were obtained using a Monte Carlo simulation with 10 000 realizations each with $K = 400$, $L = 32$ and a polynomial of degree $p-1 = 8$. A zero-mean Gaussian noise $\xi_n \sim \mathcal{N}(0, \sigma^2)$, $\sigma = \{1.5; 2.25; 3.5\}$ was added to signals before insertion at rate $R = 0.125$. In figure 1, theoretical and empirical results are almost identical.

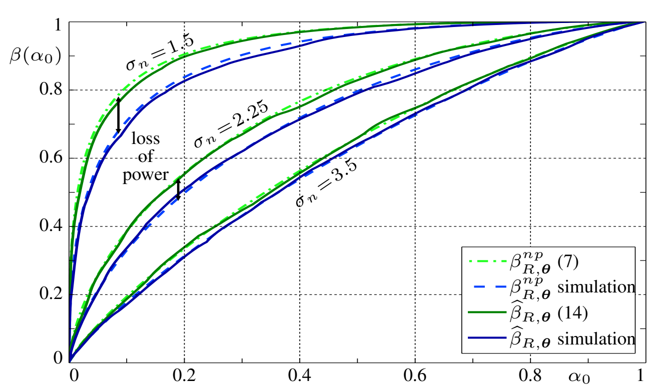

Fig. 1: Numerical simulation of theorem 1: comparison between theoretical and empirical detection performance.

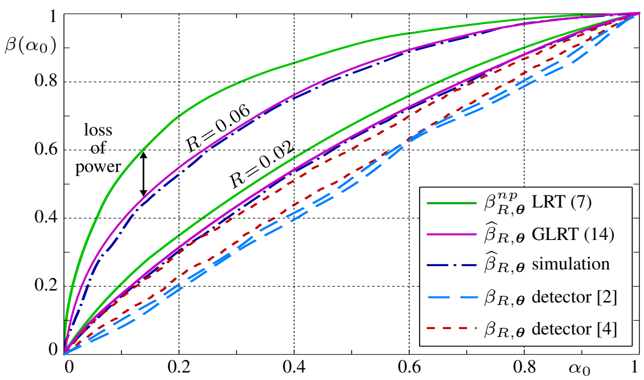

Fig. 2: Detection performance of LRT, proposed test and detectors proposed in [4, 5] for heartbeat sound.

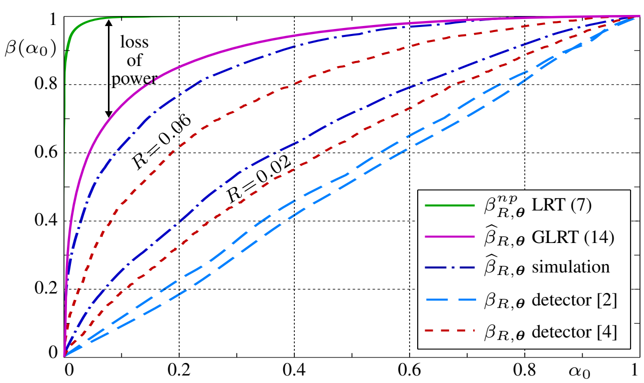

Fig. 3: Detection performance of LRT, proposed test and detectors proposed in [4, 5] for textured image.

Figures 2 and 3 present a comparison between theoretical and empirical performance of proposed test through ROC curves. For those two experimentation, a Monte Carlo simulation with 10 000 realizations was performed by adding a noise $\xi_n \, \mathcal{N}(0, \sigma^2)$ to the inspected medium. The results presented in figure 2 were obtained using an uncompressed heartbeat sound modeled with $L = 32$, $p - 1 = 6$ and $\sigma_n 2$; it shows that empirical results complies with theory. Note that theorem 1 assumes that the inspected medium follows the polynomial model (10) which might be doubtful in practice. For instance, results of figure 3 were obtained with a textured image modeled with parameters $L = 16$, $p = 12$ and $\sigma_n = 0.5$; it clearly shows that even with a high degree, empirical power if widely less than theoretically expected.
Together, figures 2 and 3 show that the polynomial model (9)–(10) proposed is this paper is not very accurate for natural image for which a more sophisticated model, as the one given in [11] could be used. Figures 2 and 3 also present the results of two state-of-the-art detectors, namely ALE [5] and Adjacent HCFCOM [4]. For both the heartbeat sound and textured image, proposed test outperforms these two competitors.

6. CONCLUSION

This paper made a first step in the statistical detection of information hidden with the LSB matching scheme. A general methodology is presented whose two main contribution are the followings: first, the use a local linear parametric model of medium content permits to reject the nuisance parameters and, second, the statistical performance of proposed test are analytically established.
This paper proposes to use a rather simple polynomial model which is efficient for rather simple media but inaccurate for complex media. However, even for more complex media the proposed test outperforms state-of-the-art detectors.

7. REFERENCES

[1] R. Böhme, *Advanced Statistical Steganalysis*, 1st ed. Springer Publishing Company, Incorporated, 2010.

[2] G. Cancelli, & al., “Detection of ±1 LSB steganography based on the amplitude of histogram local extrema,” in *IEEE ICIP*, oct. 2008, pp. 1288 –1291.

[3] S. Lyu and H. Farid, “Steganalysis using higher-order image statistics,” *Information Forensics and Security, IEEE Transactions on*, vol. 1, no. 1, pp. 111 – 119, march 2006.

[4] A. Ker, “Steganalysis of LSB matching in grayscale images,” *Signal Processing Letters, IEEE*, vol. 12, no. 6, pp. 441 – 444, june 2005.

[5] J. Zhang, & al., “Steganalysis for LSB matching in images with high-frequency noise,” in *IEEE Multimedia Signal Processing*, oct. 2007, pp. 385 –388.

[6] R. Cogranne, & al., “Statistical decision by using quantized observations,” in *IEEE International Symposium on Information Theory*, 2011, pp. 1135 – 1139.

[7] R. Cogranne, & al., “Statistical detection of LSB matching using hypothesis testing theory,” in *Information Hiding*, Springer, 2012.

[8] A. Foi, & al., “Practical poissonian-gaussian noise modeling and fitting for single-image raw-data,” *Image Processing, IEEE Transactions on*, vol. 17, no. 10, pp. 1737–1754, Oct. 2008.

[9] E. Lehman and J. Romano, *Testing Statistical Hypotheses, Second Edition*, 3rd ed. Springer, 2005.

[10] V. Katkovnik, & al., *Local Approximation Techniques in Signal and Image Processing*. SPIE Press, 2006.

[11] R. Cogranne, & al., “Reliable detection of hidden information based on a non-linear local model,” in *IEEE Statistical Signal Processing*, 2011, pp. 493 – 496.