Content-Adaptive Steganography by Minimizing Statistical Detectability

Vahid Sedighi, *Member, IEEE*, Rémi Cogranne, *Member, IEEE*, and Jessica Fridrich, *Senior Member, IEEE*

***Abstract*—Most current steganographic schemes embed the secret payload by minimizing a heuristically defined distortion. Similarly, their security is evaluated empirically using classifiers equipped with rich image models. In this paper, we pursue an alternative approach based on a locally-estimated multivariate Gaussian cover image model that is sufficiently simple to derive a closed-form expression for the power of the most powerful detector of content-adaptive LSB matching but, at the same time, complex enough to capture the non-stationary character of natural images. We show that when the cover model estimator is properly chosen, state-of-the-art performance can be obtained. The closed-form expression for detectability within the chosen model is used to obtain new fundamental insight regarding the performance limits of empirical steganalysis detectors built as classifiers. In particular, we consider a novel detectability-limited sender and estimate the secure payload of individual images.**

***Index Terms*—Adaptive steganography and steganalysis, hypothesis testing theory, information hiding, multivariate Gaussian, optimal detection.**

I. Introduction

Historically, the design of steganographic schemes for digital images has heavily relied on heuristic principles. The current trend calls for constraining the embedding changes to image segments with complex content. Such adaptive steganographic schemes are typically realized by first defining the cost of changing each pixel and then embedding the secret message while minimizing the sum of costs of all changed pixels. Efficient coding methods [1] can embed the desired payload with an expected distortion near the minimal possible value prescribed by the corresponding rate–distortion bound.

Although this paradigm based on the concepts of pixel costs and distortion gave birth to a multitude of content-adaptive data hiding techniques with markedly improved security [2]–[6], the entire design is rather unsettling because there is no formal connection between distortion and statistical detectability. As argued in [7], this connection may never be found as empirical cover sources, such as digital media, are fundamentally incognizable. Steganography designers thus primarily rely on empirical evidence to support the claims concerning the security of their embedding schemes.

The design of distortion functions that measure statistical detectability rather than distortion was identified as one of the most important open problems in the recent motivational review article [8].[1] As far as the authors of the current manuscript are aware, there are only a few examples of distortion functions that consider cover models in their design. The first is the distortion function of HUGO [2] that prefers changing pixels with the smallest impact on the empirical statistical distribution of pixel groups represented in the SPAM feature space [9]. In [10], the distortion function is first parametrized and then optimized to minimize the empirical detectability in terms of the margin between cover and stego images represented using low-dimensional features. These approaches are limited to empirical "models" that need to be learned from a database of images. Such embedding schemes may become "overoptimized" to the feature space and cover source and become highly detectable should the Warden choose a different feature representation [11].

The first attempt to design the distortion as a quantity related to statistical detectability appeared in [12]. The authors proposed to use the Kullback–Leibler divergence between the statistical distributions of cover and stego images when modeling the cover pixels as a sequence of independent Gaussian random variables with unequal variances (multivariate Gaussian or MVG). Using a rather simple pixel variance estimator, the authors showed that the empirical security of their embedding method was roughly comparable to HUGO but subpar with respect to state-of-the-art steganographic methods [3]–[5]. In [13], this approach was extended by utilizing a better variance estimator and replacing the Gaussian model with the generalized Gaussian. The authors focused on whether it

The work on this paper was supported by the Air Force Office of Scientific Research under the research grant FA9550-09-1-0147. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation there on. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied of AFOSR or the U.S. Government.

This work has also been partially funded by the conseil régional de Champagne-Ardenne within program for scholar mobility, projects STEG-DETECT and IDENT.

Vahid Sedighi and Dr. Jessica Fridrich are with the Department of Electrical and Computer Engineering, Binghamton University, NY, 13902, USA. Email: {vsedigh1,fridrich}@binghamton.edu.

Dr. Rémi Cogranne is with the Lab. for system Modeling and Dependability, ICD, UMR 6281 CNRS, Troys University of Technology, Troys, France. This work has been done while he was a visiting scholar at Binghamton University. Email: remi.cogranne@utt.fr.

[1] See Open Problems no. 2 and 9.

is possible to further improve the security by allowing a pentary embedding operation with a thicker-tail model.

While the current paper builds upon this existing art, it addresses numerous novel issues not investigated elsewhere. To clarify the main contribution of this paper, the closed-form expression for the detectability within the chosen model is used to obtain the following fundamental insight regarding the limits of empirical steganalysis detectors built as classifiers:

1) For the first time empirical detectors can be compared with optimal detectors and evaluated w.r.t. the performance bound valid within the chosen cover model. In particular, when forcing the heteroscedastic model of sensor acquisition noise to an artificial image with simple content, we observed that the difference in performance between the optimal likelihood-ratio detector and empirical detectors built as classifiers using rich media models is rather small. This indicates that in this source, current empirical steganalysis is near optimal.
2) We introduce a novel type of the so-called "detectability-limited sender" that adjusts the payload size for each image to not exceed a prescribed level of statistical detectability within the chosen model. On a database of real images, we contrast the theoretical security of this detectability-limited sender dictated by the model with the one obtained empirically using classifiers employing rich models. Despite the fact that the empirical detector can capture more complex dependencies between pixels than our MVG model, its detection power is much smaller. We attribute this suboptimality primarily to the difficulty of empirical detectors to deal with content heterogeneity of real images.
3) The availability of a closed-form expression for the power of the optimal detector allows us to compute the size of the secure payload for a given image and a chosen detectability (risk) level. We compare it with the secure payload size estimated using empirical detectors and draw several interesting and important facts about the interplay between theoretical and empirical detectors.

We now discuss in more detail the relationship between the method introduced in this paper and the prior art [12], [13]. The embedding method of [12] equipped with the enhanced variance estimator described in this paper and the ternary method of [13] with a Gaussian cover model coincide in practice with the method studied in this paper. However, the approaches are methodologically different. The methods of [12], [13] minimize the KL divergence between cover and stego distributions in the asymptotic limit of a small payload, while the current paper minimizes the power of the most powerful detector instead of the KL divergence, which is achieved without the additional assumption of a small payload. This is why we coin a new acronym MiPOD standing for **Mi**nimizing the **P**ower of **O**ptimal **D**etector. Moreover, the framework introduced in this paper allows us to consider various types of Warden, which was not possible within the prior art. Finally, in contrast with [13] we investigate the effect of the parameters of the variance estimator on content adaptivity and security of MiPOD and identify a setting that gives it the the smallest empirical detectability.

In Sections II–III, we review the MiPOD algorithm by first introducing the statistical model of cover images, the multivariate Gaussian (MVG), deriving the stego image model for content-adaptive Least Significant Bit (LSB) matching, and analytically establishing the asymptotic properties of the optimal Likelihood Ratio Test (LRT) for MiPOD. We also introduce two types of Warden depending on the available information about the selection channel (content adaptivity). In Section IV, we describe the embedding algorithm of MiPOD based on minimizing the power of the optimal detector. Section V contains a detailed description of the cover model variance estimator and studies the effect of its parameters on MiPOD's adaptivity (selection channel). The main contribution of this paper appears in Section VI, which presents all numerical results divided into the following main parts. After describing the common core of all experiments, in Section VI-B we compare MiPOD with prior art on a standard image source using detectors implemented as classifiers using state-of-the-art feature sets. In Section VI-C, we use an artificial image source in which we force a heteroscedastic cover noise model to show the tightness of the asymptotic LRT and to demonstrate that the optimal detector and empirical detectors built as classifiers with rich image models achieve a very similar level of detectability. A novel detectability-limited sender is introduced and investigated on a database of real images in Section VI-D. Finally, in Section VI-E by contrasting the secure payload size computed from the model and using empirical detectors, we discover several interesting and important facts about the interplay between theoretical and empirical detectors.

The following common notational conventions are used throughout the paper. Matrices and vectors will be typeset in boldface, sets in calligraphic font, while capital letters are reserved for random variables. The transpose of matrix $\mathbf{A}$ will be denoted $\mathbf{A}^{\mathrm{T}}$, and $\|\mathbf{x}\|$ is reserved for the L_2 norm of vector $\mathbf{x}$. A probability measure is denoted with $\mathbb{P}$. The symbol $\mathbb{Z}$ stands for the set of all integers. We also use the notation $[P]$ for the Iverson bracket $[P] = 1$ when P is true and $[P] = 0$ when P is false.

II. Image model

In this section, we describe the cover model and the embedding algorithm used by Alice and derive the statistical model for stego images.

A. Cover image model

We only consider images represented in the spatial domain. Ignoring for simplicity the effects of spatial filtering and demosaicking, the pixel values in a digital image acquired with an imaging sensor are typically corrupted by an independent Gaussian noise with variance dependent on the pixel light intensity (the shot noise), temperature and exposure (dark current), and readout and electronic noise. This common noise model [14]–[16] was previously applied in digital forensics [17] as well as in steganalysis of LSB replacement [18], [19] and LSB matching [20], [21].

The local pixel mean (the content) can be estimated with local pixel predictors as is currently commonly done when forming steganalysis features [22]. However, this estimation is never perfect, which is true especially in highly textured regions. In this paper, we include the difference between the pixel value and its estimated value (the modeling error) into the noise term, which we still model as a Gaussian.

Formally, we consider the cover pixels as an N-dimensional vector $\mathbf{z} = (z_1, \dots, z_N)$ of independent realizations of N Gaussian random variables $Z_n \sim \mathcal{N}(\mu_n, \omega_n^2)$, $n = 1, \dots, N$, quantized to discrete points $k\triangle$, $k \in \mathbb{Z}$ (for simplicity and without loss on generality, we set $\triangle = 1$). Here, μ_n is the noise-free content and ω_n^2 is the variance of the Gaussian acquisition noise. Let $\hat{\mu}_n \in \mathbb{Z}$ be an estimate of the mean of the nth pixel. The differences $x_n = z_n - \hat{\mu}_n$ will thus contain both the acquisition noise as well as the modeling error. We model x_n as independent Gaussian random variables $X_n \sim \mathcal{N}(0, \sigma_n^2)$, where $\sigma_n^2 \geq \omega_n^2$ because of the inclusion of the modeling error.

Assuming the fine quantization limit, $\triangle \ll \sigma_n$ for all n, the probability mass function (pmf) of the nth pixel is given by $\mathcal{P}_{\sigma_n} = (p_{\sigma_n}(k))_{k\in\mathbb{Z}}$ with

$$p_{\sigma_n}(k) = \mathbb{P}(x_n = k) \propto \frac{1}{\sigma_n\sqrt{2\pi}} \exp\left(-\frac{k^2}{2\sigma_n^2}\right). \quad (1)$$

Note that it is assumed that the pixels are quantized using an unbounded number of levels (bits). This assumption is adopted for the sake of simplifying the subsequent theoretical exposition. For practical embedding schemes, the finite dynamic range of pixels must be taken into account, for example by forbidding embedding changes that would lead to cover values outside of the dynamic range. The fine quantization limit does not hold in saturated (overexposed) image regions, which however does not pose a problem as any content-adaptive embedding should avoid them. This can be arranged in practice by assigning very small embedding change probabilities to pixels from such regions. Additional discussion regarding the feasibility of the fine quantization assumption appears at the end of Section V.

B. Stego image model

A widely adopted and well-studied model of data hiding is the Mutually Independent (MI) embedding in which the embedding changes Alice makes at each pixel are independent of each other. In particular, we adopt one of the simplest possible setups when the pixel values are changed by at most ±1 (the so-called LSB matching or LSBM) while noting that the framework is easily extensible to any MI embedding. Given a cover image represented with $\mathbf{x} = (x_1, \dots, x_N)$, the stego image $\mathbf{y} = (y_1, \dots, y_N)$ is obtained by independently applying the following probabilistic rules:

$$\begin{aligned} \mathbb{P}(y_n = x_n + 1) &= \beta_n, \\ \mathbb{P}(y_n = x_n - 1) &= \beta_n, \\ \mathbb{P}(y_n = x_n) &= 1 - 2\beta_n, \end{aligned} \quad (2)$$

with change rates $0 \leq \beta_n \leq 1/3$. The pmf of the stego pixels is thus given by $\mathcal{Q}_{\sigma_n,\beta_n} = (q_{\sigma_n,\beta_n}(k))_{k\in\mathbb{Z}}$ with

$$\begin{aligned} q_{\sigma_n,\beta_n}(k) = \mathbb{P}(y_n = k) = (1 - 2\beta_n)p_{\sigma_n}(k) \\ + \beta_n p_{\sigma_n}(k+1) + \beta_n p_{\sigma_n}(k-1). \end{aligned} \quad (3)$$

C. Embedding in practice

In theory, if Alice used an optimal embedding scheme, she could embed a payload of R nats:

$$R(\boldsymbol{\beta}) = \sum_{n=1}^{N} H(\beta_n), \quad (4)$$

where $H(x) = -2x\log x - (1 - 2x)\log(1 - 2x)$ is the ternary entropy function expressed in nats ("log" is the natural log). In practice, Alice needs to use some coding method, such as the syndrome-trellis codes (STCs) [1] while minimizing the following additive distortion function

$$D(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{N} \rho_n[x_n \neq y_n], \quad (5)$$

where $\rho_n \geq 0$ is the cost of changing pixel x_n tied to β_n via

$$\beta_n = \frac{e^{-\lambda\rho_n}}{1 + 2e^{-\lambda\rho_n}}. \quad (6)$$

with $\lambda > 0$ determined from the payload constraint (4). Using a specific coding scheme instead of optimal coding will introduce a small suboptimality in terms of embedding a slightly smaller payload than R for a given value of the distortion. This coding loss, however, can be made arbitrarily small at the expense of computational complexity. Therefore, in the current paper we disregard the coding loss and simulate all embedding changes using simulators that execute the embedding changes with the probabilities β_n.

III. Optimal LR test and its statistical performance

The main result of this section is a closed-form expression for the deflection coefficient of Warden's detector under the assumption that both Alice and the Warden know the standard deviations $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N)$. Without any loss of generality, we will assume that the Warden uses the change rates $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)$ that might, or might not, coincide with $\boldsymbol{\beta} = (\beta_1, \dots, \beta_N)$. In this case, when analyzing the image $\mathbf{x} = (x_1, \dots, x_N)$, the Warden's goal is to decide between the following two simple hypotheses: $\forall n \in \{1, \dots, N\}$:

$$\begin{aligned}\mathcal{H}_0 &= \Big\{ x_n \sim \mathcal{P}_{\sigma_n}, \forall \sigma_n > 0 \Big\}, \\ \mathcal{H}_1 &= \Big\{ x_n \sim \mathcal{Q}_{\sigma_n,\gamma_n}, \forall \sigma_n > 0 \Big\}.\end{aligned} \tag{7}$$

The Warden is especially interested in identifying a test, a mapping $\delta : \mathbb{Z}^N \to \{\mathcal{H}_0, \mathcal{H}_1\}$, with the best possible performance. In this paper, we will use the Neyman–Pearson criterion of optimality, that is for a given false-alarm probability $\alpha_0 = \mathbb{P}(\delta(\mathbf{x}) = \mathcal{H}_1 | \mathcal{H}_0)$ we seek a test that maximizes the power function $\pi = \mathbb{P}(\delta(\mathbf{x}) = \mathcal{H}_1 | \mathcal{H}_1)$, the correct detection probability (see [23] for details about statistical hypothesis testing).

The Neyman–Pearson Lemma ([23, Theorem 3.2.1]) states that the most powerful (MP) test (the one maximizing the power function for a prescribed false-alarm probability) is the Likelihood Ratio Test (LRT), which in our case is

$$\Lambda(\mathbf{x}, \boldsymbol{\sigma}) = \sum_{n=1}^{N} \Lambda_n = \sum_{n=1}^{N} \log \left(\frac{q_{\sigma_n,\gamma_n}(x_n)}{p_{\sigma_n}(x_n)} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau, \tag{8}$$

by the statistical independence of pixels.[2]

Under the fine quantization limit, $\triangle \ll \sigma_n$ for all n, it is shown in Appendix A that, as the number of pixels $N \to \infty$, the Lindeberg's version of the Central Limit Theorem implies

$$\begin{aligned}\Lambda^\star(\mathbf{x}, \boldsymbol{\sigma}) &= \frac{\sum_{n=1}^{N} \Lambda_n - E_{\mathcal{H}_0}[\Lambda_n]}{\sqrt{\sum_{n=1}^{N} Var_{\mathcal{H}_0}[\Lambda_n]}} \\ &\rightsquigarrow \begin{cases} \mathcal{N}(0,1) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(\varrho,1) & \text{under } \mathcal{H}_1 \end{cases},\end{aligned} \tag{9}$$

where $\rightsquigarrow$ denotes the convergence in distribution and

$$\begin{aligned}\varrho &= \frac{\sum_{n=1}^{N} \left(E_{\mathcal{H}_1}[\Lambda_n] - E_{\mathcal{H}_0}[\Lambda_n] \right)}{\sqrt{\sum_{n=1}^{N} Var_{\mathcal{H}_0}[\Lambda_n]}} \\ &= \frac{\sqrt{2} \sum_{n=1}^{N} \sigma_n^{-4} \beta_n \gamma_n}{\sqrt{\sum_{n=1}^{N} \sigma_n^{-4} \gamma_n^2}}\end{aligned} \tag{10}$$

is the deflection coefficient, which completely characterizes the statistical detectability. We note that, under the fine quantization limit, $I_n = 2/\sigma_n^4$ is the Fisher information of LSBM in quantized $\mathcal{N}(0, \sigma_n^2)$ with respect to the change rate β_n (see [12] for details).

A. Impact of Warden knowledge on detectability

In this paper, we will consider two types of Warden: an *omniscient* Warden, who knows the change rates β_n used by Alice and uses $\gamma_n = \beta_n$ for all n, and an *indifferent* Warden who is completely ignorant about Alice's actions and uses the least informative (non-adaptive) change rates $\gamma_n = \gamma$ for all n. The case of the omniscient Warden represents the worst (conservative) scenario for Alice motivated by the Kerckhoffs' principle and is frequently made in steganography design. The indifferent Warden was introduced to see how the detection is affected when the Warden does not utilize the knowledge of the selection channel – the change rates β_n. In empirical steganalysis, the indifferent Warden essentially corresponds to steganalysis that does not use the knowledge of the change rates, such as a classifier equipped with the SRM [22].

For the omniscient Warden, the deflection coefficient of the optimal LR (10) simplifies to:

$$\varrho^\star = \frac{\sqrt{2} \sum_{n=1}^{N} \sigma_n^{-4} \beta_n^2}{\sqrt{\sum_{n=1}^{N} \sigma_n^{-4} \beta_n^2}} = \sqrt{2 \sum_{n=1}^{N} \sigma_n^{-4} \beta_n^2}, \tag{11}$$

while for the indifferent Warden, the LR becomes:

$$\underline{\varrho} = \frac{\sqrt{2} \sum_{n=1}^{N} \sigma_n^{-4} \beta_n}{\sqrt{\sum_{n=1}^{N} \sigma_n^{-4}}}. \tag{12}$$

The Cauchy–Schwartz inequality implies that $\varrho^\star \geq \underline{\varrho}$, which means that the indifferent Warden's detector will always be suboptimal w.r.t. the omniscient Warden.

Formally, the statistical properties of the LRT based on $\Lambda^\star(\mathbf{x}, \boldsymbol{\sigma})$ are given in the following proposition.

Proposition 1. *It follows from the limiting distribution of the LR under $\mathcal{H}_0$ (9) that for any $\alpha_0 \in (0,1)$ the decision threshold $\tau^\star$ given by:*

$$\tau^\star = \Phi^{-1}(1 - \alpha_0), \tag{13}$$

where Φ and Φ^{-1} denote the cumulative distribution function (cdf) of the standard Gaussian distribution and its inverse, asymptotically as $N \to \infty$, guarantees that the false-alarm probability of the LRT does not exceed α_0.

It also follows from the limiting distribution (9) that the power $\pi = \pi(\varrho^\star)$ of the LRT is given by:

$$\pi(\varrho^\star) = 1 - \Phi\left(\tau^\star - \varrho^\star\right) = 1 - \Phi\left(\Phi^{-1}(1 - \alpha_0) - \varrho^\star\right). \tag{14}$$

Proof: Immediately follows from (9) and the properties of Gaussian random variables. ■

B. Detectability-limited sender

A detectability-related distortion allows us to introduce a novel "detectability-limited sender" which adapts the payload for a given cover so that the embedding does not exceed a prescribed detectability level. One frequently used measure of security in practical steganalysis is the total probability or error under equal priors $P_{\mathrm{E}} = \min_{\alpha_0}(\alpha_0 + 1 - \pi_0(\alpha_0))/2$. Since the optimal LR test is essentially a test between two shifted Gaussian distributions, it is immediate that

$$P_{\mathrm{E}} = 1 - \Phi\left(\varrho^\star / 2\right). \tag{15}$$

The steganographers can adjust the embedding to guarantee that a Warden who uses the optimal test will always

[2] Note the false-alarm probability α_0 is not specified as it does not change the LRT up to the decision threshold τ.

have her $P_{\mathrm{E}} \leq P_{\mathrm{E}}^{\star}$ for any given $0 < P_{\mathrm{E}}^{\star} \leq 1/2$ by making sure that the deflection coefficient $\varrho^{\star}$ (11) satisfies:[3]

$$\varrho^{\star} \leq 2 \cdot \Phi^{-1}\left(1 - P_{\mathrm{E}}^{\star}\right). \tag{16}$$

Of course, this detectability guarantee is only valid within the chosen model. In particular, if the Warden uses a more accurate cover model than the steganographers, e.g., by considering higher-order dependencies among pixels, the bounds (15) and (16) may not be satisfied.

IV. Steganography by Minimizing the Performance of Optimal Detector (MiPOD)

In this section, we study steganography design based on the MVG cover model under the omniscient Warden who uses the optimal LRT since it will provide her with the highest possible power within the model. We also describe the embedding process using a pseudo-code to explain how to implement MiPOD in practice.

To present the theoretical foundation of the proposed approach, we will assume for now that Alice knows exactly the variance of each pixel, σ_n^2. In reality the variance will have to be estimated using the variance estimator described in Section (V). Hence, maximizing the security under the omniscient Warden means that Alice should select change rates β_n that minimize the deflection coefficient $\varrho^{\star}$ (11) or, equivalently, its square:

$$\varrho^{\star^2} = 2\sum_{n=1}^{N} \sigma_n^{-4}\beta_n^2 \tag{17}$$

under the payload constraint (4). This can be easily established using the method of Lagrange multipliers. The change rates β_n and the Lagrange multiplier $\lambda > 0$ that minimize (17) must satisfy the following $N+1$ non-linear equations for $N+1$ unknowns, which are λ and the change rates $\beta_1, \ldots, \beta_N$:

$$\beta_n \sigma_n^{-4} = \frac{1}{2\lambda} \ln \frac{1-2\beta_n}{\beta_n},\ n = 1, \ldots, N, \tag{18}$$

$$R = \sum_{n=1}^{N} H(\beta_n), \tag{19}$$

with the last equation being the payload constraint with R expressed in nats. This system can easily be solved numerically. Details of the solution can be found in the prior art [12]. Once the change rates are computed, they need to be converted to costs so that the actual message embedding can be executed with the well established framework of syndrome-trellis codes. The costs can be obtained by inverting the relationship between β_n and ρ_n (6):

$$\rho_n = \ln(1/\beta_n - 2). \tag{20}$$

To further clarify the embedding procedure, in Algorithm (1) we provide a pseudo-code that describes the individual phases of the embedding scheme.

Note that the change rates (and costs for practical embedding) of MiPOD are determined by minimizing the impact of embedding on the cover model. In contrast, all current content-adaptive steganographic schemes (with the exception of our prior work [12], [13]) use pixel cost computed in some heuristic manner by quantifying the impact of an embedding change on the local pixel neighborhood (see Figure 1 and [3]–[5]). Also notice that MiPOD costs (20) depend on the payload.

Finally, we wish to point out that in practice nothing prevents the Warden from selecting a more accurate model of pixels and improve the detection beyond that of the LRT, which is optimal only within the MVG cover model. Naturally, this possibility will always be available to the Warden and this is also what drives the current research in steganography.

V. Estimating pixel variance

The question of which variance estimator will lead to the most secure embedding scheme when evaluating security using empirical detectors is far from being simple and needs to be considered within the context of the entire steganographic channel. If the Warden was able to completely reject the content and isolate only the indeterministic acquisition noise, Alice's best choice would be to use the best possible denoising filter to estimate the pixels' variance. However, current state-of-the-art steganalyzers for adaptive LSB matching [22], [24]–[26] use feature representations of images based on joint distributions of quantized noise residuals computed using local pixel predictors. As long as the Warden stays within this established framework, Alice's "best" variance estimator should avoid rejecting the content too much or too little. In this paper, we give the variance estimator a modular structure that can be adjusted to minimize the detection using current best empirical detectors.

In particular, we use a variance estimator that consists of two steps. Assuming the cover image is an 8-bit grayscale with the original pixel values $\mathbf{z} = (z_1, \ldots, z_N)$, $z_n \in \{0, \ldots, 255\}$, we first suppress the image content using a denoising filter F: $\mathbf{r} = \mathbf{z} - F(\mathbf{z})$. This can be interpreted as subtracting from each pixel its estimated expectation. The residual $\mathbf{r}$ will still contain some remnants of the content around edges and in complex textures. To further remove the content, and to give the estimator a modular structure that can be optimized for a given

Algorithm 1 Pseudo-code for MiPOD embedder.

```
Estimate pixel residual variances $\sigma_n^2$ using the estimator described in Section V.
Numerically solve Eqs. (18) and (19) and determine the change rates $\beta_n,\ n = 1, \ldots, N$ and the Lagrange multiplier $\lambda$.
Convert the change rates $\beta_n$ to costs $\rho_n$ using Eq. (20).
Embed the desired payload $R$ using STCs with pixel costs $\rho_n$ determined in the previous step.
```

[3] Note that since the LR test remains the same for any prescribed false-alarm probability α_0, up to the decision threshold, this LR test also has the lowest achievable P_{E}.

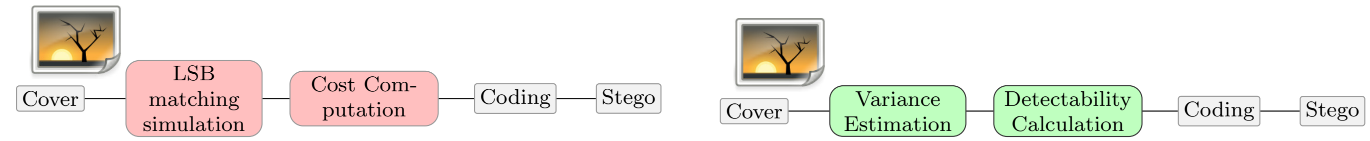


Figure 1. Simplified flowchart of a typical prior-art adaptive embedding scheme (left) and the proposed MiPOD (right).

source and detector in practice, as the second step we fit a local parametric model to the neighbors of each residual value to obtain the final variance estimate. At this point, we openly acknowledge that this is certainly not the only or the best approach one can adopt. There likely exist other estimator designs that can produce comparable or even better security. We opted for the current approach because of its modularity and because it gave us the best results out of all estimators we experimented with. This estimator can also be efficiently implemented and it produced respectable results in steganalysis [18], [20] and in image processing in general [27], [28].

Formally, this second step of the estimator design is a blockwise Maximum Likelihood Estimation (MLE) of pixel variance using a local parametric linear model [28]. We model the remaining pixel expectation within small $p \times p$ blocks as follows:

$$\mathbf{r}_n = \mathbf{G}\mathbf{a}_n + \boldsymbol{\xi}_n. \tag{21}$$

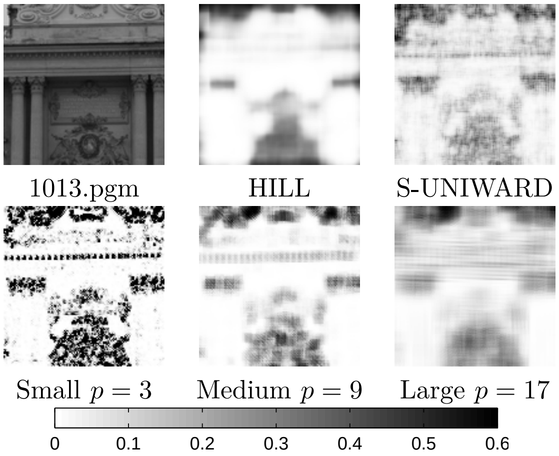
Here $\mathbf{r}_n$ represents the values of the residual $\mathbf{r}$ inside the $p \times p$ block surrounding the nth residual put into a column vector of size $p^2 \times 1$, $\mathbf{G}$ is a matrix of size $p^2 \times q$ that defines the parametric model of remaining expectations, $\mathbf{a}_n$ is a vector of $q \times 1$ of parameters, and $\boldsymbol{\xi}_n$ is the signal whose variance we are trying to estimate. We note that $\boldsymbol{\xi}_n$ is a mixture of the acquisition noise as well as the modeling error.

It is well known that for a linear model corrupted by Gaussian noise, the MLE of the parameter $\mathbf{a}_n$ from the residuals $\mathbf{r}_n$ is given by:

$$\widehat{\mathbf{a}}_n = \left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{r}_n, \tag{22}$$

which also coincides with the ordinary least squares estimator by the Gauss–Markov theorem. Hence, the estimated expectation of the residuals $\mathbf{r}_n$ is given by:

$$\widehat{\mathbf{r}}_n = \mathbf{G}\widehat{\mathbf{a}}_n = \mathbf{G}\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{r}_n. \tag{23}$$



1013.pgm HILL S-UNIWARD

Small $p = 3$ Medium $p = 9$ Large $p = 17$

Figure 2. First row, left to right: A 128 × 128 crop of '1013.pgm' from BOSSbase 1.01 and the embedding probability for payload 0.4 bpp using HILL and S-UNIWARD. Second row, left to right: MiPOD with three different settings showing an extreme, medium, and low content adaptivity obtained by changing the parameters of the variance estimator. See the text for more details.

Finally, assuming that the pixels within the n-th block have the same or similar variances, from (23) the MLE estimation of the central pixel variance in the n-th block is:

$$\widehat{\sigma}_n^2 = \frac{\left\|\mathbf{r}_n - \widehat{\mathbf{r}}_n\right\|^2}{p^2 - q} = \frac{\left\|\mathbf{P}_{\mathbf{G}}^{\perp}\mathbf{r}_n\right\|^2}{p^2 - q}, \tag{24}$$

where $\mathbf{P}_{\mathbf{G}}^{\perp} = \mathbf{I}_n - \mathbf{G}\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}$ represents the orthogonal projection onto the $p^2 - q$ dimensional subspace spanned by the left null space of $\mathbf{G}$ ($\mathbf{I}_n$ is the $n \times n$ unity matrix).

We would like to stress that this method of variance estimation is applied "pixelwise" instead of blockwise, which means that the estimated value of the variance is attributed only to the central pixel of the considered block. To obtain the variance, e.g., for the right neighbor, the block is translated by one pixel to the right, etc. Mirror padding is applied at the image boundaries to obtain the variance estimates for all pixels.

The proposed variance estimator can attain many different forms based on the employed denoising filter and the local parametric model. After experimenting with polynomial and DCT parametric models as well as numerous denoising filters, we determined that a good trade-off between complexity and empirical security was obtained with a simple two-dimensional Wiener filter implemented in Matlab as `wiener2(X,[`w w`])`, where $w > 1$ is an integer, and a parametric model with two-dimensional (discrete) trigonometric polynomial functions similar to those used in the two-dimensional DCT:

$$\mathbf{G} = \Big(\mathbf{1}, \cos(\mathbf{u}), \cos(\mathbf{v}), \cos(\mathbf{u})\cdot\cos(\mathbf{v}), \cos(2\mathbf{u}), \cos(2\mathbf{v}), \cos(2\mathbf{u})\cdot\cos(2\mathbf{v}), \ldots, \cos(l\mathbf{u}), \cos(l\mathbf{v})\Big). \tag{25}$$

In (25), the dot stands for the element-wise product, $\mathbf{1} \in \mathbb{R}^{p^2}$ is a column vector of ones, and the vectors $\mathbf{u} \in \mathbb{R}^{p^2}$ ($\mathbf{v} \in \mathbb{R}^{p^2}$) are obtained by unfolding the matrix $\mathbf{U}$

$$\mathbf{U}=\begin{pmatrix}\frac{\pi}{2p} & \frac{3\pi}{2p} & \cdots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p}\\ \frac{\pi}{2p} & \frac{3\pi}{2p} & \cdots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p}\\ \frac{\pi}{2p} & \frac{3\pi}{2p} & \cdots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p}\\ \vdots & \vdots & \ldots & \vdots & \vdots\\ \frac{\pi}{2p} & \frac{3\pi}{2p} & \cdots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p}\end{pmatrix} \qquad (26)$$

($\mathbf{V}=\mathbf{U}^{\mathrm{T}}$) into a column vector [18], [20], [27]. Thus, our parametric model has $q = l(l+1)/2$ parameters, where l is the degree of the two-dimensional cosine polynomial. The adaptivity of MiPOD can be adjusted by selecting different values for the parameters w, p, and l. We determined experimentally that it is advantageous to use a larger block size p but keep the Wiener filter width w small. In this paper, we fixed the value to $w = 2$. The profound effect of p and l on the embedding adaptivity is shown in Figure 2 contrasting the change rates of HILL and S-UNIWARD with those of MiPOD with three parameter configurations: 1) small blocks with $p = 3$ and $l = 3$ ($q = 6$), 2) medium blocks with $p = 9$ and $l = 9$ ($q = 45$), and large blocks with $p = 17$ and $l = 12$ ($q = 78$).

Finally, we wish to make an additional comment on the fine quantization assumption. It is true that at pixels whose estimated variance is small, the fine quantization limit is not satisfied. However, since Eq. (18) implies that $-\beta_n \ln \beta_n \le \sigma_n^4/(2\lambda)$, we have $\beta_n \to 0$ as $\sigma_n \to 0$ for any fixed payload (λ). Thus, even though the change rate obtained by solving (18) will be imprecise when the fine quantization is violated, the change rate will be too small to have any effect on security. Indeed, pixels with $\hat{\sigma}_n^2 \approx 0$ lie in a smooth image region and should have a small probability of change anyway. In practice, for numerical stability, we introduce a finite floor for the estimated variance:

$$\hat{\sigma}_n^2 \leftarrow \max\{0.01, \hat{\sigma}_n^2\}. \qquad (27)$$

VI. Experiments and comparison to prior art

A. Common core of all experiments

Unless mentioned otherwise, our experiments are carried out on BOSSbase 1.01 [29] containing 10,000 grayscale 512×512 images. The detectors were trained as binary classifiers implemented using the FLD ensemble [30] with default settings. We note, however, that in most experiments, the ensemble classifier was used within the framework of hypothesis testing as proposed in [31], [32] because this implementation of the FLD ensemble permits obtaining the LR values instead of binary outputs, which is crucial in order to measure the detection power for a given level of the false-alarm rate to plot Receiver Operating Characteristic (ROC) curves.

The two feature sets used are the Spatial Rich Model (SRM) [22] and its recent selection-channel-aware version called the maxSRMd2 [26], which is particularly interesting in the context of this paper as it uses the knowledge of change rates. All tested embedding algorithms are simulated at their corresponding payload–distortion bound for payloads $R \in \{0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$ bpp (bits per pixel). The statistical detectability is empirically evaluated using the original version of the FLD ensemble [30] using the minimal total probability of error under equal priors P_{E} averaged over ten 5000/5000 database splits, denoted as $\overline{P}_{\mathrm{E}}$.

We selected four content-adaptive steganographic techniques that appear to be the state of the art as of writing this paper (April 2015): WOW [3], S-UNIWARD implemented with the stabilizing constant $\sigma = 1$ as described in [4], HUGO-BD [2] implemented using the Gibbs construction with bounding distortion [33], and the HIgh-Low-Low embedding method called HILL [5]. For HILL, we used the KB high-pass filter and the 3 × 3 and 15 × 15 low-pass averaging filters for L_1 and L_2 as this setting provided the best security as reported in [5]. Finally, we also included the steganographic technique proposed in [12], which inspired the present work and which is also based on minimizing detectability for a multivariate Gaussian (MG) cover model, to show the rather dramatic improvement of this scheme when using the variance estimator described in Section V.

B. Comparison to prior art

We first tested MiPOD implemented with the three settings described in Section V to see the influence of the variance estimator. Table I shows the average total probability of error $\overline{P}_{\mathrm{E}}$ and its standard deviation for a range of payloads for all MiPOD versions and also for four steganographic schemes described in the previous section. Note that, among the three MiPOD versions, the one using the medium block size offers the best security. It also outperforms HUGO-BD, WOW, as well as S-UNIWARD with both feature sets. In the rest of this paper, we always use MiPOD with the medium block size.

Figure 3 is a graphical representation of the table with MiPOD's medium block variance estimator. Note the lower detection errors when steganalyzing with the selection-channel-aware maxSRMd2 feature set in comparison to errors obtained with the SRM. With the more advanced detector, HILL and MiPOD have comparable security with HILL being slightly better for large payloads. At this point, we note that the security of MiPOD can be increased above that of HILL by applying a step similar to what was proposed in [5], [6] by smoothing the Fisher information $I_n = 2/\sigma_n^4$ in MiPOD. In order not to disrupt the flow of this paper, we postpone this to Section VI-F. Finally, we would like to point out a very significant improvement of MiPOD over the MG scheme, which is also based on the multivariate Gaussian cover model but uses a rather simple variance estimator.

C. Experiment on artificial image source

In this section, we justify using the asymptotic approximation of the LR (9) instead of the LR (8) for detection. To this end, we executed an experiment using Monte Carlo simulation on an artificial image source for which

Table I

Detectability in terms of $\overline{P}_{\mathrm{E}}$ versus embedded payload size in bits per pixel (bpp) for three versions of MiPOD and prior art on BOSSbase 1.01 using the FLD ensemble classifier with two feature sets.

Feature	Embedding Method	0.05	0.1	0.2	0.3	0.4	0.5
SRM	WOW	.4572 ± .0026	.4026 ± .0028	.3210 ± .0038	.2553 ± .0028	.2060 ± .0022	.1683 ± .0023
	S-UNIWARD	.4533 ± .0026	.4024 ± .0019	.3199 ± .0027	.2571 ± .0016	.2037 ± .0032	.1640 ± .0024
	HUGO-BD	.4255 ± .0016	.3716 ± .0013	.2871 ± .0016	.2255 ± .0015	.1796 ± .0014	.1450 ± .0010
	HILL	.4691 ± .0017	.4364 ± .0034	.3611 ± .0024	.2996 ± .0022	.2482 ± .0030	.2055 ± .0024
	MiPOD, Small Blocks	.4204 ± .0039	.3477 ± .0023	.2484 ± .0018	.1879 ± .0020	.1420 ± .0025	.1105 ± .0025
	MiPOD, Medium Blocks	.4513 ± .0021	.4065 ± .0043	.3300 ± .0036	.2698 ± .0018	.2210 ± .0022	.1833 ± .0028
	MiPOD, Large Blocks	.4416 ± .0023	.3888 ± .0025	.3105 ± .0039	.2534 ± .0026	.2071 ± .0018	.1719 ± .0034
	MG	.3689 ± .0019	.2953 ± .0026	.2146 ± .0028	.1658 ± .0024	.1357 ± .0030	.1119 ± .0029
maxSRMd2	WOW	.3539 ± .0024	.2997 ± .0023	.2339 ± .0041	.1886 ± .0036	.1543 ± .0036	.1306 ± .0021
	S-UNIWARD	.4180 ± .0025	.3660 ± .0040	.2886 ± .0025	.2360 ± .0022	.1908 ± .0025	.1551 ± .0019
	HUGO-BD	.3652 ± .0023	.3130 ± .0025	.2431 ± .0018	.2020 ± .0015	.1635 ± .0014	.1326 ± .0007
	HILL	.4232 ± .0029	.3771 ± .0019	.3091 ± .0018	.2573 ± .0033	.2184 ± .0037	.1814 ± .0030
	MiPOD, Small Blocks	.3826 ± .0014	.3105 ± .0023	.2220 ± .0018	.1651 ± .0019	.1303 ± .0038	.1022 ± .0028
	MiPOD, Medium Blocks	.4300 ± .0028	.3747 ± .0014	.3030 ± .0019	.2481 ± .0027	.2038 ± .0039	.1678 ± .0038
	MiPOD, Large Blocks	.4195 ± .0029	.3657 ± .0026	.2962 ± .0029	.2390 ± .0036	.1948 ± .0022	.1634 ± .0030
	MG	.2315 ± .0027	.1653 ± .0019	.1161 ± .0016	.0936 ± .0015	.0813 ± .0018	.0715 ± .0018

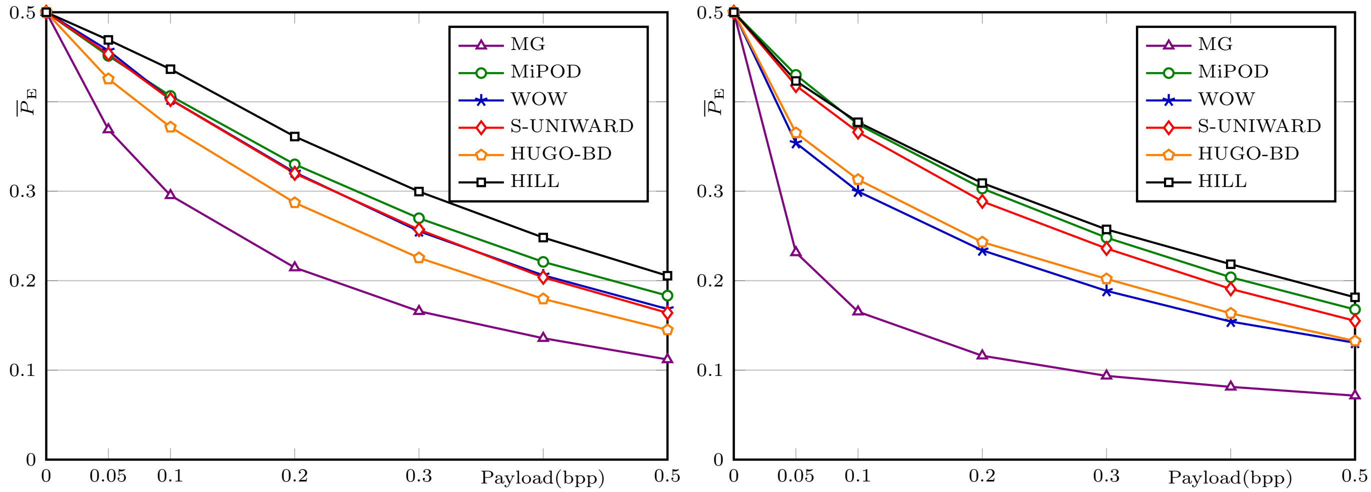


Figure 3. Detection error for different embedding schemes when steganalyzing with SRM [22] (left) and the selection-channel-aware maxSRMd2 [26] (right) which uses the knowledge of change rates. the plot correspond to the results given in Table I.

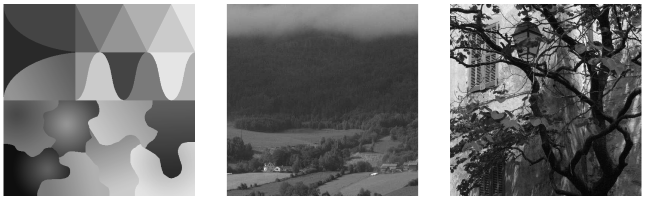


Figure 4. Artificial image (left) and two test images used in the experiment in Section VI-E.

the assumptions of our framework are better satisfied. We started with the image shown in Figure 4 (left) and then superimposed a non-stationary Gaussian noise to it to obtain a source whose noise is known.

The noise variance was selected to be scene dependent based on the heteroscedastic sensor noise model [15], [17] $\sigma_n^2 = a \cdot z_n + b$, where $z_n \in \{0, \ldots, 255\}$ is the nth pixel grayscale value and $a = 6/255$, $b = 2$ are constants. According to [15], [17], these values are fairly typical for a variety of imaging sensors at ISO 200. In other words, in this experiment we made the MVG noise component mimic just the sensor acquisition noise. This was repeated 10,000 times each time with a different realization of the noise to obtain 10,000 cover and the same number of stego images embedded with a fixed payload of 0.2 bpp. Knowing the pixel variances allowed us to compute the ROC curve of the asymptotic LRT (9). Having 10,000 images, we could also sample the LR under both hypotheses and obtain the ROC curve for the sampled LRT (8). We did this for both the omniscient and indifferent wardens.

Figure 5 shows the results when giving the knowledge of the variances to both the sender and the LRT. The close match between the ROC curve of the asymptotic LRT (9) and the sampled LR (8) testifies about the sharpness of our asymptotic analysis. Also observe that the difference in ROC curves between the omniscient and indifferent Warden ($\varrho^{\star}$ (14) vs. $\underline{\varrho}$ (12)) is not significant. In other words, the knowledge of the selection channel does not provide a substantial advantage to the Warden for the tested MiPOD. This is mainly because in our artificial image source MiPOD adapts only to the superimposed heteroscedastic noise as there is almost no modeling error. This makes the embedding only weakly adaptive because $2 \leq \sigma_n^2 \leq 8$.

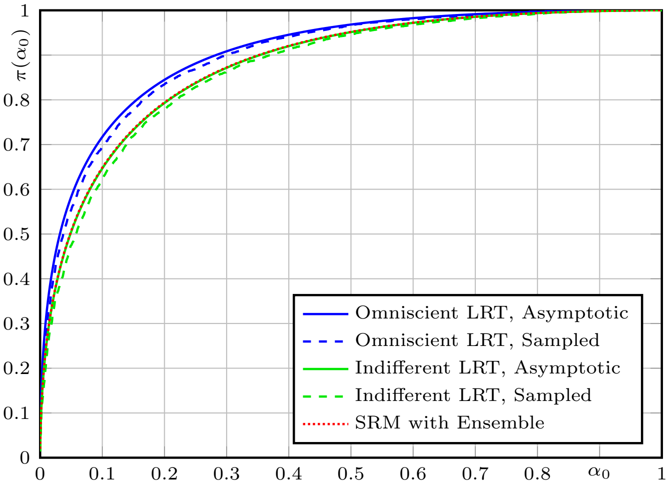


Figure 5. Comparison between the theoretical and empirical detection for a single artificial image ($\alpha = 0.2$ bpp). In this case, both MiPOD and the LR tests know the exact variance of each pixel.

Finally, to see how the optimal LRT compares with empirical detectors, we applied the FLD ensemble with the SRM feature set[4] to the database of 10,000 cover and stego images and drew the ROC curve, also shown in the figure. Remarkably, the empirical detector achieves virtually the same performance as the optimal LR test! This is not obvious at all because both detectors are built very differently. It indicates that, at least in sources with simple content and the heteroscedastic noise model, empirical steganalysis detectors are near optimal.

D. Detectability-limited sender

In this section, we investigate MiPOD's security on BOSSbase for the detectability-limited sender implemented with $\varrho^\star = 2$ as the security level. When embedding, the payload size was iteratively adjusted for each BOSSbase image so that MiPOD induced the prescribed value of $\varrho^\star$. Both the LRT and MiPOD used the same variance estimator (Section V) with the medium block size. Figure 6 shows the ROC curves for the optimal LRT that knows the embedding change rates β_n (the omniscient Warden) and when steganalyzing using the FLD ensemble classifier as described in [31] using SRM and maxSRMd2. In contrast with the results of Figure 5 obtained for a homogeneous source, the SRM now performs significantly worse than the optimal LRT because it has to deal with content diversity across images. The fact that the ROC of the optimal LRT bounds those obtained using empirical detectors indicates that the proposed variance estimators are conservative. In general, however, one cannot claim that the LRT will bound the empirical detectors because the considered MVG cover model is only an approximation.

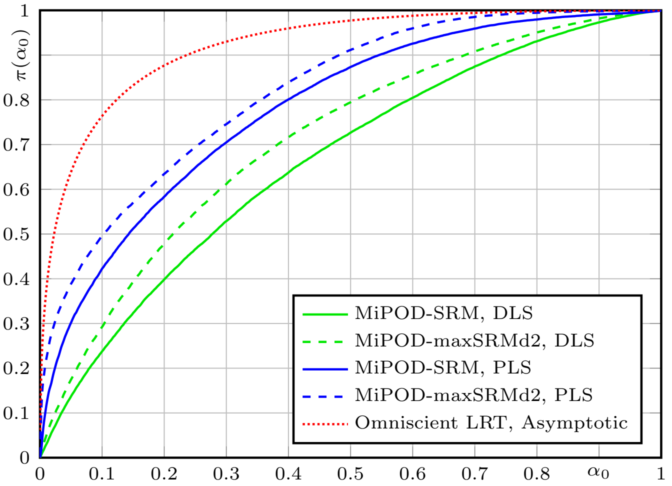


Figure 6. ROC curves for the detectability-limited MiPOD for $\varrho^\star = 2$ (asymptotic LRT, omniscient Warden) and two empirical detectors – the FLD ensemble with SRM and maxSRMd2 feature sets. For comparison, we also show the ROCs for the payload-limited sender (PLS) with payload fixed at the average payload of the DLS (0.2562 bpp).

The DLS could be used for batch steganography [34] to spread the payload among multiple covers to minimize the overall detectability. To see the gain of the DLS over a payload-limited sender (PLS), in the same figure we added the ROCs for the FLD ensemble with the SRM and maxSRMd2 features for a PLS that embeds the same payload in each image so that the average payload per image is the same as for the DLS. When comparing the corresponding ROCs for both senders, one can see a markedly lower detectability of the DLS over the PLS.

E. Determining the secure payload size

Having the distortion related to detectability gives us one more rather interesting possibility to determine, for each image, the size of the secure payload of MiPOD for a given level of risk. Here, we adopt the approach introduced by Ker [35], who proposed to measure the risk by a pair of false-alarm and correct-detection probabilities, α_0, π_0, of the Warden's detector. The steganographers are at (α_0, π_0)-risk if the Warden's detector can simultaneously satisfy $\alpha_0^{\mathrm{War}} < \alpha_0$ and $\pi_0 < \pi_0^{\mathrm{War}}$. Using the analytic expression for the performance of the optimal LRT (14), it immediately follows that:

$$\begin{aligned} & \Phi^{-1}(1-\pi_0) = \Phi^{-1}(1-\alpha_0) - \varrho^\star \\ \Leftrightarrow \quad & \varrho^\star = \Phi^{-1}(1-\alpha_0) - \Phi^{-1}(1-\pi_0) \\ \Leftrightarrow \quad & \varrho^\star = \Phi^{-1}(\pi_0) - \Phi^{-1}(\alpha_0). \end{aligned} \tag{28}$$

Hence, the steganographers are not at (α_0, π_0)-risk if the deflection coefficient $\varrho^\star$ (11) satisfies:

$$\varrho^\star \le \Phi^{-1}(\pi_0) - \Phi^{-1}(\alpha_0). \tag{29}$$

We define the secure payload that corresponds to risk (α_0, π_0) as the largest payload for which the inequality (29)

[4] The ROC curve for the maxSRMd2 features is not shown for better readability because its performance is almost identical to that of SRM because MiPOD is only weakly adaptive due to the properties of the added noise. The detection gain when using the selection-channel-aware maxSRMd2 is thus correspondingly smaller.

is satisfied. In this paper, we use two types of fundamentally different detectors – optimal detectors in the form of the likelihood ratio and empirical detectors constructed as classifiers trained on cover and stego features. We first describe how to determine the secure payload for LR tests and then for an empirical detector.

Once the pixels' variances are known (estimated), the performance of the LRT for a single image can be captured using its ROC curve, which can drawn by first computing the deflection coefficient using either (11) or (12), depending on the Warden type, and then drawing the ROC using formula (14). To estimate the size of the secure payload for a given risk, (α_0, π_0), the payload size can be iteratively adjusted so that the LRT's ROC curve goes through the pair (α_0, π_0).

To estimate the secure payload for a specific image using empirical detectors, we create a database of 10,000 images by denoising the image and then superimposing to the denoised image 10,000 different realizations of MVG noise with the estimated variances $\hat{\sigma}_n^2$. Given a payload R, one can create a database of 10,000 stego images embedded with payload R, train an empirical detector for the given cover and stego sources, and draw its ROC curve. The secure payload for a given risk (α_0, π_0) is again determined iteratively by adjusting R to force the empirical ROC curve to go through the pair (α_0, π_0).

In order to proclaim the secure payload determined from our model as an accurate estimate for an image acquired using an imaging sensor rather than an artificial image, we need a close match between our adopted model and the reality. Because BOSSbase images were processed using demosaicking (which is a form of content-driven filtering) and resizing, they are too complex to closely follow our model. Consequently, secure payload estimates obtained using our simplified model would most likely be inaccurate. Thus, for the experiments in this section we used two raw BOSSbase images, sampled them only at the red color filter (the red channel), and then centrally cropped to 512×512 pixels. The processing was executed using the 'convert' linux script from ImageMagick (version 6.7.7-10), for resizing and extracting the red color channel, and using ufraw version 0.18, which uses dcraw version 9.06, for conversion from RAW to the PPM format, see [36] for more details. Thus, in these images the pixel values were processed using only point-wise operations, which included gain and gamma adjustment.[5] Because the images were not resized (in contrast to BOSSbase images), their content is much smoother (Figure 4 middle and right). The noise variance is thus mostly affected by the acquisition noise, which follows the MVG model but not the heteroscedastic model because of the gamma correction.

In all experiments below, the variances σ_n^2 were estimated using MiPOD's variance estimator with the medium block size. They were given to the LR detectors of both the omniscient and indifferent Wardens as well as to MiPOD.

In order to see how the image content affects the secure payload estimate, we carried out experiments on two images from BOSSbase shown in Figure 4 middle and right. Figure 7 shows the secure payload on the y axis as a function of π_0 for selected values of α_0 (different risks) for BOSSbase images '1310.pgm' and '1289.pgm'. As expected, if the steganographers desire perfect security by setting $\alpha_0 = \pi_0$, the secure payload tends to zero. On the other hand, if the steganographers do not set any constraints on the security by choosing $\pi_0 = 1$, the secure payload tends to 1.

Notice that the secure payload estimates are higher for image '1289.pgm' because it has more complex content and larger differences in pixel intensity. The estimates using the sampled and asymptotic LR are close for both images and both Wardens. Because the omniscient Warden can detect embedding more reliably than the indifferent one, the secure payload determined using the omniscient Warden is understandably always smaller than the one obtained with the indifferent Warden. This difference is slightly larger for the busier image. Because of the lower detection power of the empirical detector with SRM features, its secure payload size is always overestimated. For the smoother image, however, the empirical estimates and the ones obtained using the indifferent Warden are quite close, which again validates our model. As our final note, we point out that the secure payload size for the maxSRMd2 feature set is not shown in the figures because it is similar to that of the SRM.

F. Improving MiPOD's security by smoothing the Fisher information

Recently, it has been shown that the empirical security of steganographic schemes can be improved by smoothing the embedding costs using a low-pass filter [5], [6]. This can be explained intuitively by observing that the smoothing spreads high costs of pixels into their neighborhood making the embedding more conservative. Moreover, and most importantly, it evens out the costs and thus increases the entropy of embedding changes (the payload) in highly textured regions where empirically built detectors fail to detect embedding because the changes affect mostly the marginal bins in co-occurrences of SRM noise residuals [22].

In MiPOD, the linear parametric model is applied pixelwise, which makes variance estimations of neighboring pixels (and the associated pixel costs) strongly correlated. Therefore, the smoothing is at least partially an inherent property of MiPOD rather than artificially forced. Note in Figure 2 that the embedding change probability for the medium-size-block MiPOD is much smoother than that of S-UNIWARD. On the other hand, it is not as smooth as for HILL. Thus, we decided to investigate whether additional smoothing might further boost MiPOD's security. Since in MiPOD we do not natively work with the concept of a

[5] These operations are in fact performed on the CMOS sensor.

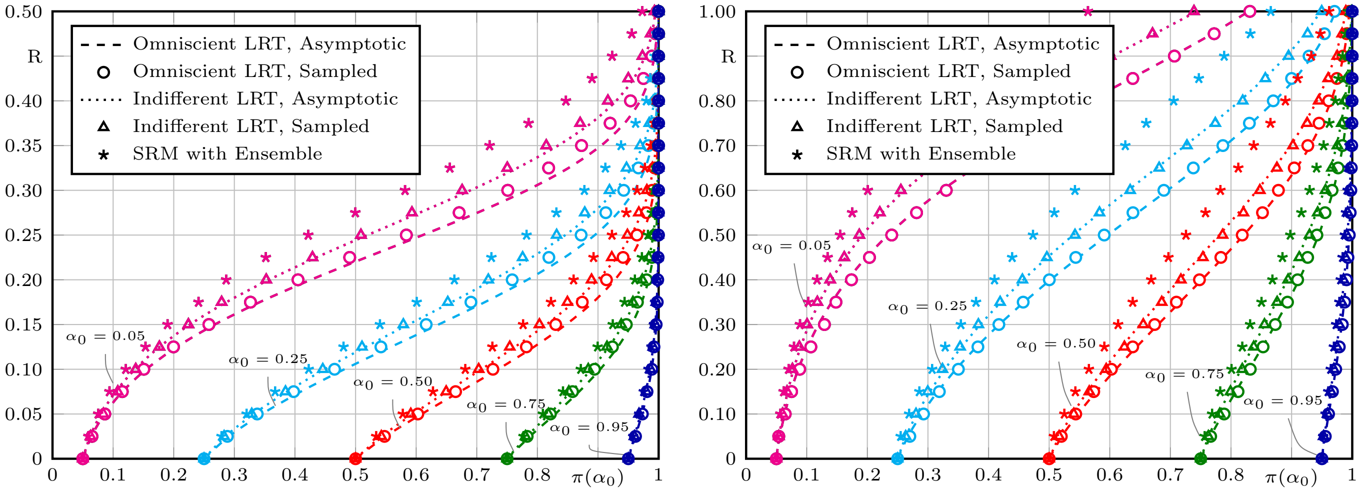


Figure 7. Secure payload determined from the asymptotic and sampled LR for both Wardens and an empirical detector implemented with FLD ensemble and SRM. The secure payload is shown for various risk levels as a function of π_0 for $\alpha_0 = \{0.05; 0.25; 0.5; 0.75; 0.95\}$ for BOSSbase image '1310.pgm' (left) and '1289.pgm' (right) with superimposed MVG noise with the variance of each pixel computed using the estimator described in Section V. Note the different y-axis scales between the figures.

pixel cost (we only need to revert to it when implementing an actual embedding scheme using codes), we decided to apply the smoothing to the Fisher information, $I_n = 2/\sigma_n^4$. Because the pixel cost of MiPOD (20) is positively correlated with I_n, smoothing I_n will have a similar effect as smoothing the costs. The result will, however, be different because the relationship between I_n and the cost is nonlinear (see Eqs. (18)–(20)).

We performed a search over the size of a simple square averaging kernel applied to I_n and determined that, in our image source, the 7×7 support gave the overall best results, boosting the detection error $\overline{P}_{\mathrm{E}}$ by up to 2.4% when detecting with the maxSRMd2 feature set. We summarize the results in Table II and Figure 8.

Table II
MiPOD's detectability $\overline{P}_{\mathrm{E}}$ when smoothing the Fisher information (BOSSbase 1.01, maxSRMd2).

Payload	HILL	MiPOD Medium Blocks	MiPOD Medium Blocks Smooth FI
0.05	.4232 ± .0029	.4300 ± .0028	.4380 ± .0012
0.1	.3771 ± .0019	.3747 ± .0014	.3939 ± .0022
0.2	.3091 ± .0018	.3030 ± .0019	.3237 ± .0021
0.3	.2573 ± .0033	.2481 ± .0027	.2717 ± .0045
0.4	.2184 ± .0037	.2038 ± .0039	.2243 ± .0055
0.5	.1814 ± .0030	.1678 ± .0038	.1845 ± .0030

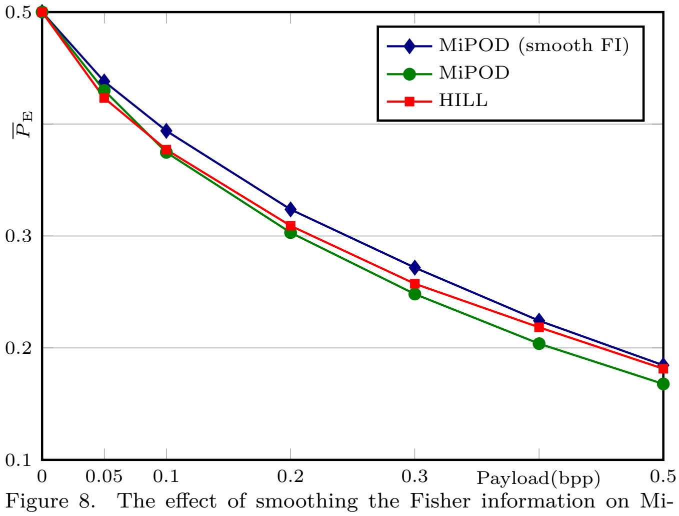


Figure 8. The effect of smoothing the Fisher information on MiPOD's security w.r.t. the maxSRMd2 feature set. The plot corresponds to the results given in Table II.

VII. Conclusions

Model based steganography has been around for almost fifteen years since the introduction of OutGuess. What makes our approach different is the dimensionality of the parameter space, which allows us to capture the non-stationary character of images, as well as the fact that we do not attempt to preserve the model but rather minimize the impact of embedding. We model the image noise residual as a sequence of independent quantized Gaussian variables with varying variances. By working with the residual, besides the acquisition noise we managed to include in the model the content-dependent modeling error, which has a strong effect on steganalysis. On the other hand, the assumption of independence and the simplicity of the Gaussian distribution allowed us to derive a closed-form expression for the power of the most powerful detector of content-adaptive LSB matching within the selected model. This allows us to achieve the following novel insights into both steganography design and steganalysis.

First, we use our approach to design steganography that minimizes the power of the optimal detector rather than a heuristically assembled distortion. By adjusting the parameters of the model variance estimator, our embedding scheme called MiPOD rivals the security of the most advanced steganographic schemes today. Further improvement is likely possible by optimizing the local variance estimator. Here, we point out a caveat that such an optimization will necessarily be limited to a given image source and empirical detector (classifier choice and the feature space).

Second, we used the closed-form expression for the theoretical detectability to reveal new fundamental insight into the complex interplay between empirical detectors constructed as classifiers and detectors derived as optimal within the chosen model. In particular, when the cover noise model was forced onto an artificial image with simple content, we observed that empirical detectors built as classifiers in rich feature spaces closely matched the detection performance of optimal detectors, despite their extremely different nature. On real image sources, however, the empirical detectors were markedly suboptimal with respect to the theoretically optimal detectors. We attributed this to the difficulty of empirical detectors to deal with the heterogeneity of natural images.

Third, we also performed experiments aimed at estimating the size of the secure payload with respect to a given level of risk as defined by Ker. Here, we used the red channel of a raw image acquired by an imaging sensor quantized to 8 bits that has undergone only gain and gamma adjustment. Because such images closely follow our model, one can compute their secure payload from the deflection coefficient of the asymptotic likelihood ratio once the variances are estimated. Such estimate was contrasted with the secure payload determined using empirical detectors (FLD classifiers) trained on a database of 10,000 cover images (and the corresponding stego images) obtained by denoising the image and superimposing 10,000 realizations of multivariate Gaussian noise estimated from the original image. For images with simple content, both estimates appear quite close while for images with more complex content the empirical detector overestimates the payload due to its lower detection power.

We intend to pursue several extensions of this work. On the steganography side, we plan to investigate models that capture dependencies among spatially adjacent pixels, e. g., by considering pairs of neighboring pixels as jointly Gaussian random variables. This may lead to schemes that adjust the direction of the embedding change based on the changes made to adjacent pixels. The detectability-limited sender and the asymptotic LRT could both be used to further investigate the difficult and open problem of batch steganography and pooled steganalysis.

Acknowledgements

The code for all steganographic methods, feature extractors, and classifiers used in this paper is available from http://dde.binghamton.edu/download/.

Appendix

In this appendix, without loss of generality, we analytically establish the performance of the Generalized Likelihood Ratio Test (GLRT) for the case in which the sender changes each pixel with probabilities $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_N)$ while the Warden uses estimated change rates $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N)$. Note also that we use the term GLRT for convenience here as generally γ_n may not be MLE estimates.

This appendix is divided into two parts, first, the GLRT is presented and a simple asymptotic expression is obtained under the fine quantization limit and for a large number of pixels. Then, the statistical performance of this GLRT is analytically established.

A. Asymptotic expression for the GLRT

Using the corresponding expressions for $p_{\sigma_n}(k)$ (1) and $q_{\sigma_n,\beta_n}(k)$ (3), the LR (8) for one observation Λ_n can be written as

$$\frac{(1-2\gamma_n)\exp\!\left(\frac{-x_n^2}{2\sigma_n^2}\right)+\gamma_n\exp\!\left(\frac{-(x_n+1)^2}{2\sigma_n^2}\right)+\gamma_n\exp\!\left(\frac{-(x_n-1)^2}{2\sigma_n^2}\right)}{\exp\left(-\frac{x_n^2}{2\sigma_n^2}\right)}, \tag{30}$$

which can be simplified as follows:

$$\begin{aligned}\Lambda_n &= 1-2\gamma_n+\gamma_n\exp\left(\frac{-(x_n+1)^2+x_n^2}{2\sigma_n^2}\right)\\ &\quad+\gamma_n\exp\left(\frac{-(x_n-1)^2+x_n^2}{2\sigma_n^2}\right) \qquad (31)\\ &= 1-2\gamma_n+\gamma_n\left(\exp\left(\frac{x_n-1/2}{\sigma_n^2}\right)+\exp\left(\frac{-x_n-1/2}{\sigma_n^2}\right)\right). \qquad (32)\end{aligned}$$

Under the fine quantization assumption, $\sigma_n^2 \gg 1$, we can further simplify using the second-order Taylor expansion around $\sigma_n^{-2} = 0$:

$$\Lambda_n \approx 1-2\gamma_n+\gamma_n\left(2-\frac{1}{\sigma_n^2}+\frac{x_n^2+1/4}{\sigma_n^4}\right) \tag{33}$$

$$= 1+\gamma_n\left(-\frac{1}{\sigma_n^2}+\frac{x_n^2+1/4}{\sigma_n^4}\right). \tag{34}$$

Using the fine quantization assumption again, we replace the log-LR, $\log\Lambda_n$, with its first-order Taylor approximation:

$$\log\Lambda_n = \gamma_n\left(-\frac{1}{\sigma_n^2}+\frac{x_n^2+1/4}{\sigma_n^4}\right). \tag{35}$$

Since the term involving $1/4$ can be removed from the test statistic (it stays the same under both hypotheses) the log-LR can be further simplified:

$$\log\Lambda_n = \gamma_n\left(-\frac{1}{\sigma_n^2}+\frac{x_n^2}{\sigma_n^4}\right). \tag{36}$$

B. Analytic expression of GLRT performance

We now compute the mean and variance of the log-LR (36) under both hypotheses. Because under $\mathcal{H}_0$, $\frac{x_n}{\sigma_n} \sim$

$\mathcal{N}(0,1)$, we have $\frac{x_n^2}{\sigma_n^2} \sim \chi_1^2$. Since $\mathbb{E}[\chi_1^2] = 1$ and $\mathbb{V}ar[\chi_1^2] = 2$, and because $\frac{x_n^2}{\sigma_n^4} = \frac{1}{\sigma_n^2}\frac{x_n^2}{\sigma_n^2}$,

$$\mathbb{E}_0\left[\frac{x_n^2}{\sigma_n^4}\right] = \frac{1}{\sigma_n^2}, \tag{37}$$

$$\mathbb{V}ar_0\left[\frac{x_n^2}{\sigma_n^4}\right] = \frac{2}{\sigma_n^4}. \tag{38}$$

Finally, it follows from the expression for the log-LR (36) that

$$\mathbb{E}_0\left[\log \Lambda_n\right] = 0, \tag{39}$$

$$\mathbb{V}ar_0\left[\log \Lambda_n\right] = \frac{2\gamma_n^2}{\sigma_n^4}. \tag{40}$$

Under hypothesis $\mathcal{H}_1$, the calculation of the log-LR's moments is slightly more complex because the pmf of stego pixels is a mixture of three different cases: $s_n = x_n$, $s_n = x_n + 1$, and $s_n = x_n - 1$. In particular,

$$\begin{aligned}
\mathbb{E}_1[x_n^2] &= (1-2\beta_n)\mathbb{E}_0[x_n^2] + \beta_n\mathbb{E}_0[(x_n-1)^2] + \beta_n\mathbb{E}_0[(x_n+1)^2]\\
&= (1-2\beta_n)\mathbb{E}_0[x_n^2] + \beta_n\mathbb{E}_0[x_n^2+1] + \beta_n\mathbb{E}_0[x_n^2+1]\\
&= (1-2\beta_n)\mathbb{E}_0[x_n^2] + 2\beta_n\mathbb{E}_0[x_n^2+1]\\
&= (1-2\beta_n)\sigma_n^2 + 2\beta_n(\sigma_n^2+1)\\
&= \sigma_n^2 + 2\beta_n.
\end{aligned}$$

Thus, $\mathbb{E}_1\left[x_n^2/\sigma_n^4 - 1/\sigma_n^2\right] = 2\beta_n/\sigma_n^4$, which implies $\mathbb{E}_1[\log \Lambda_n] = 2\gamma_n\beta_n/\sigma_n^4$. As for the variance, we use fact that $\mathbb{V}ar[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ for any random variable X and that

$$\begin{aligned}
\mathbb{E}_1[x_n^4] &= (1-2\beta_n)\mathbb{E}_0[x_n^4] + \beta_n\mathbb{E}_0[(x_n-1)^4] + \beta_n\mathbb{E}_0[(x_n+1)^4]\\
&= (1-2\beta_n)\mathbb{E}_0[x_n^4] + 2\beta_n\mathbb{E}_0[x_n^4+6x_n^2+1]\\
&= (1-2\beta_n)3\sigma_n^4 + 2\beta_n(3\sigma_n^4 + 6\sigma_n^2 + 1)\\
&= 2\beta_n + 3\sigma_n^4 + 12\beta_n\sigma_n^2.
\end{aligned}$$

After some simple arithmetic and keeping only the leading term:

$$\begin{aligned}
\mathbb{V}ar_1\left[\log \Lambda_n\right] &= \mathbb{E}_1\left[\gamma_n^2\left(\frac{x_n^2}{\sigma_n^4} - \frac{1}{\sigma_n^2}\right)^2\right] - \mathbb{E}_1\left[\gamma_n\left(\frac{x_n^2}{\sigma_n^4} - \frac{1}{\sigma_n^2}\right)\right]^2\\
&\approx \frac{2\gamma_n^2}{\sigma_n^4}(1 + \mathcal{O}(\sigma_n^{-2})).
\end{aligned}$$

Therefore, the final result under the alternative hypothesis is

$$\mathbb{E}_1\left[\log \Lambda_n\right] = \frac{2\beta_n\gamma_n}{\sigma_n^4}, \tag{41}$$

$$\mathbb{V}ar_1\left[\log \Lambda_n\right] \approx \frac{2\gamma_n^2}{\sigma_n^4} = \mathbb{V}ar_0[\log \Lambda_n]. \tag{42}$$

We are now ready to compute the detectability of LSBM. To this end, we study the properties of the log-LR of all pixels, which, from the statistical independence of pixels, is given by $\Lambda(\mathbf{x}) = \prod_{n=1}^{N} \Lambda_n$, or, after taking the logarithm, $\log \Lambda(\mathbf{x}) = \sum_{n=1}^{N} \log \Lambda_n$. From the Lindeberg's version of the Central Limit Theorem, we have under $\mathcal{H}_0$ from the moments of the log-LR (37)–(38):

$$\frac{\log \Lambda}{\sqrt{2\sum_{n=1}^{N} \sigma_n^{-4}\gamma_n^2}} \rightsquigarrow \mathcal{N}(0,1), \tag{43}$$

with $\rightsquigarrow$ denoting the convergence in distribution. Similarly, under the alternative hypothesis $\mathcal{H}_1$ one immediately gets from the moments of the log-LR (41)–(42):

$$\frac{\log \Lambda}{\sqrt{2\sum_{n=1}^{N} \sigma_n^{-4}\gamma_n^2}} \rightsquigarrow \mathcal{N}(\varrho,1) \tag{44}$$

with

$$\varrho = \frac{2\sum_{n=1}^{N} \sigma_n^{-4}\gamma_n\beta_n}{\sqrt{2\sum_{n=1}^{N} \sigma_n^{-4}\gamma_n^2}}. \tag{45}$$

References

[1] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE TIFS*, vol. 6, pp. 920–935, September 2011.

[2] T. Pevný, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Information Hiding, 12th International Conference*, vol. 6387 of *LNCS*, (Calgary, Canada), pp. 161–177, Springer-Verlag, New York, June 28–30, 2010.

[3] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Proc. IEEE WIFS*, (Tenerife, Spain), December 2–5, 2012.

[4] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion design for steganography in an arbitrary domain," *EURASIP Journal on Information Security, Special Issue on Revised Selected Papers of the 1st ACM IH and MMS Workshop*, vol. 2014:1, 2014.

[5] B. Li, M. Wang, and J. Huang, "A new cost function for spatial image steganography," in *Proceedings IEEE ICIP*, (Paris, France), October 27–30, 2014.

[6] B. Li, S. Tan, M. Wang, and J. Huang, "Investigation on cost assignment in spatial image steganography," *IEEE TIFS*, vol. 9, pp. 1264–1277, August 2014.

[7] R. Böhme, *Advanced Statistical Steganalysis*. Berlin Heidelberg: Springer-Verlag, 2010.

[8] A. D. Ker, P. Bas, R. Böhme, R. Cogranne, S. Craver, T. Filler, J. Fridrich, and T. Pevný, "Moving steganography and steganalysis from the laboratory into the real world," in *1st ACM IH&MMSec. Workshop* (W. Puech, M. Chaumont, J. Dittmann, and P. Campisi, eds.), (Montpellier, France), June 17–19, 2013.

[9] T. Pevný, P. Bas, and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," *IEEE TIFS*, vol. 5, pp. 215–224, June 2010.

[10] T. Filler and J. Fridrich, "Design of adaptive steganographic schemes for digital images," in *Proceedings SPIE, Electronic Imaging, Media Watermarking, Security and Forensics III* (A. Alattar, N. D. Memon, E. J. Delp, and J. Dittmann, eds.), vol. 7880, (San Francisco, CA), pp. OF 1–14, January 23–26, 2011.

[11] J. Kodovský, J. Fridrich, and V. Holub, "On dangers of overtraining steganography to incomplete cover model," in *Proceedings of the 13th ACM Multimedia & Security Workshop* (J. Dittmann, S. Craver, and C. Heitzenrater, eds.), (Niagara Falls, NY), pp. 69–76, September 29–30, 2011.

[12] J. Fridrich and J. Kodovský, "Multivariate Gaussian model for designing additive distortion for steganography," in *Proc. IEEE ICASSP*, (Vancouver, BC), May 26–31, 2013.

[13] V. Sedighi, J. Fridrich, and R. Cogranne, "Content-adaptive pentary steganography using the multivariate generalized Gaussian cover model," in *Proceedings SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics 2015*, (San Francisco, CA), February 9–11, 2015.

[14] J. R. Janesick, *Scientific Charge-Coupled Devices*, vol. Monograph PM83. Washington, DC: SPIE Press - The International Society for Optical Engineering, January 2001.

[15] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, “Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data,” *IEEE TIP*, vol. 17, pp. 1737–1754, Oct. 2008.

[16] G. E. Healey and R. Kondepudy, “Radiometric CCD camera calibration and noise estimation,” *IEEE TPAMI*, vol. 16, pp. 267–276, March 1994.

[17] T. H. Thai, R. Cogranne, and F. Retraint, “Camera model identification based on the heteroscedastic noise model,” *Image Processing, IEEE Transactions on*, vol. 23, pp. 250–263, Jan 2014.

[18] R. Cogranne, C. Zitzmann, L. Fillatre, F. Retraint, I. Nikiforov, and P. Cornu, “A cover image model for reliable steganalysis,” in *Information Hiding, 13th International Conference*, vol. 7692 of *LNCS*, (Prague, Czech Republic), pp. 178–192, May 18–20, 2011.

[19] C. Zitzmann, R. Cogranne, F. Retraint, I. Nikiforov, L. Fillatre, and P. Cornu, “Statistical decision methods in hidden information detection,” in *Information Hiding, 13th International Conference*, vol. 7692 of *LNCS*, (Prague, Czech Republic), pp. 163–177, May 18–20, 2011.

[20] R. Cogranne and F. Retraint, “An asymptotically uniformly most powerful test for LSB Matching detection,” *IEEE TIFS*, vol. 8, no. 3, pp. 464–476, 2013.

[21] R. Cogranne and F. Retraint, “Application of hypothesis testing theory for optimal detection of LSB matching data hiding,” *Signal Processing*, vol. 93, pp. 1724–1737, July, 2013.

[22] J. Fridrich and J. Kodovský, “Rich models for steganalysis of digital images,” *IEEE TIFS*, vol. 7, pp. 868–882, June 2011.

[23] E. Lehmann and J. Romano, *Testing Statistical Hypotheses, 2nd edition.* Springer, 2005.

[24] L. Chen, Y. Shi, P. Sutthiwan, and X. Niu, “A novel mapping scheme for steganalysis,” in *Proc. IWDW*, vol. 7809 of *LNCS*, pp. 19–33, Springer Berlin Heidelberg, 2013.

[25] W. Tang, H. Li, W. Luo, and J. Huang, “Adaptive steganalysis against WOW embedding algorithm,” in *2nd ACM IH&MMSec. Workshop*, (Salzburg, Austria), June 11–13, 2014.

[26] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich, “Selection-channel-aware rich model for steganalysis of digital images,” in *Proc. IEEE WIFS*, (Atlanta, GA), December 3–5, 2014.

[27] R. Cogranne and F. Retraint, “Statistical detection of defects in radiographic images using an adaptive parametric model,” *Signal Processing*, vol. 96-B, no. 3, pp. 173–189, 2014.

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

[29] P. Bas, T. Filler, and T. Pevný, “Break our steganographic system – the ins and outs of organizing BOSS,” in *Information Hiding, 13th International Conference* (T. Filler, T. Pevný, A. Ker, and S. Craver, eds.), vol. 6958 of *LNCS*, (Prague, Czech Republic), pp. 59–70, May 18–20, 2011.

[30] J. Kodovský, J. Fridrich, and V. Holub, “Ensemble classifiers for steganalysis of digital media,” *IEEE TIFS*, vol. 7, no. 2, pp. 432–444, 2012.

[31] R. Cogranne, T. Denemark, and J. Fridrich, “Theoretical model of the FLD ensemble classifier based on hypothesis testing theory,” in *Proc. IEEE WIFS*, (Atlanta, GA, USA), December 3–5 2014.

[32] R. Cogranne and J. Fridrich, “Modeling and extending the ensemble classifier for steganalysis of digital images using hypothesis testing theory,” *IEEE TIFS*, 2015 (to appear).

[33] T. Filler and J. Fridrich, “Gibbs construction in steganography,” *IEEE TIFS*, vol. 5, no. 4, pp. 705–720, 2010.

[34] A. D. Ker, “Batch steganography and pooled steganalysis,” in *Information Hiding, 8th International Workshop* (J. L. Camenisch, C. S. Collberg, N. F. Johnson, and P. Sallee, eds.), vol. 4437 of *LNCS*, (Alexandria, VA), pp. 265–281, Springer-Verlag, New York, July 10–12, 2006.

[35] A. D. Ker, “A capacity result for batch steganography,” *IEEE Signal Processing Letters*, vol. 14, no. 8, pp. 525–528, 2007.

[36] M. Goljan, R. Cogranne, and J. Fridrich, “Rich model for steganalysis of color images,” in *Proc. IEEE WIFS*, (Atlanta, GA), December 3–5, 2014.

Vahid Sedighi received his B.S. degree in electrical engineering in 2005 from Shahed University, Tehran, Iran, and his M.S. degree in electrical engineering in 2010 from Yazd University, Yazd, Iran. He is currently pursuing the Ph.D degree in the Department of Electrical and Computer Engineering at Binghamton University, State University of New York. His research interests include statistical signal processing, steganography, steganalysis, and machine learning.

Rémi Cogranne holds the position of Associate Professor at Troyes University of Technology (UTT). He has received his PhD in Systems Safety and Optimization in 2011 and his engineering degree in computer science and telecommunication in 2008 both from UTT. He has been a visiting scholar at Binghamton University in 2014-2015. During his studies, he took a semester off to teach in a primary school in Ziguinchor, Senegal and studied one semester at Jöonköping University, Sweden. His main research interests are in hypothesis testing, steganalysis, steganography, image forensics and statistical image processing.

Jessica Fridrich holds the position of Professor of Electrical and Computer Engineering at Binghamton University (SUNY). She has received her PhD in Systems Science from Binghamton University in 1995 and MS in Applied Mathematics from Czech Technical University in Prague in 1987. Her main interests are in steganography, steganalysis, digital watermarking, and digital image forensic. Dr. Fridrich’s research work has been generously supported by the US Air Force and AFOSR. Since 1995, she received 19 research grants totaling over $9 mil for projects on data embedding and steganalysis that lead to more than 160 papers and 7 US patents. Dr. Fridrich is a member of IEEE and ACM.