

# Content-Adaptive Pentary Steganography Using the Multivariate Generalized Gaussian Cover Model

Vahid Sedighi,<sup>+</sup> Jessica Fridrich,<sup>+</sup> and Rémi Cogramne,<sup>×</sup>

<sup>+</sup>Department of ECE, SUNY Binghamton, NY, USA, {vsedigh1,fridrich}@binghamton.edu

<sup>×</sup>ICD - ROSAS - LM2S, Troyes University of Technology, Troyes, France, remi.cogramne@utt.fr

## ABSTRACT

The vast majority of steganographic schemes for digital images stored in the raster format limit the amplitude of embedding changes to the smallest possible value. In this paper, we investigate the possibility to further improve the empirical security by allowing the embedding changes in highly textured areas to have a larger amplitude and thus embedding there a larger payload. Our approach is entirely model driven in the sense that the probabilities with which the cover pixels should be changed by a certain amount are derived from the cover model to minimize the power of an optimal statistical test. The embedding consists of two steps. First, the sender estimates the cover model parameters, the pixel variances, when modeling the pixels as a sequence of independent but not identically distributed generalized Gaussian random variables. Then, the embedding change probabilities for changing each pixel by 1 or 2, which can be transformed to costs for practical embedding using syndrome-trellis codes, are computed by solving a pair of non-linear algebraic equations. Using rich models and selection-channel-aware features, we compare the security of our scheme based on the generalized Gaussian model with pentary versions of two popular embedding algorithms: HILL and S-UNIWARD.

## 1. MOTIVATION

Currently, the most successful approach to designing adaptive steganography is based on minimizing a distortion between the cover and the corresponding stego object. A popular form of the distortion is obtained by first assigning a cost of changing each cover element (e.g., pixel or DCT coefficient) and then computing the total distortion as a sum of costs of all modified pixels. When larger costs are assigned to pixels where the detection is expected to be easier (smooth segments, such as blue sky), the embedding changes concentrate in textured and/or noisy regions where the modifications will be harder to detect by an adversary. Coding methods, such as the syndrome-trellis codes,<sup>9</sup> can be used to realize such steganographic schemes in practice in a near-optimal fashion w.r.t. the theoretical rate-distortion bound.

The fundamental assumption made here is that the distortion should be related to statistical detectability. However, most distortion functions are assembled in an ad hoc or empirical manner, often assigning the pixel costs by quantifying the impact of making an embedding change on outputs of one or more high-pass filters (noise residuals).<sup>15,16,21,22</sup> Recently, a qualitatively different and entirely model-driven approach has been proposed in Refs [5,12]. It starts with imposing a model on the cover object and estimating the model parameters, the local pixel variances at each pixel. Then, the costs (the probabilities of modifying each pixel or change rates) are computed *analytically* from the estimated model in order to minimize the Kullback-Leibler divergence between the cover and stego distributions<sup>12</sup> or, equivalently, the power of an optimal statistical test.<sup>5</sup> By assuming that pixels form a sequence of independent but not necessarily identically distributed random variables, it becomes tractable to compute the optimal embedding change probabilities (costs) using the method of Lagrange multipliers.

In this paper, we use the same framework with two innovative elements. First, we explore the possibility to further improve the empirical security by using a more general cover model. While in the prior work<sup>5,12</sup> the multivariate Gaussian distribution was used to model pixels, here we employ the Multivariate Generalized Gaussian (MVG) model. Since the generalized Gaussian can have thicker tails, it makes sense to allow embedding changes with amplitude larger than 1 to embed a larger payload in pixels from textured areas. Pentary embedding has already been studied and shown beneficial in the past.<sup>8</sup> Our study shows that within our model-driven framework when measuring the security empirically using classifiers trained on examples of cover and

stego images, pentary embedding schemes indeed improve security. This gain becomes larger for larger payloads. Surprisingly, despite the fact that the shape parameter of the generalized Gaussian model affects the properties of the selection channel in a major manner, we see little effect on empirical security. This suggests that the way current empirical detectors equipped with rich media features incorporate the knowledge of the selection channel is suboptimal.

In the next section, we introduce the cover model, the embedding operation, and derive the stego source model. In Section 3, we formulate steganalysis as a hypothesis testing problem and derive the asymptotic statistical distribution of the optimal detector in the fine-quantization and small-payload limit for a large number of pixels. In Section 3, we show that the optimal change rates controlling the embedding changes by  $\pm 1$  and  $\pm 2$  must satisfy a set of two non-linear equations that need to be solved for each pixel together with the global payload constraint (for the payload-limited sender). The cover parameter estimation routine is described in Section 5. The proposed MVGG embedding scheme is tested in Section 6, where we report the detection results with the Spatial Rich Model (SRM)<sup>11</sup> and its selection-channel-aware version called maxSRMd2<sup>7</sup> on a standard image database. We investigate the effect of the shape parameter in the MVGG model on detection and compare the schemes to existing state-of-the-art embedding schemes. The paper is summarized in Section 7.

## 2. COVER AND STEGO IMAGE MODELS

In this section, we introduce the model we use to describe images represented in the spatial domain. We also define the embedding operation and derive the distribution of pixels in the stego image.

### 2.1 Cover model

The light intensity values registered by an imaging sensor are corrupted by the shot noise, dark current, readout, and electronic noise. The superposition of these noise sources can be well modeled using a Gaussian distribution.<sup>13,14</sup> For simplicity and tractability, we will ignore the effects of the subsequent processing, such as demosaicking and filtering, that make the noise in spatially adjacent pixels dependent. Formally, we will consider the pixels in a digital image to be a sequence of  $N$  independent realizations  $(z_1, \dots, z_N)$  of quantized Gaussian random variables  $(Z_1, \dots, Z_N)$ ,  $Z_n \sim \mathcal{N}(\mu_n, \omega_n^2)$ . When estimating the pixel mean using denoising, the difference  $x_n = \mu_n - \hat{\mu}_n$  will inevitably contain also the modeling error besides the Gaussian acquisition noise. The modeling error will be especially significant in textured regions of the image and may exhibit thick tails (non-Gaussianity). We include the modeling error into the noise component and adopt a Generalized Gaussian (GG) model for  $\mathbf{x} = (x_1, \dots, x_N)$ , which we consider as  $N$  independent realizations of quantized zero-mean generalized Gaussian random variables  $(X_1, \dots, X_N)$  distributed according to

$$X_n \sim f(x; b_n, \nu) = \frac{\nu}{2b_n\Gamma(1/\nu)} \exp\left(-\left|\frac{x}{b_n}\right|^\nu\right), \quad (1)$$

where  $\nu$  is the shape parameter and  $b$  is the width parameter directly related to the variance,  $\sigma_n^2$ , via the relationship  $\sigma_n^2 = b_n^2\Gamma(3/\nu)/\Gamma(1/\nu)$ . Notice that, for simplicity, we assume that the shape parameter of the GG is constant across the image and only the variances vary. When estimating the cover model, only the variances will be estimated at every pixel as explained in Section 5. We will experiment with several values for the shape parameter  $\nu$  to see which value provides the best empirical security.

As already mentioned, we assume that the above distributions are quantized with a quantizer with centroids  $k\Delta$ ,  $k \in \mathbb{Z}$ . For simplicity and without loss on generality, in this paper we set  $\Delta = 1$ . Assuming the fine quantization limit,  $\Delta \ll \sigma_n$  for all  $n$ , the probability mass function (pmf) of the  $n$ th pixel is given by  $\mathcal{P}_0 = (p_0(k; \sigma_n, \nu))_{k \in \mathbb{Z}}$

$$p_0(k; \sigma_n, \nu) = \mathbb{P}(x_n = k) \propto \frac{\nu}{2b_n\Gamma(1/\nu)} \exp\left(-\frac{|k|^\nu}{b_n^\nu}\right). \quad (2)$$

Note that in our model, we do not truncate the pixel values and instead consider them unbounded. This is intentional to simplify the subsequent theoretical exposition. The fine quantization limit,  $\Delta \ll \sigma_n$ , is violated whenever the pixel values almost do not vary, such as saturated patches due to overexposed light sources. Since the embedding changes in such regions would be easily detectable, any content-adaptive embedding should avoid them, which can be arranged by making the costs very large rather than deriving them from the pixel model.

## 2.2 Stego image model

Since the cover model is a sequence of independent random variables, it makes sense to embed with a mutually independent embedding operation.<sup>10</sup> In the pentary setup, the pixel values are changed by at most  $\pm 2$ . Thus, given the vector representing the cover image  $\mathbf{x} = (x_1, \dots, x_N)$ , the stego image  $\mathbf{y} = (y_1, \dots, y_N)$  is obtained via the following random process:

$$\begin{aligned} \mathbb{P}(y_n = x_n + 1) &= \beta_n, & \mathbb{P}(y_n = x_n + 2) &= \theta_n, & \mathbb{P}(y_n = x_n) &= 1 - 2\beta_n - 2\theta_n, \\ \mathbb{P}(y_n = x_n - 1) &= \beta_n, & \mathbb{P}(y_n = x_n - 2) &= \theta_n, \end{aligned} \quad (3)$$

with  $0 \leq \beta_n, 0 \leq \theta_n, \beta_n + \theta_n \leq 2/5$  being the so-called change rates. Thus, the pmf of the stego pixels is a mixture  $\mathcal{P}_{(\beta_n, \theta_n)} = (p_{(\beta_n, \theta_n)}(k; \sigma_n, \nu))_{k \in \mathbb{Z}}$  with  $\mathbb{P}(y_n = k) = p_{(\beta_n, \theta_n)}(k; \sigma_n, \nu)$

$$\begin{aligned} p_{(\beta_n, \theta_n)}(k; \sigma_n, \nu) &= (1 - 2\beta_n - 2\theta_n)p_0(k; \sigma_n, \nu) + \beta_n p_0(k + 1; \sigma_n, \nu) + \beta_n p_0(k - 1; \sigma_n, \nu) \\ &\quad + \theta_n p_0(k + 2; \sigma_n, \nu) + \theta_n p_0(k - 2; \sigma_n, \nu). \end{aligned} \quad (4)$$

Thus, Alice can embed a payload of  $R$  nats given by

$$R(\boldsymbol{\beta}, \boldsymbol{\theta}) = \sum_{n=1}^N H(\beta_n, \theta_n), \quad (5)$$

where  $H(x, y) = -2x \log x - 2y \log y - (1 - 2x - 2y) \log(1 - 2x - 2y)$  is the pentary entropy function expressed in nats (“log” is the natural log). Equation (5) shows the highest payload that can be theoretically embedded using the mutually independent embedding operation(3) if Alice used an optimal coding scheme.

In practice, Alice embeds her payload while minimizing the following additive distortion function

$$D(\mathbf{x}, \mathbf{y}) = 2 \sum_{n=1}^N \left( \rho_n^{(1)} [x_n = y_n \pm 1] + \rho_n^{(2)} [x_n = y_n \pm 2] \right), \quad (6)$$

where  $[P]$  is the Iverson bracket defined as  $[P] = 1$  when the statement  $P$  is true and 0 otherwise, and  $\rho_n^{(l)} \geq 0$  are the costs of changing pixel  $x_n$  by  $\pm l$ ,  $l = 1, 2$ , related to  $\beta_n$  and  $\theta_n$  correspondingly via the relationships

$$\beta_n = \frac{e^{-\lambda \rho_n^{(1)}}}{1 + 2e^{-\lambda \rho_n^{(1)}} + 2e^{-\lambda \rho_n^{(2)}}}, \quad \theta_n = \frac{e^{-\lambda \rho_n^{(2)}}}{1 + 2e^{-\lambda \rho_n^{(1)}} + 2e^{-\lambda \rho_n^{(2)}}}. \quad (7)$$

The expected value of the distortion is  $\mathbb{E}[D] = 2 \sum_{n=1}^N \rho_n^{(1)} \beta_n + \rho_n^{(2)} \theta_n$ . The constant  $\lambda > 0$  is determined from the payload constraint (5). In practice, the actual embedding is typically realized using a specific coding scheme, which will introduce suboptimality – either a given payload will be embedded with a slightly larger distortion than  $\mathbb{E}[D]$  or one will embed a slightly smaller payload,  $\tilde{R}(\boldsymbol{\beta}) < R(\boldsymbol{\beta})$ , when bounding the distortion by  $\mathbb{E}[D]$ . The difference  $R(\boldsymbol{\beta}) - \tilde{R}(\boldsymbol{\beta}) > 0$  is called the coding loss. With the use of advanced coding, such as the syndrome-trellis codes,<sup>9</sup> the coding loss can be made arbitrarily small at the expense of computational complexity. This is why in conceptual studies, such as the present text, instead of actual embedding schemes, we use embedding simulators that execute the embedding changes with the probabilities  $\beta_n$  and  $\theta_n$  disregarding thus the coding loss for simplicity.

## 3. OPTIMAL DETECTOR

When analyzing the image  $\mathbf{x} = (x_1, \dots, x_N)$ , the Warden’s goal is to decide between the following hypotheses:  $\forall n \in \{1, \dots, N\}$ :

$$\begin{aligned} \mathcal{H}_0 &: x_n \sim \mathcal{P}_0, \quad \forall n, \\ \mathcal{H}_1 &: x_n \sim \mathcal{P}_{(\beta_n, \theta_n)}, \quad \forall n. \end{aligned} \quad (8)$$

With the Neyman–Pearson criterion of optimality,<sup>19</sup> the Warden seeks a mapping  $\delta : \mathbb{Z}^N \rightarrow \{\mathcal{H}_0, \mathcal{H}_1\}$  that satisfies a given bound on the false-alarm probability  $\alpha_0 = \mathbb{P}(\delta(\mathbf{x}) = \mathcal{H}_1 | \mathcal{H}_0)$  while maximizing the power function  $\pi = \mathbb{P}(\delta(\mathbf{x}) = \mathcal{H}_1 | \mathcal{H}_1)$ , the correct detection probability.

We will assume that the standard deviations of all pixels  $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N)$  are known to the Warden because the embedding changes are subtle and will not change the estimated values much. Furthermore, we will assume the case of the so-called omniscient Warden who knows the change rates  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_N)$  and  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$  used by the sender. This assumption deserves more discussion because it implicitly assumes that the Warden knows the payload size and that the change rates are not sensitive to the embedding changes themselves. Both assumptions can be considered as the worst case scenario for the sender. Such conservative assumptions are commonly made in security studies. Furthermore, as the prior work<sup>7,23</sup> shows, the accuracy of steganalysis detectors built from features that consider the embedding change probabilities do not require a precise knowledge of the payload and their performance degrades gracefully when the Warden misjudges the payload size. Finally, for most modern spatial-domain steganographic algorithms the costs are rather insensitive to the embedding changes. We thus summarize that, under our assumptions listed above, the hypothesis testing problem (8) is reduced to a test between two simple hypotheses.

It follows from the Neyman–Pearson Lemma, see Theorem 3.2.1 in Ref. [19], that the most powerful test that maximizes the power for a prescribed false-alarm probability is the Likelihood Ratio Test (LRT) given, from the statistical independence of pixels, by the following decision rule:

$$\Lambda(\mathbf{x}, \boldsymbol{\sigma}, \nu) = \sum_{n=1}^N \Lambda_n(\mathbf{x}, \boldsymbol{\sigma}, \nu) = \sum_{n=1}^N \log \left( \frac{p(\beta_n, \theta_n)(x_n; \sigma_n, \nu)}{p_0(x_n; \sigma_n, \nu)} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau, \quad (9)$$

where  $\tau$  is a decision threshold determined by a prescribed false-alarm rate.

For a large number of pixels,  $N \rightarrow \infty$ , small payload (small and known change rates  $\beta_n, \theta_n$ ), and the fine quantization limit, the normalized LRT is asymptotically Gaussian due to Lindeber’s version of the Central Limit Theorem:

$$\Lambda(\mathbf{x}, \boldsymbol{\sigma}, \nu) = \frac{\sum_{n=1}^N \Lambda_n - E_{\mathcal{H}_0}[\Lambda_n]}{\sqrt{\sum_{n=1}^N \text{Var}_{\mathcal{H}_0}[\Lambda_n]}} \rightsquigarrow \begin{cases} \mathcal{N}(0, 1) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(\varrho, 1) & \text{under } \mathcal{H}_1, \end{cases} \quad (10)$$

where  $\rightsquigarrow$  denotes the convergence in distribution,  $\varrho^2$  is the deflection coefficient

$$\varrho^2 = \sum_{n=1}^N (\beta_n, \theta_n) \begin{pmatrix} I_n^{(11)} & I_n^{(12)} \\ I_n^{(21)} & I_n^{(22)} \end{pmatrix} \begin{pmatrix} \beta_n \\ \theta_n \end{pmatrix}, \quad (11)$$

and

$$I_n^{(11)} = -4 + \frac{\nu}{2b_n\Gamma(1/\nu)} \int_{\mathbb{R}} \exp(|x/b_n|^\nu) \{ \exp(-|(x-1)/b_n|^\nu) + \exp(-|(x+1)/b_n|^\nu) \}^2 dx \quad (12)$$

$$I_n^{(22)} = -4 + \frac{\nu}{2b_n\Gamma(1/\nu)} \int_{\mathbb{R}} \exp(|x/b_n|^\nu) \{ \exp(-|(x-2)/b_n|^\nu) + \exp(-|(x+2)/b_n|^\nu) \}^2 dx \quad (13)$$

$$I_n^{(12)} = I_n^{(21)} = -4 + \frac{\nu}{2b_n\Gamma(1/\nu)} \int_{\mathbb{R}} \exp(|x/b_n|^\nu) (\exp(-|(x-1)/b_n|^\nu) + \exp(-|(x+1)/b_n|^\nu)) \\ \times (\exp(-|(x-2)/b_n|^\nu) + \exp(-|(x+2)/b_n|^\nu)) dx \quad (14)$$

are the elements of the  $2 \times 2$  Fisher information matrix for pentary embedding in GG noise (see Appendix A for the proof).

It follows from the limiting distribution of the LR under  $\mathcal{H}_0$  (10) that for any false-alarm probability  $\alpha_0 \in (0, 1)$  the decision threshold  $\tau$  given by  $\tau = \Phi^{-1}(1 - \alpha_0)$ , where  $\Phi$  and  $\Phi^{-1}$  denote the cumulative distribution function (cdf) of the standard Gaussian distribution and its inverse, asymptotically as  $N \rightarrow \infty$ , guarantees that the false-alarm probability of the LRT does not exceed  $\alpha_0$ . It also follows from the limiting distribution (10) that the power  $\pi = \pi(\varrho)$  of the LRT is given by  $\pi(\varrho) = 1 - \Phi(\tau - \varrho) = 1 - \Phi(\Phi^{-1}(1 - \alpha_0) - \varrho)$ .

#### 4. OBTAINING THE CHANGE RATES AND PIXEL COSTS

Having derived the deflection coefficient for pentary embedding within the MVGG cover model, the sender will minimize the detectability when she selects the change rates  $\beta_n$  and  $\theta_n$  that minimize the deflection coefficient  $\varrho^2$  (11) under the payload constraint (5). This can be easily established using the method of Lagrange multipliers:

$$\frac{\partial}{\partial \beta_n} \left( \frac{1}{2} \sum_{i=1}^N I_n^{(11)} \beta_i^2 + 2I_n^{(12)} \beta_i \theta_i + I_n^{(22)} \theta_i^2 - \frac{1}{\lambda} \left[ \sum_{i=1}^N H(\beta_i, \theta_i) - \alpha N \right] \right) = 0, \quad (15)$$

$$\frac{\partial}{\partial \theta_n} \left( \frac{1}{2} \sum_{i=1}^N I_n^{(11)} \beta_i^2 + 2I_n^{(12)} \beta_i \theta_i + I_n^{(22)} \theta_i^2 - \frac{1}{\lambda} \left[ \sum_{i=1}^N H(\beta_i, \theta_i) - \alpha N \right] \right) = 0. \quad (16)$$

Considering the derivative of the pentary entropy  $\partial H(x, y)/\partial x = 2 \log((1 - 2x - 2y)/x)$ , equations (15) and (16) together with the payload constraint form a set of  $2N + 1$  equations for  $2N + 1$  unknowns formed by  $N$  change rates  $\beta_n$ ,  $N$  change rates  $\theta_n$ , and the Lagrange multiplier  $\lambda$ :

$$I_n^{(11)} \beta_n + I_n^{(12)} \theta_n = \frac{1}{\lambda} \log \frac{1 - 2\beta_n - 2\theta_n}{\beta_n}, \quad n = 1, \dots, N, \quad (17)$$

$$I_n^{(12)} \beta_n + I_n^{(22)} \theta_n = \frac{1}{\lambda} \log \frac{1 - 2\beta_n - 2\theta_n}{\theta_n}, \quad n = 1, \dots, N,$$

$$R(\boldsymbol{\beta}, \boldsymbol{\theta}) = \sum_{n=1}^N H(\beta_n, \theta_n). \quad (18)$$

Equations 17–18 can be solved quickly for the whole image using a parallelized Newton method over all pixels and a simple binary search in order to find  $\lambda$  satisfying the payload constraint. When implementing the embedding using syndrome-trellis codes, we need the corresponding pixel costs of changing pixel  $n$  by  $\pm 1$  and  $\pm 2$ ,  $\rho_n^{(1)}$  and  $\rho_n^{(2)}$ , respectively. These can be obtained by solving equations (7) for  $\rho_n^{(1)}$  and  $\rho_n^{(2)}$ .

At this point, we would like to stress that the proposed embedding scheme fundamentally differs from all previously proposed content-adaptive steganographic algorithms because the pixel cost is not derived by quantifying the impact of a specific embedding change on the local pixel neighborhood as in, e.g., WOW,<sup>15</sup> UNIWARD,<sup>16</sup> and HILL<sup>21</sup> schemes, or by quantifying its impact in an ad hoc feature space as in HUGO,<sup>22</sup> but directly from the impact of making an embedding change on the statistical detectability for the chosen *cover model*. Indeed, prior to embedding we first estimate cover model, the local variance at each cover pixel, and then the costs of changing the pixel by various amounts are derived *from the model itself*. Interestingly, this also makes our costs naturally dependent on the payload.

#### 5. VARIANCE ESTIMATOR

It is very important to realize that current best adaptive steganography “hides the embedding changes behind content” (behind the modeling error) in the sense that it capitalizes on the fact that the Warden cannot separate the content from the indeterministic noise in highly textured regions. Indeed, state-of-the-art detectors of adaptive embedding<sup>2,7,11,23</sup> use feature representations of images based on joint distributions of quantized noise residuals extracted using a diverse set of pixel predictors. As long as the Warden stays within this established framework, the “best” variance estimator for the proposed scheme should avoid rejecting the content too much or too little. Because the modeling error is rather hard to grasp analytically, we give our variance estimator a modular structure whose parameters are determined experimentally to minimize the detection using current best empirical detectors.

Assuming the cover image is an 8-bit grayscale with the original pixel values  $\mathbf{z} = (z_1, \dots, z_N)$ ,  $z_n \in \{0, \dots, 255\}$ , we first compute a noise residual using a simple denoising filter and then model the remaining content in the residual using a block-wise Least Square Estimation (LSE):<sup>17</sup>

1. Using a denoising filter  $F$ , compute the noise residual  $\mathbf{r} = \mathbf{z} - F(\mathbf{z})$ .
2. For each pixel  $n$ , execute Steps 3–5.
3. Model the remainder of the content in  $\mathbf{r}$  using a linear parametric model:  $\mathbf{r}_n = \mathbf{G}\mathbf{a}_n + \boldsymbol{\xi}_n$ . Here,  $\mathbf{r}_n$  are the values of the residual  $\mathbf{r}$  inside the  $p \times p$  neighborhood of pixel  $n$  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,  $\mathbf{a}_n$  is a vector of  $q \times 1$  of parameters, and  $\boldsymbol{\xi}_n$  is a  $p^2 \times 1$  noise term.
4. Using least square estimation,  $\hat{\mathbf{a}}_n = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{r}_n$  and  $\hat{\mathbf{r}}_n = \mathbf{G}\hat{\mathbf{a}}_n = \mathbf{G}(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{r}_n$ .
5. Assuming the pixels within the  $n$ th  $p \times p$  block have the same or similar variance, the least-square estimation of the variance of the  $n$ th pixel (the central pixel in the  $n$ -th block) is:

$$\hat{\sigma}_n^2 = \frac{\|\mathbf{r}_n - \hat{\mathbf{r}}_n\|^2}{p^2 - q}. \quad (19)$$

We would like to emphasize that Steps 3–5 are applied “pixelwise” instead of blockwise, meaning the estimated value of the variance is attributed only to the central pixel of the considered block (the  $n$ th pixel in the image). The block is then translated by one pixel to obtain the variance estimates for the neighboring pixels, etc. To obtain the estimates for pixels at the image boundary, we apply mirror-padding. We also put a lower bound on the estimated variance for numerical stability:

$$\hat{\sigma}_n^2 = \max\{0.01, \hat{\sigma}_n^2\} \quad (20)$$

The modular structure of the variance estimator allows many combinations of denoising filters and local parametric models. In this paper, we use the two-dimensional Wiener filter for the first denoising step (implemented in Matlab as `wiener2(X, [w w])`) due to its overall good performance and a low computational cost. Inspired by the prior work using local parametric models in steganalysis,<sup>3,4,6</sup> we used two-dimensional (discrete) trigonometric polynomial functions for the matrix  $\mathbf{G}$  similar to those used in the two-dimensional DCT:

$$\mathbf{G} = \left( \mathbf{1}, \cos(\mathbf{u}), \cos(\mathbf{v}), \cos(\mathbf{u}) \cdot \cos(\mathbf{v}), \cos(2\mathbf{u}), \cos(2\mathbf{v}), \cos(2\mathbf{u}) \cdot \cos(2\mathbf{v}), \dots, \cos(l\mathbf{u}), \cos(l\mathbf{v}) \right), \quad (21)$$

where  $\mathbf{u} \cdot \mathbf{v}$  represents the elementwise product,  $\mathbf{1} \in \mathbb{R}^{p^2}$  is a column vector of ones,  $\mathbf{u} \in \mathbb{R}^{p^2}$  and  $\mathbf{v} \in \mathbb{R}^{p^2}$  are obtained by putting into column vectors the following matrices ( $\mathbf{U} = \mathbf{V}^T$ ):

$$\mathbf{U} = \begin{pmatrix} \frac{\pi}{2p} & \frac{3\pi}{2p} & \dots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p} \\ \frac{\pi}{2p} & \frac{3\pi}{2p} & \dots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p} \\ \frac{\pi}{2p} & \frac{3\pi}{2p} & \dots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p} \\ \frac{\pi}{2p} & \frac{3\pi}{2p} & \dots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \frac{\pi}{2p} & \frac{3\pi}{2p} & \dots & \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-1)}{2p} \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \frac{\pi}{2p} & \frac{\pi}{2p} & \dots & \frac{\pi}{2p} \\ \frac{3\pi}{2p} & \frac{3\pi}{2p} & \dots & \frac{3\pi}{2p} \\ \frac{\pi}{2p} & \frac{\pi}{2p} & \dots & \frac{\pi}{2p} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\pi(2p-3)}{2p} & \frac{\pi(2p-3)}{2p} & \dots & \frac{\pi(2p-3)}{2p} \\ \frac{\pi(2p-1)}{2p} & \frac{\pi(2p-1)}{2p} & \dots & \frac{\pi(2p-1)}{2p} \end{pmatrix}. \quad (22)$$

Note that here  $l$  represents the degree of the two-dimensional cosine polynomial, which is related to the number of parameters (columns or basis vectors) in matrix  $\mathbf{G}$  by:  $q = (l+1)(l+2)/2$ . By changing the block size of the Wiener filter  $w$ , the neighborhood size  $p$ , and the polynomial degree  $l$ , we can change the estimator and thus the properties of the proposed embedding scheme. We determined these parameters experimentally by inspecting the testing error on the BOSSbase 1.01 database<sup>1</sup> embedded with payload 0.5 bits per pixel (bpp) when steganalyzing with the FLD ensemble<sup>18</sup> and the maxSRMd2 features.<sup>7</sup> Without providing the full details of our search, we merely state that the best estimator we found did not depend on the shape parameter  $\nu$  and corresponded to  $w = 2$ ,  $p = 9$ , and  $l = 8$  ( $q = 45$ ).

In Figure 1, we show the total embedding change probability (the sum of probabilities of changes by  $\pm 1$  or  $\pm 2$ ) and the actual embedding changes for the pentary versions of HILL, S-UNIWARD (see the description in

the next section), and the proposed scheme with  $\nu = 1, 2$  and the variance estimator described in this section. Out of the four embedding algorithms, HILL has the highest change rate of  $\pm 2$  changes while the MVGG with  $\nu = 2$  has the lowest. Also notice the differences in the selection channel between the MVGG Laplacian and Gaussian models.

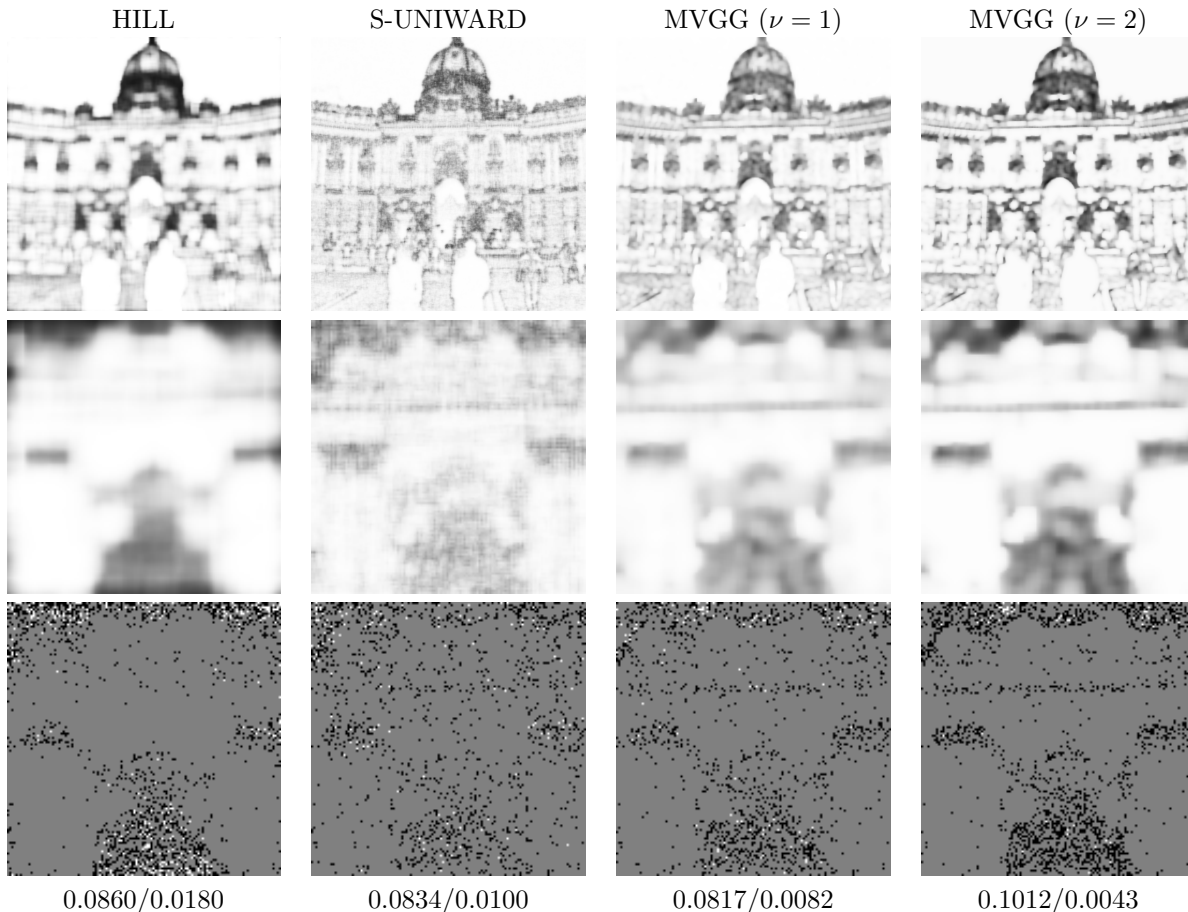


Figure 1. First and second row: total embedding change probability in BOSSbase image '1013.pgm' (and its  $128 \times 128$  crop) for payload 0.5 bpp using pentary versions of HILL, S-UNIWARD, and MVGG with  $\nu = 1$  and  $\nu = 2$ . The probabilities were scaled so that white corresponds to embedding probability zero while black corresponds to the maximal total change probability of  $4/5$ . Third row: the actual changed pixels. Pixels changed by  $\pm 1$  and  $\pm 2$  are in black and white, medium gray corresponds to no changes. The last row shows the fraction of pixels changed by  $\pm 1/\pm 2$ .

We note that once the pixel variances have been estimated, we can proceed with computing the Fisher information matrix (12)–(14) for each pixel. Before computing the change rates in Section 4, however, we smooth each term of the matrix using a  $7 \times 7$  averaging filter. This step was introduced based on the experimental evidence reported in Refs. [20, 21] that smoothing the pixel costs increases the empirical security. This can be understood intuitively as smoother costs lead to more even change rates and thus higher entropy (higher payload). Here, we are capitalizing on the fact that current steganalysis is unable to detect embedding changes in sufficiently textured areas. The averaging also spills high embedding costs into the neighboring pixels, which makes the embedding more conservative. Since in our framework we do not work with costs, in order not to depart from our embedding paradigm, we apply the smoothing to all four Fisher information terms, which has a very similar effect because the Fisher information is strongly positively correlated with costs. We admit that the smoothing is rather artificial and cannot be justified from our adopted model. We hypothesize that, as steganalysis improves, the smoothing will not be necessary and may, in fact, decrease the security.

## 6. EXPERIMENTS

All experiments in this paper were carried out on BOSSbase 1.01.<sup>1</sup> The detectors were trained as binary classifiers implemented using the FLD ensemble<sup>18</sup> with default settings. A separate classifier was trained for each embedding algorithm, and payload to see the security across different payloads. The ensemble by default minimizes the total classification error probability under equal priors  $P_E = \min_{P_{FA}} \frac{1}{2}(P_{FA} + P_{MD})$ , where  $P_{FA}$  and  $P_{MD}$  are the false-alarm and missed-detection probabilities. We evaluate the security by averaging  $P_E$  measured on the testing set over ten 5000/5000 database splits and denote it as  $\bar{P}_E$ .

In our first set of experiments, we investigated the effect of the shape parameter of the generalized Gaussian model on security when using the Spatial Rich Model (SRM) and its selection-channel-aware version maxSRMd2. Figure 2 shows the effect of the shape parameter  $\nu$  on empirical security. Although changing the parameter  $\nu$  from 1 to 2 considerably changes the corresponding embedding change probabilities (see Figure 1) as well as the change rate of  $\pm 2$  changes, which is almost twice as large for  $\nu = 1$  than for  $\nu = 2$ , the detection error  $\bar{P}_E$  stays almost constant for both feature sets in Figure 2. This rather intriguing result indicates the inability of current state-of-the-art steganalysis to distinguish different change patterns in highly textured areas of the image. Considering the results of the experiment shown in Figure 2, we set  $\nu = 2$  in our MVGG model for all remaining experiments.

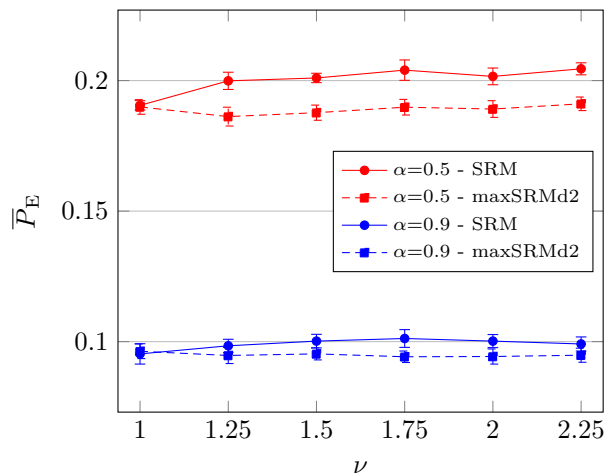


Figure 2. Average detection error  $\bar{P}_E$  of MVGG as a function the shape parameter  $\nu$  using SRM and maxSRMd2 features for two different payloads.

Our next experiment contrasts the performance of ternary and pentary versions of our model based approach with HILL and S-UNIWARD using two different feature sets. The pentary version of HILL and S-UNIWARD used the costs of  $\pm 2$  changes defined in the same way as the costs of  $\pm 1$  changes. In particular,  $\rho_n^{(\pm 2)} = D(\mathbf{x}, x_n \pm 2 \mathbf{x}_{\sim n})$ , where  $D$  is the distortion function of the corresponding embedding algorithm and  $x_n \pm 2 \mathbf{x}_{\sim n}$  denotes the cover image in which only the  $n$ th pixel was modified by  $\pm 2$  while the remaining pixels were kept unmodified.

Figure 3 summarizes the results. Pentary embedding outperforms ternary and its effect becomes more pronounced with increasing payload ( $\approx 1.5\%$  for  $\alpha = 1$ ). Another interesting result is the drop in the performance of pentary HILL for small payloads using maxSRMd2, which is probably due to the fact that it executes too many  $\pm 2$  changes for small payloads.

Figure 4 compares the performance of pentary versions of three embedding schemes. Using the SRM feature set, HILL exhibits a better performance for small payloads ( $\alpha \leq 0.5$  bpp) and then merges with MVGG in the large-payload range ( $\alpha > 0.5$  bpp). MVGG outperforms both HILL and S-UNIWARD using the selection-channel-aware maxSRMd2 for all payloads.

In all our previous experiments, we used the maxSRMd2 feature set with the exact knowledge of the embedded payload ( $\hat{\alpha} = \alpha$ ). In Ref. [7], the authors showed that by deviating from the exact knowledge of the embedding



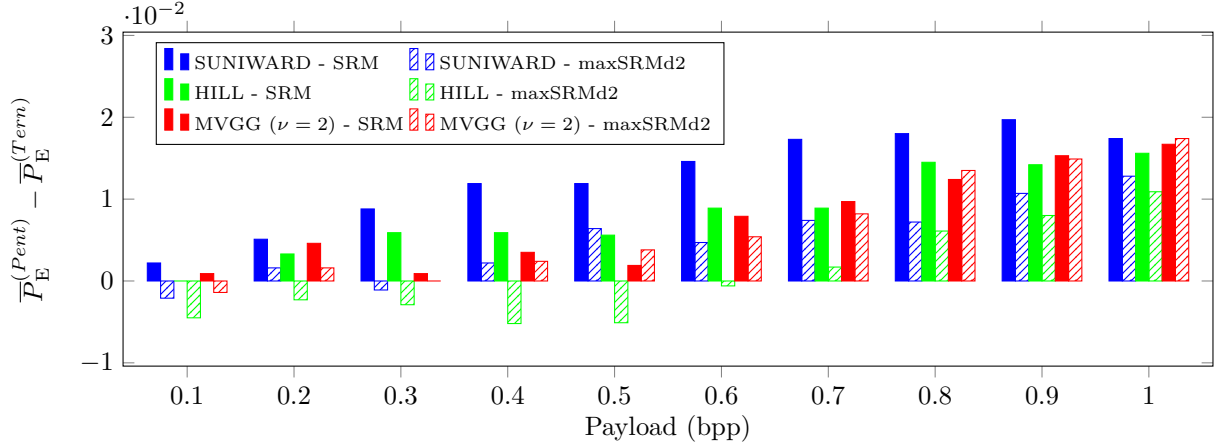


Figure 3. Average difference in detection error  $\bar{P}_E$  between pentary and ternary embedding as a function of payload for S-UNIWARD, HILL, and MVGG ( $\nu = 2$ ) using SRM and maxSRMd2 features.

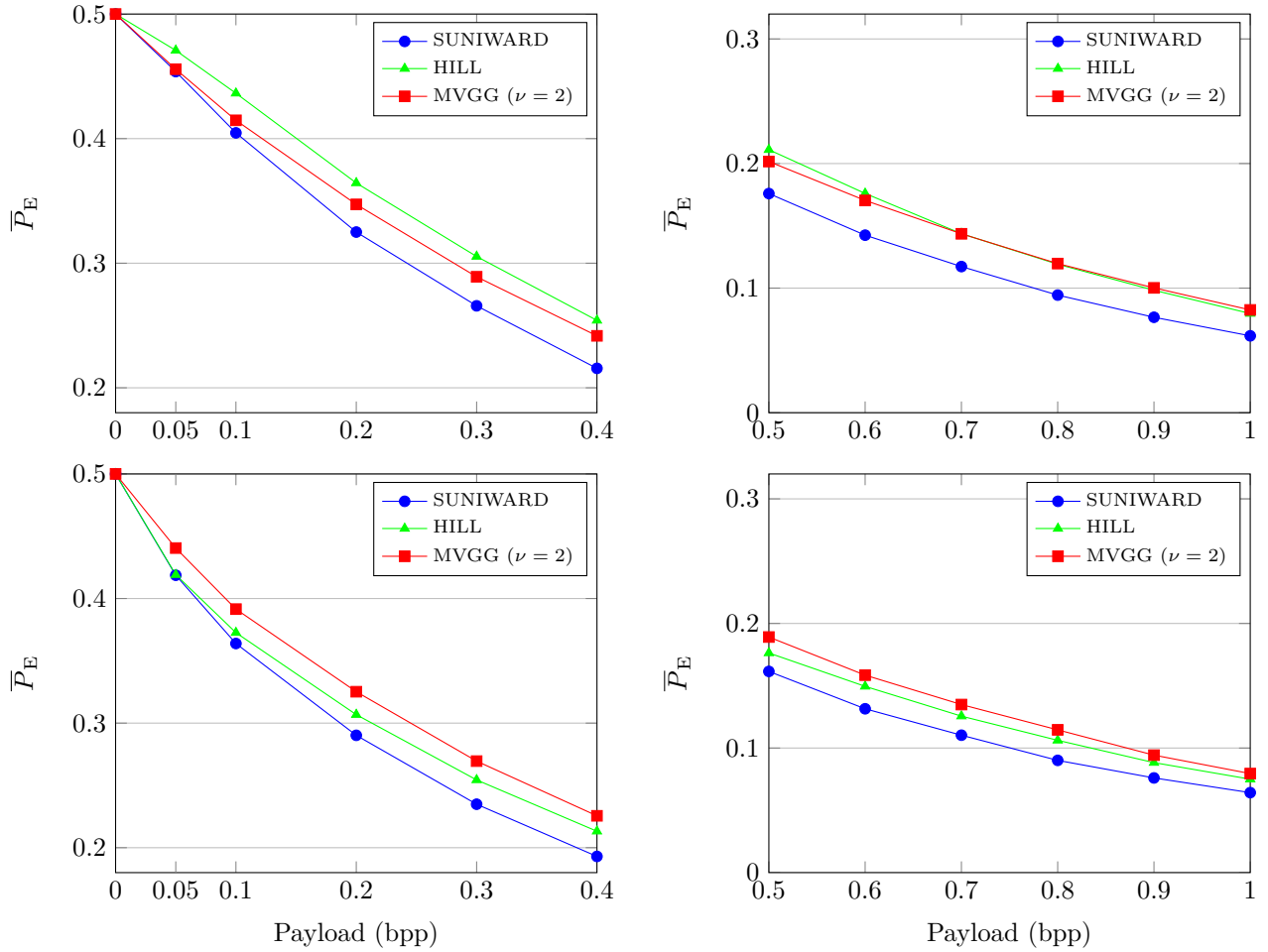


Figure 4. Average detection error  $\bar{P}_E$  for pentary versions of S-UNIWARD, HILL, and MVGG ( $\nu = 2$ ) using SRM (top) and maxSRMd2 features (bottom).

payload and setting  $\hat{\alpha} = 0.1$  bpp for extracting the features, the average performance drop in the small payload range ( $\alpha \leq 0.5$  bpp) is below 1%. Since we are using a broader payload range ( $\alpha \leq 1$  bpp), in our last experiment

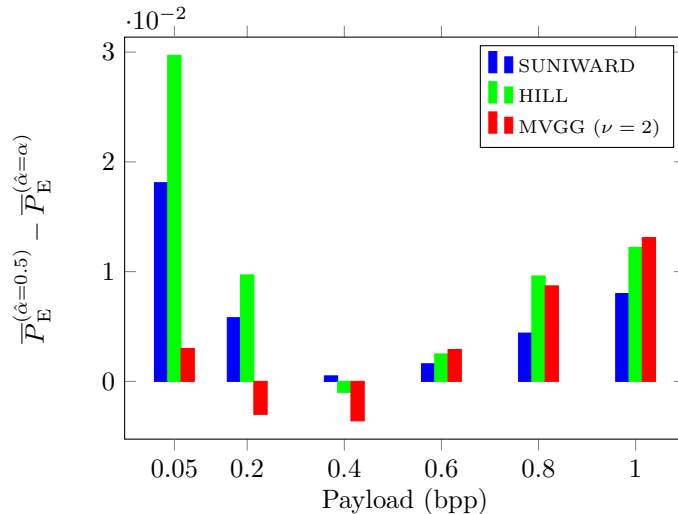


Figure 5. Drop in detection accuracy  $\bar{P}_E$  when steganalyzing with maxSRMd2 with a fixed payload  $\hat{\alpha} = 0.5$  bpp and the real payload  $\hat{\alpha} = \alpha$  bpp for S-UNIWARD, HILL, and MVGG ( $\nu = 2$ ).

we focused on assessing the loss of detection accuracy due to a mismatch between the real payload and the one used for extracting the maxSRMd2 features ( $\hat{\alpha} = 0.5$  bpp). The results summarized in Figure 5 show that the average performance drop over the entire payload range for all of the embedding schemes is about 1.5%.

## 7. CONCLUSIONS

Empirical security of steganographic schemes can likely be improved by allowing embedding changes with a larger amplitude in highly textured regions. Previously proposed schemes assign costs of embedding changes by evaluating their impact on selected noise residuals. In contrast, in this paper we introduce an alternative approach that is well theoretically founded as the embedding scheme is designed to minimize the power of the most powerful detector within a chosen cover model. The embedding in a specific cover image starts with estimating the variance of the superposition of the acquisition noise and the modeling error modeled as a sequence of independent generalized Gaussian random variables with unequal variances. By properly adjusting the variance estimator, we were able to obtain state-of-the-art results with the Gaussian model when compared to pentary versions of S-UNIWARD and HILL. By going pentary, we were able to improve statistical detectability w.r.t. the corresponding ternary versions by 1 – 2% when measured using classifiers and selection-channel-aware rich models. As expected, the improvement is larger for larger payloads.

The generalized Gaussian shape parameter can be used to adjust the balance between embedding changes by  $\pm 1$  and  $\pm 2$ . Surprisingly, however, despite the differences in the ensuing selection channels, the empirical security of the steganographic schemes constructed this way stays approximately invariant to the shape parameter. This indicates that current feature-based steganalysis is unable to distinguish rather large differences in the selection channel.

Among possible future directions, we mention the possibility to extend the proposed framework to cover models able to capture dependencies among pixels as well as embedding changes.

## 8. ACKNOWLEDGMENTS

The work on this paper was supported by Air Force Office of Scientific Research under the research grant number FA9950-12-1-0124. 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.

## APPENDIX A. ASYMPTOTIC DISTRIBUTION OF THE LRT

To derive the expression for the deflection coefficient (11)–(14), we need to compute the expectations and variances of the LRT for each pixel under both hypotheses  $l \in \{0, 1\}$ :

$$\mu_n^{(l)} = E_{\mathcal{H}_l}[\Lambda_n] = E_{\mathcal{H}_l} \left[ \log \frac{p_{(\beta_n, \theta_n)}(X_n; \sigma_n, \nu)}{p_0(X_n; \sigma_n, \nu)} \right], \quad v_n^{(l)} = \text{Var}_{\mathcal{H}_l}[\Lambda_n] = \text{Var}_{\mathcal{H}_0} \left[ \log \frac{p_{(\beta_n, \theta_n)}(X_n; \sigma_n, \nu)}{p_0(X_n; \sigma_n, \nu)} \right]. \quad (23)$$

For brevity, in this appendix we drop the dependence of the cover and stego distributions on the parameters  $\sigma_n$  and  $\nu$ . For the means

$$E_{\mathcal{H}_l} \left[ \log \frac{p_{(\beta_n, \theta_n)}(X)}{p_0(X)} \right] = \begin{cases} \sum_k p_0(k) \log \frac{p_{(\beta_n, \theta_n)}(k)}{p_0(k)} = -D_{\text{KL}}(p_0 \| p_{(\beta_n, \theta_n)}) & \text{under } \mathcal{H}_0, \\ \sum_k p_{(\beta_n, \theta_n)}(k) \log \frac{p_{(\beta_n, \theta_n)}(k)}{p_0(k)} = D_{\text{KL}}(p_{(\beta_n, \theta_n)} \| p_0) & \text{under } \mathcal{H}_1. \end{cases} \quad (24)$$

It can be easily verified that the leading term in the Taylor expansion of both KL divergence terms is the same and equal to the following quadratic form:

$$D_{\text{KL}}(p_0 \| p_{(\beta_n, \theta_n)}) \approx D_{\text{KL}}(p_{(\beta_n, \theta_n)} \| p_0) \approx \frac{1}{2} (\beta_n, \theta_n) \mathbb{I}_n (\beta_n, \theta_n)^T, \quad (25)$$

where

$$\mathbb{I}_n = \begin{pmatrix} I_n^{(11)} & I_n^{(12)} \\ I_n^{(22)} & I_n^{(21)} \end{pmatrix}, \quad (26)$$

is the  $2 \times 2$  Fisher information matrix with:

$$I_n^{(11)} = \sum_k \frac{1}{p_0(k)} \left( \frac{\partial p_{(\beta_n, \theta_n)}(k)}{\partial \beta} \right)^2, \quad I_n^{(22)} = \sum_k \frac{1}{p_0(k)} \left( \frac{\partial p_{(\beta_n, \theta_n)}(k)}{\partial \theta} \right)^2 \quad (27)$$

$$I_n^{(12)} = I_n^{(21)} = \sum_k \frac{1}{p_0(k)} \frac{\partial p_{(\beta_n, \theta_n)}(k)}{\partial \beta} \frac{\partial p_{(\beta_n, \theta_n)}(k)}{\partial \theta}, \quad (28)$$

where all partial derivatives are evaluated at  $\beta_n = \theta_n = 0$ .

For the variances,

$$v_n^{(l)} = \text{Var}_{\mathcal{H}_l} \left[ \log \frac{p_{(\beta_n, \theta_n)}(X_n)}{p_0(X_n)} \right] = E_{\mathcal{H}_l} \left[ \left( \log \frac{p_{(\beta_n, \theta_n)}(X_n)}{p_0(X_n)} \right)^2 \right] - E_{\mathcal{H}_l} \left[ \log \frac{p_{(\beta_n, \theta_n)}(X_n)}{p_0(X_n)} \right]^2. \quad (29)$$

We already know that the second term in (29) is of the fourth order in terms of the vector parameter  $(\beta_n, \theta_n)$ . The first term can be computed using the Taylor expansion for  $p_{(\beta_n, \theta_n)}$  around  $p_{(\beta_n, \theta_n)} = p_0$  and for the logarithm:

$$\log \frac{p_{(\beta_n, \theta_n)}(X_n)}{p_0(X_n)} \approx \log \left( 1 + \frac{(\beta_n, \theta_n) \cdot \nabla p_{(\beta_n, \theta_n)}(X_n)}{p_0(X_n)} \right) \approx \frac{(\beta_n, \theta_n) \cdot \nabla p_{(\beta_n, \theta_n)}(X_n)}{p_0(X_n)}, \quad (30)$$

where we denoted the dot product of two two-dimensional vectors with  $\cdot$  and the column vector of first derivatives of  $p_{(\beta_n, \theta_n)}$  w.r.t. the parameters as  $\nabla p'_{(\beta_n, \theta_n)}$ . Keeping only the leading terms, we obtain

$$v_n^{(1)} \approx E_{\mathcal{H}_l} \left[ \left( \log \frac{p_{(\beta_n, \theta_n)}(X_n)}{p_0(X_n)} \right)^2 \right] \approx E_{\mathcal{H}_l} \left[ \left( \frac{(\beta_n, \theta_n) \cdot \nabla p_{(\beta_n, \theta_n)}(X_n)}{p_0(X_n)} \right)^2 \right] \quad (31)$$

$$\approx \sum_k \frac{1}{p_0(k)} \left( (\beta_n, \theta_n) \cdot \nabla p_{(\beta_n, \theta_n)}(k) \right)^2 \quad \text{under both } \mathcal{H}_0 \text{ and } \mathcal{H}_1 \quad (32)$$

$$= \sum_k \frac{1}{p_0(k)} (\beta_n, \theta_n) \cdot \nabla p_{(\beta_n, \theta_n)}(\nabla p_{(\beta_n, \theta_n)})^T (\beta_n, \theta_n)^T \quad (33)$$

$$= (\beta_n, \theta_n) \mathbb{I}_n (\beta_n, \theta_n)^T. \quad (34)$$

Note that the variance is the same under both hypotheses,  $v_n^{(0)} = v_n^{(1)}$ . Using (24) and (34), the square of the expectation of the normalized logLR under  $\mathcal{H}_1$  (10) becomes

$$\varrho^2 = \frac{\left(\sum_{n=1}^N (\mu_n^{(1)} - \mu_n^{(0)})\right)^2}{\sum_{n=1}^N v_n^{(0)}} = \sum_{n=1}^N (\beta_n, \theta_n) \mathbb{I}_n(\beta_n, \theta_n)^T. \quad (35)$$

To find the formulas for the Fisher information matrix (27) and (28), we first work out the partial derivatives of the stego mixture (4):

$$\begin{aligned} \frac{\partial p_{(\beta_n, \theta_n)}(x)}{\partial \beta} &= -2f(x; \nu, b_n) + f(x-1; \nu, b_n) + f(x+1; \nu, b_n), \\ \frac{\partial p_{(\beta_n, \theta_n)}(x)}{\partial \theta} &= -2f(x; \nu, b_n) + f(x-2; \nu, b_n) + f(x+2; \nu, b_n), \end{aligned}$$

and substitute into (27) and (28). To further minimize the writing and clutter, we write the GG density as

$$f(x; \sigma_n, \nu) = A_n \exp(-a_n |x|^\nu), \quad (36)$$

where  $A_n = \frac{\nu}{2b_n \Gamma(1/\nu)}$  and  $a_n = 1/b_n^\nu$ . Since in our case, we assume that the quantization step is 1, the corresponding sums for  $I_n^{(11)}$  and  $I_n^{(22)}$  can be approximated with integrals ( $l \in \{1, 2\}$ ):

$$\begin{aligned} I_n^{(ll)} &\approx \int_{\mathbb{R}} \frac{A_n^2}{A_n \exp(-a_n |x|^\nu)} (-2 \exp(-a_n |x|^\nu) + \exp(-a_n |x-l|^\nu) + \exp(-a_n |x+l|^\nu))^2 dx \\ &= A_n \int_{\mathbb{R}} \exp(a_n |x|^\nu) \left\{ 4 \exp(-2a_n |x|^\nu) + \exp(-2a_n |x-l|^\nu) + \exp(-2a_n |x+l|^\nu) - \right. \\ &\quad \left. - 4 \exp(-a_n |x|^\nu - a_n |x-l|^\nu) - 4 \exp(-a_n |x|^\nu - a_n |x+l|^\nu) + 2 \exp(-a_n |x+l|^\nu - a_n |x-l|^\nu) \right\} dx \\ &= -4 + A_n \int_{\mathbb{R}} \exp(a_n |x|^\nu) \left\{ \exp(-2a_n |x-l|^\nu) + \exp(-2a_n |x+l|^\nu) + 2 \exp(-a_n |x+l|^\nu - a_n |x-l|^\nu) \right\} dx \\ &= -4 + A_n \int_{\mathbb{R}} \exp(a_n |x|^\nu) \left\{ \exp(-a_n |x-l|^\nu) + \exp(-a_n |x+l|^\nu) \right\}^2 dx. \end{aligned}$$

And for the mixed indices

$$\begin{aligned} I_n^{(12)} &\approx \int_{\mathbb{R}} \frac{A_n^2}{A_n \exp(-a_n |x|^\nu)} (-2 \exp(-a_n |x|^\nu) + \exp(-a_n |x-1|^\nu) + \exp(-a_n |x+1|^\nu)) \\ &\quad \times (-2 \exp(-a_n |x|^\nu) + \exp(-a_n |x-2|^\nu) + \exp(-a_n |x+2|^\nu)) dx \\ &= A_n \int_{\mathbb{R}} (-2 + \exp(a_n |x|^\nu - a_n |x-1|^\nu) + \exp(a_n |x|^\nu - a_n |x+1|^\nu)) \\ &\quad \times (-2 \exp(-a_n |x|^\nu) + \exp(-a_n |x-2|^\nu) + \exp(-a_n |x+2|^\nu)) dx \\ &= -4 + A_n \int_{\mathbb{R}} \exp(a_n |x|^\nu) (\exp(-a_n |x-1|^\nu) + \exp(-a_n |x+1|^\nu)) \\ &\quad \times (\exp(-a_n |x-2|^\nu) + \exp(-a_n |x+2|^\nu)) dx \end{aligned}$$

## REFERENCES

1. P. Bas, T. Filler, and T. Pevný. Break our steganographic system – the ins and outs of organizing BOSS. In T. Filler, T. Pevný, A. Ker, and S. Craver, editors, *Information Hiding, 13th International Conference*, volume 6958 of Lecture Notes in Computer Science, pages 59–70, Prague, Czech Republic, May 18–20, 2011.

2. L. Chen, Y. Q. Shi, P. Sutthiwan, and X. Niu. A novel mapping scheme for steganalysis. In Y. Q. Shi, F. Pérez-González, and H. J. Kim, editors, *Proc. 11th International Workshop on Digital-Forensics and Watermarking*, volume 7809 of Lecture Notes in Computer Science, pages 19–33, Shanghai, China, October 31–November 3, 2012. Springer-Verlag, New York.
3. R. Cogranne and F. Retraint. An asymptotically uniformly most powerful test for LSB Matching detection. *IEEE Transactions on Information Forensics and Security*, 8(3):464–476, 2013.
4. R. Cogranne and F. Retraint. Statistical detection of defects in radiographic images using an adaptive parametric model. *Signal Processing*, 96-B(3):173–189, 2014.
5. R. Cogranne, V. Sedighi, T. Denemark, and J. Fridrich. Using statistical image model for steganographic distortion function based on detectability. *IEEE Trans. on Inf. Forensics and Sec.*, 2014. Under review.
6. R. Cogranne, C. Zitzmann, L. Fillatre, F. Retraint, I. Nikiforov, and P. Cornu. A cover image model for reliable steganalysis. In T. Filler, T. Pevný, A. Ker, and S. Craver, editors, *Information Hiding, 13th International Conference*, Lecture Notes in Computer Science, pages 178–192, Prague, Czech Republic, May 18–20, 2011.
7. T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich. Selection-channel-aware rich model for steganalysis of digital images. In *Sixth IEEE International Workshop on Information Forensics and Security*, Atlanta, GA, USA, December 3–5 2014.
8. T. Filler and J. Fridrich. Steganography using Gibbs random fields. In J. Dittmann, S. Craver, and P. Campisi, editors, *Proceedings of the 12th ACM Multimedia & Security Workshop*, pages 199–212, Rome, Italy, September 9–10, 2010. ACM.
9. T. Filler, J. Judas, and J. Fridrich. Minimizing additive distortion in steganography using syndrome-trellis codes. *IEEE Transactions on Information Forensics and Security*, 6(3):920–935, September 2011.
10. T. Filler, A. D. Ker, and J. Fridrich. The Square Root Law of steganographic capacity for Markov covers. In N. D. Memon, E. J. Delp, P. W. Wong, and J. Dittmann, editors, *Proceedings SPIE, Electronic Imaging, Media Forensics and Security*, volume 7254, pages 08 1–11, San Jose, CA, January 18–21, 2009.
11. J. Fridrich and J. Kodovský. Rich models for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security*, 7(3):868–882, June 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. G. E. Healey and R. Kondepudy. Radiometric CCD camera calibration and noise estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(3):267–276, March 1994.
14. G. C. Holst. *CCD Arrays, Cameras, and Displays*. JCD Publishing & SPIE Press, USA, 2nd edition, 1998.
15. V. Holub and J. Fridrich. Designing steganographic distortion using directional filters. In *Fourth IEEE International Workshop on Information Forensics and Security*, Tenerife, Spain, December 2–5, 2012.
16. V. Holub and J. Fridrich. Universal distortion design for steganography in an arbitrary domain. *EURASIP Journal on Inf. Sec., Spec. Issue on Revised Selected Papers of 1st ACM IH&MMS Workshop*, 2014:1, 2014.
17. V. Katkovnik, K. Egiazarian, and J. Astola. *Local Approximation Techniques in Signal and Image Processing*. SPIE Press, Monograph, 2006.
18. J. Kodovský, J. Fridrich, and V. Holub. Ensemble classifiers for steganalysis of digital media. *IEEE Transactions on Information Forensics and Security*, 7(2):432–444, 2012.
19. E. L. Lehmann and J. P. Romano. *Testing Statistical Hypotheses*. Springer, 3rd edition, 2005.
20. B. Li, S. Tan, M. Wang, and J. Huang. Investigation on cost assignment in spatial image steganography. *IEEE Transactions on Information Forensics and Security*, 9(8):1264–1277, August 2014.
21. B. Li, M. Wang, and J. Huang. A new cost function for spatial image steganography. In *Proceedings IEEE, International Conference on Image Processing, ICIP*, Paris, France, October 27–30, 2014.
22. T. Pevný, T. Filler, and P. Bas. Using high-dimensional image models to perform highly undetectable steganography. In R. Böhme and R. Safavi-Naini, editors, *Information Hiding, 12th International Conference*, volume 6387 of Lecture Notes in Computer Science, pages 161–177, Calgary, Canada, June 28–30, 2010. Springer-Verlag, New York.
23. W. Tang, H. Li, W. Luo, and J. Huang. Adaptive steganalysis against WOW embedding algorithm. In A. Uhl, S. Katzenbeisser, R. Kwitt, and A. Piva, editors, *2nd ACM IH&MMSec. Workshop*, Salzburg, Austria, June 11–13, 2014.