Statistical Model of Quantized DCT Coefficients : Application in the Steganalysis of Jsteg Algorithm

Thanh Hai Thai, Rémi Cogranne, Florent Retraint

***Abstract*—The goal of this paper is to propose a statistical model of quantized Discrete Cosine Transform (DCT) coefficients. It relies on a rigorous mathematical framework of studying the image processing pipeline of a typical digital camera instead of fitting the empirical distribution with a variety of popular models proposed in the literature. To highlight the accuracy of the proposed model, the paper uses it for the detection of hidden information in JPEG images. The paper proposes and designs a statistical test for the steganalysis of Jsteg algorithm. A ML estimator for embedding rate is also derived based on the proposed model of DCT coefficients. Numerical results on a large database also emphasize the accuracy of the proposed model.**

***Index Terms*—Digital Image Model, Discrete Cosine Transform, JPEG Compression, Steganalysis, Hypothesis Testing.**

I. Introduction

DIGITAL image processing has remarkably developed during the past decades with dramatic advancement in computing and network technologies [1]. Many applications of this field involve the storage and transmission of digital images. The JPEG compression has gained widespread popularity for image storage and transmission because of its standardization and cost effectiveness. This image format has been extensively studied in various domains such as image segmentation [2], image coding [3], image restoration or reconstruction [4], pattern recognition [5], digital watermarking [6], steganography [7], and image tampering detection [8]. Such applications require the knowledge of a model of digital images in JPEG format. This paper aims to propose a novel statistical model of JPEG images and applies this model for the hidden information detection problem to emphasize its accuracy.

Thanh Hai Thai, Rmi Cogranne and Florent Retraint are with the ICD - LM2S, Troyes University of Technology (UTT), UMR 6281, CNRS, Troyes, France.

Correspondance should be addressed to remi.cogranne@utt.fr or to florent.retraint@utt.fr

Research partially funded by Troyes University of Technology (UTT) strategic program COLUMBO.

A. State of the Art

The model of Discrete Cosine Transform (DCT) coefficients has been considerably studied in the literature. Many researches focus on comparing the empirical data with a variety of popular statistical distributions by conducting the goodness-of-fit (GOF) test, e.g. the Kolmogorov-Smirnov (KS) or χ^2 test. Firstly, the Gaussian distribution for the DCT coefficients was conjectured in [1]. The Laplacian distribution was verified in [9] by performing the KS test. This Laplacian distribution remains a dominant choice in image processing because of its simplicity and relative accuracy. Other possible distributions such as Gaussian mixture [10] and Cauchy [11] were also proposed. In order to model the DCT coefficients more accurately, the previous distributions were extended to the generalized versions consisting of the Generalized Gaussian [12] and the Generalized Gamma (GΓ) [13] distributions. Two main drawbacks of those researches are that the choice of distribution for the DCT coefficients is not based on a mathematical analysis and the empirical use of GOF test on a few standard images. Thus, this can not guarantee a good fitting of the chosen model to a wide range of images, which leads to a lack of robustness of the model.

The first mathematical analysis given in [14] relied on the doubly stochastic model due to the variability of block variance. By considering that the block variance follows the exponential or the half-Gaussian distribution, the Laplacian model was finally obtained for DCT coefficients [14]. This analysis was incomplete due to the lack of mathematical justification for the block variance model. Besides, a collection of models was proposed for block variance in [15] without justification. On the contrary, it was established in our previous work [16] that the block variance can be approximately modeled by the Gamma distribution. Then, following the same framework proposed in [14], a statistical model of DCT coefficients, which outperforms the Laplacian or the GΓ model was established; see details in [16].

The JPEG image, which relies on the DCT operation, has been exploited in many applications. Recently, the JPEG format has received the most attention from law enforcement agencies and academic researchers. Several steganographic methods embed a secret message within a JPEG compressed image and create a so-called stego-image. The stego-image is

then transmitted to the receiver via an insecure channel without raising suspicion of an adversary. The first steganographic algorithm for JPEG images was Jsteg [17]. The Jsteg is based on the well-known embedding method called Least Significant Bit (LSB) replacement. It replaces the LSB of quantized DCT coefficients that differ from 0 and 1 with bits of message. Despite of its relative insecurity, the Jsteg algorithm remains popular in downloadable steganography softwares due to its simplicity and high embedding payload. However, the steganalysis of Jsteg algorithm still remains an open problem. In spite of lots of existing methods proposed for the steganalysis of Jsteg algorithm in the literature, an improvement of the existing steganalysis methods is still desirable.

Despite the fact that the secret content is not visually revealed, the modification of cover image changes its statistical properties and creates artifacts that can be detected statistically. If one possesses a model that perfectly captures statistical properties of cover image, the insertion is detectable following the information theoretic sense [18] or hypothesis testing theory [19]. Therefore, the steganalysis of Jsteg algorithm requires a very accurate model of DCT coefficients, which allows to detect any small change in the cover image due to the insertion of secret message. This model-based approach was exploited in [20] for the steganalysis of Jsteg using the Generalized Cauchy distribution of DCT coefficients. The quantized Laplacian model-based steganalysis was also presented in [21]. However, a considerable loss of power was revealed since the Laplacian model was not sufficiently accurate to model DCT coefficients. Other approaches involve structural detectors [22], [23], WS detectors [24], and universal blind detectors [25].

In an operational context, for instance a steganalysis tool for law enforcement or intelligence agencies, the design of an accurate detector might not be sufficient. The most important and challenging problem is to provide a detector with analytically predictable results in order to guarantee a prescribed false alarm probability. The existing detectors can provide overall acceptable detection performance. However, their statistical performance remains analytically unestablished in practice. It is only evaluated on a large database. Besides, as in all applications of machine learning, the main difficulties for blind detectors are the choice of appropriate feature set and the analytic establishment of detection performance. The latter remains an open problem in the framework of statistical learning [26].

B. Contributions and Organization of the Paper

The contribution of this paper is twofold:

1) The paper establishes a mathematical framework of studying statistical distribution of quantized DCT coefficients under some mild assumptions: RAW pixels are statistically independent, the pixels are identically distributed in each 8×8 block, and the correlation between DCT coefficients is negligible. The study is carried out by following the image processing pipeline of a digital camera. The estimates of the model parameters are given by the method of Maximum Likelihood (ML). Numerical experiments show that the proposed model is more accurate than the recent empirical ones in the literature (Laplacian and Generalized Gamma model).
2) To highlight the efficiency of proposed model, the paper exploits it for the steganalysis of Jsteg algorithm. By formulating the hidden information detection as a hypothesis testing problem, the paper designs a most powerful Likelihood Ratio Test (LRT) assuming that all model parameters are known. The statistical performance of the LRT is analytically established. The test satisfies the Constant False Alarm Rate (CFAR) property, i.e. the threshold is set independently of the image content, and maximizes the detection probability. A ML estimator for embedding rate is also derived based on the proposed model of DCT coefficients.

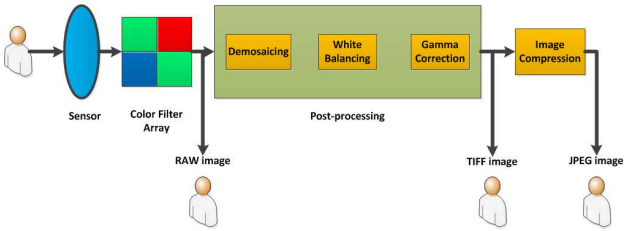

Fig. 1: Image processing pipeline of a digital camera.

The paper is organized as follows. Section II presents the image processing pipeline of a digital camera. Section III proposes the statistical distribution of quantized DCT coefficients. Section IV provides ML estimates of model parameters. Section V applies the proposed model in the steganalysis of Jsteg algorithm. Section VI presents numerical results of the comparison between the proposed model and the state-of-the-art GΓ model and the popular Laplacian model based on the χ^2 GOF test. The performance obtained from the LRT based on these three models to detect hidden messages using the Jsteg algorithm is also studied on a large database of JPEG images. Finally, Section VII concludes the paper.

II. Image Processing Pipeline of a Digital Camera

The image processing pipeline of a digital camera is shown in Fig. 1. The image processing pipeline involves the steps by which an image is rendered from the measured light intensity of each pixel. Each stage affects the final output image. It should be noted that the sequence of operations differs from manufacturer to manufacturer. The reader is referred to [27], [28] for the general structure of a digital camera and to [29] for the image processing pipeline.

Usually, a digital camera records an image by using the photosites of an image sensor. These photosites enable to convert light energy to electrical energy. The output signals of the image sensor are analog. These signals are then converted to digital signals by an analog-to-digital (A/D) converter inside the camera. The RAW image is obtained at this stage. Depending on the analog-to-digital circuit of the camera, the RAW image is recorded with 12, 14 or even 16 bits. One key advantage is that the RAW image contains exactly information

recorded by the image sensor and it has not yet undergone post-acquisition operations. This offers more flexibility for further adjustments.

Since the photosites are insensitive to color, the digital camera samples the color spectrum using the Color Filter Array (CFA) such that each pixel samples only one color band usually red, green or blue. Although the use of the CFA allows to reduce the cost of the camera, this requires to estimate the missing color values at each pixel location in order to render a full-color image. This estimation process is commonly referred as CFA demosaicing; see [30] for a review on some demosaicing methods. Among available demosaicing methods, the bilinear interpolation might be the simplest and most computationally efficient one. It estimates missing color values with weighted averages of their neighboring values. Let $\mathbf{Z}$ be a matrix representing the RAW image of size $M \times N$ whereas R, G and B denote respectively the red, green and blue channels of the image. A sub-image of color channel $c = \{R, G, B\}$ extracted from the RAW image $\mathbf{Z}$ is denoted by $\mathbf{Z}^c$. The bilinear interpolation can be written as a linear filtering

$$\mathbf{Z}^c_{DM} = \mathbf{H}^c \circledast \mathbf{Z}^c, \tag{1}$$

where $\mathbf{H}^c$ is the linear filter for the color channel c, $\circledast$ denotes the 2-D convolution and $\mathbf{Z}^c_{DM}$ denotes the demosaiced image of the color channel c.

Besides, the RAW image requires to undergo the white balance process [29]. The white balance aims to compensate the color shifts caused by different color temperatures of light sources so that a captured white object is rendered white in the image. It is assumed that the white balance is implemented after the demosaicing process. One of popular white balance algorithms is the Gray World [29]. This algorithm relies on the assumption that the average value of three color channels will average to a common gray value

$$\overline{\theta}^R_{DM} = \overline{\theta}^G_{DM} = \overline{\theta}^B_{DM}, \tag{2}$$

where $\overline{\theta}^c_{DM}$ denotes the average intensity of the demosaiced image $\mathbf{Z}^c_{DM}$. In this scheme, the white-balanced image of color channel c, denoted $\mathbf{Z}^c_{WB}$, is given by

$$\mathbf{Z}^c_{WB} = \lambda^c \times \mathbf{Z}^c_{DM} \quad \text{with} \quad \lambda^c = \frac{\overline{\theta}^G_{DM}}{\overline{\theta}^c_{DM}}. \tag{3}$$

It is worth noting that other white-balancing algorithms may be also modeled as a weighting of color channels but using different weight factors. Another fundamental post-acquisition process is γ-correction [29], which involves a non-linear input-output mapping. This process is necessary for contrast display purposes. It is defined by the following power-law expression

$$\mathbf{Z}^c_{GM} = |\mathbf{Z}^c_{WB}|^{\frac{1}{\gamma}}, \tag{4}$$

where $|\cdot|$ denotes the absolute value and γ is the correction factor (typically, $\gamma = 2.2$). After going through these post-acquisition processes, a full-color image, referred as TIFF image in this paper, with higher quality is rendered.

In order to be stored or transmitted easily on telecommunication networks, the image has to be compressed to reduce its size. The JPEG standard [31] is the most popular compression technology in digital computing. The use of the JPEG compression is a balancing act between storage size and image quality. An image which is compressed with a high degree of compression requires little storage space, but it will probably be reconstructed with a poor quality.

The JPEG compression scheme works in the different color space, typically $YCbCr$ color space, rather than the RGB color space. The Y channel represents the brightness of a pixel, and the Cb and Cr channels represent the chrominance. Therefore, prior to the JPEG compression, the TIFF image is converted from the RGB color space into the $YCbCr$ color space using a linear transformation. In the JPEG compression scheme, each channel Y, Cb and Cr is processed separately. The index of the channel Y, Cb, Cr can be omitted for the sake of simplicity. Let $\widetilde{\mathbf{Z}}$ be a matrix representing the sub-image of the channel Y, Cb, Cr. The JPEG compression involves two key steps : the DCT and the quantization. The DCT operation is performed in each 8×8 block of $\widetilde{\mathbf{Z}}$ as follows

$$I_{p,q} = \frac{1}{4} T_p T_q \sum_{m=0}^{7} \sum_{n=0}^{7} \tilde{z}_{m,n} \times \cos\left(\frac{(2m+1)p\pi}{16}\right) \cos\left(\frac{(2n+1)q\pi}{16}\right), \tag{5}$$

where $\tilde{z}_{m,n}$ denotes a pixel within a 8×8 block of $\widetilde{\mathbf{Z}}$, $0 \leq m \leq 7$, $0 \leq n \leq 7$ and $I_{p,q}$ denotes the two-dimensional DCT coefficient and

$$T_p = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } p = 0 \\ 1 & \text{for } p > 0. \end{cases} \tag{6}$$

The term T_q can be easily derived from T_p as well. The coefficient at location $(0, 0)$, called the Direct Current (DC) coefficient, represents the mean value of pixels in the 8×8 block. The remaining 63 coefficients are called the Alternating Current (AC) coefficients.

Then, the DCT coefficients have to undergo the quantization operation. It is carried out by simply dividing each coefficient by the corresponding quantization step, and then rounding to the nearest integer

$$V_{p,q} = \text{round}\left(\frac{I_{p,q}}{\Delta_{p,q}}\right), \tag{7}$$

where $V_{p,q}$ is quantized DCT coefficient, $\Delta_{p,q}$ denotes an element of the 8×8 quantization matrix and round denotes the rounding operation. The goal of the quantization operation is to discard information which is not visually significant [31]. It is the principal lossy operation in the JPEG compression technology. It should be noted that the final processing step is entropy coding, which is a form of lossless data compression. It arranges quantized DCT coefficients into the zig-zag sequence and then employs the run-length encoding (RLE) algorithm and Huffman coding. Since the entropy coding is perfectly reversible, the statistical distribution of quantized DCT coefficients does not change in this step. Therefore, the entropy coding is not considered in this paper.

III. Statistical Study of Quantized AC Coefficients

A. Impact of Post-Acquisition Processes

This paper only focuses on the distribution of quantized AC coefficients. As DC coefficient represents the mean value of pixels within each 8×8 block, the distribution of DC coefficient can not be straightforwardly derived due to the heterogeneity in a natural image. For the sake of clarity, the index of pixel and AC coefficients is omitted in this section.

The RAW image can be modeled by considering the noises that contribute to the degradation of the captured image during the image acquisition process [16], [28], [32]. In fact, the photon shot noise and dark current are modeled as Poissonian random variables whereas other electronic noises are modeled as a zero-mean Gaussian one. For the sake of simplification, the normal approximation of the Poisson distribution may be exploited because of a large number of incident photons. Finally, the RAW pixel z^c follows the Gaussian distribution

$$z^c \sim \mathcal{N}(\theta^c, s^c), \tag{8}$$

where θ^c is an element of the mean matrix $\mathbf{\Theta^c}$ and s^c is an element the variance matrix $\mathbf{\Sigma^c}$ of $\mathbf{Z}^c$. As discussed in [16], [32], the Gaussian model is a suitable approximation of a RAW image acquired by a digital imaging sensor. It should be noted that the RAW image has to go through the quantization process in the image acquisition chain. The quantization step is very small compared with noise in the RAW image as it is often coded with $B \in \{12, 14, 16\}$ bits. Therefore, the quantization at this stage supposedly has no impact on the statistical distribution of RAW pixels.

Since the bilinear interpolation and Gray World algorithm are linear operations, it follows from (1) and (3) that, after those two processes the pixel z^c_{WB} also follows the Gaussian distribution

$$z^c_{WB} \sim \mathcal{N}\big(\theta^c_{WB}, s^c_{WB}\big), \tag{9}$$

where $\mathbf{\Theta}^c_{WB} = \lambda^c \cdot \mathbf{H}^c \circledast \mathbf{\Theta^c}$ and $\mathbf{\Sigma}^c_{WB} = (\lambda^c)^2 \cdot \mathbf{H}^c \cdot \mathbf{H}^c \circledast \mathbf{\Sigma^c}$. Here, $(\cdot)$ is the element-wise multiplication operator. By applying the change of variables theorem, the probability density function (pdf) of the pixel z^c_{GM} is given by

$$f_{z^c_{GM}}(x) = \frac{\gamma x^{\gamma-1}}{\sqrt{2\pi s^c_{WB}}}\Big[\exp\Big(-\frac{(x^\gamma - \theta^c_{WB})^2}{2s^c_{WB}}\Big) + \exp\Big(-\frac{(x^\gamma + \theta^c_{WB})^2}{2s^c_{WB}}\Big)\Big], \quad x \in \mathbb{R}^+, \tag{10}$$

where $f_X(x)$ denotes the pdf of a random variable X.

B. Doubly Stochastic Model of AC Coefficients

As mentioned above, the TIFF image needs to be converted into the color space $YCbCr$. Since this transformation is linear, the distribution of each channel Y, Cb, Cr does not differ fundamentally from (10). It can be easily calculated at the expense of much complicated notations. For the sake of clarity, the distribution (10) is hence considered. In order to model AC coefficients, the variability of block variance due to the heterogeneity in the image is taken into account. Based on the doubly stochastic model [14], the pdf of AC coefficient I is given by

$$f_I(x) = \int_0^\infty f_{I|\sigma^2}(x|t) f_{\sigma^2}(t)dt \quad x \in \mathbb{R}, \tag{11}$$

where σ^2 denotes the block variance. In this model, the block variance σ^2 is itself a random variable. It is assumed that the pixels $\tilde{z}_{m,n}$ are identically distributed within a 8×8 block [14]. Given a constant block variance σ^2, the AC coefficient I may be approximately distributed as zero-mean Gaussian in virtue of Central Limit Theorem (CLT) for correlated random variables [33]

$$f_{I|\sigma^2}(x|t) = \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{x^2}{2t}\Big). \tag{12}$$

The rate of convergence in the CLT for a sequence of n dependent random variables is roughly studied; see [34], [35] and references therein. If the pixels within a block are weakly spatially correlated, it is expected that the rate of convergence is of the order of $n^{-\frac{1}{2}} \log n$. A small error of approximation evidently exists with only 64 random variables. Nevertheless, the fact of "Gaussianization" is important because it allows to simplify the study of the sum of 64 random variables of which pdf (10) is rather complicated.

The equations (11) and (12) highlight the crucial importance of studying the distribution of block variance σ^2. In fact, the block variance σ^2 can be defined by

$$\sigma^2 = \frac{1}{63}\sum_{m=0}^{7}\sum_{n=0}^{7}\big(\tilde{z}_{m,n} - \bar{\tilde{z}}\big)^2, \tag{13}$$

where $\bar{\tilde{z}}$ is the average of pixels within a 8×8 block

$$\bar{\tilde{z}} = \frac{1}{64}\sum_{m=0}^{7}\sum_{n=0}^{7}\tilde{z}_{m,n}. \tag{14}$$

It is important to note that $\bar{\tilde{z}}$ is also a random variable. It follows that

$$\tilde{z}_{m,n} - \bar{\tilde{z}} = \frac{1}{64}\sum_{m'=0}^{7}\sum_{n'=0}^{7}\big(\tilde{z}_{m,n} - \tilde{z}_{m',n'}\big). \tag{15}$$

By invoking again the CLT [33], the distribution of $\tilde{z}_{m,n} - \bar{\tilde{z}}$ approaches to the zero-mean Gaussian distribution. It should be noted that the square of a standard Gaussian random variable follows the chi-square distribution of one degree of freedom. Moreover, a chi-square random variable scaled by a constant follows the Gamma distribution. Accordingly, one derives that

$$\frac{1}{63}\big(\tilde{z}_{m,n} - \bar{\tilde{z}}\big)^2 \xrightarrow{D} \mathcal{G}\Big(\frac{1}{2}, \nu_{m,n}\Big), \tag{16}$$

where $\mathcal{G}(\cdot)$ denotes the Gamma distribution, $\nu_{m,n}$ is a scale parameter depending on the variance of $\tilde{z}_{m,n} - \bar{\tilde{z}}$, and the notation $\xrightarrow{D}$ denotes the convergence in distribution. Consequently, the block variance σ^2 is considered as a sum of correlated Gamma random variables. The exact distribution of the sum of correlated Gamma variables was analytically established in [36]. However, this exact distribution is too complicated for establishing the pdf f_I (11). Besides, it follows from [36]

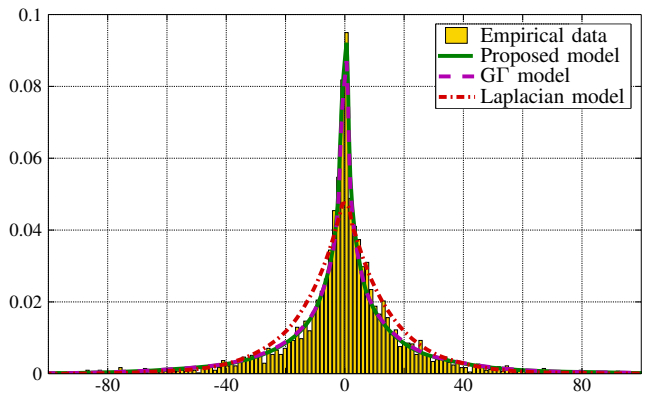

Fig. 2: Comparison between the Laplacian, GΓ and proposed model of DCT coefficients.

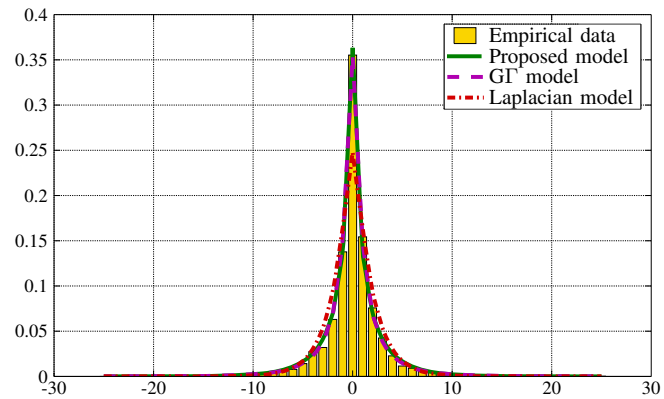

Fig. 3: Comparison between the quantized Laplacian, quantized GΓ and proposed model for quantized AC coefficient.

that the Moment-Generating Function (MGF) of σ^2 can be expressed as

$$\mathrm{M}_{\sigma^2}(t) = \Big[\det\big(\mathbf{I}_{64} - t\mathbf{D}\cdot\mathbf{Co}\big)\Big]^{-\frac{1}{2}}, \tag{17}$$

where $\det(\cdot)$ denotes the determinant operator, $\mathbf{I}_{64}$ is the 64×64 identity matrix, $\mathbf{D}$ is the 64×64 diagonal matrix with the entries $\{\nu_{m,n}\}$, and $\mathbf{Co}$ is the 64×64 covariance matrix defined by

$$\mathbf{Co} = \begin{pmatrix} 1 & \sqrt{\rho_{1,2}} & \cdots & \sqrt{\rho_{1,64}} \\ \sqrt{\rho_{2,1}} & 1 & \cdot & \sqrt{\rho_{2,64}} \\ \cdot & \cdot & \cdot & \cdot \\ \sqrt{\rho_{64,1}} & \cdots & \cdots & 1. \end{pmatrix} \tag{18}$$

Here, $\rho_{i,j}$ is the correlation coefficient between two pixels within a block. Denoting $\{\lambda_i\}_{i=1}^{64}$ the eigenvalues of the matrix $\mathbf{D}\cdot\mathbf{Co}$, the MGF $\mathrm{M}_{s^2}(t)$ is rewritten as

$$\mathrm{M}_{s^2}(t) = \prod_{i=1}^{64}\big(1 - t\lambda_i\big)^{-\frac{1}{2}}, \tag{19}$$

which has a similar form as the MGF of the sum of independent Gamma variables $\mathcal{G}\big(\frac{1}{2},\lambda_i\big)$, $i=\{1,\ldots,64\}$. The moment matching method is used to approximate the distribution of the sum of independent Gamma variables $\mathcal{G}\big(\frac{1}{2},\lambda_i\big)$ by a Gamma distribution $\mathcal{G}(\eta,\nu)$. By matching the two first moments of two distributions, the parameters (η,ν) are given by

$$\eta = \frac{\big(\sum_{i=1}^{64}\lambda_i\big)^2}{2\sum_{i=1}^{64}\lambda_i^2} \tag{20}$$

$$\nu = \frac{\sum_{i=1}^{64}\lambda_i^2}{\sum_{i=1}^{64}\lambda_i}. \tag{21}$$

As a result, the block variance σ^2 can be approximately modeled by the Gamma distribution $\mathcal{G}(\eta,\nu)$

$$f_{\sigma^2}(t) = \frac{t^{\eta-1}}{\nu^\eta\Gamma(\nu)}\exp\Big(-\frac{t}{\nu}\Big), \tag{22}$$

where η is a positive shape parameter, ν is a positive scale parameter, and $\Gamma(\cdot)$ denotes the gamma function.

The Gamma distribution of block variance is used to establish the model of AC coefficient I. It follows from (11), (12), and (22) that

$$f_I(x) = \frac{1}{\sqrt{2\pi}\nu^\eta\Gamma(\eta)}\int_0^\infty \exp\Big(-\frac{t}{\nu} - \frac{x^2}{2t}\Big)t^{\eta-\frac{3}{2}}dt. \tag{23}$$

From [37], the integral representation of the modified Bessel $K_\nu(\cdot)$ yields to

$$f_I(x) = \sqrt{\frac{2}{\pi}}\frac{\Big(|x|\sqrt{\frac{\nu}{2}}\Big)^{\eta-\frac{1}{2}}}{\nu^\eta\Gamma(\eta)}K_{\eta-\frac{1}{2}}\Big(|x|\sqrt{\frac{2}{\nu}}\Big). \tag{24}$$

The Fig. 2 illustrates the empirical distribution of the third DCT coefficient, extracted from the image in the BOSS Base [38], and the proposed model compared with the Laplacian and GΓ model.

Based on the law of total expectation, the variance of the AC coefficient I is given by

$$\mathrm{Var}_I\big[I\big] = \mathbb{E}_I\big[I^2\big] = \mathbb{E}_{\sigma^2}\Big[\mathbb{E}_{I|\sigma^2}\big[I^2|\sigma^2\big]\Big] = \mathbb{E}_{\sigma^2}\big[\sigma^2\big] = \eta\nu \tag{25}$$

where $\mathbb{E}_X$ and Var_X represents the mathematical expectation and variance with respect to a random variable X. Similarly, the kurtosis coefficient of I is defined by

$$\gamma_2 = \frac{\mathbb{E}_I\big[I^4\big]}{\mathrm{Var}_I^2\big[I\big]} = \frac{\mathbb{E}_{\sigma^2}\big[3\sigma^4\big]}{\mathbb{E}_{\sigma^2}^2\big[\sigma^2\big]} = 3\frac{\eta\nu^2(\eta+1)}{\eta^2\nu^2} = 3\Big(1+\frac{1}{\eta}\Big). \tag{26}$$

The proposed model includes the Laplacian and Gaussian as special cases. In fact, as $\eta\to\infty$, then $\gamma_2\to 3$. The AC coefficient I tends to be distributed as Gaussian variable. Similarly, as $\eta=1$, $\gamma_2=6$, then the Gamma distribution of block variance reduces back to the exponential distribution, the Laplacian model for AC coefficient I is obtained [14]. The proposed model of I outperforms the Laplacian, yet at the expense of more complex expressions and extra computational cost.

C. Impact of Quantization : Final Model of Quantized AC coefficients

Let $P_V(l)$, $l\in\mathbb{Z}$, be the probability mass function (pmf) of the quantized AC coefficient V with the corresponding

quantization step Δ. The uniform quantization operation with step Δ can be written as follows

$$V = l \Longleftrightarrow I \in \Big[\Delta\Big(l - \frac{1}{2}\Big), \Delta\Big(l + \frac{1}{2}\Big)\Big[. \tag{27}$$

Therefore, the pmf $P_V(l)$ is defined by

$$P_V(l) = \mathbb{P}\big[V = l\big] = \int_{\Delta(l-\frac{1}{2})}^{\Delta(l+\frac{1}{2})} f_I(x)dx. \tag{28}$$

Because the pmf $P_V(l)$ is symmetric, it is sufficient to consider $l \geq 0$. Let define the function $G(l)$ as

$$G(l) = \int_0^{\Delta(l+\frac{1}{2})} f_I(x)dx \quad \forall l \in \mathbb{Z}^+. \tag{29}$$

By changing the variable $x = \Delta(l + \frac{1}{2}) \cdot t$, a direct calculation from (24) and (29) yields to

$$G(l) = \sqrt{\frac{2}{\pi}} \frac{\Big(\sqrt{\frac{\nu}{2}}\Big)^{\eta-\frac{1}{2}} \Big(\Delta(l+\frac{1}{2})\Big)^{\eta+\frac{1}{2}}}{\nu^{\eta}\Gamma(\eta)} \times \int_0^1 t^{\eta-\frac{1}{2}} K_{\eta-\frac{1}{2}}\Big[t \cdot \Delta\Big(l + \frac{1}{2}\Big)\sqrt{\frac{2}{\nu}}\Big]dt. \tag{30}$$

It follows from [37] that:

$$G(l) = \frac{1}{2} g(l)\Big[K_{\eta-\frac{1}{2}}(g(l))\mathbf{L}_{\eta-\frac{3}{2}}(g(l)) + K_{\eta-\frac{3}{2}}(g(l))\mathbf{L}_{\eta-\frac{1}{2}}(g(l))\Big], \tag{31}$$

where $g(l) = \Delta(l + \frac{1}{2})\sqrt{\frac{2}{\nu}}$ and $\mathbf{L}_\nu(\cdot)$ is the modified Struve function. Finally, the pmf $P_V(l)$ is given by

$$P_V(l) = \begin{cases} G\big(|l|\big) - G\big(|l| - 1\big) & \forall l \in \mathbb{Z}_* \\ 2G(0) & l = 0. \end{cases} \tag{32}$$

The Fig. 3 illustrates the empirical data and the proposed model of quantized AC coefficients, compared with the quantized Laplacian and quantized GΓ model that are detailed in Appendix A.

The above mathematical framework is based on some assumptions that may not be realistic. In fact, the image processing pipeline that goes from the image scene to the final output (e.g. JPEG image) is complicated and so difficult to model. This paper does not aim to cover all effects occurred in the image processing pipeline. Therefore we need to simplify the reality and make assumptions in order to built a statistical model for DCT coefficients. The proposed model of DCT coefficients shows a better fit to a wide range of images than existing models in the literature (e.g. Laplacian and Generalized Gamma model); see Fig. 3 and more numerical results in Section VI. Moreover, the high efficiency of the proposed model is highlighted when applying in the steganalysis of Jsteg algorithm to detect a small change in the cover image due to the insertion of secret message.

IV. Model Parameter Estimation

For the sake of clarity, from this section, the quantized DCT coefficients are arranged into 64 vectors of coefficients. Let $V_k = (v_{k,1}, \ldots, v_{k,n})^T$, $k \in \{1, \ldots, 64\}$, be the vector of length n representing the k-th quantized DCT coefficient where $v_{k,i}$, $1 \leq i \leq n$, denotes the value of the k-th DCT coefficient in the block i and $\mathbf{U}^T$ denotes the transpose of the matrix $\mathbf{U}$. Accordingly, Δ_k denotes the quantization step associated with V_k.

The above mathematical analysis does not explain the difference in scale of the distributions across the DCT coefficients. In fact, in a natural image, the energy tends to be more concentrated in the lower frequency than in the higher frequency. The quantization step Δ_k also depends on frequency. The quantization table is often designed to preserve information in a low frequency and discard details in a high frequency. After the quantization process, larger variance is expected in the low frequency for which the quantization step is smaller. Therefore, we should treat each frequency separately. The parameters characterizing the distribution of quantized AC coefficient V_k are now denoted by (η_k, ν_k), $k = \{2, \ldots, 64\}$, with respect to the k-th quantized DCT coefficient.

Obviously, the original image can not be perfectly reconstructed because of the lossy compression. For a practical use, the model parameters (η_k, ν_k) need to be estimated from the quantized AC coefficients V_k.

A. Method of Moments (MM) Estimates

According to the theory of quantization [39], the effect of uniform quantization can be modeled by an additive noise that is uniformly distributed and uncorrelated with the input signal. The quantized AC coefficient V_k can be given by

$$V_k = \frac{I_k}{\Delta_k} + \epsilon_k, \quad \epsilon_k \sim \mathcal{U}\Big[-\frac{1}{2}, \frac{1}{2}\Big], \tag{33}$$

where $\mathcal{U}$ represents the uniform distribution and I_k is the unquantized AC coefficient. Since the distribution of I_k and V_k is symmetric, their odd moments vanish. Based on the definitions of the expectation, the second and fourth moments of V_k are therefore given by

$$\begin{aligned} \mathbb{E}_{V_k}\big[V_k^2\big] &= \frac{1}{\Delta_k^2}\mathbb{E}_{I_k}\big[I_k^2\big] + \mathbb{E}_{\epsilon_k}\big[\epsilon_k^2\big] \\ &= \frac{\eta_k \nu_k}{\Delta_k^2} + \frac{1}{12} \end{aligned} \tag{34}$$

$$\begin{aligned} \mathbb{E}_{V_k}\big[V_k^4\big] &= \frac{1}{\Delta_k^4}\mathbb{E}_{I_k}\big[I_k^4\big] + \frac{6}{\Delta_k^2}\mathbb{E}_{I_k}\big[I_k^2\big]\mathbb{E}_{\epsilon_k}\big[\epsilon_k^2\big] + \mathbb{E}_{\epsilon_k}\big[\epsilon_k^4\big] \\ &= \frac{3}{\Delta_k^4}\eta_k\nu_k^2(\eta_k + 1) + \frac{1}{2\Delta_k^2}\eta_k\nu_k + \frac{1}{80}. \end{aligned} \tag{35}$$

It follows that the parameters (η_k, ν_k) can be expressed as

$$\eta_k = \frac{\Big(\mathbb{E}_{V_k}\big[V_k^2\big] - \frac{1}{12}\Big)^2}{\frac{1}{3}\mathbb{E}_{V_k}\big[V_k^4\big] - \mathbb{E}_{V_k}^2\big[V_k^2\big] + \frac{1}{360}} \tag{36}$$

$$\nu_k = \frac{\Delta_k^2\Big(\mathbb{E}_{V_k}\big[V_k^2\big] - \frac{1}{12}\Big)}{\eta_k}. \tag{37}$$

The MM estimates of (η_k, ν_k) are then derived as

$$\hat{\eta}_k^{MM} = \frac{\left(m_{k,2} - \frac{1}{12}\right)^2}{\frac{1}{3}m_{k,4} - m_{k,2}^2 + \frac{1}{360}} \tag{38}$$

$$\hat{\nu}_k^{MM} = \frac{\Delta_k^2\left(m_{k,2} - \frac{1}{12}\right)}{\hat{\eta}_k^{MM}}, \tag{39}$$

where $\hat{m}_{k,2}$ and $\hat{m}_{k,4}$ are the empirical second and fourth moments of V_k

$$\hat{m}_{k,2} = \frac{1}{n}\sum_{i=1}^{n} v_{k,i}^2 \qquad \hat{m}_{k,4} = \frac{1}{n}\sum_{i=1}^{n} v_{k,i}^4. \tag{40}$$

B. ML Estimates

By definition, the ML estimates of (η_k, ν_k) are defined as the solution of a maximization problem

$$\left(\hat{\eta}_k^{ML}, \hat{\nu}_k^{ML}\right) = \underset{(\eta_k,\nu_k)}{\arg\max} \sum_{i=1}^{n} \log P_V\left(v_{k,i}\right). \tag{41}$$

The ML estimates $\left(\hat{\eta}_k^{ML}, \hat{\nu}_k^{ML}\right)$ can not be analytically provided because the maximization problem (41) has no closed-form solution. It is proposed to resolve the maximization problem numerically by using the Nelder-Mead optimization method [40]. The MM estimates $\left(\hat{\eta}_k^{MM}, \hat{\nu}_k^{MM}\right)$ is taken as initial solution in the optimization algorithm. Even though the convergence to the global solution can not be ensured in a practical context, this procedure can be used to obtain a heuristic solution for ML estimates $\left(\hat{\eta}_k^{ML}, \hat{\nu}_k^{ML}\right)$. By contrast, the ML estimates of the parameters of quantized Laplacian and quantized GΓ model are obtained by taking ML estimates in [21] and [13], respectively, as initial solution.

V. Application in the Steganalysis of Jsteg Algorithm

A. Problem Statement

As described above, the Jsteg algorithm is based on the LSB replacement scheme in the DCT domain. Let $\mathbf{C}$ be a matrix representing the cover image that is composed of 64 vectors of quantized DCT coefficients C_k, $k = \{1, \ldots, 64\}$. The pmf of the quantized DCT coefficient C_k is denoted by $P_{\boldsymbol{\theta}_k,\Delta_k}$ characterized by the parameter vector $\boldsymbol{\theta}_k$ and the corresponding quantization step Δ_k. It should be noted that in our proposed model, $\boldsymbol{\theta}_k = (\eta_k, \nu_k)$. Let us assume that $\mathcal{M}$ represents the encrypted secret message of length L. In the Jsteg algorithm, each hidden bit, that is either 0 or 1, is statistically independent of the cover coefficients. Moreover, the probability of insertion is equal for every coefficient. The Jsteg does not embed in the coefficients that are equal to 0 and 1 since artifacts caused by such insertion can be easily detected. For the same reason, the DC coefficient is not used for insertion as well. The number of usable coefficients in each vector C_k, $k = \{2, \ldots, 64\}$, is represented by a random variable $n_k \leq n$. The number of usable coefficients n_k depends on the image content and the quantization matrix (hence the quality factor). Without loss of generality, the n_k first components of the vector are usable and the remaining $n - n_k$ components are excluded. Accordingly, the insertion rate R is defined as the ratio of the length L and the number of usable coefficients in the whole cover image $\mathbf{C}$

$$R = \frac{L}{\sum_{k=2}^{64} n_k}. \tag{42}$$

The message is embedded with rate R in the cover image $\mathbf{C}$ to create a stego-image $\mathbf{S} = (S_1, \ldots, S_{64})$. The pmf of S_k, denoted $Q_{R,\boldsymbol{\theta}_k,\Delta_k}$, is well described in [41], [42]: $\forall l \neq \{0, 1\}$

$$Q_{R,\boldsymbol{\theta}_k,\Delta_k}(l) = \left(1 - \frac{R}{2}\right)P_{\boldsymbol{\theta}_k,\Delta_k}(l) + \frac{R}{2}P_{\boldsymbol{\theta}_k,\Delta_k}(\bar{l}), \tag{43}$$

where $\bar{l}$ indicates the integer l with LSB flipped $\bar{l} = l + (-1)^l$. As described above, coefficients with value 0 and 1 are not used for security reason. Hence, the pmf $P_{\boldsymbol{\theta}_k,\Delta_k}$ does not change for $l = \{0, 1\}$ after insertion

$$Q_{R,\boldsymbol{\theta}_k,\Delta_k}(0) = P_{\boldsymbol{\theta}_k,\Delta_k}(0) \tag{44}$$

$$Q_{R,\boldsymbol{\theta}_k,\Delta_k}(1) = P_{\boldsymbol{\theta}_k,\Delta_k}(1). \tag{45}$$

When inspecting the image $\mathbf{V}$ that is either a cover image $\{\mathbf{V} = \mathbf{C}\}$ or a stego-one $\{\mathbf{V} = \mathbf{S}\}$, the goal of the test is to decide between two hypotheses defined as $\forall k = \{2, \ldots, 64\}$, $\forall i = \{1, \ldots, n\}$

$$\begin{cases} \mathcal{H}_0 = \left\{v_{k,i} \sim P_{\boldsymbol{\theta}_k,\Delta_k}, \boldsymbol{\theta}_k \in \mathbb{R}_+^2\right\} \\ \mathcal{H}_1 = \left\{v_{k,i} \sim Q_{R,\boldsymbol{\theta}_k,\Delta_k}, \boldsymbol{\theta}_k \in \mathbb{R}_+^2, R \in (0, 1]\right\}. \end{cases} \tag{46}$$

The DC coefficients V_1 are excluded in the problem (46) because they are not used for the insertion of secret message. The problem (46) involves two main difficulties. First, two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are composite because the embedding rate R is unknown in practice. Second, there is the presence of nuisance parameters $\boldsymbol{\theta}_k$ that is also unknown. In fact, the nuisance parameters $\boldsymbol{\theta}_k$ do not contain any information about the presence of hidden bits.

This paper aims to design an statistical test for the problem (46) assuming that the embedding rate R and the parameters $\boldsymbol{\theta}_k$ are known. In practice, when the embedding rate R is not known in advance, one can rely on the Locally Asymptotically Uniformly Most Powerful (LAUMP), which was exploited in [42], [43]. Meanwhile, the Generalized Likelihood Ratio Test (GLRT) [44] allows to deal with unknown nuisance parameters $\boldsymbol{\theta}_k$. However, the statistical performance of GLRT can not be easily established due to the difficulty of studying statistical properties of the parameters $\boldsymbol{\theta}_k$. Moreover, the proposed test only considers first order statistics of DCT coefficients. The use of higher order statistics (e.g. the correlation between DCT coefficients) is beyond the scope of this paper.

B. Most Powerful Likelihood Ratio Test

As previously explained, this paper focuses on guaranteeing a prescribed false-alarm probability. Hence, let define

$$\mathcal{K}_{\alpha_0} = \left\{\delta : \mathbb{P}_0\left[\delta(\mathbf{V}) = \mathcal{H}_1\right] \leq \alpha_0\right\}$$

the class of tests whose false alarm probabilities are upper-bounded by α_0. Here, $\mathbb{P}_j[\cdot]$ denotes the probability under hypothesis $\mathcal{H}_j$, $j = \{0, 1\}$. Among all the tests in the class

$\mathcal{K}_{\alpha_0}$, it is aimed at finding a test δ which maximizes the power function defined by

$$\beta_\delta = \mathbb{P}_1\Big[\delta(\mathbf{V}) = \mathcal{H}_1\Big].$$

In virtue of the Neyman-Pearson lemma [19, theorem 3.2.1], the most powerful test over the class $\mathcal{K}_{\alpha_0}$ is the LRT given by the following decision rule

$$\delta = \begin{cases} \mathcal{H}_0 & \text{if} \quad \Lambda(\mathbf{V}) = \sum_{k=2}^{64} \Lambda(V_k) < \tau \\ \mathcal{H}_1 & \text{if} \quad \Lambda(\mathbf{V}) = \sum_{k=2}^{64} \Lambda(V_k) \geq \tau \end{cases} \tag{47}$$

where the decision threshold τ is the solution of the equation

$$\mathbb{P}_0\Big[\Lambda(\mathbf{V}) \geq \tau\Big] = \alpha_0 \tag{48}$$

to ensure that $\delta \in \mathcal{K}_{\alpha_0}$ and $\Lambda(V_k) = \sum_{i=1}^{n_k} \Lambda(v_{k,i})$. Here, the LR for one observation $\Lambda(v_{k,i})$ is defined by

$$\begin{aligned} \Lambda(v_{k,i}) &= \log \frac{Q_{R,\boldsymbol{\theta}_k,\Delta_k}(v_{k,i})}{P_{\boldsymbol{\theta}_k,\Delta_k}(v_{k,i})} \\ &= \log\left[1 - \frac{R}{2} + \frac{R}{2}\frac{P_{\boldsymbol{\theta}_k,\Delta_k}(\overline{v}_{k,i})}{P_{\boldsymbol{\theta}_k,\Delta_k}(v_{k,i})}\right]. \end{aligned} \tag{49}$$

where $\overline{v}_{k,i} = v_{k,i} + (-1)^{v_{k,i}}$ is the coefficient $v_{k,i}$ with flipped LSB. Accordingly, $\Lambda(v_{k,i})$ can be interpreted as a function of the integer $v_{k,i}$.

Let define the function $d_k(m)$ as

$$d_k(m) = \log\left[1 - \frac{R}{2} + \frac{R}{2}\frac{P_{\boldsymbol{\theta}_k,\Delta_k}(m + (-1)^m)}{P_{\boldsymbol{\theta}_k,\Delta_k}(m)}\right]. \tag{50}$$

Based on the definitions of the mathematical expectation and variance, one can derive the expectation and variance of $\Lambda(v_{k,i})$ under hypothesis $\mathcal{H}_0$

$$\mu_{k,0} = \mathbb{E}_0\Big[\Lambda(v_{k,i})\Big] = \sum_{m\in\mathbb{Z}} d_k(m) P_{\boldsymbol{\theta}_k,\Delta_k}(m) \tag{51}$$

$$\sigma_{k,0}^2 = \mathrm{Var}_0\Big[\Lambda(v_{k,i})\Big] = \sum_{m\in\mathbb{Z}} \big(d_k(m) - \mu_{k,0}\big)^2 P_{\boldsymbol{\theta}_k,\Delta_k}(m), \tag{52}$$

where $\mathbb{E}_j[\cdot]$ and $\mathrm{Var}_j[\cdot]$ respectively denote the expectation and variance under hypothesis $\mathcal{H}_j$, $j = \{0, 1\}$. Meanwhile, the random variable n_k corresponds to the number of coefficients that are different from 0 and 1. Accordingly, n_k follows the binomial distribution $\mathcal{B}(n, p_k^*)$ where

$$p_k^* = 1 - P_{\boldsymbol{\theta}_k,\Delta_k}(0) - P_{\boldsymbol{\theta}_k,\Delta_k}(1) \tag{53}$$

is the success probability. This success probability remains identical under every hypothesis since the insertion is not performed in coefficients that are equal to 0 and 1. It follows from the Wald's identity [45] that the expectation and variance

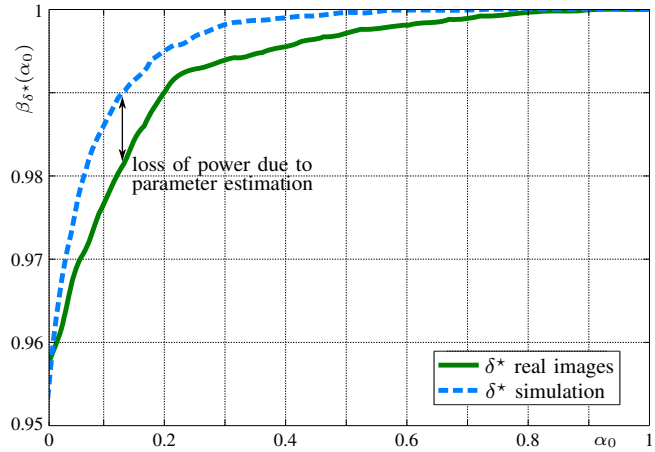

Fig. 4: Detection performance of the test $\delta^\star$ based on the proposed model with embedding rate $R = 0.05$ on the simulated images and real images.

of the random sum $\Lambda(V_k)$ are defined by

$$\begin{aligned} \mathbb{E}_0\Big[\Lambda(V_k)\Big] &= \mathbb{E}_0[n_k]\mathbb{E}_0\Big[\Lambda(v_{k,i})\Big] \\ &= np_k^*\mu_{k,0} \end{aligned} \tag{54}$$

$$\begin{aligned} \mathrm{Var}_0\Big[\Lambda(V_k)\Big] &= \mathbb{E}_0[n_k]\mathrm{Var}_0\Big[\Lambda(v_{k,i})\Big] \\ &\quad + \mathbb{E}_0^2\Big[\Lambda(v_{k,i})\Big]\mathrm{Var}_0[n_k] \\ &= np_k^*\sigma_{k,0}^2 + np_k^*(1-p_k^*)\mu_{k,0}^2. \end{aligned} \tag{55}$$

In virtue of the Lindeberg CLT [19, theorem 11.2.5], $\Lambda(V_k)$ is normally distributed with mean $np_k^*\mu_{k,0}$ and variance $np_k^*\sigma_{k,0}^2 + np_k^*(1-p_k^*)\mu_{k,0}^2$. Finally, because of the linearity property of the Gaussian distribution, the decision function $\Lambda(\mathbf{V})$ also follows the Gaussian distribution under hypothesis $\mathcal{H}_0$

$$\Lambda(\mathbf{V}) \xrightarrow{D} \mathcal{N}(\mu_0, \sigma_0^2) \tag{56}$$

where

$$\mu_0 = \sum_{k=2}^{64} \mathbb{E}_0\Big[\Lambda(V_k)\Big] = \sum_{k=2}^{64} np_k^*\mu_{k,0} \tag{57}$$

$$\sigma_0^2 = \sum_{k=2}^{64} \mathrm{Var}_0\Big[\Lambda(V_k)\Big] = \sum_{k=2}^{64} \Big[np_k^*\sigma_{k,0}^2 + np_k^*(1-p_k^*)\mu_{k,0}^2\Big]. \tag{58}$$

Similarly, under hypothesis $\mathcal{H}_1$, the expectation and variance of $\Lambda(v_{k,i})$ are given by

$$\mu_{k,1} = \sum_{m\in\mathbb{Z}} d_k(m) Q_{R,\boldsymbol{\theta}_k,\Delta_k}(m) \tag{59}$$

$$\sigma_{k,1}^2 = \sum_{m\in\mathbb{Z}} \big(d_k(m) - \mu_{k,1}\big)^2 Q_{R,\boldsymbol{\theta}_k,\Delta_k}(m). \tag{60}$$

Accordingly, $\Lambda(\mathbf{V})$ follows the Gaussian distribution under hypothesis $\mathcal{H}_1$

$$\Lambda(\mathbf{V}) \xrightarrow{D} \mathcal{N}(\mu_1, \sigma_1^2) \tag{61}$$

with mean and variance defined by

$$\mu_1 = \sum_{k=2}^{64} np_k^* \mu_{k,1} \tag{62}$$

$$\sigma_1^2 = \sum_{k=2}^{64} \left[np_k^* \sigma_{k,1}^2 + np_k^*(1-p_k^*)\mu_{k,1}^2 \right]. \tag{63}$$

Since natural images are heterogeneous, it is proposed to normalize the LR $\Lambda(\mathbf{V})$ and use the test $\delta^\star$ defined as follows

$$\delta^\star = \begin{cases} \mathcal{H}_0 & \text{if} \quad \Lambda^\star(\mathbf{V}) < \tau^\star \\ \mathcal{H}_1 & \text{if} \quad \Lambda^\star(\mathbf{V}) \geq \tau^\star \end{cases} \tag{64}$$

with

$$\Lambda^\star(\mathbf{V}) = \frac{\Lambda(\mathbf{V}) - \mu_0}{\sigma_0}. \tag{65}$$

Therefore, $\Lambda^\star(\mathbf{V}) \xrightarrow{D} \mathcal{N}(0,1)$ under hypothesis $\mathcal{H}_0$. The fact of normalizing the LR makes the test applicable to any natural image since the LR follows the standard Gaussian distribution under hypothesis $\mathcal{H}_0$. This also allows to set the decision threshold independently of the image content. The decision threshold $\tau^\star$ and the power function $\beta_{\delta^\star}$ of the test $\delta^\star$ are given in the following theorem:

Theorem 1. *Assuming that the embedding rate R and the parameters $\boldsymbol{\theta}_k$ are known, the decision threshold and the power function of the test $\delta^\star$ are given by*

$$\tau^\star = \Phi^{-1}(1-\alpha_0) \tag{66}$$

$$\beta_{\delta^\star} = 1 - \Phi\left(\frac{\mu_0 - \mu_1 + \tau^\star \sigma_0}{\sigma_1} \right) \tag{67}$$

where $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ denotes respectively the cumulative distribution function of the standard Gaussian random variable and its inverse.

The main strength of the proposed test $\delta^\star$ is the guaranteeing of a prescribed false alarm rate and the analytic establishment of the detection performance. It can be noted that the scenario studied by the test $\delta^\star$ may not be realistic because the parameters $\boldsymbol{\theta}_k$ can not be known in advance in a real image. An usual approach in practice is to replace the parameters $\boldsymbol{\theta}_k$ by ML estimates $\hat{\boldsymbol{\theta}}_k^{ML}$. Thus, the detection performance of the test $\delta^\star$ depends on the accuracy of the proposed model of DCT coefficients and ML estimates $\hat{\boldsymbol{\theta}}_k^{ML}$. The Fig. 4 shows the detection performance of the proposed test $\delta^\star$ on 10000 simulated images in which DCT coefficients perfectly follow the proposed model and there is no correlation between DCT coefficients, and 10000 real images from the BOSSBase database [38]. Here, the accuracy of the proposed model of DCT coefficients is highlighted as the empirical detection power fits almost perfectly the theoretical one.

The fact of replacing the unknown parameters $\boldsymbol{\theta}_k$ by ML estimates $\hat{\boldsymbol{\theta}}_k^{ML}$ shows that the test $\delta^\star$ seems to coincide with the GLRT. However, the test $\delta^\star$ does not consider the variability of ML estimates $\hat{\boldsymbol{\theta}}_k^{ML}$. Because the ML estimates $\hat{\boldsymbol{\theta}}_k^{ML}$ are numerically derived by the optimization method, their statistical properties can not be easily studied. This leads to a difficulty of establishing analytically the statistical performance of the GLRT. The study of the GLRT lies out of the scope of the paper.

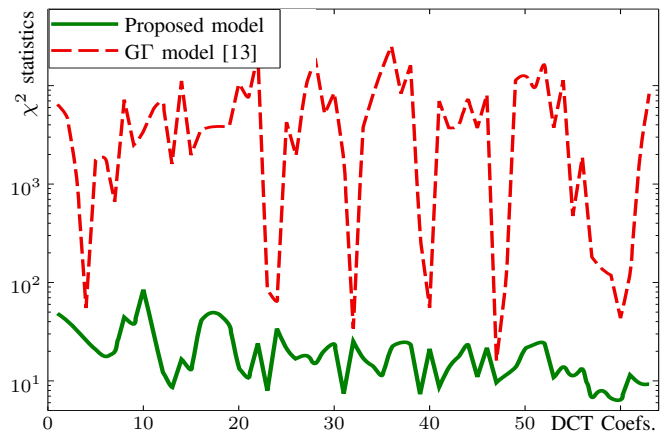

Fig. 5: Averaged χ^2 GOF test statistics of GΓ and proposed model for 63 AC coefficients

C. Embedding Rate Estimation

In this section, the problem of embedding rate estimation is formulated into the ML framework using the proposed model of DCT coefficients. As discussed in [46], ML estimators [41], [47] are more statistically rigorous, but their performance is weak due to lack of accurate models for cover images. An extension for ML framework is derived in [22] that is based on the concept of a precover introduced in [46] and the Generalized Cauchy distribution for unquantized DCT coefficients [7]. On the contrary, this paper exploits the model of quantized DCT coefficients to estimate the embedding rate R. Given an inspected image $\mathbf{V}$, the ML estimate $\hat{R}$ is given by

$$\hat{R} = \underset{0 \leq R \leq 1}{\arg\max} \sum_{k=2}^{64} \sum_{i=1}^{n_k} \log Q_{R,\boldsymbol{\theta}_k,\Delta_k}(v_{k,i}) \tag{68}$$

where $Q_{R,\boldsymbol{\theta}_k,\Delta_k}(v_{k,i})$ is defined in (43). Here again, the DC coefficients and the coefficients that are equal to 0 and 1 are excluded. The maximization problem (68) is solved numerically by the Nelder-Mead method [40].

VI. Numerical Results

A. Comparison Between the Proposed Model and the GΓ Model

Since the Laplacian model is a special case of the proposed model and GΓ model, these two models would result in a better fit to the empirical data than Laplacian. The Fig. 2 and Fig. 3 show that the Laplacian model is not relevant to model accurately DCT coefficients. However, they do not show clearly the difference between the proposed model and the GΓ model. It is proposed to use a GOF test to compare these two models. The χ^2 GOF test is preferable for the comparison than the KS test because the deviation from an assumed pdf is more interesting than deviation from a cumulative distribution function, as noted in [12]. The model whose the χ^2 value is smaller is more relevant to characterize the distribution of

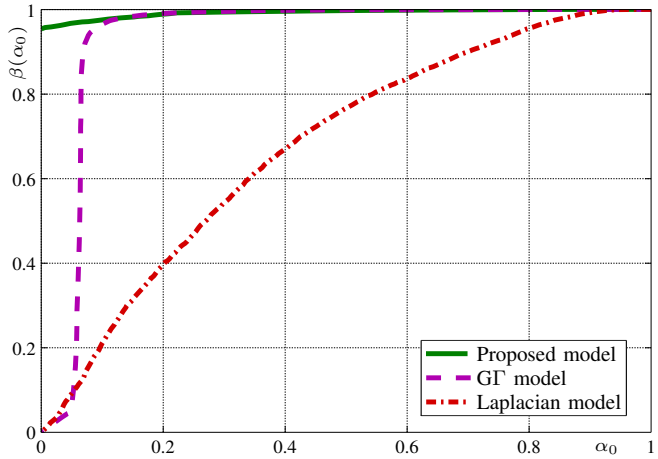

Fig. 6: Detection performance of the test $\delta^\star$ based on the quantized Laplacian, quantized GΓ and proposed model on the BOSSBase with embedding rate $R = 0.05$.

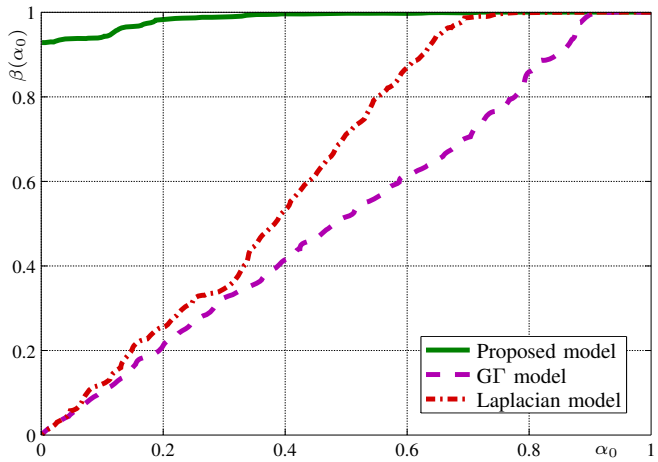

Fig. 7: Detection performance of the test $\delta^\star$ based on the quantized Laplacian, quantized GΓ and proposed model on the subset of 1000 images from the BOSSBase with embedding rate $R = 0.05$.

DCT coefficients. The experiments are conducted on all the images of the Dresden Image Database [48]. The averaged χ^2 GOF test statistics for 63 AC coefficients illustrated in Fig. 5. These results obviously show the relevance of the proposed model.

B. Steganalysis of Jsteg Algorithm

To illustrate the detection performance of the test $\delta^\star$ based on the proposed model, the reference BOSSBase database [38] containing 10000 grayscale images of size 512×512 in PGM format is chosen to conduct experiments. The embedding rate R is set at 0.05 for the Jsteg algorithm. The embedded message is drawn from a binomial distribution $\mathcal{B}(1, 1/2)$, i.e. each hidden bit can be 0 or 1 with the same probability. The coefficients in which secret bits are embedded are randomly chosen. All the PGM images are converted to JPEG format using *imagemagick* with quality factor of 70. The 63 vectors of AC coefficients are extracted from every image. The quantization matrix is given in the header of each image file. The parameters $\boldsymbol{\theta}_k = (\eta_k, \nu_k)^T$ are estimated for each image and for each frequency based on the ML approach. The estimates $\hat{\boldsymbol{\theta}}_k$ are used to generate each vector V_k such that the simulated data involves the same model parameters as the ones estimated from real images. Therefore, the simulated data perfectly follows the proposed model and there is no correlation between simulated DCT coefficients. The test $\delta^\star$ is performed on 10000 simulated images and 10000 real JPEG images. It is desirable to evaluate the loss of power of the proposed test $\delta^\star$ in the practical context. The power functions obtained from simulated images and real images are shown in Fig. 4. A small loss of power is obviously revealed between the two power functions, which may be caused by the following reasons. Firstly, the accuracy of the proposed model may be affected by assumptions required for above mathematical analysis (e.g. the pixels are identically distributed within 8×8 block) and the small error of approximation in the CLT. Secondly, the small correlation between DCT coefficients exists in real JPEG images. Finally, it can be caused by the estimation of model parameters, which are numerically provided by the optimization method. It should be noted that the proposed model is more accurate than other empirical ones in the literature (e.g. Laplacian and Generalized Gamma model). The loss of power of the test $\delta^\star$ is considerably smaller than the test proposed in [21] based on the quantized Laplacian distribution.

The proposed test $\delta^\star$ can be used with any cover image model. The more accurate the cover image model is, the better the detection performance is. The Fig. 6 illustrates the detection performance of the test $\delta^\star$ based on quantized the Laplacian, quantized GΓ and proposed model on 10000 real JPEG images. The detection performances are illustrated using the Receiver Operating Characteristic (ROC) curve which presents the detection power β as a function of the false alarm probability α_0. The test $\delta^\star$ based on the GΓ and proposed model shows a higher detection probability than the Laplacian model-based test. Moreover, it appears that the GΓ model fails for a subset of about 1000 images, which leads to its very low power function for $0 \leq \alpha_0 \leq 0.1$ while the proposed model still shows a high detection performance. The detection performance for this subset is illustrated in Fig. 7. These results also show that the proposed model is more robust and accurate for the DCT coefficients. The term "robust" means the accuracy of the model for a wide range of images. Besides, the test $\delta^\star$ based on the proposed model is nearly perfect for embedding rate $R = 0.1$, i.e. $\beta_{\delta^\star} \cong 1$, for any false alarm probability α_0.

Potentially, there are many detectors in the literature could be compared with the proposed test. The ZP detector [49] was known as the first quantitative attack on Jsteg. The well-known WS detector [24, Eq. (9)] is also included in the comparison because of its efficiency and low computational complexity. The recent quantitative structural detector ZMH-Sym [22] based on the zero message hypothesis (ZMH) framework and the exploitation of the natural symmetry of DCT coefficients in the cover image was shown as the best detector among histogram-based attacks. The quantized Laplacian model-

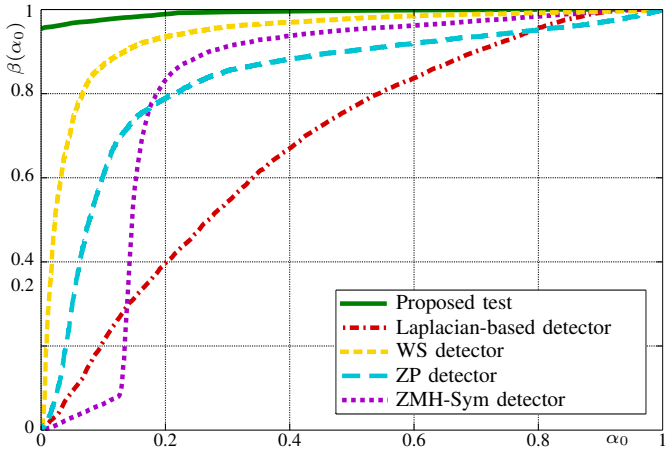

Fig. 8: Comparison between the proposed test $\delta^\star$, ZMH-Sym detector, ZP detector, WS detector and quantized Laplacian model-based test.

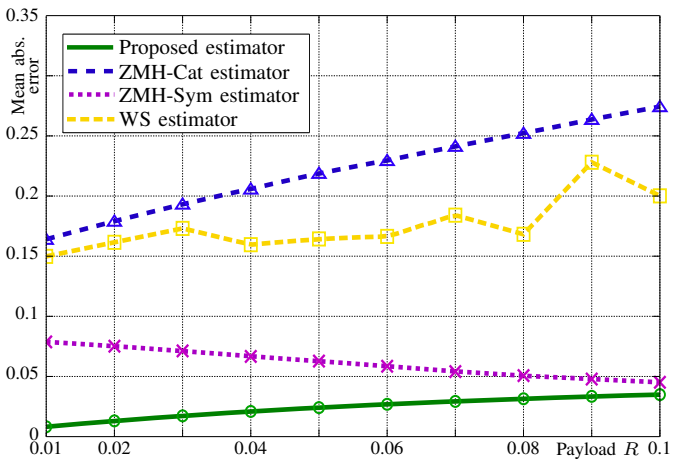

Fig. 9: Mean absolute error for all estimators.

based test [21] is also performed because it is based on the same framework of hypothesis testing theory. On the contrary to the support vector regression-based detector [25] that needs an expensive training phase, all above detectors, including the proposed test, work solely on an image-by-image basis. The Fig. 8 shows the comparison between the proposed test $\delta^\star$ and those detectors. We have also performed the ZMH-Cat structural detector [22] but do not report it in Fig. 8 because its power function is considerably worse than the detector ZMH-Sym's one. Obviously, the proposed test outperforms other detectors, whatever the false alarm probability. It should particularly be noted that for very low-false alarm rate the proposed test performs much better than the others, which is the most important in practice since the false-alarm probability must be set very low.

In terms of embedding rate estimation, the accuracy of all estimators is evaluated for embedding rate R ranging from 0.01 to 0.1 using the Mean Absolute Error (MAE) criteria: $\frac{1}{N}\sum_{i=1}^{N}|\hat{R}_i - R|$ where $N = 10000$ is the number of images. It should be noted that the ZMH-Sym and ZMH-Cat estimators [22] were proposed to estimate the change rate β that is defined as the relative portion of modified DCT coefficients with respect to the number of DCT coefficients in the image that are not equal to 0 or 1. Under assumption that no matrix embedding is used, which is the case in this paper, the expected value of the embedding rate R is 2β. Therefore, the output value of those estimators is simply multiplied by a factor of 2 to obtain an estimator of embedding rate R. The Fig. 9 shows the MAE for all estimators. The proposed ML estimator (68) outperforms other estimators. The ZMH-Sym estimator has a comparable accuracy to the proposed ML estimator but it has more outliers, which leads to the degradation of the ROC curve. The very high detection performance of the test $\delta^\star$ (Fig. 8) and high accuracy of embedding rate estimation (Fig. 9) highlight the accuracy of the proposed model of quantized DCT coefficients.

VII. Conclusion

This paper proposes a novel statistical model of quantized DCT coefficients based on a mathematical framework that has not been provided yet in the literature. Numerical results show that the proposed model is more accurate than the other ones including the popular Laplacian and Generalized Gamma. Based on this high accurate model, the paper exploits it for the steganalysis of Jsteg algorithm to detect a small change in the cover image due to the insertion of secret message. A statistical test is designed to warrant a prescribed false alarm probability and a ML estimator for embedding rate is also proposed. Further research apply the proposed model in different fields such as digital forensics.

Appendix
Review on Existing Models of AC Coefficients

This appendix only reviews the Laplacian and GΓ model of AC coefficients. These models are used for the comparison with the proposed model in experiments. The Laplacian pdf of the coefficient I is given by

$$f_I(x) = \frac{\lambda}{2}\exp\big(-\lambda|x|\big) \tag{69}$$

Accordingly, the pmf of the quantized coefficient V is derived as [21]

$$P_V(l) = \begin{cases} \exp\big(-\lambda\Delta l\big)\sinh\Big(\dfrac{\lambda\Delta}{2}\Big) & \text{for} \quad l \neq 0 \\ 1 - \exp\Big(-\dfrac{\lambda\Delta}{2}\Big) & \text{for} \quad l = 0 \end{cases} \tag{70}$$

Meanwhile, the GΓ pdf is given by [13]

$$f_I(x) = \frac{\gamma\beta^\eta}{2\Gamma(\eta)}|x|^{\eta\gamma-1}\exp\big(-\beta|x|^\gamma\big) \tag{71}$$

One obtains the pmf of V for $l \neq 0$

$$P_V(l) = \frac{\gamma_c\Big(\eta, \beta\Big[\Delta\Big(|l| + \frac{1}{2}\Big)\Big]^\gamma\Big) - \gamma_c\Big(\eta, \beta\Big[\Delta\Big(|l| - \frac{1}{2}\Big)\Big]^\gamma\Big)}{2\Gamma(\eta)} \tag{72}$$

and

$$P_V(0) = \frac{\gamma_c\Big(\eta, \beta\big(\frac{\Delta}{2}\big)^\gamma\Big)}{\Gamma(\eta)} \tag{73}$$

where $\gamma_c(s,x) = \int_0^x t^{s-1}e^{-t}dt$ is the incomplete gamma function.

REFERENCES

[1] W. K. Pratt, *Digital Image Processing*. New york: Wiley, 1978.

[2] P. K. Singh, "Segmentation of regions in JPEG compressed medical images," in *Image Processing, International Conference on*, vol. 5, Oct. 2004, pp. 3483 – 3486.

[3] M. F. Sabir, H. R. Sheikh, R. W. Heath, and A. C. Bovik, "A joint source-channel distortion model for JPEG compressed images," *Image Processing, IEEE Transactions on*, vol. 15, no. 6, pp. 1349 – 1364, Jun. 2006.

[4] J. R. Price and M. Rabbani, "Biased reconstruction for JPEG decoding," *Signal Processing Letters, IEEE*, vol. 6, no. 12, pp. 297 – 299, Dec. 1999.

[5] Z. Daidi and I. Defee, "Pattern recognition in compressed DCT domain," in *Image Processing, International Conference on*, vol. 3, Oct. 2004, pp. 2031 – 2034.

[6] D. Y. Gang, C. Ying, W. L. Feng, and Y. Zheng, "Distributions of the DCT coefficient for watermark detection," in *Computer Application and System Modeling (ICCASM), International Conference on*, vol. 13, Oct. 2010, pp. 96 – 98.

[7] P. Sallee, "Model-based steganography," in *Digital Watermarking*, vol. 2939, Oct. 2003, pp. 254 – 260.

[8] M. C. Stamm and K. J. R. Liu, "Forensic detection of image manipulation using statistical intrinsic fingerprints," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 3, pp. 492 – 506, Sept. 2010.

[9] R. Reininger and J. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *Communications, IEEE Transactions on*, vol. 31, no. 6, pp. 835 – 839, Jun. 1983.

[10] T. Eude, R. Grisel, H. Cherifi, and R. Debrie, "On the distribution of the DCT coefficients," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, vol. 5, Apr. 1994, pp. 365 – 368.

[11] J. E. Eggerton and M. D. Srinath, "Statistical distributions of image DCT coefficients," *Computers and Electrical Engineering*, vol. 12, no. 3-4, pp. 137 – 145, Jan. 1986.

[12] F. Muller, "Distribution shape of two-dimensional DCT coefficients of natural images," *Electronics Letters*, vol. 29, no. 22, pp. 1935 – 1936, Oct. 1993.

[13] J.-H. Chang, J.-W. Shin, N. S. Kim, and S. K. Mitra, "Image probability distribution based on generalized gamma function," *Signal Processing Letters, IEEE*, vol. 12, no. 4, pp. 325 – 328, Apr. 2005.

[14] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *Image Processing, IEEE Transactions on*, vol. 9, no. 10, pp. 1661 – 1666, Oct. 2000.

[15] S. Nadarajah, "Gaussian DCT coefficients models," *Acta. Appl. Math.*, vol. 106, no. 3, pp. 455 – 472, 2009.

[16] T. H. Thai, R. Cogranne, and F. Retraint, "Statistical Model of Natural Images," in *International Conference on Image Processing*, Sep. 2012, pp. 2525 – 2528.

[17] D. Upham, "Steganographic algorithm Jsteg [online]. Available : http://zooid.org/paul/crypto/jsteg," 1993.

[18] C. Cachin, "An information-theoretic model for steganography," in *Information Hiding*, vol. 1525. Springer, Apr. 1998, pp. 306 – 318.

[19] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, 3rd ed. New york: Springer, 2005.

[20] X. Yu, Y. Wang, and T. Tan, "Model based steganalysis," in *Image Processing, International Conference on*, vol. 4, Oct. 2004, pp. 2625 – 2628.

[21] C. Zitzmann, R. Cogranne, L. Fillatre, I. Nikiforov, F. Retraint, and P. Cornu, "Hidden information detection based on quantized laplacian distribution," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, Mar. 2012, pp. 1793 – 1796.

[22] J. Kodovsky and J. Fridrich, "Quantitative structural steganalysis of Jsteg," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 4, pp. 681 – 693, Dec. 2010.

[23] K. Lee, A. Westfeld, and S. Lee, "Generalized category attack : Improving histogram-based attack on JPEG LSB embedding," in *Proc. 9th Int. Workshop Information Hiding*, vol. 4567. Springer, Jun. 2007, pp. 378 – 392.

[24] R. Böhme, "Weighted stego-image steganalysis for JPEG covers," in *Proc. 10th Int. Workshop Information Hiding*, vol. 5284. Springer, May 2007, pp. 178 – 194.

[25] T. Pevny, J. Fridrich, and A. Ker, "From blind to quantitative steganalysis," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 2, pp. 445 – 454, Apr. 2012.

[26] C. Scott, "Performance measures for Neyman-Pearson classification," *Information Theory, IEEE Transactions on*, vol. 53, no. 8, pp. 2852 –2863, aug. 2007.

[27] J. Nakamura, *Image Sensors and Signal Processing for Digital Still Cameras*. CRC Press, 2005.

[28] G. E. Healey and R. Kondepudy, "Radiometric ccd camera calibration and noise estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, pp. 267 – 276, 1994.

[29] R. Ramanath *et al.*, "Color image processing pipeline," *Signal Processing Magazine, IEEE*, vol. 22, no. 1, pp. 34 – 43, Jan. 2005.

[30] ——, "Demosaicking methods for bayer color arrays," *Journal of Electronic Imaging*, vol. 11, no. 3, pp. 306 – 615, Jul. 2002.

[31] W. Pennebaker and J. Mitchell, *Jpeg Still Image Compression Data*. Springer, 1992.

[32] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical poissonian-gaussian noise modeling and fitting for single-image raw-data," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1737 – 1754, Oct. 2008.

[33] M. Blum, "On the central limit theorem for correlated random variables," *Proceedings of the IEEE*, vol. 52, no. 3, pp. 308 – 309, Mar. 1964.

[34] S. Louhichi, "Rates of convergence in the central limit theorem for some weakly dependent random variables," *Theory of Probability and Its applications*, vol. 46, no. 2, pp. 297 – 315, 1998.

[35] L. Chen, "The rate of convergence in a central limit theorem for dependent random variables with arbitrary index set," *Annals of probability*, 1987.

[36] M.-S. Alouini, A. Abdi, and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over nakagami-fading channels," *Vehicular Technology, IEEE Transactions on*, vol. 50, no. 6, pp. 1471 – 1480, Nov. 2001.

[37] I. M. Ryzhik and I. S. Gradshteyn, *Tables of Integrals, Series, and Products*. United Kingdom: Elsevier, 2007.

[38] P. Bas, T. Filler, and T. Pevný, "Break our steganographic system — the ins and outs of organizing boss," in *Information Hiding, 13th International Workshop*, ser. Lecture Notes in Computer Science. Prague, Czech Republic: Springer-Verlag, New York, May 2011.

[39] B. Widrow, I. Kollar, and M.-C. Liu, "Statistical theory of quantization," *Instrumentation and Measurement, IEEE Transactions on*, vol. 45, no. 2, pp. 353 – 361, Apr. 1996.

[40] J. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, pp. 308 – 313, 1965.

[41] O. Dabeer, K. Sullivan, U. Madhow, S. Chandrasekaran, and B. Manjunath, "Detection of hiding in the least significant bit," *Signal Processing, IEEE Transactions on*, vol. 52, no. 10, pp. 3046 – 3058, Oct. 2004.

[42] R. Cogranne, C. Zitzmann, L. Fillatre, F. Retraint, I. Nikiforov, and P. Cornu, "Statistical decision by using quantized observations," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, Aug. 2011, pp. 1210 – 1214.

[43] C. Zitzmann, R. Cogranne, F. Retraint, I. Nikiforov, L. Fillatre, and P. Cornu, "Statistical decision methods in hidden information detection," in *Information Hiding*, vol. LNCS 6958, Prague, Czech Republic, May 2011, pp. 163 – 177.

[44] D. Birkes, "Generalized likelihood ratio tests and uniformly most powerful tests," *The American Statistician*, vol. 44, no. 2, pp. 163 – 166, 1990.

[45] A. Wald, "Some generalizations of the theory of cumulative sums of random variable," *The Annals of Mathematical Statistics*, pp. 287 – 293, 1945.

[46] A. D. Ker, "A fusion of maximum likelihood and structural steganalysis," in *Proceedings of the 9th international conference on Information hiding*, vol. 4567. Berlin, Heidelberg: Springer-Verlag, Jun. 2007, pp. 204 – 219.

[47] M. T. Hogan, N. J. Hurley, G. C. M. Silvestre, F. Balado, and K. M. Whelan, "Ml detection of steganography," in *Proc. SPIE Security, Steganography, and Watermarking of Multimedia Contents VII*, vol. 5681, Jan. 2005.

[48] T. Gloe and R. Bohme, "The 'dresden image database' for benchmarking digital image forensics," *Proc. ACM SAC*, vol. 2, pp. 1585–1591, 2010.

[49] T. Zhang and X. Ping, "A fast and effective steganalytic technique against jsteg-like algorithms," in *ACM SAC '03*, Mar. 2003, pp. 307 – 311.