Nonparametric Steganalysis of QIM Data Hiding using Approximate Entropy

Hafiz Malik\textsuperscript{a,b}, K. P. Subbalakshmi\textsuperscript{b}, and R. Chandramouli\textsuperscript{b}

\textsuperscript{a}Department of Electrical and Computer Engineering, University of Michigan - Dearborn, Dearborn, MI 48127, USA;
\textsuperscript{b}Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA

ABSTRACT

This paper proposes a nonparametric steganalysis method for quantization index modulation (QIM) based steganography. The proposed steganalysis method uses irregularity (or randomness) in the test-image to distinguish between the cover- and the stego-image. We have shown that plain-quantization (quantization without message embedding) induces regularity in the resulting quantized-image; whereas message embedding using QIM increases irregularity in the resulting QIM-stego image. Approximate entropy, an algorithmic entropy measure, is used to quantify irregularity in the test-image. Simulation results presented in this paper show that the proposed steganalysis technique can distinguish between the cover- and the stego-image with low false rates (i.e. $P_{fp} < 0.1$ & $P_{fn} < 0.07$ for dither modulation stego and $P_{fp} < 0.12$ & $P_{fn} < 0.002$ for QIM-stego).

Keywords: Steganography, Steganalysis, Quantization Index Modulation, Dither Modulation, Entropy, Complexity, Approximate Entropy, Algorithmic Entropy

1. INTRODUCTION

Quantization based data hiding schemes are based on the Costa’s seminal work\textsuperscript{1} to achieve the theoretical capacity of the Gaussian channel by using communication with side information. The ideal Costa scheme (ICS) gives theoretical upper bound for the capacity of all data hiding schemes under additive white Gaussian noise attack. However, the ICS requires an infinite length random codebook which makes ICS impractical.\textsuperscript{2} The practical realizations of ICS include, quantization index modulation (QIM), scalar Costa scheme (SCS), dither modulation (DM), and quantization projection (QP).\textsuperscript{2} The QIM-based data hiding schemes are commonly used for steganography due their high embedding capacity and controlled embedding distortion and robustness tradeoff.

Steganalysis refers to the act of analyzing a given multimedia data (e.g. images, video, audio etc.) for the presence of the hidden message, with limited or no access to information regarding the embedding algorithm used, hidden message, and the original media. Existing steganalysis techniques may be classified into passive- or active-steganalysis\textsuperscript{3} depending on whether the aim of the steganalyst is to detect the presence/absence of the hidden data only or to extract the hidden message as well.

To date, there appears to have been limited investigation of issues related to steganalysis of QIM-based steganography. Guillon et al in\textsuperscript{4} proposed a framework for steganalysis of message embedding using SCS by modeling QIM steganography as an additive noise channel. Recently Sullivan et al in\textsuperscript{5} have proposed steganalysis scheme for QIM steganography. The steganalysis scheme proposed in\textsuperscript{5} uses supervised learning for steganalysis of QIM steganography. Detection performance of the scheme proposed in\textsuperscript{5} is constrained by the limitations of learning-based steganalysis techniques, that is, 1) they cannot combat zero-day attack,\textsuperscript{6} 2) they require separate classifier training for every new steganographic algorithm, 3) their detection performance depend on the selection of features used to train the classifier, etc.\textsuperscript{7} Main contribution of this paper is to address aforementioned limitations of existing steganalysis schemes for QIM steganography by designing a nonparametric steganalysis framework under stego-only attack scenario, that is, only a stego-image is available to the steganalyst for steganalysis.

Further author information: Send correspondence to Hafiz Malik, E-mail: hafiz@umd.umich.edu, Tel.: 1 313 593 5677
This paper proposes a nonparametric steganalysis scheme for QIM steganography which uses measure of randomness (or irregularity) in the test-image to distinguish between the cover- and the stego-image. The proposed scheme exploits the fact that plain-quantization introduces regularity in the resulting quantized-image, whereas message embedding using QIM actually increases irregularity in the resulting stego-image. For example, consider a quantized-cover image, in the discrete cosine transform (DCT) domain, obtained using uniform quantizer with quantization step-size, $\Delta$. Due to high spatial correlation in natural images, coefficients in the neighboring blocks tend to map to a single quantization grid-point (say $k\Delta$, where $k$ is an integer) especially when quantization step-size is relatively large and/or in low-texture regions of the cover-image. The quantization process therefore induces regularity (or predictability) in the resulting quantized-image coefficients. Whereas, in case of message embedding using QIM, the corresponding coefficients will map to more than one quantization grid-points. Therefore, a sequence consisting of stego image coefficients, obtained by embedding message using QIM, tends to exhibit higher level of randomness than the sequence consisting of coefficients of the corresponding quantized-image. The relative irregularity (or complexity) in the sequence, generated using test-image, can be used to distinguish between the cover- and the stego-image.

In this paper, we assume stationary Generalized Gaussian model for the cover-image\(^8\) which is a special case of stochastic image model based on the Markov Random Field (MRF) model.\(^9\) We also assume gray-scale cover-image of size $N_1 \times N_2$, where $64 \leq N_1, N_2 \leq 256$ and equally probable binary message is embedded in discrete cosine transform (DCT) domain. Moreover, stego-only attack scenario is assumed which mean prior probabilities of the underlying source symbols are not known to the steganalyst. Therefore, entropy measures presented in\(^9–15\) cannot be used blindly to quantify time-series, generated using test-image, based on the estimated irregularity. This is mainly because entropy definitions\(^9–15\) in the literature on information theory cannot be related to each other.\(^13\)

For example, KS entropy is an algorithmic entropy measure\(^10,11\) which uses rate of information generation to classify deterministic dynamical systems. But the KS entropy methods fail to quantify time-series representing output of a stochastic or mixed processes.\(^13,14\) Moreover, KS complexity is very sensitive to small amount of noise or outliers. These inabilities of KS complexity to quantify irregularity in stochastic processes or noisy data can be attributed to its non-statistical framework used to calculate complexity in the time-series. Therefore, a blind application of KS complexity to practical time-series, i.e., DCT coefficient of the test-image, will only evaluate noise not the properties of the underlying sources. As we assume stationary Generalized Gaussian model for the cover-image\(^8\) here. In addition, KS complexity requires large amount of data (theoretically infinite sequence) to converge.\(^16\) Therefore KS complexity cannot be used to quantify smaller sequences, generated using test-images, based on their estimated KS complexities.

Shannon in\(^9\) proposed entropy as a measure of randomness (or irregularity) in the output of a probabilistic source that generates an infinite sequence of symbols. Entropy characterizes the irregularity of a given source by the probabilities of symbols and blocks of symbols. Ergodicity assumption about the source representing a Markov process, i.e., observed frequencies and prior probabilities of symbols and block of symbols are approximately equal, is exploited to estimate entropy of a finite sequence. Shannon’s probabilistic entropy\(^9\) requires prior probabilities of the underlying source symbols or block of symbols to estimate irregularity in a given sequence. However it cannot be used to quantify sequences generated from the test-image based on irregularity, as we assume stego-only attack scenario therefore probabilities of the symbols and block of symbols are not available to the steganalyst.

Pincus proposed an algorithmic entropy method, known as approximate entropy (ApEn) in\(^13–15\) to measure irregularity (or complexity) in the finite sequences when prior probabilities of symbols and blocks of symbols are missing. The ApEn makes no prior assumption on the sequence symbols or the source generating it. The ApEn is motivated by the Shannon’s information-theoretic entropy of a Markov process rather than by the conditional complexity of algorithmic information theory.\(^13–15\) The ApEn is very useful in discriminating finite sequences
based on their relative irregularity. Approximate entropy (ApEn) is a statistical tool designed to quantify irregularity in the time-series.\textsuperscript{13-15} Mathematically, ApEn is a natural information theoretical parameter, i.e., the rate of entropy, for an approximating Markov chain to a process.\textsuperscript{14,17} The ApEn provides both noise filtering and artifacts suppression capabilities through suitable filtering threshold selection.\textsuperscript{14} In addition, despite algorithmic similarities, the ApEn is not an approximate value of the KS entropy\textsuperscript{10,11} rather it is a family of statistics parameterized by the filtering threshold, $r$, and embedding dimension, $\xi$.\textsuperscript{13,14,18} Following salient features of the ApEn,

- ApEn is an algorithmic entropy measure,
- its robustness to noise as long as noise is below specified filtering threshold,
- it is applicable to short sequences, for example, it requires 100 or more points to estimate regularity with good confidence level,
- ApEn is finite for stochastic, noisy deterministic, and mixed processes,
- a change in the estimated ApEn corresponds to change in the complexity of the underlying process, and
- ApEn allows a direct computable alternative to severely noncomputable approach, i.e., KS complexity,

make it an attractive candidate to access irregularity in the real-world practical finite or periodic sequences. The proposed steganalysis scheme uses ApEn to estimate irregularity in the test-image. Estimated irregularity in the test-image is used to detect the stego-image.

Rest of the paper is organized as follow: Existing steganalysis framework generally used to attack QIM steganography is discussed in Section 2. An overview of the approximate entropy (ApEn) estimation is provided in Section 3. The proposed steganalysis scheme to attack QIM steganography and its detection performance to detect QIM-stego and DM-stego are provided in Section 4. Concluding remarks and future directions are discussed in Section 4.

### 2. STEGANALYSIS OF QIM STEGANOGRAPHY

A key issue in QIM steganalysis is to distinguish between the following cases:

1. the quantized-cover, $x_q$, (quantized image obtained using plain-quantization) and the QIM-stego, $x_{QIM}$, (stego-image obtained using QIM), and
2. the cover, $S$, and the DM-stego, $x_{DM}$, (stego obtained using DM).

Therefore, to design a parametric hypothesis test the probability mass functions of $S$, $x_q$, $x_{QIM}$, and $x_{DM}$ are required. Let $P_s(s)$ denote the probability mass function \textit{pmf} of the cover-image, $S \in \mathcal{R}$, in DCT domain. We assume $S$ to be an independent and identically distributed (i.i.d.) random vector. In the case of plain-quantization, the quantizer output, say $x_k$, is an integer multiple of the quantization step-size, $\Delta^*$, i.e., $x_k = k\Delta^*$. The probability of quantizer output is determined by the unquantized DCT coefficients, $s_i, i = 1, \cdots , N_k$, falling in the range $S_q(t) \triangleq (t - \frac{k\Delta^*}{2}, t + \frac{k\Delta^*}{2}]$, i.e.,

$$P_{x_q}(x_k) = \sum_{s_i \in S_q(t)} P_s(s_i) \quad (1)$$

where $N_k$ is the number of coefficients in the range $S_q(t)$, here $k$ is a positive integer.

In the QIM-stego case, two identical quantizers are used to encode binary message, $M \in \{0, 1\}^N$, where $N$ is the message size. Each quantizer is designed with a step-size $\Delta = 2\Delta^*$ and is offset (shifted) from the other by $\Delta/2$. That is, $Q_0(x) = Q_1(x) \pm \Delta/2$, here $Q_0(\cdot)$ and $Q_1(\cdot)$ denote quantizers used to embed message bit '0' and
For QIM the probability of a given output, $x_k$, can be expressed as,

$$P_{x_{QIM}}(x_k) = \frac{1}{2} \sum_{s_i \in S_{QIM}(t)} p_s(s_i)$$  \hspace{1cm} (2)

where $S_{QIM}(t) \triangleq (t - \Delta_k/2, t + \Delta_k/2]$ and $x_k = k\Delta$. Here we assume equally probable message symbols, that is, $P[r = 0] = P[r = 1] = \frac{1}{2}$.

In case of dither modulation, two dither quantizers are used to embed message using dither modulation. A dither quantizer is obtained by adding (or subtracting) a dither value $d_{ui}$ to the quantizer output, $x_k$, where $d_{ui}$ is uniformly distributed noise over $[-\Delta/4, \Delta/4]$. Therefore, the quantizer output covers whole range of the cover-image. In this range, $P_{x_{DM}}(d_{ui}) = 2\epsilon/\Delta$, where $\epsilon$ is the granularity of the data. Data hiding based on DM can be expressed as,

$$x_i(S, M) = Q_m(s_i + d_{ui}) - d_{ui}, \quad i = 0, 1, \ldots, N - 1.$$  \hspace{1cm} (3)

where $S_{DM}(t) \triangleq (t - \Delta_k/2, t + \Delta_k/2]$.

The stego detection rule based on the likelihood ratio test (LTR), for both embedding scenarios, can be expressed as,

$$L(x) \triangleq \frac{P_{x_{QIM}}(x)}{P_{x_{QIM}}(x)} \geq \tau \quad \text{(to detect QIM-stego)}$$  \hspace{1cm} (4)

$$\triangleq \frac{P_{x_{DM}}(x)}{P_{x_{QIM}}(x)} \geq \tau \quad \text{(to detect DM-stego)}$$  \hspace{1cm} (5)

where the decision threshold, $\tau$, can be minimized using Neyman-Pearson rule which maximizes the probability of detection, $P_d$, for a given probability of false alarm, $P_f$.\textsuperscript{19}

Substituting $P_{x_{QIM}}(x)$ and $P_{x_{QIM}}(x)$ from Eq. (1 & 2) in Eq. (4) yields

$$L(x) = \prod_{i=1}^{N} \left( \frac{1}{2} \frac{\sum_{s_i \in (x_i - \Delta/2, x_i + \Delta/2]} p_s(s_i)}{\sum_{s_i \in (x_i - \Delta/4, x_i + \Delta/4]} p_s(s_i)} \right)$$  \hspace{1cm} (6)

Here Eq. (6) shows that the likelihood statistic is a function of the cover pmf, $p_s(s)$, and under stego-only attack scenario $p_s(s)$ is not available at the stego detector. Therefore, parametric detection based on Neyman-Pearson rule cannot be used to detect the QIM-stego image.
Similarly, to detect DM-stego substituting $P_{\text{DM}}(x)$ from Eq. (3) in Eq. (5) yields

$$L(x) = \prod_{i=1}^{N} \left( \frac{1}{\Delta} \sum_{s \in \{x_i - \Delta/2, x_i + \Delta/2\}} P_s(s) \right)$$

(7)

Eq. (7) shows that the likelihood statistic is also a function of the cover pmf, $P_s(s)$, therefore parametric detector cannot be used for DM-stego detection either. An important observation however can be made from Eq. (6 & 7) that message embedding using QIM or DM introduce smoothness in the resulting stego pmf. Relative smoothness in the pmf of the test-image can be using to distinguish between the cover and the stego. To highlight this claim further, we analyzed empirical pmfs of the quantized-cover and the QIM-stego images, here empirical pmf is estimated using histogram. The empirical pmfs of DCT coefficients of the QIM-stego for $\Delta = \{0.5, 4, 8\}$ are plotted in Fig. 2. In addition, to differentiate smoothing effect on the cover pmf due two different types of quantization further, the empirical pmf of the quantized-cover and the QIM-stego are plotted in Fig. 3. Some

![Figure 2](image1.png)

Figure 2. Empirical pmf based on histogram of DCT coefficients of the cover (top-left) and quantized DCT coefficients of the QIM-stego obtained with $\Delta = \{0.5, 4, 8\}$ (top-right and the bottom-row)

![Figure 3](image2.png)

Figure 3. Empirical pmf of the quantized-cover (top-row) and the corresponding QIM-stego (bottom-row) of the experimental observations on the difference between the QIM-stego and the quantized-cover images based on their empirical pmfs are summarized below.

Firstly, we note that the quantization (with and without message embedding) introduces smoothness in the pmf of the resulting quantized images. It can be observed from Fig. 2 that as $\Delta$ increases the empirical pmf of the resulting QIM-stego tends to change from a super-Gaussian like pmf (e.g. Laplacian pmf) to a more Gaussian like pmf. Secondly, quantization step-size, $\Delta$, controls the amount of smoothness introduced in the pmf of the quantized-image. Finally, quantization with message embedding (e.g. QIM) introduces more smoothness than the plain-quantization. It can be observed from Fig. 3 that for same $\Delta$, the QIM introduces more smoothness.
than plain-quantization. Moreover, for large $\Delta$ ($\Delta \geq 4$) message embedding using QIM splits the peak around zero in the cover pmf into three peaks (e.g. peaks $P_{-\Delta}$, $P_0$, and $P_{\Delta}$ around $-\Delta$, 0, and $\Delta$ respectively), which can be used to distinguish between the quantized-cover and the QIM-stego. However, such visual attack might not yield consistent result for smaller quantization step-size or if the cover-image exhibit smoothly varying pmf.

Learning-based steganalysis techniques have been proposed in the past to distinguish between the quantized-cover and the QIM-stego, but as noted earlier, there are some inherent disadvantages with these steganalysis schemes.

To address limitations of learning-based steganalysis schemes for QIM steganography, a nonparametric steganalysis scheme based on measure of randomness in the test-image is proposed here. The proposed scheme exploits relative randomness in the test-image to distinguish between the cover- and the stego-image. As discussed earlier that the QIM-stego exhibits more randomness than the quantized-cover, though both quantized images are obtained using same quantization step-size. This fact is illustrated in Fig. 4.

![Figure 4. Illustration of quantization noise: quantized-cover and quantization noise (left); QIM-stego and the corresponding quantization noise (right)](image)

It can be observed from Fig. 4 that the distortion due to message embedding using QIM is relatively more irregular (random) than the distortion due to plain-quantization (especially in low-texture regions). This implies that coefficients of the quantized-cover image are relatively more predictable (regular) than the corresponding coefficients in the QIM-stego. The proposed steganalysis scheme uses relative irregularity in the test-image to distinguish between the cover, $(S, x_q)$, and the stego, $(x_{QIM}, x_{DM})$, images. The proposed schemes uses $ApEn$ to access randomness in the test-image. An algorithm to calculate $ApEn$ from a finite-length sequence and its mathematical interpretation are discussed next.

### 2.1 Approximate Entropy Estimation

$Approximate$ $entropy$ is a regularity statistic that quantifies irregularity or fluctuations in a time-series, $\{x\}_n^1$, where $n$ is the number of observations of the time-series. The $ApEn$ reflects the likelihood that similar patterns of observations will not be followed by additional observations. A time-series containing many repetitive patterns (e.g. a regular sequence) exhibits a relatively small $ApEn$ value, whereas a time-series consisting of less predictable patterns (or a more irregular sequence) exhibits higher $ApEn$ value. A detailed analysis of algorithm for computing $ApEn$ and its statistical properties can be found in $^{13-15,20-22}$ and references therein.

#### 2.1.1 Definition of ApEn

Given a time-series sequence, $\{x\}_n^1$, consisting of $n$ measurements equally spaced in time i.e. $x_1, x_2, \cdots, x_n$. For a fixed-positive integer $\xi$ and a positive real number $r$, consider embedding vectors $u_{(1)}, u_{(2)}, \cdots, u_{(n-\xi+1)}$ in $\mathbb{R}^{\xi}$, where $u_{(i)} = [x_i, x_{i+1}, \cdots, x_{i+\xi-1}]$. Let define the correlation measure, $C^\xi_i(r)$, for every $i, 1 \leq i \leq n-\xi+1$,

$$C^\xi_i(r) = \frac{\text{[# of } j \leq n-\xi+1 \mid d(u_i, u_j) \leq r]}{n-\xi+1}$$

(8)
where \(d(\mathbf{u}_i, \mathbf{u}_j)\) is the \(L_\infty\) norm between vectors \(\mathbf{u}_i\) and \(\mathbf{u}_j\), which can be expressed as,

\[
d(\mathbf{u}_i, \mathbf{u}_j) = \max_{k=1, \cdots, \xi} |(u(i + k - 1) - u(j + k - 1))|
\]

(9)

here the quantity \(C_\xi^k(r)\) is a fraction of patterns of length \(\xi\) that resemble the pattern of same length that begins at index \(i\). In other words, quantity \(C_\xi^k(r)\) measures regularity (or frequency) of patterns similar to a given pattern of window length \(\xi\) and a tolerance \(r\).

The approximate entropy, \(ApEn(\xi, r, n)\), of a sequence \(\{x\}_n^1\), with parameters \(\xi\), \(r\), and \(n\) is defined as,

\[
\text{ApEn}(\xi, r, n) = \left[ \Phi^\xi(r) - \Phi^{\xi+1}(r) \right],
\]

(10)

where

\[
\Phi^\xi(r) = \frac{\sum_{i=1}^{n-\xi+1} \ln C_i^\xi(r)}{n - \xi + 1}
\]

(11)

and,

\[
\Phi^{\xi+1}(r) - \Phi^{\xi+1}(r) = E_i \{\log (Pr[|(u(j + \xi) - u(i + \xi)| \leq r]) | (|u(j + k) - u(i + k)| \leq r))\}
\]

(12)

where \(k = 0, 1, \cdots, \xi - 1\), \(E_i\) denotes average over \(i\), and \(Pr[\cdot]\) is conditional probability.

The \(ApEn(\xi, r, n)\) measures likelihood that run of patterns that are close for \(\xi\) observations remain close during next incremental comparisons. Smaller value of \(ApEn\) implies regularity in the time-series, that is, similar patterns are highly predictable from additional similar measurements. Whereas, a large value of \(ApEn\) indicates that the underlying time-series is highly irregular. In addition, for a given application \(ApEn(\xi, r, n)\) should be consider as a family of statistics and for time-series comparison fix values of \(\xi\) and \(r\) should be used.

**3. STEGANALYSIS USING \(ApEn\)**

Measure of irregularity in the test-image is used to detect the stego-image. Irregularity in the test-image is measured in terms of the \(ApEn\) estimate. In order to calculate \(ApEn\) from the test-image (\(S\), \(x_n\), \(x_{QIM}\), or \(x_{DM}\)) using the \(ApEn(\xi, r, n)\) algorithm outlined in the previous Section, the test-image must be transformed into finite sequences. To this end, the test-image is segmented into non-overlapping blocks, each of size 8x8 pixels, and calculated two-dimensional (2D) DCT for each block. Each block in DCT domain is then converted into one-dimensional (1D) vector using zigzag ordering (commonly used during baseline JPEG compression\(^{23}\)). These 1D blocks of the test-image are used to generate 64 sequences, \(x_n^i, i = 0, \cdots, 63\), each of length \(n\). Here \(n = \left\lfloor \frac{N}{8} \right\rfloor \times \left\lfloor \frac{N}{8} \right\rfloor\) where \([x]\) denote the largest integer not exceeding \(x\). Fig. 5 illustrates the finite-length sequence generation process from the test-image used in the paper.

![Figure 5. Finite-length sequence generation from the test-image](image)

Finite sequences are then analyzed to estimate randomness in the test-image. To estimate randomness (or irregularity) in the test-image, finite sequences, \(x_n^i, i = 1, \cdots, 63\) are analyzed using Eq. (10) which generates 63-dimensional vector of \(ApEn\) estimates, i.e.,

\[
ApEn_i = ApEn(x_n^i, \xi, r, n), i = 1, \cdots, 63
\]

(13)

Estimated \(ApEn\) vector, \(ApEn\), represents randomness in the test-image which is used to distinguish between the cover- and the stego-image.
3.1 Steganalysis of QIM-stego using $ApEn$

To investigate effect of message embedding using QIM on irregularity in the QIM-stego image, the $ApEn(\xi, r, n)$ is calculated from $S$, $x_q$, and $x_{QIM}$. To this end, two quantized images, one quantized-cover and the other QIM-stego, were generated from an uncompressed cover-image, $S$, of size 256x256 using uniform quantizers with $\Delta = 2$ and $\Delta^* = 1$. To obtained quantized images, we used image number 47 of the image database downloaded from\textsuperscript{2} as a cover-image. The cover-image was resized to 256x256 and converted to gray-scale. To embed binary message into the gray-scaled cover-image using QIM, the cover-image was first segmented into non-overlapping blocks, each of 8x8 pixels and then 2D DCT transform was applied to each block followed by message embedding using QIM. A 64 KB binary message was embedded in the cover-image using binary QIM which yielded the QIM-stego image. Similarly, the corresponding quantized-cover image was obtained. Both the quantized-cover and the QIM-stego images were then transformed into 64 1D sequences each. The $ApEn$ was estimated from 1D sequences corresponding to AC coefficients of $S$, $x_q$, and $x_{QIM}$ in DCT domain with parameter settings $\xi = 4$ and $r = 0.1 \times \sigma_x$. Fig. 6 shows plots of the estimated $ApEn$ from $S$, $x_q$, and $x_{QIM}$ in DCT domain. In Fig. 6 the horizontal axis represents sequence number (or AC coefficients number) and vertical axis represents estimated $ApEn$ (or level of randomness each sequence).

![Figure 6. Plots of the estimated $ApEn$ from $S$, $x_q$, and $x_{QIM}$ in DCT domain](image)

Following observations can be made from Fig. 6:

- The estimated $ApEn$ from $S$ remains approximately constant for all sequences which implies that all sequences exhibit approximately same level of randomness.
- In general, estimated $ApEn$ from $x_q$ and $x_{QIM}$ decreases from low to high frequency. Here low and high frequency correspond to sequence number 1 to 32 and sequence number 32 to 63, respectively.
- For both quantized-images, i.e. $x_q$ and $x_{QIM}$, the estimated $ApEn$ decrease at a higher rate in the low frequency-coefficients than in the high frequency-coefficients.
- The estimated $ApEn$ from $x_{QIM}$ has lower gradient in both frequency regions than the estimated $ApEn$ from $x_q$.
- Let $m_{low}$ and $m_{high}$ denote gradient of the estimated $ApEn$ in low and high frequency-coefficients respectively, and change in the gradient, $\delta m$, of the estimated $ApEn$ where $\delta m$ is defined as,

$$
\delta m = (m_{low} - m_{high})/m_{low} \times 100
$$

(14)

The $\delta m$ for the quantized-cover is well below 50% (36% to be exact) and $\delta m$ is well above 50% (85% to be exact) for the QIM-stego.
- For QIM-stego, the estimated $ApEn$ is approximately constant for high frequency-coefficients.
- The estimated $ApEn$ from the QIM-stego is higher than the $ApEn$ estimated from the quantized-cover in high frequency-coefficients which implies that in high frequency-coefficients the QIM-stego is relatively more irregular than the corresponding quantized-cover. This higher $ApEn$ value in the QIM-stego than the corresponding quantized-cover can be attributed to the randomness in the embedded message $M$. 
These observations indicate that variation in the gradient of the estimated ApEn from low to high frequency-coefficients along with ApEn value in the high frequency-coefficients can be used to distinguish between the quantized-cover and the QIM-stego. The proposed steganalysis scheme however uses relative change in the gradient, $\delta m$, from low to high frequency-coefficients to detect QIM-stego image. Block diagram of the proposed steganalysis scheme used to attack QIM-steganography is given in Fig. 7.

![Block diagram of the proposed steganalysis scheme to distinguish between the quantized-cover and the QIM-stego](image)

**Figure 7.** Block diagram of the proposed steganalysis scheme to distinguish between the quantized-cover and the QIM-stego

### 3.1.1 Experimental Results

Detection performance of the proposed scheme was tested for two different message embedding scenarios,

- **all frequency embedding** (AFE), that is, message is embedded into all AC coefficients of 8x8 blocks of the cover-image (in DCT domain), and
- **mid-frequency embedding** (MFE), that is, message is embedded into mid-frequency AC coefficients i.e. AC coefficient number 5 through 32 of each 8x8 block of the cover-image.

Detection performance of the proposed steganalysis scheme is evaluated in terms of false rates, that is, false positive rate, $P_{fp}$, and false negative rate, $P_{fn}$. Image database downloaded from\(^{24}\) was used to evaluate performance of the proposed steganalysis scheme for QIM-stego detection. This image database\(^{24}\) contains 1338 uncompressed color images, however results presented in this paper are based on gray-scaled versions of first 1000 images of the database.\(^ {24}\) Moreover, these 1000 images of the database\(^ {24}\) were resized to 256x256. Two-thousand QIM-stego images were obtained by embedding 2000 random messages into first 1000 images of the database using QIM with $\Delta = 2.0$ (1000 QIM-stego images using AFE and 1000 QIM-stego images using MFE). Similarly, 1000 quantized-images were obtained by quantizing these 1000 gray-scale images using $\Delta^{*} = 1$. The proposed steganalysis scheme was then applied to resulting 3000 quantized images (1000 QIM-stego using AFE, 1000 QIM-stego using MFE, and 1000 quantized-cover). Detection performance the proposed steganalysis scheme with the decision threshold, that is, relative change in the gradient $\delta m = 50\%$ and ApEn estimation parameters $\xi = 2$, $r = 0.1 \times \sigma_x$, $n = 1024$, is listed in Table 1. Simulation results for MFE listed in the Table 1 are based on abrupt variation (sudden jump(s)) in the estimated ApEn vector from the test-image.

### 3.2 Steganalysis of the DM-Stego using ApEn

Similarly, to steganalyze the DM-stego based on randomness in the test-image image, the $\text{ApEn}(\xi, r, n)$ is calculated from the finite sequences obtained from $S$, and $x_{DM}$. The DM-stego was generated by segmenting the cover-image into non-overlapping blocks, each of 8x8 pixels, followed by 2D DCT transform. Binary message was embedded using DM (dither modulation). The DM-stego, $x_{DM}$, was obtained using $\Delta = 2$ and a dither vector $d_u \sim \mathcal{U}(0, 2^2/12)$. Fig. 8 shows plots of the estimated ApEn from the gray-scale cover-image, $S$, (image number 47 of the database downloaded from\(^ {24}\)) and the corresponding $x_{DM}$ (in DCT domain) with $\xi = 4$, $r = 0.1 \times \sigma_x$, $n = 1024$.

It can be observed from Fig. 8 that message embedding using DM makes the estimated ApEn vector smoother (or reduces variance) but it is still hard to distinguish between the cover and the DM-stego based on their estimated ApEn vectors and their corresponding variances. As under stego-only attack scenario only one plot of the estimated ApEn is available to the steganalyst to make decision. Therefore, estimated ApEn vector from the test-image cannot be used to make decision the DM-stego with high confidence. However, we have observed through extensive experimental results that the DM steganography actually increases variance of the DM-stego
coefficients. To amplify difference between the estimated \( ApEn_s \) and \( ApEn_{x_{DM}} \) from \( S \) and \( x_{DM} \) respectively, we normalized the estimated \( ApEn \) vector from the test-image by its variance, i.e., \( nApEn_x = ApEn_x / \sigma_x^2 \).

The estimated normalized \( ApEn \), \( nApEn \), vector still cannot be used to distinguish between the cover and the DM-stego, as still only one vector is available to the steganalyst to determine whether the test-image is a cover or a stego. To get away with this, a second test-image (say DM\(^2\)-stego) is generated by reprocessed the test-image. The reprocessed test-image is obtained by encoding an arbitrary message \( M \) using DM with an arbitrary dither vector \( \mathbf{d}_u \) and arbitrary step-size, \( \Delta \). It has been observed that estimated \( nApEn \) vectors from the DM\(^2\)-stego and the test-image are very close in 63-dimensional space if the test-image is a DM-stego image and are far apart otherwise. To illustrate this claim, we estimated \( nApEn \) vectors from \( S \), \( x_{DM} \), and \( x_{DM(2)} \). The plots of the estimated \( nApEn \) vectors from \( S \), \( x_{DM} \), and \( x_{DM(2)} \) are given in Fig. 9.

It can be observed from Fig. 9 that the estimated \( nApEn \) vectors from the DM\(^2\)-stego and the DM-stego are very close, and the estimated \( nApEn \) vectors from the cover and the DM-stego are far apart. This observation reveals that the mutual distance between the estimated \( nApEn \) vectors from the test-image and its corresponding reprocessed version (i.e. DM\(^2\)-stego) can be used to distinguish between the cover and the DM-stego using simple binary hypothesis that if mutual distance between the estimated \( nApEn \) vectors from the test-image and DM\(^2\)-stego is below certain threshold then the test-image is identified as a DM-stego and the cover otherwise.

To summarize the proposed steganalysis method to detect DM-stego, the test-image is reprocessed to obtained DM\(^2\)-stego by embedding an arbitrary message, \( M \), using DM with arbitrary parameters \( \mathbf{d}_u \) and \( \Delta \). The \( nApEn \) vectors are estimated from both the test-image and the corresponding DM\(^2\)-stego. The Euclidian distance, \( D \), between the estimated \( nApEn \) vectors from the test-image and the DM\(^2\)-stego, defined as,

\[
D = \sqrt{\sum_{i=1}^{63} (nApEn_{(t)}^{(i)} - nApEn_{(DM(2))}^{(i)})^2},
\]
is then used to distinguish between the cover and the DM-stego. Here, $nApEn^{(1)}$ and $nApEn^{(DM(2))}$ denote estimated normalized ApEn vectors estimated from the test-image and the corresponding DM(2)-stego image, respectively. Block diagram of the proposed steganalysis scheme to distinguish between the cover and the DM-stego is given in Fig. 10.

![Figure 10. Block diagram of the proposed steganalysis scheme to distinguish between the cover and the DM-stego](image)

### 3.2.1 Experimental Results

Detection performance of the proposed steganalysis scheme to attack DM-stego is also evaluated for the same image database which was used to evaluate performance of the QIM-stego detection. Two-thousand DM-stego images were obtained by embedding 2000 random messages into first 1000 images of the database using DM with $\Delta = 2.0$ and an independent and uniformly distributed dither vector $d_u$. Here, once again these 1000 images were resized to 256x256 and transformed to gray-scale for message embedding. The proposed steganalysis scheme was then applied to resulting 3000 test-images (2000 DM-stego images and 1000 cover-images in the DCT domain). During detection process, each test-image was reprocessed to obtain the corresponding DM(2)-stego image by embedding an independent message $M$ using randomly selected quantization step-size in the range 1.0 through 5.0, $\Delta \in \{1.0, 5.0\}$, and an independent dither vector $d_u$. The $nApEn$ vectors were estimated from each test-image and its corresponding DM(2)-stego image using $ApEn$ parameter settings, $\xi = 2$ and $r = 0.1 \times \sigma_x$. Experimental results when the proposed steganalysis scheme was applied to 3000 test-images are listed in Table 1.

Simulation results to detect DM-stego listed in Table 1 are based on quantization step-size randomly chosen in the range 1.0 through 4.0, i.e. $\Delta \in (1.0, 4.0)$, decision threshold $D = 2.0$.

### Table 1. Detection Performance

<table>
<thead>
<tr>
<th></th>
<th>$X_q$ vs $X_{QIM}$</th>
<th>$S$ vs $X_{DM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error AFE</td>
<td>0.12</td>
<td>0.002</td>
</tr>
<tr>
<td>Error MFE</td>
<td>0.08</td>
<td>0.001</td>
</tr>
<tr>
<td>$P_{FP}$</td>
<td>0.1</td>
<td>0.07</td>
</tr>
<tr>
<td>$P_{FN}$</td>
<td>0.05</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Experimental results listed in Table 1 show that the proposed nonparametric steganalysis scheme can distinguish between the quantized-cover (cover) and the QIM-stego (DM-stego) and DM-stego images with very low false rates.

### 4. CONCLUSION

This paper presents a novel steganalysis scheme for QIM steganography. The proposed steganalysis scheme is non-learning based therefore capable of addressing limitations of learning-based steganalysis schemes. The proposed scheme uses irregularity in test-image to distinguish between the quantized-cover and the QIM-stego. The proposed scheme uses normalized irregularity to distinguish between the cover and the DM-stego. Experimental results presented in this paper show that the proposed steganalysis scheme can successfully distinguish between the cover and the stego. We are currently investigating performance of the proposed steganalysis scheme to detect stego images carrying smaller messages embedded using non-sequential embedding.
REFERENCES