Application of Error Control codes in Steganography

H. Kostadinov[1], N. L. Manev[2]

[1]Institute of Mathematics and Informatics, BAS

[2]Institute of Mathematics and Informatics, BAS and University of Structural Engineering & Architecture "Lyuben Karavelov"

ACCT2014, September 7 - 13, 2014, Svetlogorsk, Russia

Basic Terms

- **Staganography** is a tool that assures undetectable communications between partners. It is a technique of altering the digital object in undetectable manner, that is, no one but the intended recipient to be able to detect this altering.

- **Watermarking** is a practice of imperceptibly altering a digital object to embed a message about that work. The goal of watermarking is to prevent piracy or to prove the ownership.

General Description

Embedding process

$$\mathcal{E}: \left| \begin{array}{lcl} (\mathfrak{P}, \mathfrak{T}, \mathfrak{K}) & \rightarrow & \mathfrak{P} \\ (\mathbf{c}_0, \mathbf{m}, \mathbf{k}) & \rightarrow & \mathbf{c}_m = \mathcal{E}(\mathbf{c}_0, \mathbf{m}, \mathbf{k}) \end{array} \right. ,$$

where $\mathbf{c}_0$, the *cover* digital object, is an element of the set $\mathfrak{P}$ of possible digital objects (e.g., images), $\mathbf{m} \in \mathfrak{T}$ is the message intended for embedding, and $\mathbf{k}$ is a key randomly taken from the space of keys $\mathfrak{K}$.

Decision function (based on a chosen ***detection metric***):

$$\mathbf{m}' = \mathcal{D}(\mathbf{c}_{mn}), \qquad \text{or} \quad \mathbf{m}' = \mathcal{D}(\mathbf{c}_{mn}, \mathbf{c}_0), \quad \text{if } \mathcal{D} \text{ needs the original.}$$

where $\mathbf{c}_{mn}$ is the received (eventually distorted) copy of $\mathbf{c}_m$.

In many cases the distortion can be considered as an additive noise $\mathbf{n}$, i.e, $\mathbf{c}_{mn} = \mathbf{c}_m + \mathbf{n}$. Very often in practice $\mathbf{c}_{mn} = \mathbf{c}_m$.

Where to embed?

Watermarks (messages) can be embedded by adding either in spatial domain, or in frequency domain.

- **Spatial domain** is the term used for the standard form of a digital object that represents the media product. For example, a gray-scale image is represented by a matrix whose entries are integers in [0, 255].
- **Frequency domain** is the term used for the object obtained after applying to the cover object discrete cosine, discrete Fourier, Hadamard, wavelet, or such other transformations.
  The message is embedded by altering the components of that transformed object, and then applying the reverse transformation.

Embedding Process

Embedding process is a composition of two mappings:

1. $\mathcal{M}$, transforms the cover image into a matrix over Z_q or $GF(q)$

$$\mathcal{M}: \ (\mathfrak{P}, \mathfrak{T}, \mathfrak{K}) \ \rightarrow \ \mathfrak{M}.$$

   (We denote the set of such matrices by $\mathfrak{M}$.)

2. The second transformation realizes the embedding algorithm (based on the use of error control codes in our case).
   It is referred to as ***embedding function*** and we shall keep the notation $\mathcal{E}()$ for it.

$\mathcal{M}$ depends on $\mathbf{c}_0$ and the chosen key $\mathbf{k}$. After careful analysis of the image some pixels are market as "wet" pixels which are not used. Only the rest ("dry") pixels are modified by the embedding function. The order of modifying these pixels also depends on $\mathbf{k}$.

The Goal of the Talk

- This talk presents a part of our research on the use of the error-control codes in watermarking and steganography.

- Herein we discuss two methods of information embedding in spatial domain:

Two Types of Embedding Algorithms

A: **Syndrome embedding** (also known as matrix embedding)
The parties agree on a parity-check matrix $\mathbf{H}$ of a linear code and the secrete message is extracted as a sequence of syndromes (with respect to $\mathbf{H}$) from $\mathbf{c}_m$.

B: **Method based on pseudo-noise patterns and the erasure capability of error correcting codes.**

$$\mathbf{c}_m = \mathcal{E}(\mathbf{c}_0, \mathbf{m}) = \mathbf{c}_0 + \alpha \mathbf{w}_m,$$

where ***message pattern*** $\mathbf{w}_m$ is a function of $\mathbf{m}$ and the set of predefined reference patterns $\mathbf{W} = \{\mathbf{w}_{r1}, \mathbf{w}_{r2}, \ldots, \mathbf{w}_{rk}\}$.
$\mathbf{W}$ consists of matrices whose entries have a given probability distribution, most often normal or uniform distribution. These patterns are pair-wise orthonormal according to the chosen detection metric

$$\delta(\ ,\ ) : \mathfrak{M} \times \mathfrak{M} \to \mathbb{R}.$$

In fact it is enough $\delta(\mathbf{w}_{ri}, \mathbf{w}_{rj})$ to be relatively small.

Algorithm

Introduced by Grandall. Many authors have studied this topic, e.g., Galand, Kabatiansky, Westfeld, Moulin, Koetter, Rifa et al. Mainly binary codes are used.

Let $\mathbf{H}$ be a $r \times n$ parity-check matrix of a linear code over $\mathbb{Z}_q$ or $GF(q)$. Let $\mathbf{D}$ be a $n \times N$ matrix over $\mathbb{Z}_q$ or $GF(q)$ obtained from the target image by the transformation $\mathcal{M}$. Let the message (randomized and eventually encrypted) be also transformed into a $r \times N$ matrix $\mathbf{m}$ over $\mathbb{Z}_q$ or $GF(q)$.

1. Compute $\mathbf{S} = \mathbf{m} - \mathbf{HD}$, i.e., $\mathbf{m} = \mathbf{S} + \mathbf{HD}$
2. Find $n \times N$ matrix $\mathbf{E}$ such that $\mathbf{S} = \mathbf{HE}$
3. Compute $\mathbf{V} = \mathbf{D} + \mathbf{E}$
4. Construct an image $\mathbf{c}_m$ such that $\mathcal{M}(\mathbf{c}_m, \mathbf{k}) = \mathbf{V}$ and send $\mathbf{c}_m$.

The receiver

1. Determines $\mathbf{V} = \mathcal{M}(\mathbf{c}_w, \mathbf{k})$
2. computes $\mathbf{HV} = \mathbf{HD} + \mathbf{HE} = \mathbf{HD} + \mathbf{S} = \mathbf{m}$

The Use of q-ary Codes

Let q' be an integer less than 256 (or 2^b, $b = 12, 16$). The matrix $\mathbf{D}' \equiv \mathbf{c}_0 \pmod{q'}$ is the start point of the embedding procedure. The used q-ary codes are with $q \leq q'$. If binary codes are used then it is convenient q' to be a power of 2. In our experiments we have tested even the extremal case $q = q' = 256$.

Embedding efficiency in the case of syndrome encoding using a code with $r \times n$ parity check matrix is

$$E_f = \frac{r \log_2 q'}{n} \quad \text{bits/pixel}$$

The Use of q-ary Codes

We apply codes over $\mathbb{Z}_q$ that correct errors of type $\pm e$ for small integer e. We have constructed many such codes for integer coded modulation. These codes have simple decoding algorithms. Soft decoding (trellis) can be applied to them, too.

The representative of the coset that corresponds to a given syndrome has to be chosen with the minimum possible Lee weight. Also its entries have to be with minimum absolute value.
For example, the $\mathbb{Z}_4$-vectors $(0, 3, 3)$ is preferred to $(2, 0, 0)$ ($3 = -1$ in $\mathbb{Z}_4$).

Our approach simplifies the process of determining the matrix $\mathbf{E}$. The experiments show that the codes with small parameters, and thus with a simple decoding, are very effective. Without the original even with the use of a specialized statistical software is very difficult to disclosed the modification.

Pseudo-noise patterns embedding

A-1. Let C be a binary $[n, n-r]$ linear code. Encode the source message into a binary sequence $\mathbf{m}$.

A-2. Starting with a given state (used as password) of the random number generator generate t reference patterns of size $a \times b$: $\mathbf{W} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_t\}$.

A-3. Divide (in some way) the cover work into N blocks, $\mathbf{c}_1, \ldots, \mathbf{c}_N$, each of size $a \times b$.

A-4. (optional) Replace the set $\mathbf{W}$ by the set of patterns $\{\mathbf{h}_i\}$ which are orthogonal to all blocks $\mathbf{c}_1, \ldots, \mathbf{c}_N$.

Pseudo-noise patterns embedding

A-5. In each block $\mathbf{c}_j$, $j = 1, 2, \ldots, N$, embed t bits of the message sequence $\mathbf{m}$ by

$$\mathbf{c}_{jw} = \mathbf{c}_j + \frac{\alpha_j}{\sqrt{t}} \left(\epsilon_1 \mathbf{w}_1 + \epsilon_2 \mathbf{w}_2 + \cdots + \epsilon_t \mathbf{w}_t \right),$$

where $\epsilon_i = \begin{cases} 1, & m_{ji} = 1 \\ -1, & m_{ji} = 0 \end{cases}$.

The scale constant α_j controls the trade-off between visibility and robustness of the hiding data.

$$E_f = \frac{(n-r)tN}{nab} \text{ bits/pixel.}$$

Detection and Decoding

A-6. The recipient divides the received image into N blocks $\{\tilde{\mathbf{c}}_j\}$ and knowing the reference patterns (or the key to generate them) calculate

$$\delta(\tilde{\mathbf{c}}_j, \mathbf{w}_i),\ j = 1, 2, \ldots, N,\ \ i = 1, \ldots, r,$$

where $\delta(,)$ is the chosen detection measure. (Indeed $\tilde{\mathbf{c}}_j$ are noise versions of $\mathbf{c}_{jw}$) Then recover the message:

$$\tilde{\mathbf{m}}_{ji} = \begin{cases} 1, & \text{if} \quad \delta(\tilde{\mathbf{c}}_j, \mathbf{w}_i) > \tau \\ 0, & \text{if} \quad \delta(\tilde{\mathbf{c}}_j, \mathbf{w}_i) < -\tau \\ \text{an erasure} & \text{if} \quad -\tau \leq \delta(\tilde{\mathbf{c}}_j, \mathbf{w}_i) \leq \tau \end{cases},$$

Detection and Decoding

A-7. The error control code decoder corrects errors and erasures. Its output can be

- "there is no watermarking or hidden message", when the number of erasures is $> N/2$;
- a decoded message (a sequence of bits);
- a decoded message with warning "errors are possible".

Detection measures

The majority of systems proposed in the literature fall into the class of ***correlation-based watermarking systems***.

Linear correlation:

$$lc(\mathbf{x}, \mathbf{y}) = \frac{1}{N}\mathbf{x} \cdot \mathbf{y} = \frac{1}{N}\sum_{i=1}^{N} x_i y_i.$$

Normalized correlation:

$$nc(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}| \cdot |\mathbf{y}|} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2}\sqrt{\sum_{i=1}^{N} y_i^2}}.$$

Correlation coefficient:

$$cc(\mathbf{x}, \mathbf{y}) = \frac{\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}}{|\hat{\mathbf{x}}| \cdot |\hat{\mathbf{y}}|} = \frac{\sum_{i=1}^{N} \hat{x}_i \hat{y}_i}{\sqrt{\sum_{i=1}^{N} \hat{x}_i^2}\sqrt{\sum_{i=1}^{N} \hat{y}_i^2}},$$

where $\hat{\mathbf{x}} = \mathbf{x} - E[\mathbf{x}]$ and $\hat{\mathbf{y}} = \mathbf{y} - E[\mathbf{y}]$.

The case of normalized correlation

We introduced in this case a new non-additive embedding:

A-5

$$\mathbf{c}_{jw} = \mathbf{c}_j \cos\varphi + \epsilon \frac{|\mathbf{c}_j|}{|\mathbf{h}_i|}\mathbf{h}_i \sin\varphi,$$

where $\epsilon = +1$ or -1, when the embedded bit m_i is 1 or 0, respectively. The parameter φ controls the tradeoff between visibility and robustness of the watermark.

Error Analysis

The expected value of $\delta(,)$ is $\mu_1 = \mu = \frac{\alpha}{\sqrt{r}}$ and $\mu_0 = -\mu = -\frac{\alpha}{\sqrt{r}}$ when $m_i = 1$ and $m_i = 0$ is embedded, respectively (see A-5 and A-6). Let us assume that $\delta(,)$ is normal distributed. Then the variance is $\sigma^2 = \sigma^2_{\boldsymbol{w}_i}(\sigma^2_{\boldsymbol{c}} + \sigma^2_{\boldsymbol{n}})$, where $\sigma^2_{\boldsymbol{w}_{ri}} = 1$, and $\sigma^2_{\boldsymbol{c}}$ and $\sigma^2_{\boldsymbol{n}}$ are the variance the cover work and the channel noise, respectively. Usually $\sigma^2_{\boldsymbol{c}} \approx (60/255)^2$. Let p_c, p_{er} and p_{es} be the probability of correct detection, of error, and of an erasure, respectively. Then in both cases, when $m_i = 1$ and $m_i = 0$ is embedded:

$$p_c = \frac{1}{2}\operatorname{erfc}\left(\frac{\tau - \mu}{\sigma\sqrt{2}}\right); \qquad p_{er} = \frac{1}{2}\operatorname{erfc}\left(\frac{\tau + \mu}{\sigma\sqrt{2}}\right);$$

$$p_{es} = 1 - p_c - p_{er} = \frac{1}{2}\left[\operatorname{erf}\left(\frac{\tau - \mu}{\sigma\sqrt{2}}\right) + \operatorname{erf}\left(\frac{\tau + \mu}{\sigma\sqrt{2}}\right)\right].$$

The case when no watermark is embedded

The probability of **a false positive decision** ($\mu = 0$) is given by

$$P_{fp} = \sum_{i=0}^{\lfloor \frac{N}{2} \rfloor} \binom{N}{i} p^{N-i}(1-p)^i,$$

where

$$p = \operatorname{erfc}\left(\frac{\tau}{\sigma\sqrt{2}}\right)$$

The case when no watermark is embedded

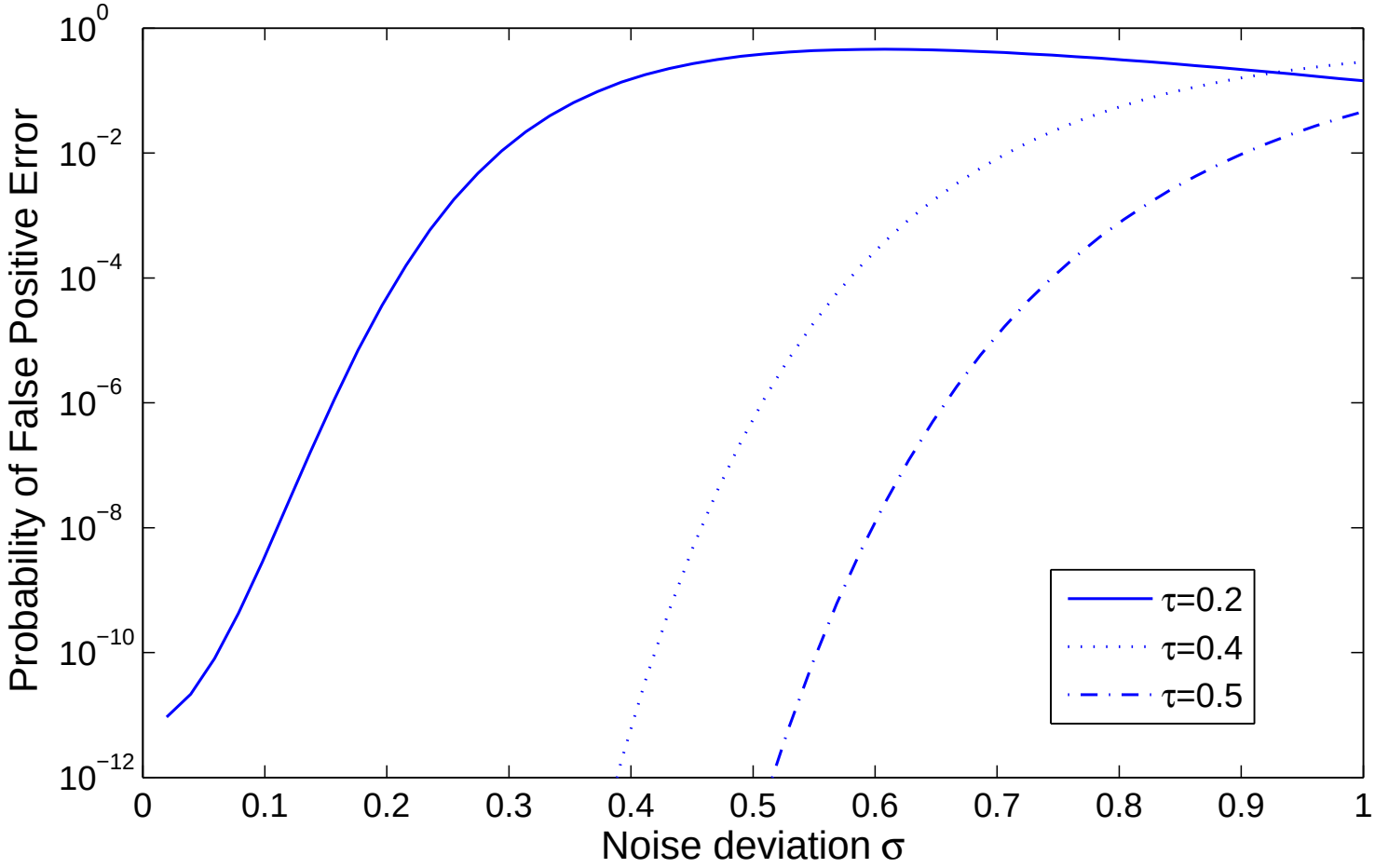

Figure: False Positive Error for $N = 13.16 = 208$ blocks

The case when a watermark is embedded

(a) The embedded n bits can be correctly decoded with a probability

$$P_{corr} = P_1 + P_2 + P_3, \qquad \text{where}$$

$$P_1 = \sum_{s=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{s} p_{es}^s \left(\sum_{t=0}^{\lfloor \frac{d-1-s}{2} \rfloor} \binom{n-s}{t} p_{er}^t p_c^{n-t-s} \right),$$

$$P_2 = \sum_{s=\lfloor \frac{d-1-s}{2} \rfloor + 1}^{d-1} \binom{n}{s} p_c^{n-s} p_{es}^s, \qquad P_3 = \sum_{s=d}^{n} \binom{n}{s} p_c^{n-s} (q - p_c)^s,$$

where d is the minimum distance of the code and $q = \frac{1}{2} \operatorname{erfc}\left(\frac{-\mu}{\sigma\sqrt{2}} \right)$ (this is p_c with $\tau = 0$) is the probability of positive (resp. negative) value of $\delta(\tilde{\mathbf{c}}_{wn}, \mathbf{w}_i)$.

The case when a watermark is embedded

(**b**) The probability of correct decoding in the case when the codes are used only in error correcting mode is given by

$$Q_{corr} = \sum_{t=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{t} q^{n-t}(1-q)^t.$$

(**c**) If more than $N/2$ erasures are marked for a given watermark pattern $\mathbf{w}_i$ then the detector outputs "there is no watermark", that is, it makes **false negative decision**. The probability, P_{fn}, for such an output is given by

$$P_{fn} = \sum_{i=0}^{\lfloor \frac{N}{2} \rfloor} \binom{N}{i} p_{es}^{N-i}(1-p_{es})^i. \quad (1)$$

An Example: [13, 8, 4] code

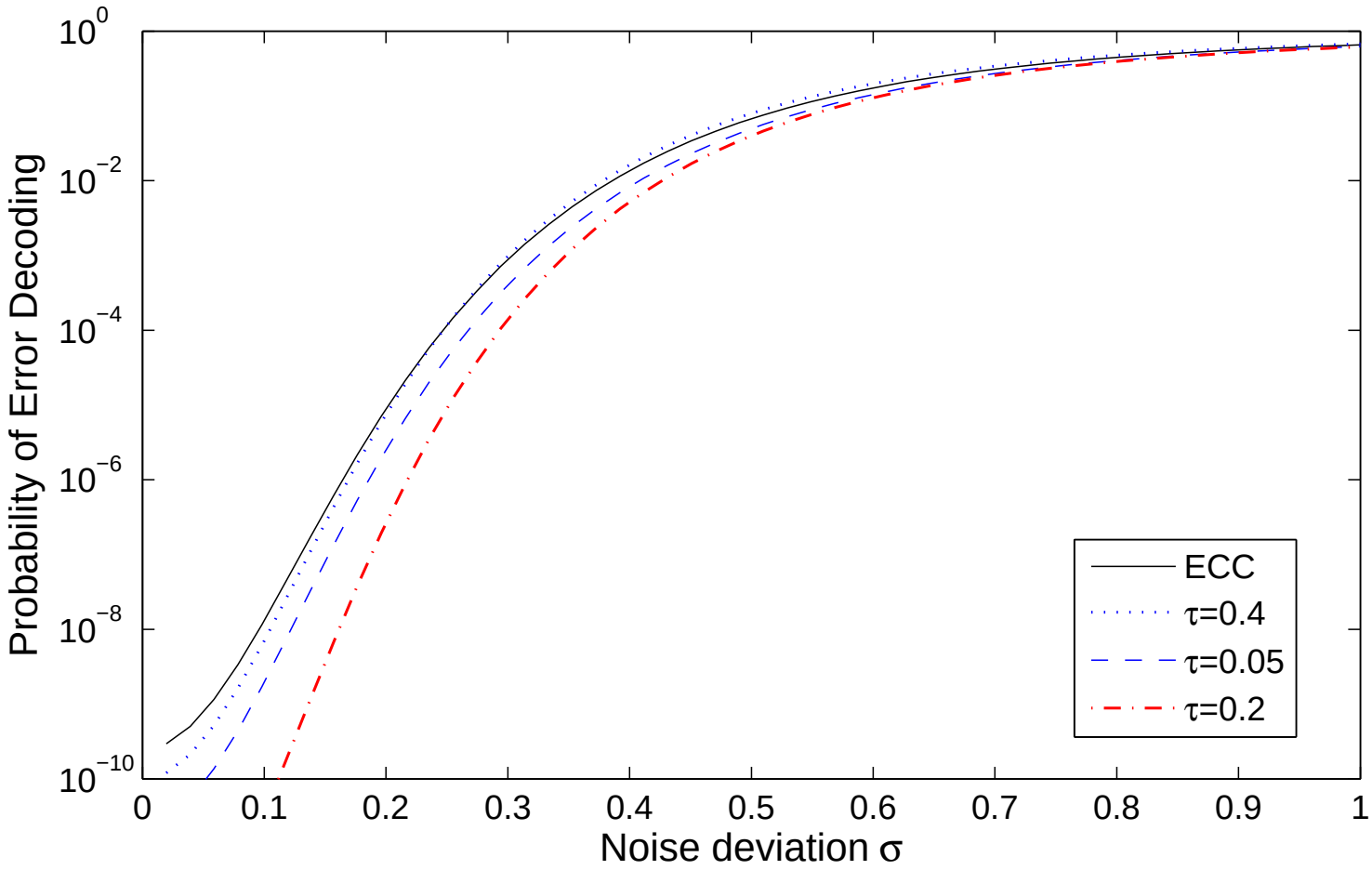

Figure: The probability of error decoding: $1 - P_{corr}$.

Conclusions

We have made numerous experiments with picture from several galleries (with grey-scale and color images) and many error control codes for both type of algorithms. Also, we have tested the cover images with the available in the internet software for stego-analysis. Although the channel is assumed noiseless and detection blind (i. e., the receiver and the opponent don't have the original), we have compered the results of stego-analysis on both original and cover objects.
Our observations show that q-ary codes are good choice in the case of syndrome embedding. For both algorithms it is better to use not very long codes with simple decoding and leave security issues to $\mathcal{M}$.

THANK YOU

FOR YOUR ATTENTION!