

USING of POLAR CODES in STEGANOGRAPHY

I.DIOP, S.M.FARSSI, B.DIOUF

Polytechnic school of Cheikh Anta Diop University Dakar Senegal

idydiop@yahoo.fr, farsism@yahoo.com

Abstract—In this paper, we propose a new steganography scheme based on the polar codes. The scheme works according to two steps. The first offers a stego vector from given cover vector and message. The stego vector provided by the first method can be the optimal; in this case, the insertion is successful with a very low complexity. Otherwise, we formalize our problem in a linear program form with initial solution the stego vector given by the first method, to converge to the optimal solution. Our scheme works with the case of a constant profile as well with any profile; it is then adapted to the case of wet paper. Tests of the scheme on multiple images in grey scale have showed its good performance in terms of minimizing the embedding impact.

Keywords— linear programming, matrix embedding, polar codes, steganography, wet paper codes.

I. INTRODUCTION

The steganography is a technique that allows hiding information in an unsuspected medium (image, sound or video) so that it was undetectable. To reach this objective it is indispensable to use a technique in order to reduce the distortion induced by the hiding of the secret message. The matrix embedding technique introduced by Crandall [1] has allowed the definition of steganography schemes that minimize the embedding impact. The first implementation was created with the work of Westfeld [2] in which the Hamming code has been used. Afterwards the BCH codes [3], [4], the Reed-Solomon codes [5] and the STC [6] are used in steganography. The combination of the techniques of LSB, of matrix embedding and wet paper has allowed realizing more effective and more reliable steganography schemes. Our work is a contribution to schemes of embedding impact minimization. We propose in this paper a new steganography scheme based on the polar codes. The scheme is applied to the cases of constant profile and of wet paper.

This paper is organized as following. Section II describes the concepts of matrix embedding and minimization of embedding impact. In Section III we study the linear programming. The polar codes used for the implementation of our scheme are presented in Section IV. In Section V we propose our scheme based on the polar codes. Section VI show the results obtained when the scheme is applied on images. Section VII concludes the paper.

II. STEGANOGRAPHY AND MATRIX EMBEDDING

A. Steganography

Steganography or the art of secret communication aims to hide a message in an apparently innocuous cover medium.

Steganography schemes are characterized by different parameters. The most used to evaluate the performance of a steganography scheme is the **embedding efficiency**. It is the number of bits of the message by distortion.

B. Distortion measure with the PSNR

The PSNR (Peak Signal Noise Ratio) is a distortion measure between two images. It is calculate from MSE (Mean Square Error) and is expressed in *dB*. Let I_o and I_r be, respectively, the original and reconstructed images of same length $M \times N$.

The PSNR and the MSE are given by [7]:

$$PSNR(I_o, I_r) = 10 \log_{10} \frac{D^2}{MSE(I_o, I_r)} \quad (1)$$

$$MSE(I_o, I_r) = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N [I_o(i, j) - I_r(i, j)]^2 \quad (2)$$

where D is the dynamic (the maximum value of a pixel). If the pixels are coded with 8 bits then $D = 2^8 - 1 = 255$.

More the value of PSNR is greater; more the images compared are similar. A PSNR of more than 35 dB between two images means that there is no visible difference between these two images [7]. If the PSNR is less than 20 dB the two images are very different.

C. Principle of matrix embedding

We will consider, thought all the paper, the cover vector $\mathbf{v}$ made up of the LSBs of the cover image, the stego vector $\mathbf{y}$, the vector of changes $\mathbf{e}$ ($\mathbf{y} = \mathbf{v} + \mathbf{e}$), the secret message $\mathbf{m}$ and the parity check matrix $\mathbf{H}$ of the used errors correcting code. The principle of matrix embedding is to find the stego vector $\mathbf{y}$ closest to $\mathbf{v}$ such that $\mathbf{yH}^T = \mathbf{m}$. By replacing $\mathbf{y}$ by $\mathbf{v} + \mathbf{e}$, we will have $\mathbf{eH}^T = \mathbf{m} - \mathbf{vH}^T$.

The objective of the sender is to find the vector $\mathbf{e}$ of minimum weight in the coset $\mathcal{C}(\mathbf{m} - \mathbf{vH}^T)$ (the set of the vectors of size n and syndrome $\mathbf{m} - \mathbf{vH}^T$). At the reception, the decoding is just done by the matrix product $\mathbf{m} = \mathbf{yH}^T$.

D. Minimization of embedding impact

We consider the vectors defined above. Assuming that the changes do not interact with each other, the total embedding impact is the sum of the embedding impact at each pixel [6]:

$$D(\mathbf{v}, \mathbf{y}) = \sum_{i=1}^n \rho_i |v_i - y_i|, \quad (3)$$

where $0 \leq \rho_i \leq \infty$ is the cost of change of the pixel v_i into y_i .

The goal is, for the sender, to insert its message $\mathbf{m}$ so that the distortion D is minimized.

The insertion and extraction functions are defined by:

$$Emb(v, m) = \arg \min_{y \in \mathcal{C}(m)} D(v, y) \quad (4)$$

$$Ext(y) = yH^T = m \quad (5)$$

III. LINEAR PROGRAMMING

The linear programming is a central domain of optimization. It can be formulated as follows [8]:

$$\begin{aligned} \min \text{ or } \max f(x) \\ \text{s.t.} \\ x \in \mathcal{C} \end{aligned}$$

where f is the *objective function*, x the *variable* and $\mathcal{C}$ the *feasible set*.

A linear program can be written either in the canonical form or in the standard form (obtained from the canonical form).

Canonical form	Standard form
$\min f_{s,t}(x) = c^T x$	$\min f_{s,t}(x) = c'^T x'$
$\begin{cases} Ax \geq b \\ x \geq 0 \end{cases}$	$\begin{cases} A'x' = b \\ x' \geq 0 \end{cases}$

Before solving a linear programming problem, we must put it in standard form. A linear program can be solved by simplex method or interior point methods.

A. Simplex method

This method was developed by G. Danzig to solve linear programs. To avoid calculating the solutions of all linear systems $m \times m$ extracted from A , we may use the simplex algorithm. This algorithm is based on the following approach [8]: *starting from a vertex (the initial solution), we move from extreme point (vertex) to extreme point along the polyhedron frontier (the set of feasible solutions)*. We determine if the current vertex is optimal and if not the case, we move to adjacent vertex that optimizes the objective function.

B. Methods of interiors points

The publication of the work of Karmarkar [9] in 1984 gave rise to interior points methods which are intended to reduce the complexity observed in the simplex algorithm. The interior point methods *start from an interior point (initial solution) to the feasible solutions domain* and, by using a fixed strategy, determines an approximate value of the optimal solution [10]. The advantages of these methods, compared to the simplex method, are polynomial complexity and fast convergence for the real problems of large sizes.

A method of solving a linear program is even faster than the initial solution is close to the optimal value.

IV. POLAR CODES

A. Usual notations

Let $W: \mathcal{X} \rightarrow \mathcal{Y}$ the channel B-DMC (Binary input-Discrete Memoryless Channel) on which the transmission takes place. The sets $\mathcal{X}$ and $\mathcal{Y}$ are respectively the input and output alphabets of W . $W(y|x)$ is the transition probability such that $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The operation $\oplus$ defines the modulo-2 addition and $\otimes$ Kronecker product. a_1^n denote the line vector $(a_1, \dots, a_n)$ where $n = 2^p$, p being a positive integer.

B. Definitions

The symmetric capacity (bits/s) of W is defined by [11]:

$$I(W) \triangleq \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \frac{1}{2} W(y|x) \log_2 \frac{W(y|x)}{\frac{1}{2} W(y|0) + \frac{1}{2} W(y|1)} \quad (6)$$

and the Bhattacharyya parameter or reliability of W is :

$$Z(W) \triangleq \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}. \quad (7)$$

The polar coding is based on these two parameters.

Two types of usual symmetrical channels are Binary Erasure Channel (BEC) and Binary Symmetric Channel (BSC). A B-DMC W is a:

- BSC if $\mathcal{Y} = \{0,1\}$, $W(0|0) = W(1|1)$ and $W(1|0) = W(0|1)$;
- BEC if $W(y|0)W(y|1) = 0$ or $W(y|0) = W(y|1)$.

The parameters $I(W)$ and $Z(W)$ take their values in $[0,1]$ and verify [11]:

$$\begin{aligned} I(W) \approx 1 &\Leftrightarrow Z(W) \approx 0 && \text{(perfect canal)} \\ I(W) \approx 0 &\Leftrightarrow Z(W) \approx 1 && \text{(completely noisy)} \end{aligned}$$

From these two equivalences, we can say that to know the properties of a B-DMC, it suffices to study one of the two parameters. We will focus on $Z(W)$ more easy to manipulate.

C. Channel polarization and transformation

1) Channel polarization: it is the base of the construction of polar codes. It consists in synthesizing of n independent copies of a given B-DMC W to create n others $\{W_n^{(i)} : 1 \leq i \leq n\}$. The polarization appears in the sense that $\{I(W_n^{(i)})\}$ tends to $I(W)$ or $1 - I(W)$, depending on if $I(W_n^{(i)})$ is closer to 0 or 1. The operation of channel polarization is constituted of two steps: the channels combination and channel splitting.

- *Channels combination:* this is to group n copies of a given B-DMC W in a channel W_n . The combination, for the level $p = 1$, associate $n = 2^1$ independent copies of $W_1 = W$ to form the channel W_2 .

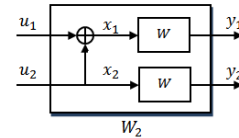


Figure 1. A construction of the channel W_2 .

From Figure 1, we have $x_1 = u_1 \oplus u_2$ and $x_2 = u_2$. Thus $x_1^2 = u_1^2 G_2$, with $G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

The generalization of the channels combination procedure, with any ≥ 1 ($n = 2^p$), is given by a combination of two independent copies of the channel $W_{n/2}$ to create W_n . One transforms the input u_1^n of W_n to s_1^n such that $s_{2i-1} = u_{2i-1} \oplus u_{2i}$ and $s_{2i} = u_{2i}$ for $1 \leq i \leq n/2$. The permutation matrix R_n transforms s_1^n into $v_1^n = (s_1, s_3, \dots, s_{n-1}, s_2, s_4, \dots, s_n)$ which is the input for the two copies of $W_{n/2}$. The relation of polar coding is:

$$x_1^n = u_1^n G_n \quad (10)$$

$G_n = B_n G_2^{\otimes p}$ is called generator matrix with B_n a permutation matrix and $G_2^{\otimes p}$ the Kronecker product of p copies of G_2 .

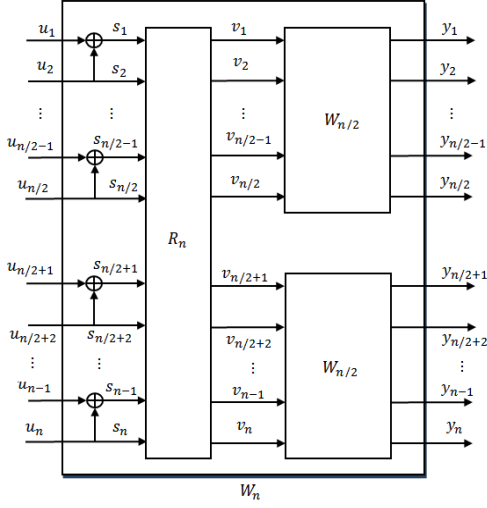


Figure 2. Construction of the channel W_n from two copies of $W_{n/2}$.

- **Channel splitting:** it consist of subdivision of the channel W_n into n channels $W_n^{(i)}$, $1 \leq i \leq n$.

If u_1^n is uniform on $\mathcal{X}^n$ then $W_n^{(i)}$ is the channel really seen by u_i ; u_i borrows the channel $W_n^{(i)}$ to give (y_1^n, u_1^{i-1}) .

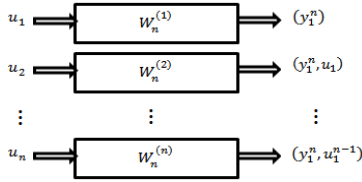


Figure 3. Equivalent scheme of polar coding.

2) Recursive channel transformation: The process of transformation can be generalized recursively:

$$(W_n^{(i)}, W_n^{(i)}) \xrightarrow{\text{we construct}} (W_{2n}^{(2i-1)}, W_{2n}^{(2i)})$$

The recursion relations, used for the construction of the polar codes, are [11]:

$$Z(W_{2n}^{(2i-1)}) \leq 2Z(W_n^{(i)}) + Z(W_n^{(i)}), \quad (11)$$

$$Z(W_{2n}^{(2i)}) = Z(W_n^{(i)})^2. \quad (12)$$

with equality in (11) if W is a BEC.

D. Polar coding

The principle of polar coding is to create a system of coding allowing to access to each channel $W_n^{(i)}$ and send the data through those more reliable. Let A be a subset of dimension k and A^c its complementary in $\{1, \dots, n\}$. We will fix A and u_{A^c} . u_A is called the information vector and u_{A^c} frozen vector. In general, we choose $u_{A^c} = 0_{1^{n-k}}$. The set A is chosen such that $Z(W_n^{(i)}) \leq Z(W_n^{(j)})$ for any $i \in A$ and $j \in A^c$.

1) Construction of polar codes: to construct a polar code, we need, as inputs, the B-DMC W on which the code is applied, the block length n and dimension k of the code. The algorithm of construction provide as outputs the information set $A \subset \{1, \dots, n\}$ of size k and the channels $W_n^{(i)}$.

2) Polar encoding: we use the relation (10) to encode a data word u_1^n in a codeword x_1^n . The generator matrix is [11]:

$$G_n = R_n(G_2 \otimes G_{n/2}) = B_n G_2^{\otimes p} \quad (13)$$

where the permutation matrix

$$B_n = R_n(I_2 \otimes B_{n/2}) \quad (14)$$

We will use these two recursion formulas to find B_n and G_n .

3) Decoding of polar codes: several types of decoding of polar codes exist, namely Successive Cancellation (SC) [11], Linear Programming (LP) decoding [12] and Belief Propagation (BP) decoding. Among these types of decoding the most powerful is the SC but, because of the probability of the channels involved in its implementation, its application in steganography is not yet possible. Therefore, we use the LP decoding for the formulation of our steganography scheme.

V. STEGANOGRAPHIC SCHEME BASED ON POLAR CODES

A. Different steps of the construction of polar codes

The construction of polar codes for the steganography can be summed up in three steps, as shown by Figure 4.

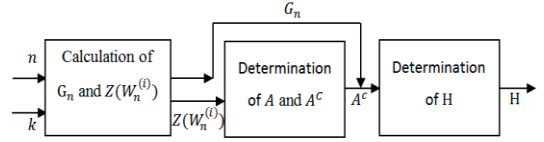


Figure 4. The scheme of the construction of a polar code.

Step 1: the calculation of reliability parameters.

For a symmetric channel, we have relations (11) and (12), with equality in (11) for the case of a BEC. Steganographic channel is BSC; consider that our channel W is chosen such that equality in (11) is satisfied. Thus, we can recursively calculate all reliability parameters. Relations (11) and (12) become:

$$Z(W_n^{(j)}) = 2Z(W_{n/2}^{((j+1)/2)}) - Z(W_{n/2}^{((j+1)/2)})^2 \text{ if } j \text{ is even}$$

$$Z(W_n^{(j)}) = Z(W_{n/2}^{(j/2)})^2 \text{ if } j \text{ is odd}$$

The initial value is calculated with the relation (7):

$$Z(W_1^{(1)}) = Z(W) = 2\sqrt{W(0|0)W(0|1)} = 2\sqrt{p_e(1-p_e)}$$

where p_e is the error probability of the channel W , $p_e = W(0|1) = W(1|0)$ and $1 - p_e = W(0|0) = W(1|1)$.

Step 2: determination of information and redundancy bits.

We select channels with the parameters of the lowest reliabilities for data bits. The indices of these channels form the information bits A . Its cardinality is equal to the dimension k of the considered polar code. The other channels (there are $n - k$) carry redundancy bits. Their indices constitute A^c .

Step 3: generation of the parity check matrix H .

A common element in the construction of any steganography scheme is the parity check matrix H . Indeed, it is used both for insertion and extraction of the messages. To determine the parity check matrix of a polar code, we can use the lemma given by Goela and al. [12, Lemma 1] which says that if the bits of redundancy vector u_{A^c} are set to 0 then the parity check matrix is given by the columns of the generator matrix G_n of the polar code whose indices are in A^c .

B. Steganographic scheme with a polar code

In this part of the paper, we consider that the change of any pixel produces the same distortion (constant profile). Thus we are interested in the number of bits changed during insertion. The research of the stego vector will follow two phases.

1) First method proposed: we will draw our inspiration from the work of T. Filler and al. [6] in which the proposed steganography scheme is strongly related to the form of the parity check matrix of the STC. Indeed, by observing the parity check matrix H of a polar code and its transpose H^T , we can make the following remarks:

- a) the columns of H are *pairwise independent*;
- b) if we scan the columns of H^T , the position at which one meets the *first non-zero coefficient* (equal to 1) differs from that of other columns of H^T ;
- c) more, starting from the last column and starting with the first line, the position at which one meets the *first non-zero coefficient is the first met on this line*.

With these remarks, we will define a first steganography scheme. Consider the matrix product $yH^T = m$. The decomposition of this system gives $y_1H_{1j}^T + y_2H_{2j}^T + \dots + y_nH_{nj}^T = m_j$, for $j = 1, \dots, n - k$.

Let i be the position of the first 1 met on the column j , the system of equations above becomes:

$$y_i = y_{i+1}H_{(i+1),j}^T + \dots + y_nH_{nj}^T + m_j \quad (15)$$

To determine y_i , in the above equation, we must find y_{i+l} such that $H_{(i+l),j}^T = 1$. We will assume that these positions correspond to the locked positions of v . In this case $v_{i+l} = y_{i+l}$. Therefore, before calculating the elements y_i of the initial stego vector y , we first assign it, as initial value, the cover vector $y \leftarrow v$. The changes of certain positions y_i of y will occur as and when we travel over the columns.

At the end of this process, we have a stego vector y arising from d modifications of the cover vector v .

Consider $v = (v_1, \dots, v_8)$, $m = (m_1, \dots, m_4)$ and $y = (y_1, \dots, y_8)$. To apply the method described above, we use a polar code of length $n = 8$ and dimension $k = 4$. Then, $A = \{4, 6, 7, 8\}$ and $A^c = \{1, 2, 3, 5\}$. The parity check matrix is:

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } H^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (16).$$

With the relation (15), we have the following system:

$$\begin{cases} y_2 = y_4 + y_6 + y_8 + m_4 \\ y_3 = y_4 + y_7 + y_8 + m_3 \\ y_5 = y_6 + y_7 + y_8 + m_2 \\ y_1 = y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + m_1 \end{cases}$$

The vector y must be initialized to the cover vector v . The calculation of the y_2 requires knowledge of y_4 , y_6 and y_8 . We consider that these positions as fixed. Then $y_4 = v_4$, $y_6 = v_6$ and $y_8 = v_8$. In the same, we fix the coefficient $y_7 = v_7$ to calculate y_3 by using y_4 and y_8 that have been fixed at the previous step. The calculation of y_5 is done using y_6 , y_7 and y_8 previously fixed. The last coefficient y_1 is determined with coefficients, either already fixed or already calculated. Hence

the necessity to start from the last column of H^T otherwise, we could not calculate y_1 without fixing all y_i , $2 \leq i \leq 8$.

The application of the method described above gives us a solution satisfying $yH^T = m$ but it is not necessarily the best. This solution is optimal if the number of changes is less than $(n - k)/2$ and not necessary optimal otherwise. In order to ensure to find the optimal solution, we will define, from the first solution, a method that offers the optimal solution.

2) Second method: the objective is to find the cover vector y the closest to v by using the polar code $C(n, k)$. Let y_p be the stego vector found by applying the first approach ($y_p = v + e_p$). The question is how to find, from y_p , the optimal vector y_{opt} verifying $yH^T = m$. In other words, the algorithm should, from e_p , provide the errors vector corresponding to e_{opt} . So, we have to create an algorithm that, initialized to e_p , converges to e_{opt} . Thus we look out the error vector e of minimum weight satisfying $eH^T = m - vH^T$.

Recapitulation:

- We have a **starting solution** $e_p \mapsto$ initial solution;
- we search **e of minimal weight** $\mapsto$ minimization problem;
- verifying **$eH^T = m - vH^T = s$** $\mapsto$ constraints.

Considering these three points, we have an optimization problem, particularly a minimization problem with initial solution e_p . Find the vector e of minimum weight amounts to find the vector realizing the minimum of the scalar product with the vector $c = \{1\}^n$. From there, we can define our optimization problem as follows:

$$\begin{aligned} \arg \min_{e} \quad & f(e) = \langle c, e \rangle = c^T e \\ \text{s.t.} \quad & \mathfrak{C} = \begin{cases} e \in \{0, 1\}^n \text{ a binary vector} \\ eH^T = m - vH^T = s \\ e_p \text{ initial solution} \Leftrightarrow e_p H^T = s \end{cases} \end{aligned}$$

This problem is a linear optimization, with constraints equalities, written in standard form. It can be solved by the simplex or the interior points methods explained in Section III.

The application of the optimization method provides the optimal solution of our steganography problem.

The scheme can be summarized by the following figure.

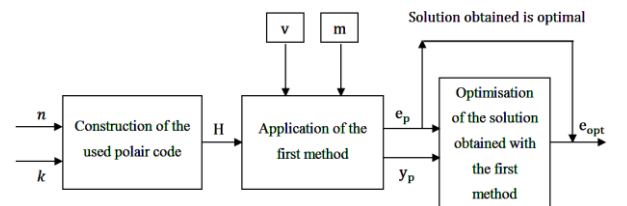


Figure 5. Representation of the proposed steganography scheme.

3) Optimality condition of the scheme: the proposed scheme is not suitable for messages of size less than or equal to $p = \log_2(n)$, with n the block length of the polar code. To have a minimum number of changes, the message must verify *size* (m) = $m > p$. This criterion is logic since, for $m \leq p$ some columns of the parity check matrix H are identical. That favours more changes. Note that even messages of size $m \leq p$ can be inserted and extracted at reception but the minimum number of changes is no longer guaranteed.

C. Steganographic scheme with wet paper

In the case where the costs of embedding changes $\rho_i = 1$, the minimization of the distortion D amounts to minimize the number of modifications of the cover vector. On the other hand, if the ρ_i are arbitrary in $[0, \infty]$ we have to define the scheme for the minimization of the number of changed positions by taking into account constraints of the cost values ρ_i . We have to modify the pixels where the changes are less perceptible. We will develop a technique of steganography to minimize the embedding impact for a given distortion.

Consider an arbitrary profile defined by $\rho = \{\rho_i\}_{1 \leq i \leq n}$ [13]. Our goal is to adapt the scheme proposed previously to this general case of the distortion profiles. We will use the technique, proposed by Filler and al. [6], which consists to lock the positions at which the changes will be most visible. They have shown that this approach is more adapted in practice.

Since our scheme consists of two parts or methods, the second improves the first, it is thus necessary to see how each of these two methods is applied to the case of wet paper.

The first method is independent of the considered profile. Thus, it is applied identically to the case of constant profile.

Concerning the second method, changes should be made. Since the problem is the same as the case of constant profile (minimization problem), we use the same principle of linear programming to find our optimal solution. The initial solution and the constraints have not been changed. But the objective function will change. Indeed, for an arbitrary profile, the goal is to minimize the distortion function (3). Rewrite the relations of (3) depending on the changes vector e :

$$D(e) = \sum_{i=1}^n \rho_i e_i,$$

where $|v_i - y_i| = e_i$ and $0 \leq \rho_i \leq \infty$ the modification cost of the LSB v_i of a pixel into y_i .

$$Emb(v, m) = \arg \min_{e \in \mathcal{C}(s)} D(e)$$

$$Ext(y) = yH^T = m \Leftrightarrow eH^T = s = m - vH^T$$

We can see that our objective function $f(e) = \langle c, e \rangle$, which must be written as a scalar product between the cost vector of the linear program and the variable e , appears well in the expression of $D(e)$. Consequently, the elements of the cost vector are represented by the costs of pixels change. Let $c = \rho = \{\rho_i\}_{1 \leq i \leq n}$, we have still a linear program which can be solved in the same way that the constant profile case.

VI. EXPERIMENTAL RESULTS

To verify the efficiency of our scheme and the invisibility of the hidden messages using this scheme, we have tested it on different images in grey scale *pgm* format size **(512×512) pixels**. The images are taken from the *BOSS*¹ database.

To make the message less detectable, we choose to permute the pixels of the cover image before making the insertion. Because our images are of size 512×512 and 512 is a power of 2, we can use the permutation matrix B_n with $n = 512$, for the permutation. Thus the changes will be spread over isolated pixels of the image making less detectable the secret message inserted and thus allowing a more secure insertion. After

insertion, it is necessary to find the original order of pixels of the cover image. To achieve this, we still use the matrix B_n . Since it is invertible and equal to its own inverse, it suffices just to repeat the same operation as in the permutation.

We insert, first, a message of size **3 ko** or **4576 bits** in an image (10.pgm) of *BOSS* database.

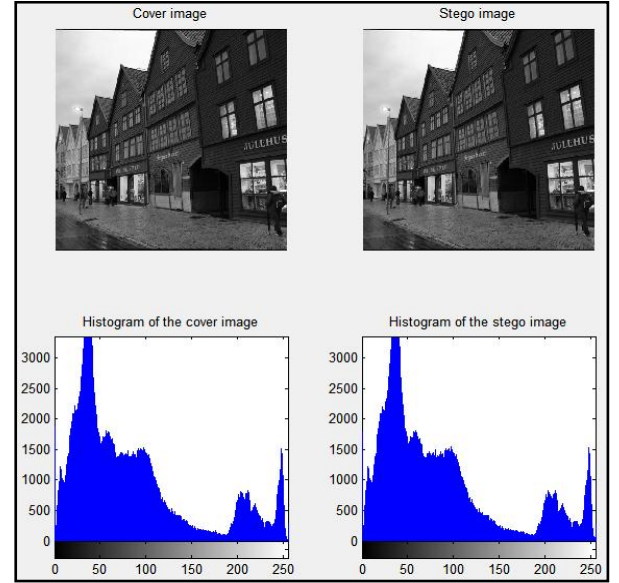


Figure 6. The cover image and stego image and their histograms.

The first and most simple evaluation is the visual imperceptibility. The changes in the stego image are invisible to the naked eye, as shown in Figure 6. Hence the first and main goal of steganography is achieved. If we are confronted with a passive attacker we have clearly seen that by comparing the cover and stego images on the one hand and their histogram on the other hand, for an attacker to "semi-active", the distinction between the two images is almost impossible. Because there is a very small difference between the histograms of the cover and stego images that is very difficult to perceive. This difference is more perceptible if the size of the message to be inserted increases. The scheme is even more secure that the attacker has only the stego image to see if it contains a message or not. He should therefore use much more sophisticated means to reach to detect the presence of message.

Like any good steganography scheme, ours allows to extract the secret message in full, without any alteration.

The message inserted in the cover image (Figure7) is identical to that extracted from the stego image (Figure 8).

To evaluate the performance of our scheme we calculated the MSE and the PSNR with the equations (1) and (2). We randomly generate 10 messages of different size that we insert in 5 images (1, 10, 100, 1000 and 10000.pgm). We averaged the MSE and PNSR and the results are shown in Figure 9.

The value of PNSR vary between **65.4113 dB** for $\alpha = 1/20$ and **55.4017 dB** for $\alpha = 1/2$. These values are well above **35 dB**, a value beyond which the difference between two images (cover and stego) is very low. This shows that the proposed scheme has good performance in terms of embedding efficiency.

¹ *BOSS* (Break Our Stego-System): the competition for the attacks of the steganographic schemes.

1 Steganography is the whole techniques for to transmit a message by hiding
2 its existence. It consists to modify a medium (image, video, sound, etc.)
3 so that it contains an undetectable message. Steganography is old but was
4 really developed until the year 1998 with the matrix embedding technique
5 introduced by Crandall in [4]. This technique has allowed improving the
6 embedding efficiency of the steganography schemes. View the similarities
7 between matrix embedding and the principle of decoding error correcting
8 codes, they are exploited for steganography. The first implementation of
9 the matrix encoding concept was proposed by Westfeld in the F5 algorithm
10 [5] with a Hamming code $C(n,k,1)$ and allow reducing the modification of
11 quantized DCT coefficients. Structured BCH codes were introduced in
12 steganography by Schönfeld and Winkler [6]. The authors have proposed two
13 ways of calculating the syndrome: a classical approach to find the error
14 vector of minimum weight by using a parity check matrix H and another,
15 less complex, using a generator polynomial to find the roots of the error
16 locator polynomial. An improvement is provided by Sachnev and al. in [7]
17 with the use of lookup tables to reduce the complexity. The use of RS
18 codes in steganography is made possible with the work of Fontaine and
19 Galand [8] in which the version of the GRS (Generalized Reed-Solomon)
20 codes is used to define a wet paper steganographic scheme. More recently,
21 convolutional codes, more particularly STC codes [3] have been exploited
22 by J. Friedrich et al. in steganography. Their performances have proven to
23 be better. T. Filler et al. have used these codes in their scheme HUGO [1].
24 The combination of the techniques of LSB, embedding matrix and wet paper
25 has allowed realizing more efficient steganographic schemes. As specified
26 in [1] by T. Filler et al., the realization of better steganographic
27 schemes is possible by using either a good model or a good code. The polar
28 coding, a new paradigm of channel coding, recently discovered by Arıkan [2]
29 and whose performances have been proved, motivates our choice on the
30 proposition of a new code in steganography.

Figure 7. Inserted message

1 Steganography is the whole techniques for to transmit a message by hiding
2 its existence. It consists to modify a medium (image, video, sound, etc.)
3 so that it contains an undetectable message. Steganography is old but was
4 really developed until the year 1998 with the matrix embedding technique
5 introduced by Crandall in [4]. This technique has allowed improving the
6 embedding efficiency of the steganography schemes. View the similarities
7 between matrix embedding and the principle of decoding error correcting
8 codes, they are exploited for steganography. The first implementation of
9 the matrix encoding concept was proposed by Westfeld in the F5 algorithm
10 [5] with a Hamming code $C(n,k,1)$ and allow reducing the modification of
11 quantized DCT coefficients. Structured BCH codes were introduced in
12 steganography by Schönfeld and Winkler [6]. The authors have proposed two
13 ways of calculating the syndrome: a classical approach to find the error
14 vector of minimum weight by using a parity check matrix H and another,
15 less complex, using a generator polynomial to find the roots of the error
16 locator polynomial. An improvement is provided by Sachnev and al. in [7]
17 with the use of lookup tables to reduce the complexity. The use of RS
18 codes in steganography is made possible with the work of Fontaine and
19 Galand [8] in which the version of the GRS (Generalized Reed-Solomon)
20 codes is used to define a wet paper steganographic scheme. More recently,
21 convolutional codes, more particularly STC codes [3] have been exploited
22 by J. Friedrich et al. in steganography. Their performances have proven to
23 be better. T. Filler et al. have used these codes in their scheme HUGO [1].
24 The combination of the techniques of LSB, embedding matrix and wet paper
25 has allowed realizing more efficient steganographic schemes. As specified
26 in [1] by T. Filler et al., the realization of better steganographic
27 schemes is possible by using either a good model or a good code. The polar
28 coding, a new paradigm of channel coding, recently discovered by Arıkan [2]
29 and whose performances have been proved, motivates our choice on the
30 proposition of a new code in steganography.

Figure 8. Extracted message.

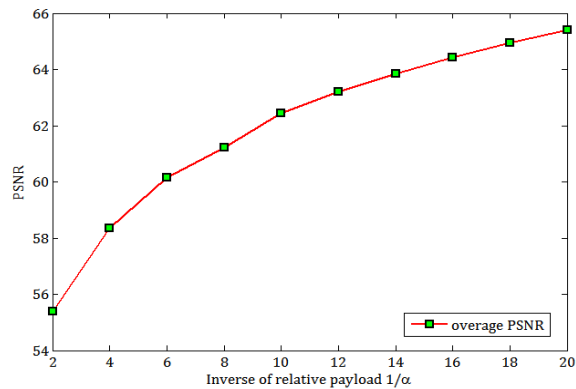


Figure 9. Evolution of PSNR depending on the relative payload.

VII. CONCLUSION

We have defined a steganography scheme based on a new type of coding called polar coding with a good embedding efficiency. The proposed scheme consists of two parts: the first part gives an initial solution and a second part ensures a

convergence to the optimal solution using linear programming. In the case where the first solution corresponds to the optimum, it is not necessary to proceed to the second method. This scheme is especially suitable for messages whose size is greater than $p = \log_2 n$, with n the size of the cover vector. However, the size of the message can go up to n . Our scheme is also suitable for the case of wet paper.

We have shown, by applying it to 5 different images, that the visually undetectable and even statistical, by using histograms, is reached. We also calculated the PSNR with these 5 images and their value varies between 55 dB and 66 dB. That is greater than the limit value 35 dB.

An prospect of improvement is to search a scheme in a single step by using the *LP decoding by changing the polytope* $yH^T = m$. In the definition of our scheme, we have used a generator matrix G_n constructed from a sub-matrix G_2 , $n = 2^p$. We can try to see what would give a construction of G_n from a sub-matrix G_ℓ [14], $\ell \geq 3$, $n = \ell^p$ as STC [6].

REFERENCES

- [1] R. Crandall, "Some notes on steganography", Posted on Steganography Mailing List (1998).
- [2] A. Westfeld, "High capacity despite better steganalysis (F5 – a steganographic algorithm)", In: Moskowitz, I.S. (ed.) IH 2001. LNCS, vol. 2137, pp. 289–302, Springer, Heidelberg (2001).
- [3] Schönfeld, D., Winkler, A *Embedding with syndrome coding based on BCH codes*. In: Proceedings of the 8th ACM Workshop on Multimedia and Security, pp. 214 – 223, 2006.
- [4] Rongyue Zhang, Vasilii Sachnev, Hyoung Joong Kim, *Fast BCH syndrome coding for steganography*; S. Katzenbeisser and A.-R. Sadeghi (Eds.), IH 2009, LNCS 5806, pp. 44-58, Springer-Verlag Berlin Heidelberg 2009.
- [5] F. Galand C. Fontaine, How Reed-Solomon Codes Can Improve Steganographic Schemes, In Information Hiding, Rennes, France, 2009.
- [6] Tomáš Filler, Jan Judas and Jessica Fridrich, *Minimizing Embedding Impact in Steganography using Trellis-Coded Quantization*, Department of Electrical and Computer Engineering SUNY Binghamton, Binghamton, NY 13902-6000, USA 2010.
- [7] Xuanwen Luo, Qiang Cheng, Joseph Tan, "A Lossless Data Embedding Scheme For Medical in Application of e- Diagnosis," Proceedings of the 25th Annual International Conference of the *IEEE EMBS Cancun*, Mexico. September 17-21, 2003.
- [8] Aaid Djamel, "Étude numérique comparative entre des méthodes de résolution d'un problème de transport à quatre indices avec capacités," Thèse, École Doctorale de Mathématiques, pôle de Constantine, 2010. [Online]: <http://bu.umc.edu.dz/theses/math/AAI5587.pdf>
- [9] Narendra Karmarkar (1984). "A New Polynomial Time Algorithm for Linear Programming", *Combinatorial*, Vol 4, nr. 4, p. 373–395.
- [10] Gilles Savard Gerad, "Introduction aux méthodes de points intérieurs," Extrait de notes de cours, Département de mathématiques et génie industriel Ecole Polytechnique de Montréal.
- [11] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memory-less channels", *IEEE Trans. Inform. Theory*, vol. IT-55, pp. 3051–3073, July 2009.
- [12] N. Goela, S. B. Korada, and M. Gastpar, "On LP Decoding of Polar Codes," submitted to *IEEE Trans. Inform. Theory*, 2010.
- [13] Tomáš Pevný, Tomáš Filler and Patrick Bas, "Using High-Dimensional Image Models to Perform Highly Undetectable Steganography", Czech Technical University, Prague, Czech Republic; State University, New York in Binghamton, NY, USA; CNRS-LAGIS, Lille, France, 2010.
- [14] S.B. Korada, E. Sasoglu, and R. Urbanke, "Polar Codes: Characterization of Exponent, Bounds and Construction", submitted to *IEEE Trans. Inform. Theory*, January 2009.