Polar Coding Steganographic Embedding Using Successive Cancellation

Birahime Diouf(✉), Idy Diop, Khadidiatou Wane Keita, Madiop Diouf, Sidi Mohamed Farsi, Khaly Tall, and Ousmane Khouma

Department of Computer Science, Polytechnic Institute (ESP), Cheikh Anta Diop University (UCAD), Dakar, Senegal
dioufbirall@yahoo.fr, idydiop@yahoo.fr, {wane.keita,madiop.diouf}@esp.sn, farsism@yahoo.com, khalytal@gmail.com, ousmane.khouma@ucad.edu.sn

Abstract. In this paper, we propose a practical adaptive embedding methodology based on Successive Cancellation (SC) polar coding. The new proposed SC-based Polar Coding Steganography (SC-PCS) defines message bits as frozen bits of the SC decoder and computes path metrics according to embedding costs of pixels of the cover image. Simulation results demonstrate that SC-PCS minimizes an arbitrary embedding distortion while embedding covert message.

Keywords: Adaptive embedding · Additive distortion · Polar code Steganography · Successive cancellation · Syndrome coding

1 Introduction

With the development of internet and diversity of communications media, information security becomes a necessity and even a priority. To satisfy this demand, techniques such as cryptography protect information but do not guarantee discretion. However, often the existence of the communication must be kept secret. Steganography is today positioned, deservedly, as a means to address this widely shared concern. It is to conceal secret information in others unsuspected media such as text, image (used in this paper), audio or video so that only the recipient is aware of the existence of the communication. The majority of modern steganographic techniques are based on embedding a secret message while minimizing the embedding distortion [1]. This approach consists of two complementary tasks that both participate to increase steganographic security [2]. Most of the proposed methods focus on one of these two tasks. The first is to effectively define the embedding costs of all the pixels of the cover image. To design costs several methods are proposed in current state of the art such as HUGO (Highly Undetectable steGO) [3], WOW (Wavelet Obtained Weights) [4], UNIWARD (Spatial UNIversal WAvelet Relative Distortion) [5], HILL (HIgh Low Low) [6], MiPOD (Minimizing the Power of Optimal Detector) [7]. The second task, that we are interested in, relies on minimizing the distortion defined from costs using practical coding. For practice, matrix embedding is proposed by Crandall [8]. The first implementation of matrix embedding was provided by Westfeld [9] who used

C. M. F. Kebe et al. (Eds.): InterSol 2017/CNRIA 2017, LNICST 204, pp. 189–201, 2018.
https://doi.org/10.1007/978-3-319-72965-7_18

Hamming codes in F5 algorithm. The current state of the art methods includes BCH (Bose-Chaudhuri-Hocquenghem) [10], RS (Reed Solomon) [11], LDPC (Low Density Parity Check) [12] and STC (Syndrome Trellis Codes) [13].

Our work is a contribution to methods that use codes to minimize distortion function. The codes used for this shake are polar codes (PC) introduced by Arikan [14] as the first capacity-achieving codes. Several decoding types exist for PC such as Successive Cancellation (SC) [14], SC List (SCL) [15], Linear Programming (LP) [16] and Adaptive LP (ALP) [17]. PC are demonstrated to be applicable in steganography with constant profile and wet paper codes [18] and minimize embedding impact [19–21] for these two important frameworks. The first adaptive embedding method based on polar codes is proposed in [22], in which ML-certificate ALP decoder is used. The natural decoding technique of polar code is SC which achieves the rate-distortion bound. These benefices of polar codes and their optimality to Payload-Limited Sender (PLS) problem emphasized by Filler et al. [13] motive the use of SC decoding to minimize distortion function and increase embedding efficiency.

The rest of this paper is organized as follows. In Sect. 2, we present steganography, distortion function and practical embedding coding techniques. In Sect. 3, we review polar coding and SC decoding. The proposed steganographic embedding technique based on SC is described in Sect. 4. Section 5 shows experimental results of the tests. Conclusions are drawn in Sect. 6.

2 Steganography Based on Minimizing Additive Distortion

2.1 Distortion Definition for Adaptive Embedding

In modern steganography, most of the proposed techniques are based on embedding secret message while minimizing the embedding distortion between cover $\mathrm{x} \in \mathcal{X}$ and stego $\mathrm{y} \in \mathcal{Y} = I_1 \times \cdots \times I_n$ images which can be defined in additive form [13]:

$$D(\mathrm{x}, \mathrm{y}) = \sum_{i=1}^{n} \rho_i(\mathrm{x}, y_i), \tag{1}$$

where $\rho_i(\mathrm{x}, y_i)$ is a local distortion measure and denotes the cost of replacing pixel x_i by y_i if digital images are used. In such definition, we assume that the embedding changes are mutually independent. The additive distortion can be rewritten as follows:

$$D(\mathrm{x}, \mathrm{y}) = \sum_{i=1}^{n} \rho_i \cdot [y_i \neq x_i], \tag{2}$$

where $[P]$ is the logical operator that is equal to 1 if relation P is true and 0 else. Several methods exist to calculate the embedding costs. HUGO was the first method proposed to calculate costs from features vectors difference in SPAM space. WOW and UNIWARD use directional filter to design costs. HILL cost function is designed using one high-pass filter and two low-pass filters. MiPOD is model based and computes costs by minimizing the power of the most efficient detector.

2.2 Optimal Embedding Problem and Practical Coding

For message embedding there are two methods: Distortion-Limited Sender (DLS) and Payload-Limited Sender (PLS) [13]. However, PLS that consists in embedding a fixed payload while minimizing the embedding distortion, is the most used compared with DLS that is to maximize payload while introducing an expected distortion. Let the stego image y be a random variable over $\mathcal{Y}$ and its distribution $\pi_{\mathrm{x}}(\mathrm{y}) \triangleq Pr(\mathrm{y}|\mathrm{x})$, the optimal embedding is realized when replacing each pixel x_i with probability $\pi_{\mathrm{x},i}$:

$$\pi_{\mathrm{x},i}(y_i) = \frac{\exp(-\lambda \rho_i(\mathrm{x}, y_i))}{\sum_{y_i \in I_i} \exp(-\lambda \rho_i(\mathrm{x}, y_i))} \tag{3}$$

where $\lambda \in [0, \infty[$ is a parameter obtained by solving the following constraint

$$h(\pi_{\mathrm{x}}) = -\sum\nolimits_{y \in \mathcal{Y}} \pi_{\mathrm{x}}(\mathrm{y}) \log_2 \pi_{\mathrm{x}}(\mathrm{y}) = m, \tag{4}$$

where m is the size of the message $\mathrm{m} \in \mathcal{M} = \{0,1\}^m$ and $\pi_{\mathrm{x}}(\mathrm{y}) = \prod_{i=1}^{n} \pi_{\mathrm{x},i}(y_i)$. The constraint can be rewritten as follows:

$$-\sum\nolimits_{i=1}^{n} \sum\nolimits_{y_i \in I_i} \pi_{\mathrm{x},i}(y_i) \log_2 \pi_{\mathrm{x},i}(y_i) = m, \tag{5}$$

In practice syndrome coding can be used to implement the embedding operation. In this context, let $\mathcal{C}$ be the linear code of length n and dimension $n - m$ and consider binary embedding operation, then the embedding $Emb : \mathcal{X} \times \mathcal{M} \to \mathcal{Y}$ and extraction $Ext : \mathcal{Y} \to \mathcal{M}$ functions, respectively used by the sender and the recipient are

$$\begin{aligned} Emb(\mathrm{x}, \mathrm{m}) &= \arg\min\nolimits_{LSB(\mathrm{y}) \in \mathcal{C}(\mathrm{m})} D(\mathrm{x}, \mathrm{y}) \\ Ext(\mathrm{y}) &= LSB(\mathrm{y}) \cdot \mathrm{H}^{\mathrm{T}} = \mathrm{m}, \end{aligned} \tag{6}$$

where $LSB(\mathrm{y}) = (LSB(y_1), \ldots, LSB(y_n))$, $\mathrm{H} \in \{0,1\}^{m \times n}$ is a parity-check matrix and $\mathcal{C}(\mathrm{m}) = \{\mathrm{z} \in \{0,1\}^n | \mathrm{z}\mathrm{H}^{\mathrm{T}} = \mathrm{m}\}$ denotes the coset of m. These two functions verify

$$Ext(Emb(\mathrm{x}, \mathrm{m})) = \mathrm{m} \quad \forall \mathrm{x} \in \mathcal{X}, \forall \mathrm{m} \in \mathcal{M}. \tag{7}$$

An embedding coding algorithm can be evaluated via its embedding efficiency $e(\alpha) = m/D = \alpha n/D$ (in bits/distortion unit) in comparison with the optimal embedding derived from (3), where $\alpha = m/n$ is called the relative payload. It is known that random linear code with syndrome coding is capacity-achieving for the PLS problem. However, random code is not practical due to the exponential complexity needed for the decoder. Non-random codes such as Hamming, BCH, RS, LDPC and STC are used to approach the optimal embedding. Notes that STC is currently the most used. Polar codes are known to be optimal for PLS problem as pointed out by Filler et al. [13]. Wet paper codes are used to embed data in a cover image for which some pixels are forbidden to be altered. Such pixels, called wet pixels, are characterized by infinite cost $\rho_i = \infty$ and then $I_i = \{x_i\}$. The others, called dry pixels, may be changed and have

finite cost $\rho_i < \infty$. The wet paper framework is an interesting topic in steganography and is largely addressed in our previous works [19–21].

3 Polar Codes and SC Decoding

3.1 Polar Codes, First Capacity Achieving Codes

Introduced by Arikan [14], Polar Codes (PC) are defined as the first codes that achieve channel capacity $I(W)$ (Shannon's threshold) in a large class of channel W with low encoding and decoding time complexity $O(nlogn)$, where $n = 2^p$ is the block-length. PC construction is designed using channel polarization which provides them their recursive nature. Channel polarization consists in constructing polarized channels $W_n^{(i)} : 1 \le i \le n$ from n independent copies of W. It is made up two steps: channel combining that associates n copies of W and recursively creates n-inputs channel W_n and channel splitting that subdivides W_n into n channels $W_n^{(i)}$ [19]:

$$(\mathrm{W}, \mathrm{W}, \ldots, \mathrm{W}) \xrightarrow{combining} W_n \xrightarrow{splitting} \left\{W_n^{(i)}\right\}_{1 \le i \le n} \quad \text{Channel polarization} \tag{8}$$

The main idea of polar coding is to access to each channel $W_n^{(i)}$ and send the information bits across the most reliable ones i.e. with lowest reliability parameter $Z\left(W_n^{(i)}\right)$ and the frozen (fixed) bits through the remaining channels. The information and frozen bits indices will be denoted by A and it complementary A^c, respectively. We denote by $a_1^i = (a_1, \ldots, a_i), 1 \le i \le n$, $\hat{u}_{1,e}^i$ and $\hat{u}_{1,o}^i$ the sub-vectors consisting of elements with odd and even indices. The generator matrix is defined from the Kronecker product $\mathrm{G}_2^{\otimes p}$ of p copies of $\mathrm{G}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ by $\mathrm{G}_n = \mathrm{B}_n \mathrm{G}_2^{\otimes p}$, where B_n the bit-reversal permutation matrix. The polar encoding is based on the relation $\mathrm{c} = \mathrm{uG}_n$, where u is the source word and c the codeword. When traveling through the channel, c can be changed into a received word r. The transition probability is denoted by $W(r_i|c_i)$.

3.2 Successive Cancellation Decoding and Polar Codes

SC decoding is the first and inherent decoding proposed by Arikan for PC. It is governed by the recursive nature of PC construction. First, the bit $\hat{u}_1$ is decoded given r. Then, $\hat{u}_2$ is given from $\hat{u}_1$ and r, and so far i.e. $\hat{u}_i$ is estimated given the previously decoded bits $\hat{u}_1^{i-1}$ and the received word r. For frozen bits the value is known $\hat{u}_i = u_i$. The aim of SC decoder is to provide estimated source-word $\hat{\mathrm{u}}$ given A, u_{A^c} and r. The SC decoder gives its decision using the following function h:

$$\hat{u}_i \triangleq \begin{cases} u_i, & if\ i \in A^c \\ 0 & if\ L_n^{(i)}\left(r, \hat{u}_1^{i-1}\right) \ge 1\ ,\ 1 \le i \le n \\ 1, & else \end{cases} \tag{9}$$

where $L_n^{(i)}\left(\mathrm{r}, \hat{u}_1^{i-1}\right) = \frac{W_n^{(i)}\left(\mathrm{r}, \hat{u}_1^{i-1}|0\right)}{W_n^{(i)}\left(\mathrm{r}, \hat{u}_1^{i-1}|1\right)}$, with $L_1^{(1)}(r_i) = L(r_i) = \frac{W(r_i|0)}{W(r_i|1)}$ is LR (Likelihood-Ratio). In LLR domain non-frozen bits are chosen depending on if $LL_n^{(i)}\left(\mathrm{r}, \hat{u}_1^{i-1}\right) \geq 0$ or not. Figure 1 shows the graph of the SC decoder with $n = 8$. This process starts on the right side of the graph where the received word bits LRs $L(r_i)$ are combined in pairs by moving towards the left side. The graph consists of $\log n = 3$ stages each of which contains $n = 8$ nodes. For each node of stage j, the LR is calculated from two incoming LRs L_a and L_b of stage $j - 1$ using f or g function:

$$f(L_a, L_b) = \frac{1 + L_a \cdot L_b}{L_a + L_b}, \tag{10}$$

and

$$g(L_a, L_b, \hat{u}_{sum}) = g_{\hat{u}_{sum}}(L_a, L_b) = L_a^{(1-2\hat{u}_{sum})} \cdot L_b \tag{11}$$

where $\hat{u}_{sum}$ is the binary partial sum of previously estimated bits. In LLR domain

$$f(LL_a, LL_b) = 2\tanh^{-1}(\tanh(LL_a/2) \cdot \tanh(LL_b/2)) \tag{12}$$

and

$$g_{\hat{u}_{sum}}(LL_a, LL_b) = LL_a \cdot (-1)^{\hat{u}_{sum}} + LL_b, \tag{13}$$

where $LL_a \triangleq \log(L_a)$ and $LL_b \triangleq \log(L_b)$ are LLRs values. The min-sum approximation, used with LDPC codes, can also be exploited to reduce this complexity:

$$f(LL_a, LL_b) \approx \tilde{f}(LL_a, LL_b) \triangleq sign(LL_a) \cdot sign(LL_b) \cdot \min(|LL_a|, |LL_b|). \tag{14}$$

At the left side of the graph, estimated bits $\hat{u}_i$ are provided using function h (9).

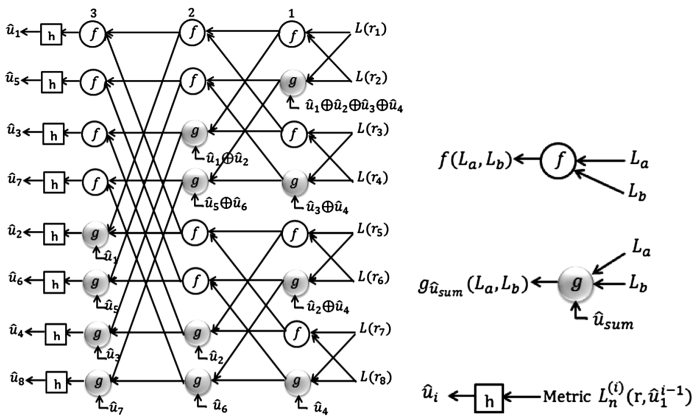

Fig. 1. Graph of SC decoding for $n = 8$.

The SC decoding can be seen as a code tree search. In this context, it begins at the root node that has two paths with labels 0 and 1 and metrics $W_n^{(i)}(\mathrm{r}|0)$ and $W_n^{(i)}(\mathrm{r}|1)$, respectively. Then, the path with highest metric is chosen. This path results, in turn, in two paths with labels 0 and 1 and metrics $W_n^{(i)}(\mathrm{r}, \hat{u}_1|0)$ and $W_n^{(i)}(\mathrm{r}, \hat{u}_1|1)$. Generally, at each level $\hat{u}_i$ is decoded by comparing the two path metrics $W_n^{(i)}(\mathrm{r}, \hat{u}_1^{i-1}|0)$ and $W_n^{(i)}(\mathrm{r}, \hat{u}_1^{i-1}|1)$ if i is not a frozen position. If i is a frozen position, $\hat{u}_i = u_i$. This procedure continues down to the leaf nodes where the last hard decision is made to estimate $\hat{u}_n$. Notice that SC decoder provides good performance if the bloc length n is sufficiently large. We will use this decoding technique to define our steganographic scheme.

4 SC-Based Adaptive Polar Coding Embedding

In channel coding, polar codes are shown to be capacity achieving. Another benefit and reason why we are interested in polar codes is their optimality for Payload Limited Sender (PLS) as emphasized out by Filler et al. [13]. Polar code construction in steganographic context is largely explained in [18]. For simplification purposes, in the rest of the paper, we will denote by x and y the LSB vectors of the cover and the stego images, respectively. The practical steganographic embedding using codes is based on syndrome coding with the extraction constraint $\mathrm{yH}^{\mathrm{T}} = \mathrm{m}$. We will first investigate the impact of this relation in the frozen bits values choice when using SC.

4.1 Definition of Secret Message Bits as Frozen Bits

We will show that the Steganographic Polar Coding (SPC) under SC encoding is equivalent to SC decoding of polar codes in channel coding when $\mathrm{u}_{A^c} = \mathrm{m}$. Indeed, it is known from [14] that polar codes are defined in terms of an invertible matrix G_n via the encoding relation $\mathrm{c} = \mathrm{uG}_n$ i.e. $\mathrm{u} = \mathrm{cG}_n^{-1} = \mathrm{cG}_n$, because $\mathrm{G}_n^{-1} = \mathrm{G}_n$. For the sake of simplification we denote $\mathrm{G} = \mathrm{G}_n$. The source word can be split into two parts $\mathrm{u} = (\mathrm{u}_A, \mathrm{u}_{A^c})$, where the information word $\mathrm{u}_A = (u_i : i \in A)$ and the frozen word $\mathrm{u}_{A^c} = (u_i : i \in A^c)$ [23]. Then, we can write $\mathrm{u} = (\mathrm{u}_A, \mathrm{u}_{A^c}) = \left(\mathrm{cG}^A, \mathrm{cG}^{A^c}\right)$ where G^A and G^{A^c} are the submatrices consisting of columns of G whore indices are in A and A^c, respectively. From the definition of polar code parity check matrix [16] whose transpose H^{T} is obtained by selecting the columns of G with indices in A^c [18, Lemma 1], we have $\mathrm{G}^{A^c} = \mathrm{H}^{\mathrm{T}}$ with $\mathrm{u}_{A^c} = \mathrm{cH}^{\mathrm{T}}$. Let a source word u such that $\mathrm{u}_{A^c} = \mathrm{m}$. Then searching y such that $\mathrm{yH}^{\mathrm{T}} = \mathrm{yG}^{A^c} = \mathrm{m}$ is equivalent to SC decoding of a polar code where secret message bits are defined as frozen bits.

In the original SC decoding, given frozen bits u_{A^c}, one decodes the source word u such that there exists a codeword $\mathrm{c} = \mathrm{uG}_n = (\mathrm{u}_A, \mathrm{u}_{A^c})\mathrm{G}_n$. In the steganographic case, we have $\mathrm{u} = (\mathrm{u}_A, \mathrm{m})$, $\mathrm{y} = \mathrm{uG}_n$ and $\mathrm{yH}^{\mathrm{T}} = \mathrm{m}$. Then, we obtain a m –coset polar code where the stego word y is a code word. Then, naturally, to find the stego word, we can

use SC polar decoding where frozen bits $u_{A^c} = m$. However, this fact does not affect the SC decoder performance because it is independent of frozen bits values choice [14]. This decoding technique outputs u rather than the code word. However, we can simply obtain the m –coset polar codeword y by applying the encoding relation uG_n. The next step is to explain how the LR and LLR-metrics will be calculated.

4.2 Metric Calculation of SC for Steganography

In digital communication the classical SC decoder task is to find the information word u from the received word r by using metrics calculated from transition probabilities $W(received|transmited) = W(r_i|c_i) = p_e$ if $r_i \neq c_i$ and $W(r_i|c_i) = 1 - p_e$ else, for Binary Symmetric Channel (BSC), where p_e is the error probability. In steganography, given the cover word x and the secret message m, the sender has to search a stego word y such that $yH^T = m$. As seen in previous subsection, the decoder does not directly output the stego word y but gives first an information word u that, encoded with uG_n, provides the searched stego word y. Since that is the stego word we seek, given the cover word, then transition probabilities in steganography will be denoted by $W(cover|stego) = W(x_i|y_i)$. This can be interpreted as the probability that the corresponding pixel changes (if $y_i = x_i$) or not (if $y_i \neq x_i$ i.e. $y_i = 1 - x_i$). The definition of $W(x_i|y_i)$ would verify the conditions of transition probabilities in SC decoding and be provided depending on the embedding costs ρ_i (cost of replacing pixel x_i by y_i.).

When all the pixels have the same sensitivity to change $\rho_i = 1$ (constant profile), all the transition probability $W(x_i|y_i) = 1/2$. Let consider an arbitrary distortion, the idea is to assign *great* values of *change probabilities* $W(x_i|1 - x_i)$ (then *small* values of *non-change probabilities* $W(x_i|x_i)$) for pixels with ***small value of embedding costs*** and *small change probabilities* (then *great non-change probabilities*) for pixels with ***great embedding costs***. Then, let

$$W(x_i|y_i) = \begin{cases} d_i & \text{if } y_i = x_i \\ 1 - d_i & \text{else} \end{cases}, \tag{15}$$

where $d_i = \rho_i/\rho_{max}$ and ρ_{max} is maximum of the costs set. It is easy to verify that this definition of transition probabilities satisfies the above concerns and, additionally $W(x_i|0) + W(x_i|1) = 1$ We can condense by:

$$W(x_i|y_i) = ([x_i = y_i])(d_i) + ([x_i \neq y_i])(1 - d_i). \tag{16}$$

The logical operator applied on binary values x_i and y_i allows rewriting as follows:

$$W(x_i|y_i) = (1 - |x_i - y_i|)(d_i) + (|x_i - y_i|)(1 - d_i). \tag{17}$$

After defining $W(x_i|y_i)$, we can obtain the transition probabilities corresponding to polarized channels $W_n^{(i)}$. In steganographic SC decoder, transition probability of

channel $W_n^{(i)}$ denotes the likelihood of u_i given the channel outputs (cover $\mathrm{x} = x_1^n$ and $i-1$ previously decoded bits u_1^{i-1}). They are expressed by [14]:

$$W_n^{(i)}\left(\mathrm{x}, u_1^{i-1}|u_i\right) = \sum\nolimits_{u_{i+1}^n}\left(\frac{1}{2^{n-1}} W_n(x_1^n|u_1^n)\right) = \sum\nolimits_{u_{i+1}^n}\left(\frac{1}{2^{n-1}}\prod\nolimits_{i=1}^{n} W(x_i|y_i)\right). \quad (18)$$

Unfortunately, this expression is not practical. That is why Arikan defined recursive formulas, that can be labeled f and g, to calculate these transition probabilities:

$$W_n^{(2i-1)}\left(x_1^n, \hat{u}_1^{2i-2}|\hat{u}_{2i-1}\right) = f(W_a, W_b) = \sum\nolimits_{\hat{u}_{2i}\in\{0,1\}} \frac{1}{2} W_a \cdot W_b, \quad (19)$$

and

$$W_n^{(2i)}\left(x_1^n, \hat{u}_1^{2i-1}|\hat{u}_{2i}\right) = g(W_a, W_b) = \frac{1}{2} W_a \cdot W_b, \quad (20)$$

where $W_a = W_{n/2}^{(i)}\left(x_1^{n/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}|\hat{u}_{2i-1} \oplus \hat{u}_{2i}\right)$, $W_b = W_{n/2}^{(i)}\left(x_{n/2+1}^n, \hat{u}_{1,e}^{2i-2}|\hat{u}_{2i}\right)$. Let $L_n^{(i)}\left(x_1^n, \hat{u}_1^{i-1}\right) = \frac{W_n^{(i)}\left(x_1^n, \hat{u}_1^{i-1}|0\right)}{W_n^{(i)}\left(x_1^n, \hat{u}_1^{i-1}|1\right)}$, then LRs can be recursively calculated using functions (10) $f(L_a, L_b) = L_n^{(2i-1)}\left(x_1^n, \hat{u}_1^{2i-2}\right)$ and (11) $g_{\hat{u}_{sum}}(L_a, L_b) = L_n^{(2i)}\left(x_1^n, \hat{u}_1^{2i-1}\right)$, where $L_a = L_{n/2}^{(i)}\left(x_1^{n/2}, \hat{u}_{1,o}^{2i-2} \oplus \hat{u}_{1,e}^{2i-2}\right)$ and $L_b = L_{n/2}^{(i)}\left(x_{n/2+1}^n, \hat{u}_{1,e}^{2i-2}\right)$. This recursion continue until the last level LRs of length 1, where the LR calculation is given directly by $L_1^{(1)}(x_i) = L(x_i) = \frac{W(x_i|0)}{W(x_i|1)}$. From (17), the LR is defined by

$$L(x_i) = \frac{W(x_i|0)}{W(x_i|1)} = \frac{(1-x_i)(d_i) + (x_i)(1-d_i)}{(x_i)(d_i) + (1-x_i)(1-d_i)} = \begin{cases} \frac{d_i}{1-d_i} & \text{if } x_i = 0 \\ \frac{1-d_i}{d_i} & \text{if } x_i = 1 \end{cases}. \quad (21)$$

In logarithmic domain, the Log-Likelihood Ratios (LLR) are defined by

$$LL(x_i) = \log(L(x_i))(1-2x_i) \cdot \log\left(\frac{d_i}{1-d_i}\right) = (1-2x_i) \cdot \log\left(\frac{\rho_i}{\rho_{max} - \rho_i}\right). \quad (22)$$

These steganographic frozen bits and the LLRs will be used in the steganographic SC.

4.3 Steganographic Successive Cancellation Algorithm

Once we have given the frozen bits values and the LRs or LLRs (metrics) for the choice of unfrozen bits, we can implement the SC decoder of polar codes for steganography which is given by Algorithm 1:

Algorithm 1. SC Polar Coding Steganography (SC-PCS)

Inputs: cover objet $\mathbf{x}$, embedding cost $\boldsymbol{\rho}$, message $\mathbf{m}$.
Output: estimated source word $\mathbf{u}$ or stego object $\mathbf{y}$.
For all $i \in \{1, 2, \cdots, n\}$
  if $i \in A^c$ then // Frozen index
    $u_i \leftarrow m_i$; // Take message bits as frozen bits
  else compute LLR metrics $LL_n^{(i)}\left(\mathrm{x}, u_1^{i-1}\right)$ from $LL(x_i)$ using f and g functions (19) and (20).
    if $\left(LL_n^{(i)}\left(\mathrm{x}, u_1^{i-1}\right) \geq 0\right)$ then
      $u_i \leftarrow 0$;
    else
      $u_i \leftarrow 1$;
  end if
end For
return $\mathbf{u}$ or $\mathbf{y} = \mathbf{u}\mathrm{G}_n$;

The running SC-PCS provides the stego medium $\mathrm{y} \leftarrow \mathrm{uG}_n$ that will be transmitted to the recipient of the secret message. Once the stego medium is received, the recipient has two ways to find de covert message. Firstly, he can encode the stego medium to find the source word by $\mathrm{u} = \mathrm{yG}_n$. Then, he selects $u_{A^c} = (\mathrm{yG}_n)_{A^c} = \mathrm{m}$. The second alternative is to use the usual matrix relation via the parity check matrix $\mathrm{m} = \mathrm{yH}^{\mathrm{T}}$. In this manner, the embedding process provides the stego objet whose syndrome is the secret message m and minimizing the additive distortion D.

5 Experimental Results

The implementation of the proposed SC-PCS is given by the following steps:

- Construct the used polar code and initialize these parameters;
- Update channel transition probabilities calculated from costs using (17);
- Apply the SC-PCS embedding (Algorithm 1) using LLR path-metrics from (22);

Example 1: Let consider a cover $\mathrm{x} = (0, 1, 1, 1, 0, 0, 1, 0)$, a secret message $\mathrm{m} = (1, 0, 1, 0)$ and the corresponding stego object y. The different steps of the SC-PCS embedding with a polar code of bloc length $n = 8$ and dimension $k = 4$ are:

- When we use the construction method of polar code then, $A = \{4, 6, 7, 8\}$, $A^c = \{1, 2, 3, 5\}$ and the parity check matrix is:

$$\mathrm{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \mathrm{H}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{23}$$

- Generate randomly an embedding cost $\rho = (39, 57, 8, 6, 54, 78, 94, 13)$, for example. In this case, the mean is $\rho_{mean} = 43.6250$, the maximum is $\rho_{max} = 94$.
- The channel transition probabilities update in LLR domain provides $LL = (-0.3438, -0.4321, 2.3749, 2.6856, 0.3001, 1.5841, -Inf, -1.8295)$.
- When applying the SC-PCS embedding, the stego objet obtained is $\mathrm{y} = (0, 1, 1, 1, 1, 0, 1, 0)$ with distortion function $D(\mathrm{x}, \mathrm{y}) = 54$. The change object is $\mathrm{e} = (0, 0, 0, 0, 1, 0, 0, 0)$.

In this example, the optimal stego medium (corresponds to the one which minimizes distortion function) is obtained by using the SC decoder even with very small length $n = 8$. The total embedding distortion is $D(\mathrm{x}, \mathrm{y}) = 54$.

We will apply our SC-PCS on 512×512 8-bit gray scale digital images coming from BOSSbase database version 1.01 (Break Our Stego System) containing 10.000 images of *pgm* format obtained by rescaling and cropping natural images of various sizes of eight different cameras. Since the SC decoder performance is good if the bloc length n is enough large, then we can consider the image entirely i.e. the cover size $n = 512 \times 512 = 262144 = 2^{18}$. This value is sufficiently large to provide good performance. Note that, with the existing cost computation methods, often the high costs crowd into smooth area and low costs in textured area of the image (Fig. 2). Then, after calculating the costs, we will use the bit-reversal permutation matrix B_n as suggested and largely explained in [21]. This will scatter the pixels of the cover image and then increase the success probability of the optimal stego image search.

(a) Cover (b) Stego

Fig. 2. Cover vs stego images of '1013' from Boss base.

Figure 2 shows the cover image '1013.pgm' and the corresponding stego image after embedding a 0.4 *bpp* (bit per pixel) payload. The secret message is generated in a random format. The pixels of some smooth areas are located at the bottom of the images. These pixels will be assigned to very high cost (for example: 10^{10} in practice). For the embedding, the images are reshaped in a single vector of size 2^{18} from which we extract the LSB vector that we make use to apply SC-PCS.

To investigate the performances of our adaptive steganographic scheme we compute the embedding efficiency $emb_{eff} = m/D(\mathrm{x}, \mathrm{y})$ in comparison with other codes w.r.t reciprocal relative payload $1/\alpha$ (Fig. 3).

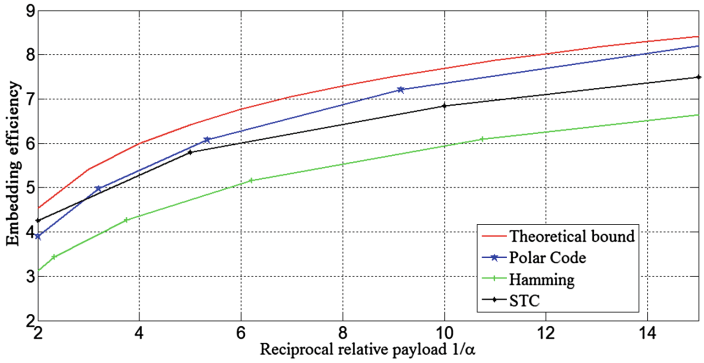

Fig. 3. Embedding efficiency of the steganographic SC polar decoder SC_PCS.

As shown by Fig. 3, SC-PSC provides similar performance that ALP-PCS [22] and better performance than Hamming and STC but for $\alpha = 1/2$ it is lower that STC's. Notice that this result is not surprising and could be better if we are referred to the theoretical results given in the literature about polar codes (capacity-achieving) when applied in PLS problem as emphasized in [13].

6 Conclusion

In modern adaptive steganographic scheme, there are two levers on which one could emphasis to increase the security of stego-system. Either focus on designing distortion measure (embedding costs), or concentrate on defining a near optimal embedding coding scheme. We have opted for the second option and have proposed, in this paper, a practical and efficient embedding method based SC polar coding. Indeed, as shown by the test results and the practical examples, this SC-PCS scheme minimizes arbitrary additive distortion function properly defined and provides better performance than STC. The main advantage of SC compared to ALP is that it always provides a valid stego medium. The simulation results show the good performance of polar codes in terms of embedding efficiency. Additionally, since SC is the natural and basic decoding technique of polar code, it is important to design steganographic embedding scheme

based on it. This allows future improvements if others decoding methods based on SC are improved.

The application of SC algorithm in steganography opens interesting perspective to improve further the embedding efficiency and move closer the optimal bound. Then, we plan to use the list version of SC called SCL [15] albeit more complex than SC. Other perspectives are to adapt the SC-PCS in JPEG domain and in non-binary embedding operation with multilayered construction.

References

1. Ker, A.D., Bas, P., Böhme, R., Cogranne, R., Craver, S., Filler, T., Fridrich, J., Pevný, T.: Moving steganography and steganalysis from laboratory to real world. In: Proceedings of the IH&MMSec 2013. ACM, Montpellier, France, June 2013
2. Holub, V.: Content Adaptive Steganography – Design and Detection. Ph.D. thesis, Binghamton University, May 2014
3. Pevný, T., Filler, T., Bas, P.: Using high-dimensional image models to perform highly undetectable steganography. In: Böhme, R., Fong, P.W.L., Safavi-Naini, R. (eds.) IH 2010. LNCS, vol. 6387, pp. 161–177. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-16435-4_13
4. Holub, V., Fridrich, J.: Designing steganographic distortion using directional filters. In: 4th IEEE International Workshop on Information Forensics and Security, Tenerife, Spain, December 2012
5. Holub, V., Fridrich, J., Denemark, T.: Universal distortion design for steganography in an arbitrary domain. EURASIP J. Inf. Secur. Special Issue on Revised Selected Papers of the 1st ACM IH and MMS Workshop, vol. 1, pp. 1–13, January 2014
6. Li, B., Wang, M., Huang, J., Li, X.: A new cost function for spatial image steganography. In: International Conference on Image Processing (ICIP), Paris, France, pp. 4206–4210, October 2014
7. Sedighi, V., Cogranne, R., Fridrich, J.: Content-adaptive steganography by minimizing statistical detectability. IEEE Trans. Inf. Forensics Secur. **11**(2), 221–234 (2016)
8. Crandall, R.: Some notes on steganography. Steganography Mailing List (1998)
9. Westfeld, A.: F5—A Steganographic Algorithm. In: Moskowitz, Ira S. (ed.) IH 2001. LNCS, vol. 2137, pp. 289–302. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45496-9_21
10. Schönfeld, D., Winkler, A.: An embedding with syndrome coding based on BCH codes. In: 8th ACM Workshop on Multimedia and Security, pp. 214–223 (2006)
11. Fontaine, C., Galand, F.: How reed-solomon codes can improve steganographic schemes. EURASIP J. Inf. Secur. **2009**, 1–10 (2009)
12. Diop, I., Farssi, S.M., Chaumont, M., Khouma, O., Diouf, H.B.: Utilisation des codes LDPC en stéganographie. COmpression et REprésentation des Signaux Audiovisuels (CORESA 2012), Lille, France, pp. 98–104, mai 2012
13. Filler, T., Judas, J., Fridrich, J.: Minimizing additive distortion in steganography using syndrome-trellis codes. IEEE Trans. Inf. Forensics Secur. **6**(3), 920–935 (2011)
14. Arikan, E.: Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. IEEE Trans. Inf. Theory IT **55**(7), 3051–3073 (2009)
15. Tal, I., Vardy, A.: List decoding of polar codes. In: Proceedings of the IEEE International Symposium on Information Theory Proceedings (ISIT), August 2011

16. Goela, N., Korada, S.B., Gastpar, M.: On LP decoding of polar codes. In: IEEE Transactions Information Theory Workshop (ITW), Dublin (2010)
17. Taranalli, V., Siegel, P.H.: Adaptive linear programming decoding of polar codes. In: IEEE Symposium on Information Theory (ISIT), pp. 2982–2986, June–July 2014
18. Diouf, B., Diop, I., Farssi, S.M., Tall, K., Fall, P.A., Diop, A.K., Sylla, K.: Using of polar codes in steganography. In: International Conference on Advances in Computer Science and Engineering (CSE), Atlantis Press, Los Angeles, vol. 42, pp. 262–266, July 2013
19. Diouf, B., Diop, I., Farssi, S.M., Khouma, O.: Minimizing embedding impact in steganography using polar codes. In: IEEE International Conference on Multimedia Computing and Systems (ICMCS 2014), Marrakesh, Morocco, pp. 105–111, April 2014
20. Diouf, B., Diop, I., Farssi, S.M., Khouma, O.: Practical polar coding method to minimize the embedding impact in steganography. In: IEEE International Science and Information (SAI) Conference, London, United Kingdom, July 2015
21. Diouf, B., Diop, I., Farssi, S.M.: Performances of polar codes in steganographic embedding impact minimization. In: ICACT Transactions on Advanced Communications Technology (ICACT-TACT), South Korea, vol. 5(5), pp. 927–935, September 2016
22. Diouf, B., Diop, I., Wone, K., Farssi, S.M., Khouma, O., Diouf, M., Tall, K.: Adaptive linear programming of polar codes to minimize additive distortion in steganography. In: Proceedings of the IEEE International Science and Information (SAI) Conference, United Kingdom, London, pp. 1086–1092, July 2016
23. Arikan, E.: Systematic polar coding. IEEE Commun. Lett. **15**(8), 860–862 (2011)