High-Performance Steganographic Coding Based on Sub-Polarized Channel *

Haocheng Fu[1,2], Xianfeng Zhao[1,2], and Xiaolei He[1,2(✉)]

[1] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100195, China
{fuhaocheng,zhaoxianfeng,hexiaolei}@iie.ac.cn
[2] School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100195, China

Abstract. Steganographic coding is the core problem of modern steganography under the minimizing distortion model. Syndrome-Trellis Codes, designed with the Viterbi algorithm of convolutional codes, have been the only coding scheme approaching the rate-distortion bound for almost a decade. Although polar codes has shown to have the potential to construct optimal coding schemes, the low diversity of steganographic coding negatively affects the security of steganography, and the performance of the state-of-the-art coding schemes is still unsatisfactory for large-scale applications. This paper proposes a high-performance steganographic coding scheme, named Sub-Polarized Steganographic Codes (SPSC), with near-optimal efficiency and lower computational complexity. By optimal embedding theory and channel polarization, we first establish a universal model of polarized steganographic channels for the coding of steganography. Based on this, the steganographic channels are divided into a combination of multiple sub-polarized channels to construct efficient steganographic codes by updating sequential bit-wise coding to segmented coding. The proposed coding scheme is also evaluated under four polarized channels with typical patterns, of which corresponding sub-coding schemes are also illustrated. Experimental results show that the proposed steganographic coding scheme improves security by increasing the embedding efficiency of steganography, and significantly decreases the computational complexity.

Keywords: Covert communication · Steganographic coding · Sub-channel polarization · Successive cancellation.

1 Introduction

Steganography is an important branch of information hiding. By embedding secret messages into naturally-looking covers, it establishes covert communication

* This work was supported by NSFC under 61972390, 61902391, 61872356 and 62272456, and National Key Technology Research and Development Program under 2022QY0101.

while not arousing the attention of others. Modern steganography prefers using digital media to perform steganographic embedding for its wide-spreading on the Internet. Based on defined distortion of modification of each cover element, the most modern schemes achieve satisfactory undetectability by minimizing total embedding distortion under Payload-Limited Sender (PLS) and Distortion-Limited Sender (DLS) problems. For digital images, several heuristic-designed cost functions, such as HUGO [23], HILL [20], UNIWARD [18], UED [13] and UERD [14], were proposed for cover in both spatial and JPEG domain, which achieve high security performance even the distortion is considered additive.

To minimize the total distortion, Crandall [4] first proposed the concept of matrix embedding by introducing the parity-check matrix of the error-correction codes. Based on this, multiple linear coding schemes, such as Hamming [26], BCH [27, 28] and LDPC [5, 10] were utilized to optimize with the constant distortion profile globally. Thereafter the wet paper codes (WPC) [11, 12] were proposed to for distortion consisting of modifiable (dry) and unmodifiable (wet) elements since they believe the modification tends to be performed on more secure areas.

In advanced steganographic schemes, coding schemes are considered to be adopted continuously instead of multi-level discrete distortion optimization. Therefore, Filler *et al.* [8, 9] proposed the first near-optimal steganographic codes, named Syndrome-Trellis Codes (STC), by the Viterbi decoding algorithm for arbitrary distortion. After that, no more practical steganographic coding scheme has been proposed for almost a decade, which might bring negative impact on the security of steganography. Recently, since polar codes [2] is proved to be able to achieve channel capacity for any symmetric binary input discrete memoryless channels (B-DMC), in [7], Diouf *et al.* constructed the first coding scheme of steganography with the Successive Cancellation (SC) decoding algorithm of polar codes, which revealed that steganographic codes based on polar codes have the ability to reach the rate-distortion bound of content-adaptive steganography. Thereafter Li *et al.* [21] designed another near-optimal coding scheme based on polar codes for cover with the Successive Cancellation List (SCL) algorithm that further improves the performance.

Steganographic coding schemes incorporated with polar codes significantly increase the diversity of steganography. Besides, the computational efficiency of the steganographic schemes is also improved due to the low complexity of the encoding and decoding of polar codes. However, since the polar codes-based steganographic coding always processes bit-by-bit [2, 25], the state-of-the-art schemes [6, 7, 21] still cannot reach the requirements of the practical steganographic application. As the channel polarization process is recursive, the coding procedure under the sub-polarized channels, can be simplified when the type of which is typical [1, 15, 24]. Therefore, it is possible to improve the overall performance of the coding scheme in steganography.

By establishing sub-polarized steganographic channels, this paper presents a high-performance steganographic codes, named Sub-Polarized Steganographic Codes (SPSC), which is close to the rate-distortion bound. In this scheme, the

computation complexity significantly decreases while the embedding efficiency of which is improved. The contributions of this paper are listed as follows.

- Based on the discrete binary symmetric channels, construct polarized and sub-polarized steganographic channels. Define the relationship between the log-likelihood ratio (LLR) and the embedding distortion of cover element.
- Propose an encoding strategy under the sub-polarized steganographic channels and implement a near-optimal steganographic coding scheme along with the existing list decoding algorithm of polar codes.
- Present efficient listed coding schemes for four identified sub-polarized channels, which are denoted as R1, DC, SPC and R0. Construct steganographic codes under these sub-polarized channels.

The rest of this paper is organized as follows. In Section 2, we first restate the preliminaries of the optimal embedding theory and channel polarization. The proposed steganographic coding scheme is elaborated in Section 3 as well as the construction of sub-polarized steganographic channel. Section 4 gives the experimental results and analysis in detail. Finally, a brief conclusion of this paper is listed in Section 5.

2 Preliminaries

In this paper, matrices and vectors are written in boldface while sets are shown in swash letters. Without loss of generality, let $\mathbf{x} = (x_1, x_2, \cdots, x_N) \in \{\mathcal{L}\}^n$ and $\mathbf{y} = (y_1, y_2, \cdots, y_N) \in \{\mathcal{I}_i\}^N$ denote the cover and stego sequence, respectively, where $\mathcal{L}$ represents the dynamic range of cover elements and $\mathcal{I}_i \subset \mathcal{L}$ stands for the operation of modification on x_i. For instance, $\mathcal{I}_i = \{x_i, \bar{x}_i\}$ is for binary embedding where $\bar{x}_i$ is the least significant bit (LSB) flipped element with respect to x_i and $\mathcal{I}_i = \{x_i - 1, x_i, x_i + 1\}$ for the ternary embedding mode.

Besides, $h(x)$ is the binary quantizer that returns 0 when $x \geq 0$ while returns 1 otherwise. And $\mathcal{P}(x) = x \mod 2$ is defined for extracting LSB of input element. $H_q(x)$ is used for representing q-ary entropy function. Typically, the binary entropy function is defined as $H_2(x) = -[x \log_2(x) + (1-x)\log_2(1-x)]$.

2.1 The Theory of Optimal Steganographic Embedding

For a given cover sequence $\mathbf{x} \in \mathcal{X} \triangleq \{\mathcal{L}\}^n$, the stego sequence can be denoted as $\mathbf{y} \in \{\mathcal{I}_i\}^n \subset \mathcal{X}$ when the modified pattern $\mathcal{I}_i \subset \mathcal{I}$ for $\mathbf{y}$ is defined. In this case, the process of steganography can be defined by modification transition distribution $\pi(\mathbf{y}) \triangleq \mathrm{P}(\mathbf{y} \mid \mathbf{x})$.

Under the additive model, if the message $\mathbf{m}$ with length of $|\mathbf{m}|$ to be embedded is given, to obtain the optimal distribution of $\pi(\mathbf{y})$, it is equivalent to solving the PLS problem which is as follows

$$\min_{\boldsymbol{\pi}} E_{\boldsymbol{\pi}}[D] = \sum_{\mathbf{y} \in \mathcal{Y}} \pi(\mathbf{y}) D(\mathbf{y}) \tag{1}$$

$$\text{subject to } H(\boldsymbol{\pi}) = -\sum_{\mathbf{y} \in \mathcal{Y}} \pi(\mathbf{y}) \log_2 \pi(\mathbf{y}) = |\mathbf{m}| \tag{2}$$

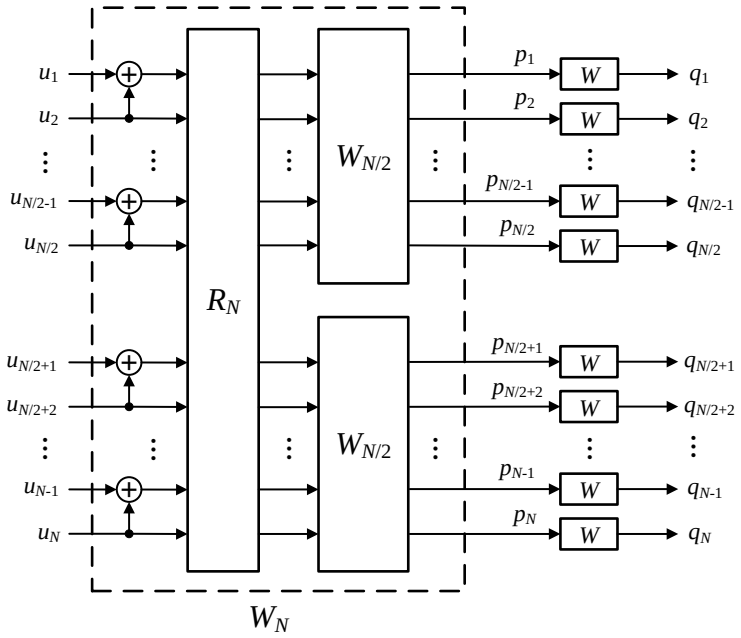


Fig. 1: Illustration of recursive construction of polarized channel W_N.

where $\boldsymbol{\rho} = (\rho(y_1), \rho(y_2), \cdots, \rho(y_n))$ denotes the pre-defined additive distortion of each element and

$$D(\mathbf{y}) = \sum_{i=1}^{n} \rho(y_i) \tag{3}$$

denotes the distortion under given modification pattern. This optimization problem can be solved with Lagrangian multiplier [10]. And the optimal distribution of modification for each element is

$$\pi_\lambda(y_i) = \frac{\exp[-\lambda\rho(y_i)]}{\sum_{y' \in \mathcal{I}_i} \exp[-\lambda\rho(y')]} \tag{4}$$

where $\lambda > 0$ is a scalar parameter determined by Equation (2). The distribution above is the best mapping from modification probability to steganographic distortion.

2.2 Channel Polarization

Channel polarization [2] is the method for constructing polar codes, the first provable capacity-achieving channel code. For distinguishing them from the notations in steganography, use $\mathcal{P} = \{0, 1\}$ and $\mathcal{Q}$ to represent the input and output

alphabets, respectively. When given a B-DMC W: $\mathcal{P} \to \mathcal{Q}$ with transition probabilities $W(q \mid p)$, $p \in \mathcal{P}$, $q \in \mathcal{Q}$, the process of channel polarization is mainly consists of two phases: Channel Combining and Channel Splitting.

After which, N synthesized bit channels $W_N^{(i)}$ are polarized and have symmetric capacity either close to 0 or close to 1 as N approaches infinity. It is proved that the channel capacity can be achieved by transmitting information bits using the noiseless channels [2].

Figure 1 gives the general recursive form of coding procedure of polar codes with the length of $N = 2^n$. Based on which the linear mapping $\mathbf{u} \to \mathbf{p}$, i.e., the encoding procedure of polar codes, is established which can be expressed as $\mathbf{p} = \mathbf{u}\mathbf{G}_N$ where $\mathbf{G}_N$ is the transform matrix that

$$\mathbf{G}_N = \mathbf{B}_N \mathbf{F}_2^{\otimes n}, \quad \mathbf{F}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{5}$$

$\mathbf{B}_N$ is the bit-reversal permutation matrix consists of the operation R_N and $\otimes$ denotes the Kronecker power.

Denote the set of indices of bit channels with smallest code rate as $\mathcal{A}^{\mathrm{c}}$, the most critical step in polar codes construction is how to determine $\mathcal{A}^{\mathrm{c}}$, i.e., the indices of frozen bits. As for B-DMC, Bhattacharyya parameter Z is generally used as a measure of quality of split bit channels, which can be calculated by

$$Z(W_{2N}^{(2i-1)}) \leq 2Z(W_N^{(i)}) - \left[Z(W_N^{(i)})\right]^2, \tag{6}$$

$$Z(W_{2N}^{(2i)}) = \left[Z(W_N^{(i)})\right]^2, \tag{7}$$

$1 \leq i \leq N$ where the equality holds if and only if W is a binary erasure channel (BEC). And the initial value of which is defined as

$$Z(W_1^{(1)}) = Z(W) = \sum_{q \in \mathcal{Q}} \sqrt{W(q \mid p = 0) \cdot W(q \mid p = 1)}. \tag{8}$$

Thus, the smaller the $Z(W_N^{(i)})$, the more reliable the $W_N^{(i)}$ is. After which a $(N, K, \mathcal{A}^{\mathrm{c}}, \mathbf{u}_{\mathcal{A}^{\mathrm{c}}})$ polar code is specified where $\mathbf{u} = (\mathbf{u}_{\mathcal{A}}, \mathbf{u}_{\mathcal{A}^{\mathrm{c}}})$ denotes the source word and $\mathbf{u}_{\mathcal{A}^{\mathrm{c}}}$ represents the frozen bits of length $N - K$.

2.3 Decoding of Polar Codes

Since each coordinate channel $W_N^{(i)}$ are successively used in polarized channels, in the decoding procedure, the unfrozen source word u_i can be calculated by previous estimated source words and received codewords, which is

$$\hat{u}_i = \arg \max_{u_i \in \{0,1\}} W_N^{(i)}(\mathbf{r}, \hat{u}_1, \hat{u}_2, \cdots, \hat{u}_{i-1} \mid u_i). \tag{9}$$

This method is SC decoding algorithm [2]. To simplify the calculation, the LLR of each coordinate channel, which is defined as

$$L_N^{(i)} = \ln \frac{W_N^{(i)}(\mathbf{r}, \hat{u}_1, \hat{u}_2, \cdots, \hat{u}_{i-1} \mid 0)}{W_N^{(i)}(\mathbf{r}, \hat{u}_1, \hat{u}_2, \cdots, \hat{u}_{i-1} \mid 1)} \tag{10}$$

is used for codeword decision [19]. Therefore, the $\hat{u}_i = 0$ if $L_N^{(i)} \leq 0$ otherwise $\hat{u}_i = 1$. The calculation of LLR can be recursively formulated by

$$L_{2N}^{(2i-1)} = F\left(L_N^{(i)}, L_N^{(i+N)}\right), \tag{11}$$

$$L_{2N}^{(2i)} = G\left(L_N^{(i)}, L_N^{(i+N)}, \hat{u}_{2i-1}\right) \tag{12}$$

which further implies that the decision of current codeword strongly depends on the previous estimated bits. $F(\cdot)$ and $G(\cdot)$ are the f and g functions defined in [2] in logarithm domain, which is

$$F(a,b) = 2\tanh^{-1}\left(\tanh\left(\frac{a}{2}\right)\tanh\left(\frac{b}{2}\right)\right), \tag{13}$$

$$G(a,b,\omega) = (1-2\omega)a + b. \tag{14}$$

To overcome the errors accumulated in the successive decoding process, in SCL decoder [25], L lists of candidate codewords, i.e., decoding paths, are reserved when estimating each bit as well as path metrics (PM) of each path. For unfrozen bit, each path generates two paths by decoding $\hat{u}_i = 0$ and $\hat{u}_i = 1$ therefore a total $2L$ paths are obtained. The metric of l-th path can be updated at the i-th decoding bit by

$$\mathrm{PM}_i^l = \sum_{k=1}^{i} \ln\left(1 + \exp\left(-\left(1-2\hat{u}_k\right)L_N^{(k)}\right)\right) \tag{15}$$

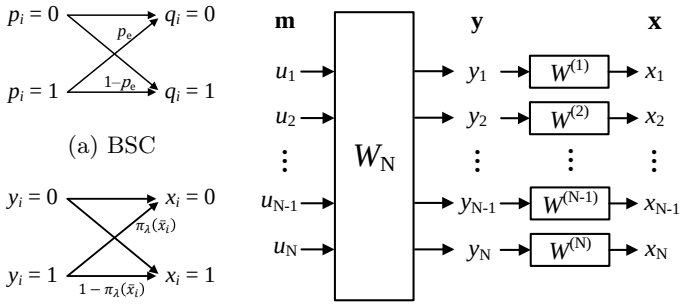
and only L paths with lowest PM are maintained for further decoding. The accumulation of errors is greatly reduced with listed SC decoder and the decoding performance is better improved.

3 Steganographic Coding on Sub-Polarized Channel

In this section, the polarized steganographic channel and its sub-channels are established through optimal embedding theory. Besides, the steganographic coding methods under the typical sub-channels are given.

3.1 Polarized Steganographic Channel

Under the optimal embedding theory, the steganographic embedding process for each cover element can be simulated as a communication process under a lossy channel [21], which is shown in Figure 2b where the cover element changes to the stego through the decoding process under the simulated channel. This model, named binary steganographic channel (BSteC), is obviously equivalent to BSC where the modification probability $\pi_\lambda(\bar{x}_i)$ is equal to the crossover probability p_e of BSC. As a result, the steganographic coding is formulated by the problem that, given the received codeword (cover) $\mathbf{x} = (x_1, x_2, \cdots, x_N)$, decoding the



(a) BSC

(b) BSteC

(c) Polarized steganographic channel

Fig. 2: Demonstration of the relationship between communication channel and embedding of steganography, where $W^{(i)}$, $1 \leq i \leq N$ represents BSteC.

codeword before transmitting (stego) $\mathbf{y}$ of which syndrome (message) is $\mathbf{m}$ under N independent BSCs with crossover probability $\pi_\lambda(\bar{x}_i)$.

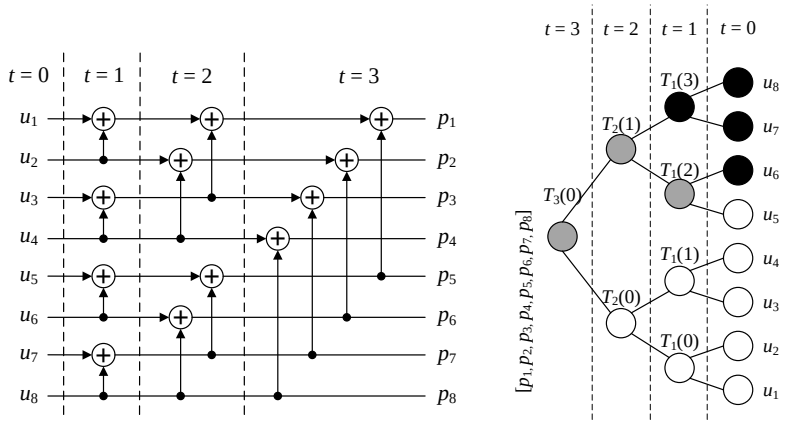
The model of polarized steganographic channel is shown in Figure 2c where multiple different steganographic channels instead of independent copies are combined and split as the modification pattern of each cover element is independent and different. Under which the steganographer can perform steganographic coding by constructing polar codes with non-uniform channel polarization [22].

However, some prior knowledge, such as embedding distortion, is required to be shared with the extractor for reconstruction of polarized channels, which is nearly impossible. As pointed out in [29], the polar codes constructed for BSC by Equation (6) and (7) with the equal sign holds still has good performance. Therefore, it can be assumed that all BSteCs are identical and treated as BECs in the construction of polarized steganographic channel. The Bhattacharyya parameter can be used as a metric of channel quality since it is currently the best for BEC and BSC as discussed in [21,29]. Besides, the initial value of Bhattacharyya parameter is discussed in Section 4.1.

3.2 Successive Cancellation on Polarized Steganographic Channel

According to the recursive form of channel polarization, $N = 2^n$ length polar codes can be represented by a binary tree $T_n(0)$ of depth n [1]. As shown in Figure 3, each node $T_t(\phi)$ corresponds to a codeword and has a left child $T_{t-1}(2\phi)$ and right child $T_{t-1}(2\phi + 1)$. As a result, each sub-tree with a node corresponds to a sub-polar codes which is constructed through a sub-polarized channel.

In the decoding process, SC and SCL decoder sequentially estimates each codeword in a depth-first order. However, partially sub-polar codes have a special form based on the position of frozen bits in the source words, which consists of a special form of codeword. Therefore a limited number of candidate codewords can be directly estimated without recursively calculating all LLRs for decoding all



(a) Channel Polarization (b) Tree Representation

Fig. 3: Polar codes and its tree representation with length of 8, where u_1 to u_5 is assigned as frozen bits. In the binary tree, all white nodes represent frozen bits while all black nodes denote unfrozen bits.

source words. For steganographic codes, the coding scheme under the identified sub-polarized channel can not only reduce the accumulation of distortions in the successive process but also greatly improve computational efficiency.

In this paper, denoted $\mathcal{V}$ as the set of nodes of sub-polar codes, 2-dim tuple $(k_i, S_i) \in \mathcal{V}$ is used to denote a node of polar code with length of S_i of which source word $(p_{k_i}, p_{k_i+1}, \cdots, p_{k_i+S_i-1}) = (u_{k_i}, u_{k_i+1}, \cdots, u_{k_i+S_i-1})\mathbf{G}_{S_i}$. The detailed steganographic coding scheme under the sub-polarized channel is listed in Algorithm 1.

3.3 Steganographic Coding Under the Typical Sub-Channel

In this subsection, the corresponding algorithms of steganographic coding will be given under four typical form of sub-polarized channels. The listed successive coding scheme is adopted for better coding efficiency. Based on an important theorem in [16, Theorem 1], the calculation of path metric defined in Equation (15) can be updated as

$$\mathrm{PM}_i^l = \sum_{k=1}^{i} \ln\left(1 + \exp\left(-\left(1 - 2\hat{p}_k\right) L_1^{(k)}\right)\right) \tag{16}$$

where $(\hat{p}_1, \hat{p}_2, \cdots, \hat{p}_S) = (\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_S)\mathbf{G}_S$. This equation will be adopted further in the steganographic coding procedure. Note that any polarized channel can be represented by the sub-channels below, since the shortest length of codeword is 2 under these sub-channels.

Algorithm 1: Binary Embedding of Steganography Under the Sub-Polarized Steganographic Channel

Input: cover $\mathbf{x}$, message $\mathbf{m}$, costs $\boldsymbol{\rho}$ and list size L.
Output: stego $\mathbf{y}$ and total distortion $D(\mathbf{y})$.

calculate Bhattacharyya parameter $Z(W_N^{(i)})$ for each bit channel by Equation (6) and (7) with initial value $Z(W_1^{(1)})$ by Equation (19);
get the set of indices $\mathcal{A}^{\mathrm{c}}$ of bit channels with largest metric $Z(W_N^{(i)})$ where $|\mathcal{A}^{\mathrm{c}}| = |\mathbf{m}|$; set the frozen bits of polarized steganographic channels as message $\mathbf{m}$, which is $\mathbf{u}_{\mathcal{A}^{\mathrm{c}}} = \mathbf{m}$;
calculate optimal probability of modification of each element $\pi_\lambda(y_i)$ by Equation (4); and further calculate the LLR $L_1^{(i)}$ of each cover element with $\mathcal{P}(\mathbf{x})$ by Equation (10);
recursively identify the type of sub-polarized channel (k_i, S_i) and storage 2-dim tuples in the set $\mathcal{V}$;
foreach *sub node* (k_i, S_i) *in set* $\mathcal{V}$ **do**
    recursively calculate the LLR $L_N^{(k_i)}$ for k_i-th source word;
    according to the type of sub node, select the corresponding coding scheme in Subsection 3.3 to generate the estimated codeword $\hat{\mathbf{p}}$;
    obtain the estimated source word $(\hat{u}_{k_i}, \hat{u}_{k_i+1}, \cdots, \hat{u}_{k_i+S_i-1})$ by polar encoding (5);
    recursively update all partial sums, i.e., intermediate codewords of all coding paths;
end
calculate $\mathcal{P}(\mathbf{y})$ with $\hat{\mathbf{u}}$ by polar encoding (5);
get the stego $\mathbf{y} = \mathbf{x} - \mathcal{P}(\mathbf{x}) + \mathcal{P}(\mathbf{y})$; calculate $D(\mathbf{y}) = \sum \rho(y_i)$.

R1 (Rate-1) Channel. Under the Rate-1 channel denoted as (k_i, S_i), all coordinate channels are used to transmit frozen bits which implies that the message transmission rate is 1. For the steganographic coding, there is only one valid codeword $\mathbf{p}_0 = (u_{k_i}, u_{k_i+1}, \cdots, u_{k_i+S_i-1})\,\mathbf{G}_{S_i}$ where $u_k \in \mathcal{A}^{\mathrm{c}}$, $k_i \le k < k_i + S$ is set by the message $\mathbf{m}$, thereafter no path splitting occurs. Besides, since the coding procedure is all conducted on frozen bits, the increment of path metrics $\Delta\mathrm{PM}^l$ are directly updated by Equation (16).

DC (Dual Candidate) Channel. The node of polar codes based on DC channel is already discussed in [24, Section IV-B]. Under the DC channels denoted by (k_i, S_i), all source words are determined as frozen bit except the last bit $u_{k_i+S_i-1}$. As a result, there are only two candidate codewords exists, which are $\mathbf{p}_0 = (u_{k_i}, \cdots, u_{k_i+S_i-2}, 0)\,\mathbf{G}_{S_i}$ and $\mathbf{p}_1 = (u_{k_i}, \cdots, u_{k_i+S_i-2}, 1)\,\mathbf{G}_{S_i}$. Each path generates two candidate paths whose PM are updated by Equation (16).

R0 (Rate-0) Channel. In Rate-0 node, all source words are unfrozen bits. The maximum likelihood (ML) decision of codeword $p_k, k_i \le k < k_i + S_i$ discussed

in [1, Lemma 1] are $\hat{p}_k = h(L_{S_i}^{(k)})$. However, $\hat{p}_k$ is not necessarily the best coding result. The other near-ML decisions have to be obtained for better performance. As for list coder with L paths, the near-ML codes can be obtained by only flipping each codeword in $\hat{p}_k$ the ascending order of corresponding LLR $L_{S_i}^{(k)}$. After each flipping, twice as many near-ML codewords are generated while at most L candidate codewords are reserved based on the path metric increment $\Delta\mathrm{PM}^l$. This operation is only performed on the codewords corresponding to the first $L-1$ smallest LLRs. After that, L paths of candidates will be generated.

SPC (Single Parity Check) Channel. For the node constructed by SPC channel denoted by (k_i, S_i), only the first coordinate channel transmits frozen bits, i.e., $u_{k_i} \in \mathcal{A}^{\mathrm{c}}$ and $u_{k_i+1}, u_{k_i+2}, \cdots, u_{k_i+S_i-1} \notin \mathcal{A}^{\mathrm{c}}$. The parity of all source word in this sub-polar codes equals to u_{k_i}.

Theorem 1. *In SPC channel, the parity of codeword of sub-polar codes satisfies*

$$P = \bigoplus_{k=k_i}^{k_i+S_i-1} p_k = u_{k_i}$$

where $\oplus$ represents modulo 2 addition.

Proof. We proof this theorem with induction. For any polar codes with $n = 1$, the codeword (p_1, p_2) equals to $(u_1, u_2)\mathbf{G}_2 = (u_1 \oplus u_2, u_2)$. The theorem holds since $p_1 \oplus p_2 = u_1 \oplus u_2 \oplus u_2 = u_1$. Now suppose the theorem stands for polar codes of SPC mode with $n = k$. For $n = k+1$, denote the source word $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2)$ where $\mathbf{u}_1 = (u_1, u_2, \cdots, u_{2^k})$ and $\mathbf{u}_2 = (u_{2^k+1}, \cdots, u_{2^{k+1}})$ are two separated words. Then for $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2) = \mathbf{u}\mathbf{B}_{2^{k+1}}\mathbf{F}_2^{\otimes k+1}$ where $\mathbf{p}_1 = (p_1, p_3, \cdots, p_{2^{k+1}-1})$ and $\mathbf{p}_2 = (p_2, p_4, \cdots, p_{2^{k+1}})$, there are

$$\begin{aligned}\mathbf{p}_1 &= (u_1, u_2, \cdots, u_{2^k})\mathbf{B}_{2^k}\mathbf{F}_2^{\otimes k} \oplus (u_{2^k+1}, \cdots, u_{2^{k+1}})\mathbf{B}_{2^k}\mathbf{F}_2^{\otimes k}\\ \mathbf{p}_2 &= (u_{2^k+1}, \cdots, u_{2^{k+1}})\mathbf{B}_{2^k}\mathbf{F}_2^{\otimes k}\end{aligned}$$

where

$$\mathbf{B}_{2^{k+1}} = \mathbf{R}_{2^k}(\mathbf{I}_2 \otimes \mathbf{B}_{2^k}), \quad \mathbf{F}_2^{\otimes k+1} = \begin{bmatrix}\mathbf{F}_2^{\otimes k} & \mathbf{0}\\ \mathbf{F}_2^{\otimes k} & \mathbf{F}_2^{\otimes k}\end{bmatrix}.$$

Since the bit-reversal permutation matrix only changes the order of codewords, the parity of $\mathbf{p}$ can be obviously calculated by codeword $(u_1, u_2, \cdots, u_{2^k})\mathbf{B}_{2^k}\mathbf{F}_2^{\otimes k}$ which is equivalent to the hypothesis above. Therefore, the theorem is valid for all positive integer n.

Similar to the node under the Rate-0 channel, the ML decision of codewords of SPC node are $\hat{p}_k = h(L_{S_i}^{(k)})$ if the parity of codewords equals to u_{k_i}. Otherwise, flip the codeword with the smallest LLR to satisfy the requirements of channel. The other $L-1$ near-ML decisions can still be obtained by sequentially flipping the codewords corresponding to the 2nd to $(L-1)$-th smallest LLRs. Besides, it is also necessary to flip the codeword corresponding to the minimum LLR to ensure the validity of the codewords.

4 Experimental Results

In this section, experiments are mainly conducted on binary embedding by various distortion profiles with randomly generated cover elements and messages. The embedding efficiency $e = |\mathbf{m}|/D(\mathbf{y})$ is used for comparing the performance with the state-of-the-art steganographic coding schemes, where $e_{\boldsymbol{\pi}} = |\mathbf{m}|/E_{\boldsymbol{\pi}}(D)$ is the theoretical upper bound of embedding efficiency. The throughput of coding, i.e., average number of cover elements processed per seconds, is also investigated to evaluate the computational efficiency. The distortion profile denoted by $\boldsymbol{\varrho} = (\varrho_1, \varrho_2, \cdots, \varrho_N)$ is defined as $\varrho_i = \varrho(i/N)$ [8,9]. The constant, linear and square profile are used for evaluation.

For comparison, STC [9] with sub-matrix height $h = 8, 10, 12$ and SPC [21] with list size $l = 1, 4, 16$ are introduced in simulations. The polar codes-based schemes are implemented in C++. The F function defined in Equation (11) is approximated when $|a| \geq 10$ or $|b| \geq 10$ by $F(a,b) \approx \operatorname{sgn}(a)\operatorname{sgn}(b)\min\{|a|,|b|\}$ to avoid overflow in exponent calculation, where $\operatorname{sgn}(\cdot)$ is the signum function.

4.1 Construction of Polarized Steganographic Channel

As discussed above, the key point of polarized steganographic channel construction is to calculate the initial value of the Bhattacharyya parameter. In this paper, three heuristically defined strategies are discussed, which are

$$\text{Type I}: \qquad Z(W_1^{(1)}) = \frac{1}{n}\sum_{i=1}^{n} 2\sqrt{\pi_\lambda(\bar{x}_i)\left(1-\pi_\lambda(\bar{x}_i)\right)}, \tag{17}$$

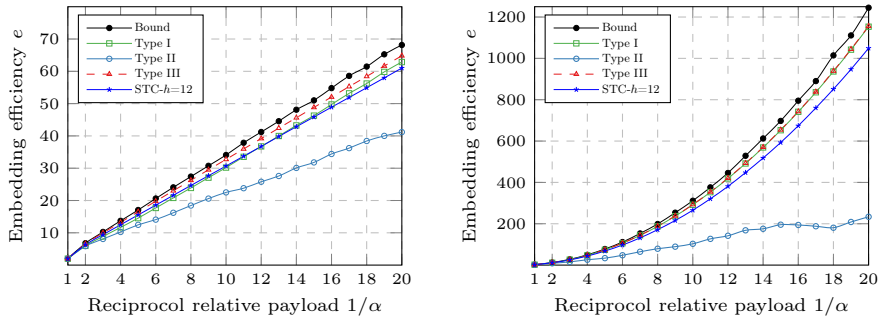
$$\text{Type II}: \qquad Z(W_1^{(1)}) = H_2\left(\frac{1}{n}\sum_{i=1}^{n}\pi_\lambda(\bar{x}_i)\right), \tag{18}$$

$$\text{Type III}: \qquad Z(W_1^{(1)}) = \frac{1}{n}\sum_{i=1}^{n} H_2(\pi_\lambda(\bar{x}_i)) = \frac{|\mathbf{m}|}{n}. \tag{19}$$

Under the linear and square distortion profile, the embedding efficiency of the three strategies in Equation (17), (18) and (19) with proposed scheme of list size of 8 are all evaluated for 100 times, which shown in Figure 4. The strategy of Type I and Type III that outperform STC both achieve near-optimal performance while Type III performs better. In the proposed scheme, we use the strategy in Type III for channel reliability initialization.

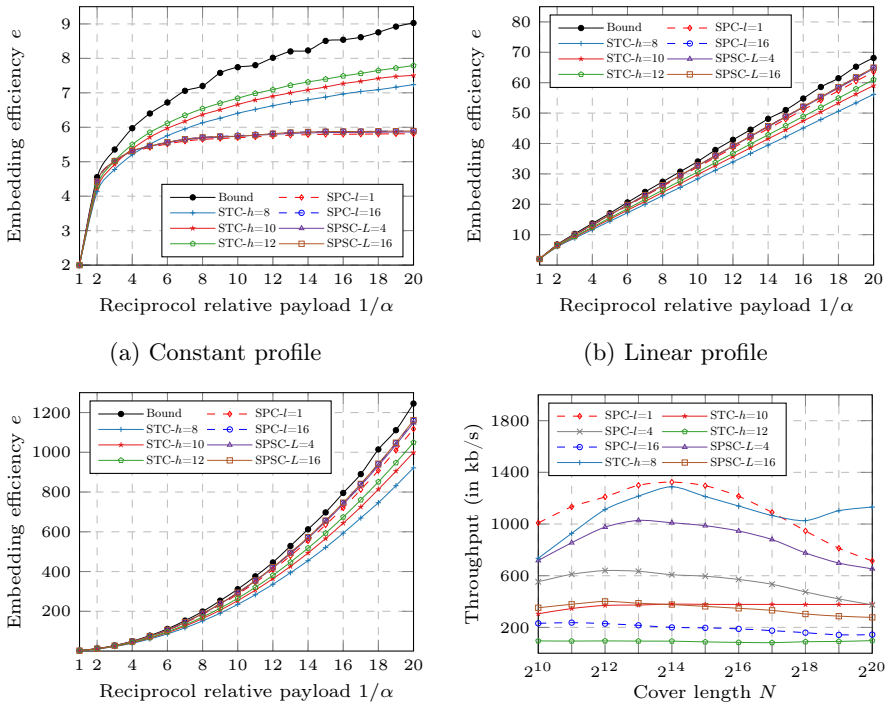
4.2 Security Evaluation under Embedding Efficiency Results

The embedding efficiencies of the proposed scheme are simulated for three typical distortion profiles, which is shown in Figure 5a, 5b and 5c. For the constant profile, the polar codes-based coding schemes are worse than STC at the small payloads. While for linear and square profiles, SPC and SPSC both perform



(a) Linear profile (b) Square profile

Fig. 4: Embedding performance of three different strategy for construction of polarized channel of which list size L is set to 8. Cover elements with length of 2^{20} are randomly generated as well as messages.



(a) Constant profile (b) Linear profile

(c) Square profile (d) Computational efficiency

Fig. 5: Embedding performance of different steganographic coding schemes.(a)-(c) Embedding efficiency for constant, linear and square distortion profiles. (d) Average throughput of different coding schemes with different parameters, which are all evaluated by executable version compiled with C++ on Intel(R) Core(TM) i7-4790 CPU @ 3.60GHz.

Table 1: The detection error rate (in %) of steganalysis tools where features are extracted by DCTR. STC, SPC and SPSC are compared with J-UNIWARD (JUNI) and UERD on BOSSBase with different quality factors.

Method	QF=75				QF=90			
	0.1	0.2	0.3	0.4	0.1	0.2	0.3	0.4
JUNI-STC-h=8	0.4099	0.2789	0.1504	0.0618	0.4520	0.3593	0.2434	0.1300
JUNI-STC-h=12	**0.4161**	0.2909	0.1663	0.0735	0.4564	0.3680	0.2589	0.1470
JUNI-SPC-l=4	0.4147	0.2935	0.1710	0.0785	0.4564	0.3715	0.2634	0.1533
JUNI-SPC-l=16	0.4155	**0.2953**	0.1719	**0.0806**	**0.4566**	**0.3725**	**0.2656**	0.1558
JUNI-SPSC-L=4	0.4153	0.2934	**0.1731**	0.0786	0.4562	0.3705	0.2650	0.1538
JUNI-SPSC-L=16	0.4158	0.2941	0.1726	0.0803	0.4553	0.3719	0.2648	**0.1562**
UERD-STC-h=8	0.4009	0.2701	0.1537	0.0708	0.4425	0.3430	0.2297	0.1258
UERD-STC-h=12	**0.4069**	0.2851	0.1699	0.0831	0.4462	0.3517	0.2434	0.1414
UERD-SPC-l=4	0.4061	0.2865	0.1725	0.0874	**0.4471**	0.3541	0.2501	0.1487
UERD-SPC-l=16	0.4062	0.2873	**0.1754**	**0.0887**	0.4460	0.3555	**0.2504**	0.1491
UERD-SPSC-L=4	0.4064	0.2872	0.1732	0.0871	0.4461	**0.3570**	0.2498	0.1485
UERD-SPSC-L=16	0.4062	**0.2879**	0.1744	0.0884	0.4461	0.3561	0.2503	**0.1505**

closer to the theoretical bound compared with STC, and SPSC slightly outperforms SPC when the list size is the same. Through the sub-polarized steganographic channel, the proposed coding scheme can reduce the error propagation of recursively coding process and improve the embedding performance.

Besides, the computational efficiency is evaluated by the throughput with payload of 1/2. As demonstrated in Figure 5d, the polar codes-based schemes are more capable of performing steganographic coding at higher embedding efficiency with lower computation complexity than STC. The throughput of SPSC with list size of 4 is 6-8 times that of STC of $h = 12$ with SIMD instructions optimization. Meanwhile, with the same list size, the throughput of SPSC is about 1.5 times compared to the SPC. Therefore, the overall performance implies the proposed scheme has the potential for large-scale applications of steganography.

4.3 Security Evaluation under Image Steganalysis

The anti-steganalysis performance is evaluated by JPEG image steganography with DCTR [17] feature extractor. Samples sized 512×512 with quality factor of 75 and 90 are generated from BOSSBase [3] of RAW format. Stego counterparts are embedded by different coding schemes with J-UNIWARD and UERD. As demonstrated in Table 1, STC still achieves high security under lower relative payload, while polar codes-based coding schemes performs better when the payload increases. Therefore, compared with other near-optimal coding schemes, the proposed scheme has comparable security performance in JPEG image steganography but requires lower computational complexity.

5 Conclusion

In this paper, a near-optimal steganographic codes is proposed with constructed polarized and steganographic channel. Four typical sub-polarized channels are introduced which further verifies the performance of the proposed scheme. Compared with the state-of-the-art steganographic codes, experimental results show that the proposed codes performs embedding with higher efficiency and half of the time consumption, since which it also enables large-scale practical steganography applications even if on portable computing devices. In our future works, the proposed scheme will be further improved with the discovery of sub-polarized channels with more specific patterns.

References

1. Alamdar-Yazdi, A., Kschischang, F.R.: A simplified successive-cancellation decoder for polar codes. IEEE communications letters **15**(12), 1378–1380 (2011)
2. Arikan, E.: Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. IEEE Transactions on information Theory **55**(7), 3051–3073 (2009)
3. Bas, P., Filler, T., Pevnỳ, T.: " break our steganographic system": the ins and outs of organizing boss. In: International workshop on information hiding. pp. 59–70. Springer (2011)
4. Crandall, R.: Some notes on steganography. Posted on steganography mailing list **1998**, 1–6 (1998)
5. Diop, I., Farss, S., Tall, K., Fall, P., Diouf, M., Diop, A.: Adaptive steganography scheme based on ldpc codes. In: 16th International Conference on Advanced Communication Technology. pp. 162–166. IEEE (2014)
6. Diouf, B., Diop, I., Fall, P.A., Dolo, B., Diop, A.K., Diouf, M., Khouma, O., Farssi, S.M., Tall, K.: Jpeg steganography based on successive cancellation decoding of polar codes. In: 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET). pp. 1–6. IEEE (2022)
7. Diouf, B., Diop, I., Keita, K.W., Diouf, M., Farsi, S.M., Tall, K., Khouma, O.: Polar coding steganographic embedding using successive cancellation. In: Innovation and Interdisciplinary Solutions for Underserved Areas, pp. 189–201. Springer (2017)
8. Filler, T., Judas, J., Fridrich, J.: Minimizing embedding impact in steganography using trellis-coded quantization. In: Media forensics and security II. vol. 7541, pp. 38–51. SPIE (2010)
9. Filler, T., Judas, J., Fridrich, J.: Minimizing additive distortion in steganography using syndrome-trellis codes. IEEE Transactions on Information Forensics and Security **6**(3), 920–935 (2011)
10. Fridrich, J., Filler, T.: Practical methods for minimizing embedding impact in steganography. In: Security, Steganography, and Watermarking of Multimedia Contents IX. vol. 6505, pp. 13–27. SPIE (2007)
11. Fridrich, J., Goljan, M., Lisonek, P., Soukal, D.: Writing on wet paper. IEEE Transactions on signal processing **53**(10), 3923–3935 (2005)
12. Fridrich, J., Goljan, M., Soukal, D.: Efficient wet paper codes. In: International Workshop on Information Hiding. pp. 204–218. Springer (2005)

13. Guo, L., Ni, J., Shi, Y.Q.: An efficient jpeg steganographic scheme using uniform embedding. In: 2012 IEEE International Workshop on Information Forensics and Security (WIFS). pp. 169–174. IEEE (2012)
14. Guo, L., Ni, J., Su, W., Tang, C., Shi, Y.Q.: Using statistical image model for jpeg steganography: uniform embedding revisited. IEEE Transactions on Information Forensics and Security **10**(12), 2669–2680 (2015)
15. Hanif, M., Ardakani, M.: Fast successive-cancellation decoding of polar codes: Identification and decoding of new nodes. IEEE Communications Letters **21**(11), 2360–2363 (2017)
16. Hashemi, S.A., Condo, C., Gross, W.J.: A fast polar code list decoder architecture based on sphere decoding. IEEE Transactions on Circuits and Systems I: Regular Papers **63**(12), 2368–2380 (2016)
17. Holub, V., Fridrich, J.: Low-complexity features for jpeg steganalysis using undecimated dct. IEEE Transactions on Information forensics and security **10**(2), 219–228 (2014)
18. Holub, V., Fridrich, J., Denemark, T.: Universal distortion function for steganography in an arbitrary domain. EURASIP Journal on Information Security **2014**(1), 1 (2014)
19. Leroux, C., Tal, I., Vardy, A., Gross, W.J.: Hardware architectures for successive cancellation decoding of polar codes. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1665–1668. IEEE (2011)
20. Li, B., Wang, M., Huang, J., Li, X.: A new cost function for spatial image steganography. In: 2014 IEEE International Conference on Image Processing (ICIP). pp. 4206–4210. IEEE (2014)
21. Li, W., Zhang, W., Li, L., Zhou, H., Yu, N.: Designing near-optimal steganographic codes in practice based on polar codes. IEEE Transactions on Communications **68**(7), 3948–3962 (2020)
22. Oliveira, R.M., de Lamare, R.C.: Non-uniform channel polarization and design of rate-compatible polar codes. In: 2019 16th International Symposium on Wireless Communication Systems (ISWCS). pp. 537–541. IEEE (2019)
23. Pevnỳ, T., Filler, T., Bas, P.: Using high-dimensional image models to perform highly undetectable steganography. In: International workshop on information hiding. pp. 161–177. Springer (2010)
24. Sarkis, G., Giard, P., Vardy, A., Thibeault, C., Gross, W.J.: Fast polar decoders: Algorithm and implementation. IEEE Journal on Selected Areas in Communications **32**(5), 946–957 (2014)
25. Tal, I., Vardy, A.: List decoding of polar codes. IEEE Transactions on Information Theory **61**(5), 2213–2226 (2015)
26. Westfeld, A.: High capacity despite better steganalysis (f5—a steganographic algorithm). In: Information Hiding: 4th International Workshop, IH 2001, Pittsburgh, PA, USA, April 25-27, 2001. Proceedings. vol. 2137, p. 289. Springer (2001)
27. Zhang, R., Sachnev, V., Botnan, M.B., Kim, H.J., Heo, J.: An efficient embedder for bch coding for steganography. IEEE Transactions on Information Theory **58**(12), 7272–7279 (2012)
28. Zhang, R., Sachnev, V., Kim, H.J.: Fast bch syndrome coding for steganography. In: International Workshop on Information Hiding. pp. 48–58. Springer (2009)
29. Zhao, S., Shi, P., Wang, B.: Designs of bhattacharyya parameter in the construction of polar codes. In: 2011 7th International conference on wireless communications, networking and mobile computing. pp. 1–4. IEEE (2011)