Secret Key Cryptosystem based on Non-Systematic Polar Codes

Reza Hooshmand
Department of Electrical Engineering, Science and Research Branch
Islamic Azad University, Tehran, Iran
r.hooshmand@srbiau.ac.ir

Abstract **- Polar codes are provably capacity achieving linear block codes. The generator matrix of these codes is specified by knowing the parameters of transmission channel, length and dimension of the used code. On the other hand, for the cryptosystems based on general decoding problem (i.e. code based cryptosystems), the generator matrix of the applied code should be properly hidden from the attacker. Moreover, in the computational security, it is assumed that an attacker with restricted processing power has unlimited access to transmission media. Thus, an attacker can construct the generator matrix of polar codes, especially for Binary Erasure Channel on which this matrix can be efficiently specified.**

In this paper, we introduce a novel method to hide the generator matrix of polar codes in such a way that an attacker cannot construct it in polynomial time even by knowledge of the channel parameters, dimension and length of the used code. By the help of this method, a secret key cryptosystem based on non-systematic polar codes over Binary Erasure Channel is proposed which provides both data security and reliability in one process simultaneously. In fact, the main goal of this research is to achieve the acceptable level of security and reliability by taking advantage of the interesting properties of polar codes. The proposed scheme resists against the typical attacks on the cryptosystems based on error correcting codes. Also, by employing some efficient methods, the key length of our scheme is decreased compared to Rao-Nam secret key cryptosystem. Moreover, our scheme benefits from high code rate, proper error performance, faster processing and efficient implementation.

Keywords **- Secret key cryptography; Code based cryptosystem; Polar codes.**

I. INTRODUCTION

Nowadays, development and rapid dissemination of wireless communication systems have increased the demand for providing data reliability and security. Channel coding is the study of techniques for achieving reliable communication between a sender and a receiver in the presence of channel errors. Also, cryptography is the study of methods for establishing secure communications in the presence of adversaries. In general, channel coding can be applied to provide two major categories of security; the information theoretic security and the computational security. Utilizing the practical channel codes such as Low Density Parity Check (LDPC) codes [1] and Polar codes [2] in the structure of wiretap channel [3] to achieve secrecy capacity is an instance of applying channel codes in establishing the information theoretic security [4, 5]. Furthermore, utilizing the various channel codes in the structure of public/secret key code based cryptosystems is an application of channel coding to provide the computational security [6, 7].

Code based cryptosystems provide security and reliability in one process to guarantee the confidentiality and the integrity of transmitted data. Also combining the security and reliability in the structure of these systems can be resulted in reducing the processing cost or providing more efficient implementation. Moreover, code based cryptosystems are one of the important classes of cryptographic systems which are believed to resist quantum computers [8]. Establishing a suitable trade-off between security and reliability is one of the important goals in designing such cryptosystems. That can properly be achieved by employing efficient linear codes in the structure of these cryptosystems. The security of some code based cryptosystems is based on the difficulty of general decoding problem [9].

For an arbitrary binary linear code C with length N and dimension K, the general decoding problem is the problem of decoding a received vector $Y = (y_1, y_2, \cdots, y_N)$ into the closest codeword $X = (x_1, x_2, \cdots, x_N)$. In this case, the Hamming distance between X and Y, $d_H(X,Y) = \left|\{i | 1 \le i \le N, x_i \ne y_j\}\right|$, is minimal [10]. It was proved that the decoding problem of arbitrary linear codes belongs to the class of NP-complete problems [9]. In fact, the general decoding problem implies that having no knowledge about the structure of generator matrix G and the parity check matrix H of an arbitrary linear code, finding an efficient decoding algorithm which decodes the received vector Y into a closest codeword X is an NP-complete problem.

A. Related Works

As mentioned before, the difficulty of the general decoding problem has been used extensively in many cryptographic applications such as public/secret key code based cryptosystems. In 1978, McEliece proposed the first public key cryptosystem which was based on Goppa codes [6]. This scheme is one of the candidates for post quantum cryptography which can be an alternative to common public key cryptosystems which are based on number theory. McEliece cryptosystem has high speed encryption/decryption algorithms compared with the other public key cryptosystems. However,

this scheme has some weaknesses such as low information rate and large key size.

Public/secret key variants of McEliece cryptosystem have already been introduced. In 1984, the first secret key code based cryptosystem, was suggested by Rao [11]. This scheme is similar to McEliece cryptosystem but the public key is kept secret. Later, it was shown that Rao's scheme is broken by chosen plaintext attacks [7, 12]. In 1986, Rao and Nam introduced a modified secret key cryptosystem which allowed the use of short length Hamming codes with high information rate while improving the security level [12]. The modified scheme was called Rao-Nam (RN) cryptosystem which is got much attention and became the reference in the secret key cryptosystems based on coding theory. The security of the RN scheme relies on the hardness of decoding general linear codes, as well as McEliece cryptosystem. The structure of RN scheme and its weaknesses will be discussed in section III.

Many modifications to RN scheme have already been proposed which are based on either applying various channel codes in its structure or modifying the set of allowed error vectors. In 1987, Struik and Tilburg proved that the RN scheme is insecure against chosen plaintext attacks for practical code lengths [13]. The same work to analyze the security of the RN scheme based on nonlinear codes can also be found in [14]. In 1993, Alencar et al. proposed a secret key cryptosystem based on burst error correcting codes. The idea of this scheme was based on the fact that single burst error correcting capacity of a code is generally larger than its random error-correcting capacity [15]. Later, Sun and Shieh gave their comment on Alencar et al.'s scheme which shows that their scheme is also insecure against chosen plaintext attacks [16].

The technique of using a class of array codes was introduced by Compello de Souza et al. [17]. However, A1 Jabri presented a chosen plaintext attack against the cryptosystem based on array codes [18]. Also, by removing the security flaws in [15, 17], Sun improved the scheme based on burst error correcting codes which is secure against chosen plaintext attacks [19]. Another scheme was suggested by Barbero and Ytrehus in which several modifications were introduced to RN cryptosystem. The goal of these modifications was to reduce the key length, while improving the security level of RN system. The complexity of this scheme was increased to provide security [20].

In recent years, some efficient and secure secret key cryptosystems based on Turbo codes [1] and LDPC codes are introduced. Turbo codes have been employed in different secure channel coding schemes to use in satellite communications [21, 22]. Also, the issue of using quasi-cyclic low-density parity-check (QC-LDPC) codes in secret key cryptosystems is addressed in [23, 24]. Due to cyclic and sparse structure of the parity check matrix of QC-LDPC codes, the key lengths of these schemes have decreased significantly compared to previous RN-like schemes.

The idea of applying polar codes to provide information theoretic security was discussed extensively in several researches [5, 25, 26]. However, in spite of interesting properties of the polar codes which will be explained in section II, these efficient codes have not been applied in the structure of cryptosystems based on general decoding problem. Recently, we introduced, for the first time to our knowledge, the application of polar codes in the structure of secret key cryptosystem over binary erasure channel [27]. In fact, this paper is continuation and extension of our previous work [27] in the context of secret key cryptosystems based on channel coding.

B. Our Contributions

In this paper, we introduce a secret key cryptosystem which makes use of non-systematic finite length polar codes in an efficient way to overcome the problems raised from insecure and unreliable communication channels. Our scheme is designed in such a way to avoid the weaknesses of RN cryptosystem and is expected to provide more security and reliability. The main contribution of this work is to propose a technique for hiding the generator matrix of the polar codes from the attacker. By the help of this method, the underlying cryptosystem can achieve a proper security level based on general decoding problem. As aforementioned, in the computational security, it is assumed that the attacker has unlimited access to transmission channel. Moreover, the generator matrix of the polar codes has a channel dependent structure. Therefore, the attacker can specify the generator matrix of these codes via the channel parameters, length and dimension of the used polar codes. Indeed, the main question is that how to hide the generator matrix of polar codes from the attacker to be able to employ these codes in the structure of cryptosystems based on hardness of general decoding problem. In this work, we propose an efficient method to solve this problem.

The proposed scheme is examined against all cryptanalytic attacks which have already been presented on the cryptosystems based on channel coding. It is shown that our scheme resists these attacks. Also, we investigate its error performance, key length and computational complexity to evaluate its efficiency. For evaluating the reliability, the upper bound on error probability of the used polar code over BEC under Successive Cancelation (SC) decoding is analyzed in detail. Moreover, it is shown that what parameters should be considered to make secure and reliable communication simultaneously. In fact, the proper trade-off between the security and reliability is attained in the proposed scheme.

To decrease the key size of this scheme, we consider several efficient approaches such as, (1) Utilizing the interesting structure of the generator matrix of polar codes to reduce its memory requirements, (2) Applying the pseudorandom number generator algorithms to reduce the memory requirements of nonsingular and permutation matrices, and (3) Exploiting the pseudorandom syndromes as frozen vectors of the non-systematic polar codes to generate error vectors and so to omit the syndrome error table of RN cryptosystem.

C. Outline

The rest of this paper is organized in the following way. Sections II & III give brief reviews of polar codes and Rao-Nam cryptosystem respectively. The concept of using non-systematic finite length polar codes in the structure of secret key cryptosystems is introduced in section IV. Also, the efficiency and security of the proposed cryptosystem are investigated in sections V & VI respectively. Finally, section VII concludes the paper with a brief discussion of future works.

II. POLAR CODES

In this section, we discuss briefly the structure of the polar codes and review an existing technique for constructing the generator matrix of them. Polar codes are a class of linear block codes that provably achieve the capacity of any symmetric Binary-input Discrete Memoryless Channel (B-DMC), such as BEC and Binary Symmetric Channel (BSC). Let $W: \mathcal{X} \to \mathcal{Y}$ be a B-DMC with input alphabet $\mathcal{X} = \{0,1\}$, output alphabet $\mathcal{Y}$ and transition probabilities $\{W(y|x), x \in \mathcal{X}, y \in \mathcal{Y}\}$. Let us consider the following parameters for a B-DMC W [2].

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

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

where $I(W) \in [0,1]$ is the mutual information between the input and the output of W with uniform distribution on the input. When W is a symmetric channel, $I(W)$ is called the capacity of W and applied as the measure of rate. Also, $Z(W) \in [0,1]$ is called the Bhattacharyya parameter of W and used as a criterion of measuring reliability. Note that $I(W) \approx 1$ iff $Z(W) \approx 0$, also $I(W) \approx 0$ iff $Z(W) \approx 1$. If the channel W is a BEC with erasure probability ϵ, denoted by BEC(ϵ), then $Z(W) = \epsilon$ and $I(W) = 1 - Z(W) = 1 - \epsilon$ [2]. In this work, we consider the transmission over BEC(ϵ), for the sake of reasons which will be explicated in sec. III.

In the polar coding schemes, a set of $N = 2^n, n \geq 1$ polarized binary-input channels $\left\{W_N^{(i)}: 1 \leq i \leq N\right\}$ with indices 'i' that can be obtained by performing a phenomenon on the N independent copies of the given B-DMC W. This phenomenon is called channel polarization and the polarized binary input channels are also called bit-channels or sub-channels [2]. By applying the channel polarization, the symmetric capacity terms $\left\{I\left(W_N^{(i)}\right), 1 \leq i \leq N\right\}$ and Bhattacharya parameters $\left\{Z\left(W_N^{(i)}\right), 1 \leq i \leq N\right\}$ of the all N bit-channels tend to 0 or 1 for N large enough [2]. In the sequel of this paper, the Bhattacharya parameter of i-th bit-channel, $Z\left(W_N^{(i)}\right)$, will be denoted by $Z_{N,i}$. Also, we consider the methods which are proposed to obtain the Bhattacharya parameters of the bit-channels. Such parameters are necessary to construct the generator matrix of polar codes.

Let $\mathcal{I} = \{i, 1 \leq i \leq N\}$ be a set of all bit-channel indices. Let A be a K-element subset of $\mathcal{I}$ which is called information set and A^c be an $(N-K)$-element subset of $\mathcal{I}$ which is complement to subset A and called frozen (fixed) set. The sets are specified in some way that $Z_{N,i} \leq Z_{N,j}$ for all $i \in A,\ j \in A^c$. In other words, it is possible to construct N bit-channels such that $NI(W)$ of them with indices in information set tend to become reliable or noiseless and $N(1 - I(W))$ of them with indices in frozen set tend to become unreliable or noisy [2, 28].

A. Constructing the Generator Matrix

Consider $N = 2^n, n \geq 1$ and $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Given the rate $R < I(W)$ and dimensionality $K = 2^n R$, a $K \times N$ generator matrix G_A is constructed for any (N, K) polar code through the following steps [29].

1. Compute the n-th kronecker product $F^{\otimes n}$ which gives an $N \times N$ matrix. Then, label the rows of $F^{\otimes n}$ from top to bottom as $1, 2, \cdots, N$.
2. Obtain the Bhattacharyya parameters of all bit-channels in the form of $Z_N = \left(Z_{N,i}, 1 \leq i \leq N\right)$ through the following recursive relation with initial condition $Z_{1,1}$.

$$Z_{2k,i} = \begin{cases} 2Z_{k,i} - Z_{k,i}^2 & 1 \leq i \leq k \\ Z_{k,i-k}^2 & k+1 \leq i \leq 2k \end{cases},\ k = 1, 2, 2^2, \cdots, 2^{n-1} \quad (1)$$

   In fact, the parameter $Z_{N,i}$ equals the erasure probability of the bit-channel $W_N^{(i)}$. Thus, even if the channel W is a BEC(ϵ), the initial condition $Z_{1,1}$ is equal to ϵ.
3. Form a permutation $\pi_N = (i_1, \ldots, i_N)$ for the set of indices of N bit-channels $\mathcal{I} = \{1, 2, \cdots, N\}$ in such a way that the inequality $Z_{N,i_j} \leq Z_{N,i_k}$ is satisfied for any $1 \leq j < k \leq N$.
4. Obtain the information set $A \subset \mathcal{I}$ which their channel indices are corresponded to K leftmost indices of the permutation π_N, i.e. $i_1, \ldots, i_K$. Then, obtain the frozen set $A^c \subset \mathcal{I}$ which their indices are corresponded to $N - K$ rightmost indices of the permutation π_N, i.e. $i_{K+1}, i_{K+2}, \ldots, i_N$.
5. Construct the generator matrix G_A by choosing $K = 2^n R$ rows of matrix $F^{\otimes n}$ which are corresponded to channel indices in information set A. If the bit-channel $W_N^{(i)}$ is chosen, then the i-th row of $F^{\otimes n}$ is selected. Also, construct $(N-K) \times N$ matrix G_{A^c} by selecting $N - K$ rows of $F^{\otimes n}$ corresponding to indices in frozen set A^c.

In short, the Bhattacharya parameters $\{Z_{N,i}, 1 \leq l \leq N\}$ of all bit-channels are generated by recursive Rel. (1) and then are sorted from the lowest to highest values. Finally, the generator matrix G_A is constructed by choosing the K rows of matrix $F^{\otimes n}$ which their indices in $\mathcal{I}$ are corresponded to K leftmost indices of π_N.

B. Non-Systematic Encoding

Polar codes in standard form are non-systematic codes. In systematic encoding, the information bits always appear in the first K bits of a codeword. But in nonsystematic encoding, the information bits do not appear as part of the codeword transparently [30, 31]. For the nonsystematic polar codes with block length $N = 2^n, n \geq 1$, an input vector $U = (u_1, u_2, \cdots, u_N) = (u_A, u_{A^c})$ consists of two subvectors. A K-bit subvector $u_A = (u_i, i \in A)$ which is called information vector. Thus, an $(N-K)$-bit subvector $u_{A^c} = (u_i, i \in A^c)$ which is called frozen (fixed) vector. The information vector u_A comprises of information data that is free to change in each process of transmission, while the frozen vector consists of the fixed values that is known to decoder [30].

The coordinates of information vector can be transmitted at rate close to 1 through the noiseless bit-channels and the coordinates of frozen (fixed) vector can be transmitted at rate close to 0 across the noisy bit-channels, So, the bit-channels are sufficient for channel coding [2]. In addition, the input vector $U = (u_A, u_{A^c})$ is encoded to binary N-tuple codeword X as follows.

$$X = u_A G_A + u_{A^c} G_{A^c} = u_A G_A + c$$

Since $c \triangleq u_{A^c}G_{A^c}$ is a fixed vector, the encoder which maps u_A to X is non-systematic [30]. The code rate is defined as $R = |u_A|/|X| = |A|/N$.

C. Successive cancellation decoding

Let X be a binary N-tuple codeword of the polar code which is transmitted across the N bit-channels. Let $Y = y_1^N = (y_1, y_2, \cdots, y_N)$ be a corresponding channel output vector which is decoded by the low complexity SC decoding algorithm. The main goal of the SC decoder is to estimate of input vector $\hat{U} = \hat{u}_1^N = (\hat{u}_1, \hat{u}_2, \cdots, \hat{u}_N)$ by knowing the information set A, frozen vector u_{A^c} and the channel output Y. The bits of input vector are estimated successively at the SC decoder by the following way [2].

$$\hat{u}_i = \begin{cases} u_i, & if\ i \in A^c \\ h_i\left(y_1^N, \hat{u}_1^{i-1}\right) & if\ i \in A \end{cases},$$

where decision functions $h_i: \mathcal{Y}^N \times \mathcal{X}^{i-1} \to \mathcal{X}, i \in A$, are computed as follows for all $y_1^N \in \mathcal{Y}^N, \hat{u}_1^{i-1} \in \mathcal{X}^{i-1}$.

$$h_i\left(y_1^N, \hat{u}_1^{i-1}\right) \triangleq \begin{cases} 0, & \text{if } \frac{w_N^{(i)}(y_1^N, \hat{u}_1^{i-1}|0)}{w_N^{(i)}(y_1^N, \hat{u}_1^{i-1}|1)} \geq 1 \\ 1, & \text{otherwise} \end{cases},$$

The information bits $u_i, i \in A$, are estimated one by one using i-th decision element after knowing the channel output vector Y and previous estimated information bits $\hat{u}_1^{i-1}$. Furthermore, the value of frozen bits, $u_i, i \in A^c$, is known for the SC decoder. So the real purpose at the SC decoder is to estimate the information vector u_A [2].

D. Upper Bound on Error Probability

It was shown that the generator matrix G_A of polar codes is obtained by choosing the rows of $F^{\otimes n}$ which are corresponded to channel indices in information set A. Also, the Hamming weight of ith row in $F^{\otimes n}$ matrix is equal to $2^{\mathrm{wt}(i)}$, where $\mathrm{wt}(i)$ is the number of ones in the binary expansion of the 'i'. The minimum Hamming distance cannot be larger than the minimum weight of the rows in the generator matrix G_A. Thus, the minimum Hamming distance of polar codes is given by $d_{min} = \min_{i \in A} 2^{\mathrm{wt}(i)}$ [32]. It can be said that the values of minimum Hamming distance and error correcting capability, $t = \left[\frac{d_{min}-1}{2}\right]$, depend on the values of indices in the information set A. Also, it is shown that for any given B-DMC W, the error probability under SC decoding is upper bounded as follows [2].

$$P_e \leq \sum_{i \in A} Z_{N,i}, \qquad (2)$$

In general, it can be said that the upper bound on error probability ($\sum_{i \in A} Z_{N,i}$) of polar code over BEC depends on the some parameters such as code length N, code rate R, information set A and erasure probability ϵ. Fig. 1 shows the variations of upper bounds on error probability for the polar codes of length $N = 2^{10}$ in terms of rates $R \in [0.5, 0.99]$. This code is transmitted over BECs with $\epsilon = 0.01, 0.05, 0.1$. It is obvious that for constant block length N, the upper bound on error probability is decreased by decreasing the rate R and erasure probability ϵ and vice versa.

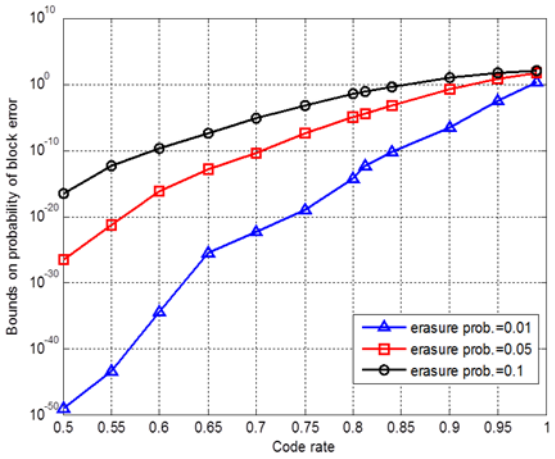

Figure 1. Upper bounds on error probability for the polar code of length $N = 2^{10}$ in terms of rates and erasure probabilities .

Another parameter that effects on $\sum_{i \in A} Z_{N,i}$ is the information set A which will be explained more about it in sec. V-A. Also, it is proved that the reliable communication over BEC using SC decoder is obtained when the following relation is satisfied [33, 34].

$$R < I(W) - N^{-1/\mu}, \qquad (3)$$

where μ is called scaling exponent. The value of scaling exponent for the polar codes is dependent on the choice of the channel. Its value for BEC is approximately equal to 3.6261. Indeed, reliable transmission by polar codes via SC decoding for any B-DMC W is obtained when the rate R is at least less than the capacity $N^{-1/\mu}$. Indeed, it is a tradeoff among rate, block length, and erasure probability for a given error probability when the SC decoder is utilized [34]. In this paper, the maximum possible code rate which fulfills the Rel. 3 is called *cutoff rate* and is denoted by R_0. Table I shows the variations of $\sum_{i \in A} Z_{N,i}$ in terms of rates $R \in [0.65, 0.9]$. The intended polar code has finite length of $N = 2^{10}$ which is used over BEC(0.05). In this case, the cutoff rate R_0 is equal to 0.8 as shown in this table.

TABLE I. Variations of upper bounds on error probability for the polar code of length $N = 2^{10}$ in terms of rates $R \in [0.65, 0.9]$.

R	K	$\sum_{i \in A} Z_{N,i}$
0.9	921	0.1379
0.85	870	0.0014
$R_0 = 0.8$	819	1.0629×10^{-5}
0.75	768	2.8925×10^{-8}
0.7	717	3.3789×10^{-11}
0.65	666	1.3275×10^{-13}

It is clear that the upper bound on error probability decreases significantly for the rates less than or equal to cutoff rate $R_0 = 0.8$.

III. THE RAO-NAM CRYPTOSYSTEM

The RN cryptosystem is an important secret key code based cryptosystem which is used as the reference for measuring the security and efficiency of secret key cryptosystems based on error correcting codes. So in this section, we explain the structure of this scheme and investigate its drawbacks in detail.

A. Secret Key

Secret key of the RN scheme consists of parameters $\{G, \mathcal{S}, \mathcal{P}, \mathcal{T}\}$ which are explicated as follows [7]:

- Let G be a $K \times N$ generator matrix of binary linear code C.
- Let $\mathcal{S}$ be a $K \times K$ random binary nonsingular matrix (scrambler).
- Let $\mathcal{P}$ be an $N \times N$ random binary permutation matrix (permutor).

In the RN cryptosystem, a set of predetermined binary N-tuple error vectors, $\mathcal{E}_{\mathcal{N}_E} = \{E_i, 1 \le i \le \mathcal{N}_E\}$, with cardinality $\mathcal{N}_E = 2^{N-K}$ is considered which has two properties. Before explaining these properties, some related definitions should be given. Let V be a binary vector of length N, the set $V + C = \{V + X | X \in C\}$ is called a coset of code C and the vector V is called a coset leader. Due to each linear code C has 2^K codeword, so each coset contains 2^K vectors and there are exactly 2^{N-K} different cosets. Also, an $(N-K)$-tuple vector $s = (V + C)H^T = VH^T$ is called the syndrome of V with respect to the code C. First property, called weight property, is that all error vectors of $E \in \mathcal{E}_{\mathcal{N}_E}$ have the average Hamming weight of half the code length, $w_H(E) \approx N/2$. The second property, called syndrome property, is that no distinct error vectors locate in the same coset of C [35]. According to these definitions, syndrome error table is defined as follows.

- Let $\mathcal{T} = \{(EH^T, E) | E \in \mathcal{E}_{\mathcal{N}_E}\}$ be a predetermined set of error vectors which is also called as syndrome-error table. This set consists of 2^{N-K} cosets and each coset has a distinct syndrome $s = EH^T$. Therefore, any set of N-bit error vectors can be selected, one from each of 2^{N-K} cosets.

B. Encryption

A binary K-tuple message M is encrypted into a binary N-tuple ciphertext $\mathcal{C} = (c_1, c_2, \cdots, c_N)$ as follows [7].

$$\mathcal{C} = (M\mathcal{S}G + E)\mathcal{P} = MG' + E\mathcal{P}, \quad G' = \mathcal{S}G\mathcal{P}.$$

where G' is a $K \times N$ encryption matrix. E denotes an N-bit error vector which is selected randomly from the syndrome error table $\mathcal{T}$ [7]. Since E is a random error vector, it is conceivable that for the message M, many ciphertexts, $\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_{\mathcal{N}_E}$, can be obtained through encryption [36].

C. Decryption

A ciphertext $\mathcal{C}$ is decrypted into a plaintext M using secret keys $\mathcal{S}^{-1}$, H^T and $\mathcal{P}^T$ as follows.

1. Compute $\mathcal{C}' = \mathcal{C}\mathcal{P}^T = M\mathcal{S}G + E = M'G + E$, $M' = M\mathcal{S}$.
2. Calculate the syndrome $s = \mathcal{C}'H^T = M'GH^T + EH^T = EH^T$, $GH^T = 0$.
3. Find the corresponding error vector E from the syndrome error table $\mathcal{T}$.
4. Obtain $M'G = \mathcal{C}' - E$ and recover M' via decoding algorithm.
5. Multiply M' by $\mathcal{S}^{-1}$ to retrieve the message M.

D. Weaknesses

The RN scheme has several drawbacks for instance,

- It needs to store the matrices $\mathcal{S}, \mathcal{P}$ and G. Also, the syndrome error table $\mathcal{T}$ should be saved to remove the errors in the decryption process. Therefore, a large amount of secret key is exchanged and stored by sender and receiver [7, 20].
- Another practical problem of this scheme is the low number of error vectors, $\mathcal{N}_E = 2^{N-K}$ for short length codes with high information rate, i.e. $\mathcal{N}_E = 2^8$ for $(72, 64)$ Hamming code. Therefore, it is vulnerable against chosen plaintext attacks [7, 13].
- In the RN scheme, there exists a tradeoff between code rate and security. In fact, the code length N will be impractical for high code rate and large number of error vectors [13].
- Another drawback is the possibility of estimating the rows of the encryption matrix G' via majority voting analysis [35, 37].

In this research, it is attempted to solve these problems by applying the interesting properties of non-systematic polar codes and other efficient methods.

IV. THE PROPOSED CRYPTOSYSTEM

In this section, we address the issue of designing secret key cryptosystem based on finite length polar codes to be able to correct channel errors as well as concealing the information from unauthorized user. We consider transmission over a BEC(ϵ), because; it was shown that BEC has the best tradeoff between rate and reliability among all the B-DMCs. It means that for a BEC, the Bhattacharyya parameter $Z(W)$ is minimized among all channels with a given capacity $I(W) = 1 - Z(W)$ [2]. Also for a general B-DMC W, no efficient algorithm is introduced so far to calculate Bhattacharya parameters except for a BEC which these parameters are constructed efficiently using recursive Rel. (1) [2]. So, unlike the other B-DMCs, the method of constructing polar code is simple for the BECs and can be performed with complexity $\mathcal{O}(N)$ [29].

As referred before, in the computational security, it is assumed that an attacker has unlimited access to the transmission media. By assuming that the transmission is over BEC, the attacker can construct easily the generator matrix of intended polar code using N, R and ϵ. Therefore, it can be said that the hardest case to establish a desirable level of computational security by polar codes is when the transmission takes place over BEC. As a sequence, in this scheme, we consider the transmission of the encrypted data over BEC to prove the high security of our scheme.

A. Hiding the generator matrix of polar codes

As aforementioned, the main question of this research is that how the generator matrix of polar codes can be hidden to use these efficient codes in the structure of cryptosystems based on general decoding problem. Here, we propose an efficient

method to solve this question. By the help of this method, an attacker cannot construct the hidden generator matrix of the polar codes over BEC(ϵ) even knowing the parameters N, R and ϵ. First, consider the set of N bit-channel indices $\mathcal{I} = \{1, 2, \cdots, N\}$, the permutation $\pi_N = (i_1, i_2, \cdots, i_{NR_0}, \cdots, i_N)$ and the cutoff rate R_0 for an (N, K) polar code, which all are defined in sec. II.

Definition 1. The NR_0 bit-channels which their Battacharya parameters are minimized (least error probability) among all the N bit-channels are called *Good bit-channels*. In other words, the indices of good bit-channels in the set $\mathcal{I}$ are corresponded to the indices $\{i_1, i_2, \cdots, i_{NR_0}\} \subset \pi_N$.

Definition 2. The $N(1-R_0)$ bit-channels which their Battacharrya parameters are maximized (most error probability) among all the N bit-channels are called *Bad bit-channels*. In other words, the indices of bad bit-channels in the set $\mathcal{I}$ are corresponded to the indices $\{i_{NR_0+1}, i_{NR_0+2}, \cdots, i_N\} \subset \pi_N$.

In the following, we explain how the generator matrix of polar codes can be private.

1. Consider the method of constructing the generator matrix for an (N, K) polar code which was discussed in sec. II-A. First, all Bhattacharya parameters of N bit-channels, $Z_{N,i}, 1 \le i \le N$, and the permutation π_N are constructed. Now, to hide the generator matrix, K indices are selected randomly from the indices of good bit-channels in the set $\mathcal{I}$. Indeed, it is equivalent to random selection of K bit-channels from NR_0 good bit-channels. Then, the random selected indices of set $\mathcal{I}$ are considered as the *secret information set*, denoted by $A(\mathcal{s})$. In fact, the secret information set $A(\mathcal{s})$ is the subset of $\mathcal{I}$ with K random selected indices of good bit-channels.

   The *secret generator matrix*, $G_{A(\mathcal{s})}$, is defined as a $K \times N$ submatrix of $F^{\otimes n}$ which its K rows are chosen corresponding to indices of secret information set $A(\mathcal{s})$. As will be shown in sec. VI-A, if the cutoff rate R_0, length N and dimension K are selected properly, the number of polar codes which are equivalent to used code is large enough and an attacker cannot find the hidden generator matrix $G_{A(\mathcal{s})}$ in the polynomial time. However, as will be shown in sec. V-A, such selection with high probability is not the best to achieve channel capacity. Indeed, it is a tradeoff between the security and efficiency that is almost inevitable in designing of code based cryptosystems.
2. The *secret frozen (fixed) set*, denoted by $A^c(\mathcal{s})$, is a subset of $\mathcal{I}$ which its elements are $N-K$ non-selected indices of set $\mathcal{I}$ in step 1. Also, the secret matrix $G_{A^c(\mathcal{s})}$ is defined as an $(N-K) \times N$ submatrix of $F^{\otimes n}$ which its $N-K$ rows are chosen corresponding to indices of secret frozen set $A^c(\mathcal{s})$.
3. To have more secure decoding, the frozen vector (sec. II-B) should be concealed from an adversary. Since the code's performance is not sensitive to the selection manner of the frozen vector, it is not determined how to choose this vector in the polar codes [2]. Therefore, in encryption/decryption process of the proposed scheme, an $(N-K)$-bit pseudorandom syndrome is generated to use as frozen vector. So, the number of possible frozen vectors is equal to $\mathcal{N}_s = 2^{N-K}$ as will be shown in sec. VI-A. As long as the length and dimensionality of used polar code are selected properly, the attacker cannot find the secret frozen vector, $u_{A^c(\mathcal{s})}$, in polynomial time.

As we recalled in sec. II-C, the inputs to SC decoder of polar codes are channel output vector, information set and frozen vector. So by hiding the information set and frozen vector through above technique, an unauthorized receiver cannot decode the channel output vector Y to estimated input vector $\widehat{U}$ in polynomial time. Figure 2(a, b) shows the proposed concept for providing security based on polar codes.

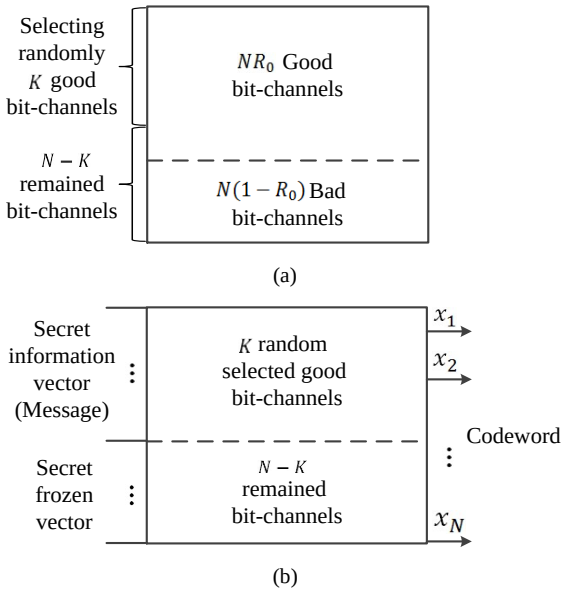

Figure 2. (a). Selecting randomly K bit-channels from NR_0 good bit-channels, (b). Transmitting the secret information vector (message) through K random selected good bit-channels.

In Fig 2(a), it is shown that in order to hide the generator matrix $G_{A(\mathcal{s})}$, the K bit-channels are selected randomly from NR_0 good bit-channels. In Fig 2(b), the secret information vector (message), $u_{A(\mathcal{s})}$, is transmitted through the K random selected good bit-channels. Also, the secret frozen vector, $u_{A^c(\mathcal{s})}$, is transmitted across $N-K$ remained bit-channels. In this case, if the parameters N, K and R_0 are selected properly, an attacker cannot recognize that on what bit-channels the secret information bits are transmitted. So the secret generator matrix $G_{A(\mathcal{s})}$ cannot be constructed by attacker even knowing the parameters of transmission channel, length and dimension of used polar code.

B. Secret Key

The set of matrices and vectors which is applied in encryption/decryption of this scheme and naturally should be held in secret. This set is $\{G_{A(\mathcal{s})}, \mathcal{S}, \mathcal{P}, \mathcal{E}\}$ where $G_{A(\mathcal{s})}$ is a $K \times N$ generator matrix of the polar codes which requires KN bits memory, $\mathcal{E}$ is also a set of N-bit error vectors which requires $N|\mathcal{E}|$ bits memory, $\mathcal{S}$ is a $K \times K$ random binary nonsingular matrix and $\mathcal{P}$ is an $N \times N$ random binary permutation matrix which both require K^2 bits and $lg_2^{N!}$ bits memory respectively. By saving these secret keys directly, the key length of the

proposed scheme will be so large. Therefore, we apply efficient methods by which the size of exchanged keys is dramatically reduced. In this case, the secret key is composed of the set $\{A(\mathcal{s}), \mathcal{IV}_{\mathcal{S}}, \mathcal{IV}_{\mathcal{P}}, \mathcal{IV}_{s}\}$ which the consisting parameters are explained as follows:

- Let $A(\mathcal{s})$ be a secret information set which consists of K random selected indices of NR_0 good bit-channels. As referred before, the secret generator matrix of an (N, K) polar code, $G_{A(\mathcal{s})}$, is defined as the $K \times N$ submatrix of $F^{\otimes n}$ which its K rows are chosen corresponding to indices of secret information set $A(\mathcal{s})$. It is required to store secret information set $A(\mathcal{s})$ instead of saving directly the generator matrix $G_{A(\mathcal{s})}$.
- Let $\mathcal{IV}_{\mathcal{S}}$ be a $(2K-4)$-bit initial value of LFSR, to generate a binary pseudorandom sequence $r_1, r_2, \cdots, r_{2K-4}, 0, 0$. The generated pseudorandom sequence is used to construct the binary nonsingular matrix $\mathcal{S}_{K\times K}$ (see sec. V-B).
- Let $\mathcal{IV}_{\mathcal{P}}$ be an $(N-1)$-bit initial value of LFSR to generate a binary pseudorandom sequence $r_1, r_2, \cdots, r_{N-2}, 0$. The generated pseudorandom sequence can be used to construct the binary permutation matrix $\mathcal{P}_{N\times N}$ (see sec. V-B).
- Let $\mathcal{IV}_{s}$ be an $(N-K)$-bit initial value of LFSR to generate binary pseudorandom syndromes s. In the non-systematic polar codes, the information bits do not appear as part of the codeword transparently [30]. Due to the non-systematic property of used polar code, the generated pseudorandom syndrome is used as secret frozen vector. So, we can consider $E = u_{A^c(\mathcal{s})}G_{A^c(\mathcal{s})} = sG_{A^c(\mathcal{s})}$ as an N-bit error vector and $\mathcal{E}_{\mathcal{N}_E} = \{E_i = s_i G_{A^c(\mathcal{s})}, 1 \le i \le \mathcal{N}_E\}$ with cardinality $\mathcal{N}_E = \mathcal{N}_s = 2^{N-K}$ as a set of N-bit error vectors. So unlike the RN cryptosystem, it doesn't need to store the syndrome error table $\mathcal{T}$.

As will be shown in sec. VI, reduction in the key size of proposed scheme by these efficient methods does not decrease the security of system.

C. Encryption

1. The sender first chooses randomly a code in a family of equivalent (N, K) polar codes over BEC(ϵ) by selecting K indices at random from NR_0 indices of good bit-channels. Then, the sender generates the $(N-K)$-bit pseudorandom syndrome for each message via usage an LFSR with the initial value $\mathcal{IV}_s$. In order to perform the decryption process properly, it is necessary to synchronize the applied LFSR between sender and receiver. By this way, the syndrome which is employed by the sender is known to the receiver synchronously [38]. Then, an N-bit error vector $E = sG_{A^c(\mathcal{s})}$ is constructed.
2. Finally, each binary K-tuple message M is encrypted into a binary N-tuple ciphertext $\mathcal{C}$ as follows.

$$\begin{aligned}\mathcal{C} &= \left(M\mathcal{S}G_{A(\mathcal{s})} + sG_{A^c(\mathcal{s})}\right)\mathcal{P} \\ &= M\mathcal{S}G_{A(\mathcal{s})}\mathcal{P} + sG_{A^c(\mathcal{s})}\mathcal{P} \\ &= MG' + E\mathcal{P},\end{aligned} \tag{4}$$

where $G' = \mathcal{S}G_{A(\mathcal{s})}\mathcal{P}$ is a $K \times N$ encryption matrix which is combinatorially equivalent to generator matrix $G_{A(\mathcal{s})}$.

D. Decryption

The ciphertext $\mathcal{C}$ is transmitted over the insecure channel and the received vector $Y = \mathcal{C} + e_{ch} = MG' + E\mathcal{P} + e_{ch}$ which is corrupted by channel error e_{ch} is decrypted by the authorized receiver as following steps.

1. The transposed of the permutation matrix, $\mathcal{P}^T$, is multiplied to the received vector Y and $Y' = {y'}_1^N = Y\mathcal{P}^T = M\mathcal{S}G_{A(\mathcal{s})} + E + e_{ch}\mathcal{P}^T$ is computed to remove the permutation matrix $\mathcal{P}$. In this case, $e_{ch}\mathcal{P}^T$ is a vector having the same Hamming weight as e_{ch} because $\mathcal{P}$ is a permutation matrix.
2. The authorized receiver uses the secret initial value $\mathcal{IV}_s$ to generate the binary $(N-K)$-tuple pseudorandom syndrome (frozen vector) $s = u_{A^c(\mathcal{s})}$. Then, the set $\{A(\mathcal{s}), u_{A^c(\mathcal{s})}, Y'\}$ is considered as the input of the SC decoder. Finally, the input vector of encoder $U = (u_1, u_2, \cdots, u_N) = \left(u_{A(\mathcal{s})}, u_{A^c(\mathcal{s})}\right) = (M\mathcal{S}, s)$ is estimated by the SC decoder as follows:

$$\hat{u}_i = \begin{cases} u_i, & if\ i \in A^c(\mathcal{s}) \\ h_i\left({y'}_1^N, \hat{u}_1^{i-1}\right) & if\ i \in A(\mathcal{s}) \end{cases}$$

which the decision function $h_i: \mathcal{Y}^N \times \mathcal{X}^{i-1} \to \mathcal{X}, i \in A(\mathcal{s})$, is defined as follows.

$$\hat{u}_i = h_i\left({y'}_1^N, \hat{u}_1^{i-1}\right) \triangleq \begin{cases} 0, & \text{if } \frac{w_N^{(i)}\left({y'}_1^N, \hat{u}_1^{i-1}\middle|0\right)}{w_N^{(i)}\left({y'}_1^N, \hat{u}_1^{i-1}\middle|1\right)} \ge 1 \\ 1, & \text{otherwise} \end{cases},\ i \in A(\mathcal{s})$$

3. After obtaining the K-bit scrambled vector $u_{A(\mathcal{s})} = M\mathcal{S} = (u_i, i \in A(\mathcal{s}))$ by use of SC decoding, the message can be recovered as $M = u_{A(\mathcal{s})}\mathcal{S}^{-1}$.

The used syndrome (frozen vector) $s = u_{A^c(\mathcal{s})}$ and the secret information set $A(\mathcal{s})$ are necessary to initiate the SC decoder and also both of them are secret, so any unauthorized user would find it hard to correct channel errors without these parameters. In fact, the general decoding problem of polar codes without knowledge of secret decoding keys $(A(\mathcal{s}), u_{A^c(\mathcal{s})})$ is an NP-complete.

Fig. 3 shows the block diagram of the proposed cryptosystem. At the first step of Fig. 1, the message is multiplied by the nonsingular matrix $\mathcal{S}$, then the binary K-tuple scrambled vector $u_{A(\mathcal{s})} = M\mathcal{S}$ is encoded to the binary N-tuple codeword $X = u_{A(\mathcal{s})}G_{A(\mathcal{s})} + sG_{A^c(\mathcal{s})}$ using the secret information set $A(\mathcal{s})$ and frozen vector $u_{A^c(\mathcal{s})}$ as other inputs of encoder. At last, the binary N-tuple ciphertext $\mathcal{C}$ is obtained by multiplying the codeword X to the permutation matrix $\mathcal{P}$. By transmitting the ciphertext through insecure channel, the received vector Y is obtained at the receiver. Then, the received vector is multiplied to the transposed of the permutation matrix and the vector $u_{A(\mathcal{s})}$ is obtained by performing the SC decoding on the N-bit vector $Y' = Y\mathcal{P}^T$. Finally, the message M is recovered by multiplying the vector $u_{A(\mathcal{s})}$ to the inverse of the nonsingular matrix $\mathcal{S}$.

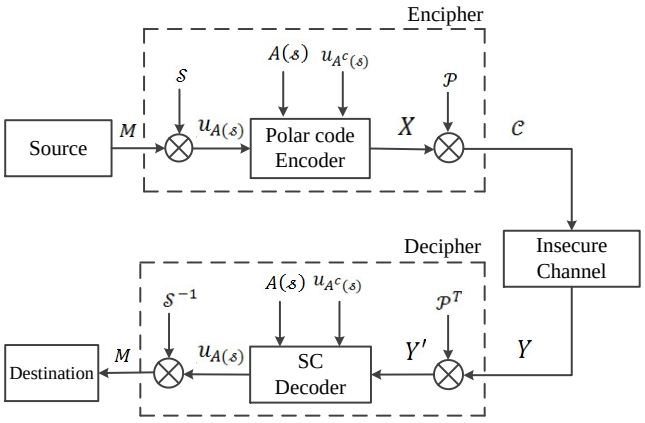

Figure 3. Block diagram of the proposed cryptosystem.

V. EFFICIENCY

To measure the efficiency of the proposed cryptosystem, we evaluate three factors: Error performance, key length and computational complexity.

A. Error Performance

In this research, two definitions for the upper bounds on error probability of used polar code are defined as follows.

Definition 3. Let A_1 be a K-element subset of $\mathcal{I} = \{1, 2, \cdots, N\}$ which its elements are corresponded to indices $\{i_1, i_2, \cdots, i_K\} \subset \pi_N$. In this scheme, the minimum upper bound on error probability under SC decoder, $P_{e_1} = \sum_{i \in A_1} Z_{N,i}$, is defined as the sum of battacharya parameters of K bit-channels which their indices are in the subset A_1. In fact, this upper bound is as the same as the standard upper bound on error probability of the polar codes under SC decoder (Rel. 2).

Definition 4. Let A_2 be a K element subset of $\mathcal{I} = \{1, 2, \cdots, N\}$ which its elements are corresponded to indices $\{i_{NR_0-K+1}, i_{NR_0-K+2}, \cdots, i_{NR_0}\} \subset \pi_N$. In the proposed scheme, the maximum upper bound on error probability under SC decoder, $P_{e_2} = \sum_{i \in A_2} Z_{N,i}$, is defined as the sum of battacharya parameters of K bit-channels which their indices are the elements of subset A_2.

In the proposed scheme, due to random selection of K bit-channels from NR_0 good bit-channels, the upper bound on error probability of used polar code can be varied from P_{e_1} to P_{e_2} depending on the sum of battacharya parameters of selected good bit-channels. In the sequel of this subsection, it will be shown that some parameters such as code length N, code rate R, erasure probability ϵ and selection manner of secret information set $A(\mathcal{S})$ can effect on the values of P_{e_1} and P_{e_2}.

As referred before, the values of P_{e_1} and P_{e_2} depend on the sum of battacharya parameters of selected good bit-channels. On the other hand, if the transmission channel is BEC, the initial value of Rel. 1 to construct the Bhattacharya parameters is $Z_{1,1} = \epsilon$. So the erasure probability ϵ is effected on error probability and should be selected in some way to be able to achieve reliable communication. In this work, we consider the condition $P_{e_2} \leq 10^{-\alpha}$ to have reliable communication which α can have different values depending on the applications of the proposed scheme. Here, we select $\alpha = 4$ and other analyzes on the error performance will be based on this selection. The erasure probabilities of BEC should be selected in which P_{e_2} is less than or equal to 10^{-4}. It is obvious that in this case, P_{e_1} is definitely less than 10^{-4}. Table II shows the simple bounds which are derived on ϵ to satisfy the condition $P_{e_2} \leq 10^{-4}$ for several finite lengths of polar codes.

TABLE II. Bounds on the erasure probabilities of BEC to satisfy $P_{e_2} \leq 10^{-4}$ for several code lengths.

N	ϵ
2^{10}	$[0, 0.07]$
2^{11}	$[0, 0.08]$
2^{15}	$[0, 0.12]$
2^{20}	$[0, 0.17]$

It is obvious that for larger code lengths, we can provide larger bounds on ϵ to achieve reliable communication. Also, the rate of used polar code should be chosen in such a way that $R < R_0$. As our goal is to employ finite length polar codes with high rate to achieve secure and reliable communication, we consider a $(1024, 832)$ polar code with $\mathrm{R} = 0.8125$ over BEC(0.01). Note that by considering the BEC with larger or smaller ϵ, It is possible to select other code rate depending on the required application. For example, for the fixed length $N = 2^{10}$, we can have $R > 0.83$ by considering $\epsilon < 0.01$. Fig. 4 shows the variations of the P_{e_1} and P_{e_2} for the polar code of length $N = 2^{10}$ over BECs with $\epsilon = 0.01, 0.05, 0.1$ in terms of rates $R \in [0.55, R_0]$.

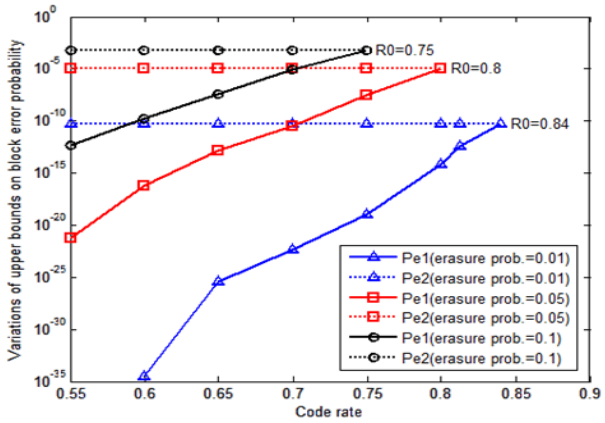

Figure 4. Variations of P_{e_1} and P_{e_2} for the polar code of length $N = 2^{10}$ in terms of rates $R \in [0.55, R_0]$ and erasure probabilities $\epsilon = 0.01, 0.05, 0.1$.

As can be seen in this figure, P_{e_1} is sensible to the variations of the rate and erasure probability. It is decreased by decreasing the rate R and erasure probability ϵ at constant block length N and vice versa. Furthermore, the cut off rate R_0 is increased by decreasing the erasure probability ϵ and vice versa. It is denoted that for the erasure probabilities $0.1, 0.05$ and 0.01, the cutoff rate R_0 is equal to $0.75, 0.8$ and 0.84 respectively. By increasing the cutoff rate R_0, it is possible to achieve reliable communication at higher code rate. On the other hand, P_{e_2} is invariable in terms of the rate but is decreased by decreasing the erasure probability.

The main reason of sensibility of P_{e_1} and insensibility of P_{e_2} to variations of rate is that the values of Battacharya parameters of the good bit-channels corresponded to set A_1 have small amount (approximately close to zero) compared to the values of Battacharya parameters of good bit-channels corresponded to set A_2. So unlike the $P_{e_1} = \sum_{i\in A_1} Z_{N,i}$, the value of $P_{e_2} = \sum_{i\in A_2} Z_{N,i}$ is invariable to variations of rate and is equal to $\sum_{i=1}^{NR_0} Z_{N,i}$. Table III shows the values of P_{e_1} and P_{e_2} for the polar code of length $N = 2^{10}$ over BEC(0.01) in terms of rates $R \in [0.65, R_0]$.

TABLE III. Values of P_{e_1} and P_{e_2} for the polar code of length $N = 2^{10}$ over BEC(0.01) in terms of rates $R \in [0.65, R_0]$.

R	K	P_{e_1}	P_{e_2}
$R_0 = 0.84$	860	5.554×10^{-11}	5.554×10^{-11}
0.8125	832	3.678×10^{-13}	5.554×10^{-11}
0.8	819	6.013×10^{-15}	5.554×10^{-11}
0.75	768	1.033×10^{-19}	5.554×10^{-11}
0.7	717	4.834×10^{-23}	5.554×10^{-11}
0.65	666	3.472×10^{-26}	5.554×10^{-11}

As it is denoted in this table, P_{e_1} is variable in terms rates but P_{e_2} is not variable and is equal to $\sum_{i=1}^{NR_0} Z_{N,i}$. For the used $(1024, 832)$ polar code, the upper bound on the error probability can be varied from $P_{e_1} = \sum_{i=1}^{832} Z_{1024,i} \approx 3.678 \times 10^{-13}$ to $P_{e_2} = \sum_{i=29}^{860} Z_{1024,i} \approx 5.554 \times 10^{-11}$.

The code length N is another parameter which effects on P_{e_1} and P_{e_2} in the proposed cryptosystem. Fig. 5 shows the variations of P_{e_1} and P_{e_2} for the polar codes of lengths $N = 2^{10}, 2^{15}$ in terms of rates $R \in [0.5, R_0]$.

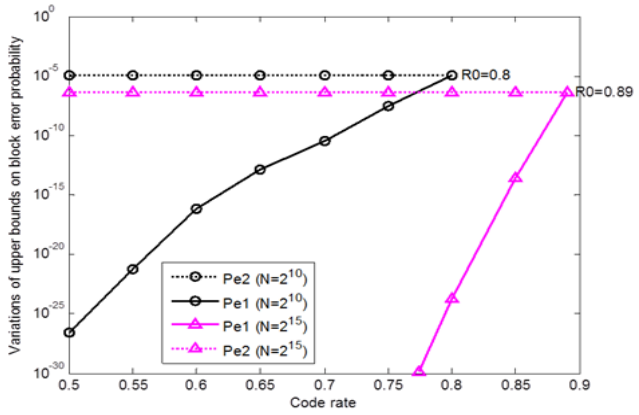

Figure 5. Variations of P_{e_1} and P_{e_2} for the polar codes of lengths $N = 2^{10}, 2^{15}$ in terms of rates $R \in [0.5, R_0]$.

It is clear that both P_{e_1} and P_{e_2} are decreased by increasing the code length and vice versa. Also, the cutoff rate R_0 is increased by increasing the code length and vice versa. It is shown that for the lengths of $N = 2^{10}$ and $N = 2^{15}$, the cutoff rate R_0 is equal to 0.8 and 0.89 respectively. In other words, it is possible to achieve reliable communication at higher code rate by increasing the code length.

As discussed in sec. II-D, the minimum Hamming distance, d_{min}, and error correcting capability, t, of the polar codes depend on the values of K selected bit-channel indices in information set A. In this scheme, since the K bit-channel indices in secret information set $A(\mathcal{s})$ are selected randomly, the minimum Hamming distance and so the error correcting capability are variable.

Therefore, it can be recalled some notes that should be considered in the proposed scheme to achieve reliable and secure communication simultaneously. These notes are summarized as follows.

- The set of K bit-channels should be chosen randomly from good bit-channels to transmit the scrambled message, $u_{A(\mathcal{s})}$, through them and also to hide the generator matrix $G_{A(\mathcal{s})}$. So depending on the selection manner of secret information set $A(\mathcal{s})$, the upper bound on error probability of the used polar code can be varied from $P_{e_1} = \sum_{i\in A_1} Z_{N,i}$ to $P_{e_2} = \sum_{i\in A_2} Z_{N,i}$.
- The length and dimension of the used polar code over BEC(ϵ) should be chosen in such a way that satisfy the Rel. 3.
- The erasure probability of BEC on which the code will be used should be satisfied the condition $P_{e_2} \leq 10^{-\alpha}$ for intended value of α.

B. Key Length

In the RN scheme, the $(72, 64)$ Hamming code was suggested to apply in its structure [39]. So a 64×64 nonsingular matrix $\mathcal{S}$, a 64×72 generator matrix G, and a 72×72 permutation matrix $\mathcal{P}$ are part of secret key. If these matrices are directly saved as keys, over 18×10^3 bits or 17.6 kbits (k = 1024bits) are required for each pair of users. In this subsection, it will be shown that although the dimensionalities of the matrices which are suggested to the proposed cryptosystem are larger than the RN cryptosystem, our scheme's key length is noticeably shorter. The memory requirement of each part of secret key is computed as follows:

- In this scheme, it does not need to save the generator matrix of polar codes directly. In fact, the set of K secret information set $A(\mathcal{s})$ is saved instead of $K \times N$ generator matrix $G_{A(\mathcal{s})}$. On the other hand, the largest possible index of bit-channels, N, may be, is one of the K selected indices in $A(\mathcal{s})$. In the proposed scheme, such bit-channel index, $N = 1024$, requires 11 bits to save in binary form. Due to random selection of channel indices, the upper bound on the required memory for storing $A(\mathcal{s})$ is computed as $\mathcal{M}_{A(\mathcal{s})} \leq 11K = 9152$ bits.
- The required memory for storing the initial value of LFSR to generate the pseudorandom syndromes (frozen vectors) u_{A^c} is $\mathcal{M}_{u_{A^c}} = N - K = 192$ bits.

Reducing the memory requirements of $\mathcal{S}$ and $\mathcal{P}$ matrices

As it was discussed earlier, the nonsingular matrix (scrambler) and the permutation matrix (permutor) are usually employed as part of secret key in the secret key code based cryptosystems. These matrices need large storage space, so this problem is one of challenges in designing of such schemes. Due to using a $(1024, 832)$ polar code in this scheme, if the nonsingular matrix $\mathcal{S}_{832\times832}$ and the permutation matrix $\mathcal{P}_{1024\times1024}$ are saved directly over 684 kbits is required which is too large to be a key. To more practical, it is attempted to solve this problem by applying efficient methods which are

particularly useful in solving the long key problem of algebraic code cryptosystems [39]. The applied methods are based on pseudorandom number generators, i.e. LFSR, with initial values (seeds) to generate the matrices $\mathcal{S}$ and $\mathcal{P}$. By this method, the short initial values $\mathcal{IV}_{\mathcal{S}}$ and $\mathcal{IV}_{\mathcal{P}}$ are saved instead of the matrices $\mathcal{S}$ and $\mathcal{P}$ respectively. This method is explained further below.

To reduce the required storage space of nonsingular and permutation matrices, two interesting algorithms were introduced in [39] to generate these matrices from the short initial values of LFSRs. These algorithms are based on a special type of matrices, called double-one (DBO) matrices in which each row and each column contains exactly two 1's [40]. Here, we explain several properties of these matrices. All the DBO matrices are singular matrices, also the rank of any $K \times K$ DBO matrix is $K-1$. The DBO matrix is called a DBO-1 matrix if all 1's in the matrix can be connected in a unique cycle alternately in the column and the row direction [40]. Fig. 6 illustrates a 5×5 DBO-1 matrix. It is clear that all the 1's in this matrix can be connected by a unique cycle. The cycle of the l's in this matrix is represented by the dashed line.

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Figure 6. A 5×5 DBO-1 matrix [40].

By adding one '1' to any entry of a $K \times K$ DBO-1 matrix, the nonsingular matrix of rank K will be resulted. Based on this interesting property, the first algorithm [39] was introduced to construct a nonsingular matrix $\mathcal{S}_{K\times K}$ from a relatively short seeds. An input of this algorithm is the LFSR's initial value, denoted by $\mathcal{IV}_{\mathcal{S}}$, of an LFSR (with maximum length $|\mathcal{IV}_{\mathcal{S}}| = 2K-4$). It is used to generate a pseudorandom sequence $r_1, r_2, \cdots, r_{2K-2}$ with 0's in the last two bits or equivalently $r_1, r_2, \cdots, r_{2K-4}, 0, 0$. These random bits are then be used to specify the location of l's in the $K \times K$ DBO-1 matrix. At the end of algorithm, one '1' is added to any entry of constructed $K \times K$ DBO-1 matrix. According to the property of DBO-1 matrices, the output matrix will be indeed a nonsingular matrix $\mathcal{S}_{K\times K}$.

In fact, this algorithm has a one-to-one mapping from the initial value $\mathcal{IV}_{\mathcal{S}}$ to the nonsingular matrix $\mathcal{S}_{K\times K}$. The time complexity of this algorithm is $\mathcal{O}(K)$ K-bit word operations. For the sake of briefness, we do not fully describe the functionality of this algorithm and refer to [39] for a detailed description.

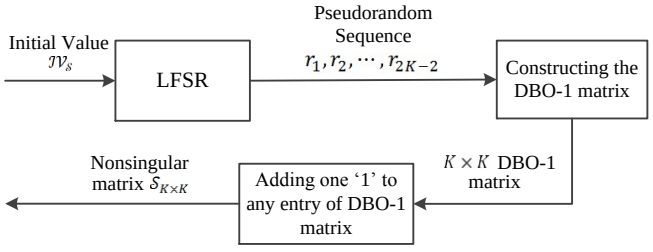

Figure 7. Block diagram of constructing the nonsingular matrix $\mathcal{S}_{K\times K}$ by an initial value $\mathcal{IV}_{\mathcal{S}}$.

Fig. 7 shows the block diagram of constructing the nonsingular matrix $\mathcal{S}_{K\times K}$ by an initial value $\mathcal{IV}_{\mathcal{S}}$. In the second algorithm presented in [38], It is shown that a binary permutation matrix $\mathcal{P}_{N\times N}$ can be generated from an $N \times N$ DBO-1 matrix by inverting the even positions of l's in its cycle, counting from any position. An input of this algorithm is the initial value, $\mathcal{IV}_{\mathcal{P}}$, of LFSR (with maximum length $|\mathcal{IV}_{\mathcal{P}}| = N-2$) to generate a pseudorandom sequence $r_1, r_2, \cdots, r_{N-1}$ with one '0' in the last bit or equivalently $r_1, r_2, \cdots, r_{N-2}, 0$. These random bits are used to specify the location of l's in the permutation matrix $\mathcal{P}_{N\times N}$. In fact, it exists a one to one mapping from the initial value $\mathcal{IV}_{\mathcal{P}}$ to the nonsingular matrix $\mathcal{P}_{N\times N}$. The time complexity of this algorithm is $\mathcal{O}(N)$ N-bit word operations. Fig. 8 shows the block diagram of constructing the permutation matrix $\mathcal{P}_{N\times N}$ by an initial vector $\mathcal{IV}_{\mathcal{P}}$.

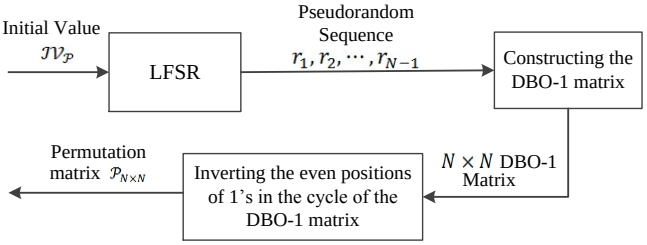

Figure 8. Block diagram of constructing the permutation matrix $\mathcal{P}_{N\times N}$ by an initial vector $\mathcal{IV}_{\mathcal{P}}$.

- By applying this method in the proposed scheme, the memory requirement for storing the nonsingular matrix $\mathcal{S}_{K\times K}$ and the permutation matrix $\mathcal{P}_{N\times N}$ is reduced respectively from K^2 bits to at most $|\mathcal{IV}_{\mathcal{S}}| = 2K-4 = 1660$ bits and from $log_2^{N!}$ bits to at most $|\mathcal{IV}_{\mathcal{P}}| = N-2 = 1022$ bits.

Therefore, the set of secret key for the proposed cryptosystem is $\mathcal{K} = \{A(s), \mathcal{IV}_s, \mathcal{IV}_{\mathcal{S}}, \mathcal{IV}_{\mathcal{P}}\}$ and the upper bound on the length of these used keys is calculated as follows.

$$\mathcal{M}_{\mathcal{K}} = \mathcal{M}_{A(s)} + \mathcal{M}_{\mathcal{IV}_s} + \mathcal{M}_{\mathcal{IV}_{\mathcal{S}}} + \mathcal{M}_{\mathcal{IV}_{\mathcal{P}}}$$
$$\leq 11K + 2N - 6 \approx 10.93\text{kbit}.$$

Table IV compares the key lengths of the proposed cryptosystem and the Rao-Nam cryptosystem. It is clear that although the length and dimesionality of used polar code is much larger than the length and dimensionality of used Hamming code in the RN cryptosystem, the key length of our proposed scheme is shorter than the key length of the RN cryptosystem.

TABLE IV. Comparing the key lengths of Rao-Nam and proposed schemes.

Cryptosystems / ***Characteristics***	***The Rao-Nam cryptosystem* [7, 12]**	***The Proposed cryptosystem***
Code	Hamming	Polar
(N, K)	$(72,64)$ [36, 39]	$(1024, 832)$
Rate	≈ 0.89	0.8125
Key length	$\mathcal{M}_{\mathcal{K}} \approx 17.6$ kbits [39]	$\mathcal{M}_{\mathcal{K}} \leq 10.93$ kbits

Also, due to table II, it can be increased the rate of used polar code by transmitting the message through BEC with smaller erasure probability.

C. Computational Complexity

The computational complexity which can be divided into two parts: Encryption complexity (C_{Enc}) and Decryption complexity (C_{Dec}). According to the Rel. 4 and Fig. 2, the encryption complexity of our scheme can be expressed as below,

$$\mathsf{C}_{Enc} = \mathsf{C}_{mul}(M\mathcal{S}) + \mathsf{C}_{enc}\left(u_{A(s)}\right) + \mathsf{C}_{mul}(X\mathcal{P}),$$

where $\mathsf{C}_{mul}(M\mathcal{S}) \leq K^2$ is the number of binary operations for multiplying binary K-tuple message M by the nonsingular matrix $\mathcal{S}_{K\times K}$. Also, $\mathsf{C}_{enc}\left(u_{A(s)}\right) = \mathcal{O}(NlogN)$ is the complexity of encoding the K-bit scrambled vector $u_{A(s)} = M\mathcal{S}$ to the codeword X by non-systematic encoding algorithm of polar codes. Finally, $\mathsf{C}_{mul}(X\mathcal{P})$ is number of required binary operations for multiplying codeword X to permutation matrix $\mathcal{P}$. The decryption complexity of this scheme can be expressed as follows,

$$\mathsf{C}_{Dec} = \mathsf{C}_{mul}(Y\mathcal{P}^T) + \mathsf{C}_{SC}(Y') + \mathsf{C}_{mul}\left(u_{A(s)}\mathcal{S}^{-1}\right),$$

where the number of binary operations for multiplying the received vector Y to the transposed of the permutation matrix $\mathcal{P}$ is computed as $\mathsf{C}_{mul}(Y\mathcal{P}^T) = N$. Also, the complexity of successive cancelation decoding of the N-bit vector Y' is $\mathsf{C}_{SC}(Y') = \mathcal{O}(NlogN)$ [2]. Furthermore, the number of required binary operations for multiplying the binary K-tuple vector $u_{A(s)}$ to the inverse matrix $\mathcal{S}^{-1}$ is obtained as $\mathsf{C}_{mul}\left(u_{A(s)}\mathcal{S}^{-1}\right) \leq K^2$.

As it was shown in Fig. 3, the encryption and encoding are combined into one process in this scheme. This subject is true for decryption and decoding which can be resulted to decrease the computational complexity, increase the speed and energy efficiency of system.

VI. SECURITY

Several cryptanalytic attacks such as, Brute Force, Majority Voting, Rao-Nam and Struik-Tilburg have been suggested so far to secret key code based cryptosystems. In this section, the cryptanalytic strength of new scheme is examined under these attacks. Our investigations indicate that the proposed scheme has proper security against them.

A. The Brute Force Attack

In the Brute Force attack, all possible keys are checked systematically until the correct key is found. In fact, this attack is impossible, only if the size key space is large enough. In the proposed cryptosystem, the number of parameters of the key set $\mathcal{K} = \{A(s), \mathcal{IV}_{\mathcal{S}}, \mathcal{IV}_{\mathcal{P}}, \mathcal{IV}_{s}\}$ is computed as follows:

- Since the sender randomly selects K bit-channels among all NR_0 good bit-channels, so the total number of possible equivalent polar codes is computed as follows.

$$\mathcal{N}_C(N,K) = \binom{NR_0}{K} = \frac{(NR_0)!}{K!(NR_0-K)!},$$

According to above relation, in this scheme, the cardinality of (N,K) equivalent polar codes increases by increasing the block length N, cutoff rate R_0 and by decreasing the dimension K. The total number of $(1024,832)$ equivalent polar codes over BEC(0.01) with $R_0 = 0.84$ is equal to $\mathcal{N}_C(1024,832) \approx 2^{174}$. Therefore, there is large enough equivalent polar codes to resist against the brute force attack.

- The number of binary nonsingular scrambling matrices $\mathcal{S}_{K\times K}$ is equal to the total number of pseudorandom sequences $r_1, r_2, \cdots, r_{2K-4}, 0, 0$ which are used to specify the locations of 1's in $K \times K$ DBO-1 matrices (sec. V-B). So the number of these binary matrices is equal to $\mathcal{N}_{\mathcal{S}} = \mathcal{N}_{\mathcal{IV}_{\mathcal{S}}} - 1 = 2^{2K-4} - 1$. Hence, for the used $(1024,832)$ polar code, there exists an impractical preliminary work for an adversary to find the nonsingular matrices.
- The number of the binary permutation matrices $\mathcal{P}_{N\times N}$ is equal to the total number of pseudorandom sequences $r_1, r_2, \cdots, r_{N-2}, 0$ which are used to specify the locations of 1's in $N \times N$ DBO-1 matrices (sec. IV-B). So the number of these binary matrices is equal to $\mathcal{N}_{\mathcal{P}} = \mathcal{N}_{\mathcal{IV}_{\mathcal{P}}} = 2^{N-2} - 1$. Therefore, finding the used permutation matrix in this scheme is infeasible in the polynomial time.
- The number of possible binary N-tuple error vectors $E = sG_{A^c(s)}$ is equal to the number of binary $(N-K)$-tuple pseudorandom syndromes, so $\mathcal{N}_E = \mathcal{N}_S = 2^{N-K} = 2^{192}$. Hence, finding the error vectors by an exhaustive search is infeasible.

Hence, the exhaustive search for finding the parameters of key set is considered hopeless due to the large number of involved parameters.

B. Rao-Nam Chosen Plaintext Attack (RN Attack)

The encryption algorithm of the proposed cryptosystem (Rel. 4) can be rewritten as follows:

$$\begin{aligned} \mathcal{C} &= M\mathcal{S}G_{A(s)}\mathcal{P} + sG_{A^c(s)}\mathcal{P} \\ &= MG' + E\mathcal{P} \\ &= MG' + E' \end{aligned} \quad (5)$$

where $G' = [g'_{ij}], i = 1, \cdots, K, j = 1,2, \cdots, N$ is an encryption matrix and $E' = (e'_1, e'_2, \cdots, e'_N)$ is the permuted error vector. In conventional secret key cryptosystem, the ciphertext $\mathcal{C}$ is constant for any given message M with a key K. In secret key code based cryptosystems, since error vector E is selected at random, an attacker can obtain all ciphertexts $\mathcal{C}_i, 1 \leq i \leq \mathcal{N}_E$ for any message M in exhaustive manner. In this way, an attacker can obtain some information about error vectors $E_i, 1 \leq i \leq \mathcal{N}_E$. This information can be used to break the cryptosystem [36]. This type of attacks was suggested by Rao and Nam [7] and rather by Struik and Tilburg [13]. The Rao-Nam (RN) attack [7] is a chosen plaintext attack which is performed as the following steps:

1. Obtaining the encryption matrix G' from a large set of plaintext-ciphertext $(M, \mathcal{C})$ pairs.
2. Recovering message M from $\mathcal{C}$ using G' obtained in the step 1.

Step 1: Let M_1 and M_2 be two binary K-tuple plaintexts differing only in the i-th, $i = 1, 2, \cdots, K$ position. Let $\mathcal{C}_1 = M_1G' + E_1\mathcal{P}$ and $\mathcal{C}_2 = M_2G' + E_2\mathcal{P}$ be two distinct binary N-tuple ciphertext vectors which are obtained from the plaintexts M_1 and M_2 respectively. The difference vector of $\mathcal{C}_1$ and $\mathcal{C}_2$ is obtained by $\mathcal{C}_1 - \mathcal{C}_2 = (M_1 - M_2)G' + (E_1 - E_2)\mathcal{P} = g_i' + (E_1 - E_2)\mathcal{P}$. So the i-th row of encryption matrix, g_i', can be computed as follows.

$$g_i' = \mathcal{C}_1 - \mathcal{C}_2 - (E_1 - E_2)\mathcal{P} \quad (6)$$

It is clear that the Hamming weight of $(E_1 - E_2)P$ is at most $2w_H(E')$, where $w_H(E')$ is the Hamming weight of permuted error vector E'. Since matrix $\mathcal{P}$ is a permutation matrix, $w_H(E') = w_H(E)$. If $2w_H(E)/N \ll 1$, the difference vector $\mathcal{C}_1 - \mathcal{C}_2$ represents one estimate of g_i'. This procedure should be repeated for all $i = 1, 2, \cdots, K$ to obtain the encryption matrix G' but this question is happened that how many operations are required to obtain the complete encryption matrix G'.

Let $\mathcal{C}_j = MG' + E_jP$ and $\mathcal{C}_k = MG' + E_kP$ be two distinct N-bit ciphertexts of the proposed scheme which are obtained from the same M. The difference between $\mathcal{C}_j$ and $\mathcal{C}_k$ is calculated as $\mathcal{C}_j - \mathcal{C}_k = (E_j - E_k)P$, so one value is obtained for $(E_j - E_k)P$. This process should be repeated until all obtained possible values of $(E_j - E_k)\mathcal{P}$ are tested in $(E_1 - E_2)\mathcal{P}$ of Rel. 6. Note that the complete solution of encryption matrix G' must be obtained and verified, because the correctness of each vector g_i' cannot be verified independently. Since, the number of distinct error vectors of this scheme was given by $\mathcal{N}_E = 2^{N-K}$, so the number of all possible pairs of (E_j, E_k) is $\binom{\mathcal{N}_E}{2} = (\mathcal{N}_E^2 - \mathcal{N}_E)/2$. As the vector g_i' should be obtained for K rows of G', so determining the encryption matrix G' from this type of chosen plaintext attack has work factor $WF \geq \frac{1}{2}\left(\frac{\mathcal{N}_E^2}{2}\right)^K$. By substituting $\mathcal{N}_E = 2^{N-K}$, the work factor is equal to $WF = \Omega\left(2^{(N-K)K}\right)$ [7]. Obviously, this attack is infeasible for the proposed cryptosystem because the number of error vectors is so large $\mathcal{N}_E = 2^{192}$.

On the other hand, Rao and Nam claimed that this attack also can be averted by applying the set of error vectors with Hamming weight $w_H(E) \approx N/2$, but later Meijers and Tilburg [35] showed that the RN cryptosystem is vulnerable to the Extended Majority Voting (EMV) attack because of having constraint on the Hamming weight of the error vectors. In fact, the predefined set of error vectors has to be chosen at random to avert the EMV attack. In the proposed cryptosystem, there is no constraint on the Hamming weight of error vectors, so the EMV attack is infeasible.

Step2: By assuming that an attacker can find the encryption matrix G', two ways are possible to determine M from $\mathcal{C}$. The first way is finding $\mathcal{S}$, $\mathcal{P}$ and $G_{A(s)}$ from G' and so to be able to recover M. Therefore, the security depends on the numbers of equivalent polar codes, the number of nonsingular and permutation matrices for a given parameters N, K and R_0. In this work, the infeasibility of finding the generator matrix via exhaustive search is clear since the number of codes which are equivalent to $(1024, 832)$ polar code. Also, the number of matrices $\mathcal{S}$ and $\mathcal{P}$ is large enough, so it is computationally infeasible to find the exact keys $G_{A(s)}$, $\mathcal{S}$ and $\mathcal{P}$ used for encryption matrix G' in an exhaustive manner.

The second way is retrieving the plaintext $M = (m_1, m_2, \cdots, m_K)$ from the ciphertext $\mathcal{C} = (c_1, c_2, \cdots, c_K, \cdots, c_N)$ without the keys. In this way, an adversary considers a set of N linear equations from Rel. 5 as follows.

$$\begin{aligned} c_1 &= m_1g_{11}' + m_2g_{21}' + \cdots + m_Kg_{K1}' + e_1' \\ c_2 &= m_1g_{12}' + m_2g_{22}' + \cdots + m_Kg_{K2}' + e_2' \\ &\vdots \\ c_N &= m_1g_{1N}' + m_2g_{2N}' + \cdots + m_Kg_{KN}' + e_N' \end{aligned}$$

where $c_i, 1 \leq i \leq N$ is the i-th coordinate of the ciphertext $\mathcal{C}$. The attacker tries to solve for a set of K unknowns $(m_1, m_2, \cdots, m_K)$ from the above set of N equations. In this case, K linearly independent equations are chosen at random. If the e_j's of K selected equations are zero, the attacker will be able to solve for M. For a fixed ciphertext $\mathcal{C}$, this procedure should be repeated until a meaningful plaintext block M is obtained. Thus, average work factor of this attack is $K^3P_K^{-1}$, where K^3 is the required operations for solving the K unknowns and $P_K = \prod_{i=0}^{K-1}\left(1 - \frac{w_H(E)}{N-i}\right)$ is the probability of obtaining a set of K linearly independent equations in which no error is occurred [6, 7]. Also, P_K^{-1} is the average number of repetitions for one message block.

It can be seen that this attack depends on the block length N, dimensionality K of used code and the Hamming weight of error vector E. In the proposed cryptosystem, due to random selection of K bit-channels from NR_0 good bit-channels, the Hamming weight of error vector E is variable. However, by using $(1024, 893)$ polar code, determining the message M from $\mathcal{C}$ without the keys is infeasible.

C. Struik-Tilburg Chosen Plaintext Attack (ST Attack)

Let $\mathcal{E}_{\mathcal{N}_E} = \{E_j, 1 \leq j \leq \mathcal{N}_E\}$ and $\mathcal{E}_{\mathcal{N}_E}^{\mathcal{P}} = \{E_j\mathcal{P}, 1 \leq j \leq \mathcal{N}_E\}$ denote a set of distinct error vectors and a set of permuted different error vectors respectively. Also, $\mathcal{E}_{\mathcal{N}_E,\Delta} = \{E_{i,j} = E_i - E_j, 1 \leq i, j \leq \mathcal{N}_E\}$ is a set of error vector differences. In the same way, $\mathcal{E}_{\mathcal{N}_E,\Delta}^{\mathcal{P}} = \{E_{i,j}\mathcal{P}, 1 \leq i, j \leq \mathcal{N}_E\}$ is a set of permuted error vector differences. There are $\mathcal{N}_E$ distinct permuted error vectors so there is a set of $\mathcal{N}_E$ distinct ciphertexts as $\mathbb{C}_{\mathcal{N}_E} = \{\mathcal{C}_j = MG' + E_j\mathcal{P}, 1 \leq j \leq \mathcal{N}_E\}$. The ST attack is described through following steps:

1. First, an arbitrary message M is enciphered until a set of $\mathcal{N}_E$ different ciphertexts $\mathbb{C}_{\mathcal{N}_E} = \{\mathcal{C}_j, 1 \leq j \leq \mathcal{N}_E\}$ is obtained.
2. A directed labeled graph $\Gamma = \left(\mathbb{C}_{\mathcal{N}_E}, \mathcal{E}_{\mathcal{N}_E,\Delta}^{\mathcal{P}}\right)$ is constructed which its vertices consist of $\mathcal{N}_E$ obtained different ciphertexts and each edge from vertex $\mathcal{C}_i$ to vertex $\mathcal{C}_j$ has a label as permuted error vector difference $\mathcal{C}_i - \mathcal{C}_j = E_{i,j}\mathcal{P}$. Afterwards, an automorphism group $\mathrm{Aut}(\Gamma)$ is constructed, i.e. all the permutations on $\mathbb{C}_{\mathcal{N}_E}$ where all the edges $e_{i,j}\mathcal{P}$ leave invariant through these permutations. Hence, the cardinality of the automorphism group is $|\mathrm{Aut}(\Gamma)| = \mathcal{N}_E$.
3. For $1 \leq i \leq K$, a message $M_i = M + u_i$ is selected where u_i is a unit vector with one '1' in its i-th position. Next, the steps 1, 2 is repeated for $M = M_i$ to construct a set of its corresponding ciphertexts $\mathbb{C}_{\mathcal{N}_E}^{(i)} = \left\{\mathcal{C}_j^{(i)} = M_iG' +\right.$

$\hat{E}_j^{(i)}\mathcal{P}, 1 \le i \le K, 1 \le j \le \mathcal{N}_E\}$ and its directed label graph $\Gamma_i = \left(\mathrm{C}_{\mathcal{N}_E}^{(i)}, \mathcal{E}_{\mathcal{N}_E,\Delta}^{\mathcal{P}}\right)$.

4. For $1 \le i \le K$, an automorphism Φ is selected randomly from the automorphism group Aut(Γ). Then, Γ_i is mapped on Γ according to selected automorphism Φ. Now, $\hat{g}_i + \tilde{E}^{(i,1)}\mathcal{P} = \mathcal{C}_1^{(i)} - \mathcal{C}_1 = M_i G' + \hat{E}_1^{(i)}\mathcal{P} - MG' + E_1\mathcal{P}$ is calculated. As there exists an automorphism Φ for which $\tilde{E}^{(i,1)} = 0$, the i-th row of encryption matrix, $\hat{g}_i$, is estimated with probability $|Aut(\Gamma)|^{-1} = \mathcal{N}_E^{-1}$.
5. Finally, by using the estimated $\hat{g}_i, 1 \le i \le K$, the encryption matrix $\widehat{G'} = (\hat{g}_1^T, \hat{g}_2^T, \cdots, \hat{g}_K^T)$ is obtained. If the solution is not correct, the steps 4, 5 must be repeated.

As aforementioned, the i-th row of encryption matrix, g_i', can be successfully estimated with probability $|Aut(\Gamma)|^{-1}$. So the attacker must construct $|Aut(\Gamma)|^K = \mathcal{N}_E^K$ encryption matrices G' to achieve the correct encryption matrix because the correctness of a row vector g_i' cannot be verified independently from the other rows. Therefore, obtaining the encryption matrix G' has the work factor of $\mathcal{O}(KN\mathcal{N}_E^K)$ operations. Obviously, if the value of $|Aut(\Gamma)| = 2^{N-K}$ is large enough, this attack will fail.

Furthermore, it was shown that the average number of required encryptions for one particular M to obtain all the error vectors is $\mathcal{N}_E \sum_{i=0}^{\mathcal{N}_E-1} 1/(\mathcal{N}_E - i) = \mathcal{O}(\mathcal{N}_E log\mathcal{N}_E)$. Since this procedure should be repeated for the K unit vectors $u_i, 1 \le i \le K$, so the total number of ciphertexts required for this attack is $\mathcal{O}(K\mathcal{N}_E log\mathcal{N}_E)$. Also, the number of N-bit error vectors is $\mathcal{N}_E$, So the preliminary work of this attack is $\mathcal{O}(KN\mathcal{N}_E^2 log\mathcal{N}_E)$ [13]. For a (1024,832) nonsystematic polar code over GF(2) in our scheme, we have an astronomical number of 2^{192} error vectors. Therefore, the ST attack fails because of the large amount of preliminary work involved.

D. The Majority Voting Analysis

Majority Voting (MV) is an efficient algorithm to analyze the cryptanalytic strength of the secret key algebraic code encryptions [37]. There are different types of the MV algorithm such as Local Majority Voting (LMV) algorithm, Global Majority Voting (GMV) algorithm and Extended Majority Voting (EMV) algorithm. Based on these algorithms, some attacks such as MV attack [37] and EMV attack [35] were proposed. These attacks will be discussed subsequently in this subsection.

An Equivalent Secret Key Cryptosystem

The secret key cryptosystem which is equivalent to the RN cryptosystem is introduced in [35, 37] to be able to examine the strength of RN scheme against MV and EMV attacks. It was referred in [35]. Due to the RN system is secret key cryptosystem, the code is not hidden by nonsingular and permutation matrices as they are done in the McEliece public key cryptosystem. Therefore, these matrices are redundant and can be omitted. Hence the equivalent cryptosystem to the RN scheme can be described as follows.

Let $G'' = [g_{ij}''], i = 1, 2, \cdots K, j = 1, 2, \cdots, N$ be a binary $K \times N$ equivalent encryption matrix with a right inverse $(G'')^{-R}$. Let H be a corresponding binary $(N-K) \times N$ parity check matrix such that $G''H^T = 0$. Also, a set of binary N-tuple error vectors, $\mathcal{E}_{\mathcal{N}_E} = \{E_i, 1 \le i \le \mathcal{N}_E\}$, are selected which satisfies the weight property and the syndrome property of the RN cryptosystem. Finally, the syndrome error table $\mathcal{T} = \{(EH^T, E) | E \in \mathcal{E}_{\mathcal{N}_E}\}$ is computed.

Encryption

A binary K-tuple message M is encrypted into a binary N-tuple ciphertext $\mathcal{C}$ by calculating $\mathcal{C} = MG'' + E, E \in \mathcal{E}_{\mathcal{N}_E}$.

Decryption

A ciphertext $\mathcal{C}$ is decrypted by performing the following steps.

- Compute the $\mathcal{C}H^T = MG''H^T + EH^T = EH^T$.
- Obtain the error vector E corresponding to EH^T using the syndrome error table $\mathcal{T}$.
- Retrieve the message $M = (\mathcal{C} + E)(G'')^{-R}$.

The structure of the proposed cryptosystem is similar to the RN scheme, so we can consider the above equivalent system to our scheme. The strength of proposed scheme is examined against MV and EMV attacks.

The Majority Voting (MV) attack

The aim of the MV attack is to recover the encryption matrix G'' of the equivalent cryptosystem by performing the following procedures [37].

1. Choose an arbitrary plaintext M, and compute a set of l distinct encryptions of M, i.e., $\mathrm{C}_l = \{\mathcal{C}_i = MG'' + E_i, 1 \le i \le l\}$. Let $\mathcal{E}_l = \{E_i, 1 \le i \le l\}$ denote the set of l distinct binary N-tuple error vectors. Then, compute $\mathcal{M}(\mathrm{C}_l) = \mathcal{M}(MG'') + \mathcal{M}(\mathcal{E}_l)$ where $\mathcal{M}(\mathrm{C}_l)$ is an $l \times N$ matrix consisting of the ciphertexts $\mathcal{C}_i, 1 \le i \le l$ in its i-th rows respectively. Also, $\mathcal{M}(MG'')$ is an $l \times N$ matrix which the N-bit vector MG'' is repeated in its all rows. Similarly, $\mathcal{M}(\mathcal{E}_l)$ is an $l \times N$ matrix consisting of the error vectors $E_i, 1 \le i \le l$ in its i-th rows respectively. The majority voting on each column of $\mathcal{M}(\mathrm{C}_l)$ yields an estimate $\widehat{MG''}$, i.e. when the number of 1's are more than the number of 0's in a column, set the corresponding bit to '1', otherwise to '0'.
2. Repeat the first step for a set of K linearly independent messages $M_1, M_2, \cdots, M_K$ and compute a set of K corresponding estimates $\widehat{M_1G''}, \widehat{M_2G''}, \cdots, \widehat{M_KG''}$.
3. Finally, obtain an estimate of encryption matrix as $\widehat{G''} = \mathcal{M}^{-1}(M)\,\mathcal{M}(\widehat{MG''})$ where $\mathcal{M}(M)$ is a $K \times K$ matrix consisting of the K-bit messages $M_i, 1 \le i \le K$ in its i-th rows respectively and $\mathcal{M}(\widehat{MG''})$ is a $K \times N$ matrix consisting of the K estimates $\widehat{M_iG''}, 1 \le i \le K$ in its i-th rows respectively.

The encryption matrix G'' via the attack can be recovered and used to break an equivalent cryptosystem. This attack requires K times l majority votes over N coordinates. So, the work factor requires an average number of $\mathcal{O}(KNl)$ bit operations. If the worst case, $l = \mathcal{N}_E$, is considered, it has a work factor of $\mathcal{O}(KN\mathcal{N}_E)$ bit operations. The RN cryptosystem is not optimally secure against an MV attack; because of its restricted number of error vectors, $\mathcal{N}_E = 2^8$. In the proposed cryptosystem, due to large number of error vectors, $\mathcal{N}_E = 2^{192}$, and the length of used polar code, $N = 2^{10}$, this attack is impractical.

An extension of MV algorithm is the LMV algorithm which considers more than one coordinate of vector MG'' simultaneously. Let $(MG'')_B$ be an s-bit subvector of the N-bit

vector MG'' corresponding to set of positions $B = \{b_1, b_2, \cdots, b_s\} \subseteq \{1, 2, \cdots, N\}, s \leq N$. Similarly, $(MG'' + E_i)_B$ is defined as an s-bit subvector of N-bit ciphertexts $\mathcal{C}_i = MG'' + E_i, 1 \leq i \leq l$ corresponding to the set of positions B. In summary, the goal of the LMV algorithm is to find an estimate $(\widehat{MG''})_B$ from l observations $(MG'' + E_i)_B, 1 \leq i \leq l$. It was shown that, for l large enough, one of the cases $(\widehat{MG''})_B = (MG'')_B$ or $(\widehat{MG''})_B = (MG'')_B + 1$ will happen with great probability [35].

On the other hand, the aim of the GMV algorithm is to perform r times of the LMV algorithm to obtain r local estimates $(\widehat{MG''})_{B_j}, 1 \leq j \leq r$ from l observations of $(MG'' + E_i)_{B_j}, 1 \leq i \leq l, 1 \leq j \leq r$, where $B_j = \{b_{j,1}, b_{j,2}, \cdots, b_{j,s}\} \subseteq \{1, \cdots, N\}$, $\{\cup_{1 \leq j \leq r} B_j\} = \{1, \cdots, N\}$ and $\forall j, \{B_j \cap B_{j+1}\} \neq 0$. It was shown that for each j and l large enough, one of the cases $(\widehat{MG''})_{B_j} = (MG'')_{B_j}$ or $(\widehat{MG''})_{B_j} = (MG'')_{B_j} + 1$ will happen with great probability. By repeating this procedure for all $B_j, 1 \leq j \leq r$, the r obtained local estimates $(MG'')_{B_j}, 1 \leq j \leq r$ cover the complete vector MG''. So, one of two cases $(\widehat{MG''}) = MG''$ or $(\widehat{MG''}) = MG'' + 1$ will happen with great probability for l large enough [35].

The Extended Majority Voting (EMV) Attack

Finally, the EMV attack [35] can be regarded as a generalized MV attack which uses the obtained estimates by the LMV and the GMV algorithms to recover the encryption matrix G''. This attack can be performed by the following steps [35].

1. Let $M_v = M + u_v$, where u_v is the v-th unit vector. Perform the GMV algorithm for the message $M_v, 1 \leq v \leq K$ to obtain the estimate $(\widehat{M_v G''}) = M_v G''$ or $(\widehat{M_v G''}) = M_v G'' + 1$ with great probability for l large enough. Then, obtain the estimate $(\widehat{MG''}) = MG''$ or $(\widehat{MG''}) = MG'' + 1$.
2. Compute the estimate of the v-th, $1 \leq v \leq K$, row vector of the encryption matrix G'' by $\widehat{g_v''} = (\widehat{M_v G''}) + (\widehat{MG''})$. Repeat this procedure for all row estimates $\widehat{g_v''}, 1 \leq v \leq K$ to obtain G''.
3. Let $\widehat{\mathcal{E}_{\mathcal{N}_E}} = M\widehat{G''} + \mathsf{C}_{\mathcal{N}_E}$, where $\widehat{G''} = (\widehat{g_1''}^T, \cdots, \widehat{g_K''}^T)^T$ is an estimate of the encryption matrix G''. Moreover, $\mathsf{C}_{\mathcal{N}_E}$ is a set of all distinct ciphertexts of a message M, i.e., $\mathsf{C}_{\mathcal{N}_E} = MG'' + \mathcal{E}_{\mathcal{N}_E}$. Further, let $\mathcal{E}_{\mathcal{N}_E} = \mathsf{C}_{\mathcal{N}_E} + MG''$ be the set of corresponding error vectors.
4. Finally, the obtained matrix G'' and the set $\mathcal{E}_{\mathcal{N}_E}$ together define an equivalent cryptosystem.

In general, the above attack requires using l encryptions to perform a successful voting. If the worst case $l = |\mathcal{E}| = \mathcal{N}_E$ is considered, the average number of required encryptions to obtain all the error vectors has the order of $\mathcal{N}_E \sum_{i=0}^{\mathcal{N}_E - 1} 1/(\mathcal{N}_E - i) = \mathcal{O}(\mathcal{N}_E log \mathcal{N}_E)$. The work factor of $\mathcal{O}(K\mathcal{N}_E log \mathcal{N}_E)$ vector operations is required for success. Similar to MV attack, the work factor of EMV attack is dependent on the number of the error vectors, $\mathcal{N}_E$, and the dimension of used code. For the proposed scheme, due to proper selection of these parameters, this attack is impractical. Another note to improve the resistance of system against EMV attack is to choose randomly the predefined set of error vectors without any weight constraint [35]. In the proposed cryptosystem, there are no constraints on the Hamming weight of error vectors to avert the EMV attack.

TABLE V. Security comparison of the Rao-Nam and the proposed cryptosystems.

Attacks \ ***Cryptosystems***		***Rao - Nam cryptosystem* [7, 12]**	***Proposed cryptosystem***
***Brute Force Attack* [6, 7]**	$\mathcal{N}_C$	$\geq K! \approx 2^{296}$	$\binom{NR_0}{K} \approx 2^{174}$
	$\mathcal{N}_\mathcal{P}$	$N! = 72! \approx 2^{345}$	$2^{N-2} = 2^{1022}$
	$\mathcal{N}_S$	$\prod_{i=0}^{K-1}(2^K - 2^i) > 2^{K^2-K} = 2^{(64)^2-64}$	$2^{2K-4} - 1 = 2^{2(832)-4} - 1$
	$\mathcal{N}_E$	$2^{N-K} = 2^8$	2^{192}
***Rao-Nam Attack* [6, 7]**	Determination of the encryption matrix [7]: $\Omega(2^{(N-K)K})$	$\Omega(2^{512})$	$\gg 2^{80}$
	Determining M from $\mathcal{C}$ by knowing the encryption matrix [6, 7]: $K^3 \times \prod_{i=0}^{K-1}\left(1 - \frac{t}{N-i}\right)$	$64^3 \times \prod_{i=0}^{63}\left(1 - \frac{1}{72-i}\right)$	$831^3 \times \prod_{i=0}^{831}\left(1 - \frac{w_H(E)}{1024-i}\right)$
***Majority Voting Analysis* [35, 37]**	MV Attack [37]: $\mathcal{O}(KN\mathcal{N}_E)$	$\approx \mathcal{O}(2^{20})$	$\mathcal{O}(2^{211})$
	EMV Attack [35]: $\mathcal{O}(K\mathcal{N}_E log \mathcal{N}_E)$	$\approx \mathcal{O}(2^{15})$	$\approx \mathcal{O}(2^{208})$
***Struik-Tilburg Attack* [13]**	Preliminary work [13]: $\mathcal{O}(KN\mathcal{N}_E^2 log \mathcal{N}_E)$	$\approx \mathcal{O}(2^{29})$	$\approx \mathcal{O}(2^{410})$
	Calculation of encryption matrix [13]: $\mathcal{O}(KN\mathcal{N}_E^K)$	$\approx \mathcal{O}(2^{524})$	$\gg 2^{80}$

As it is shown in this table, the actual security of the proposed scheme depends on some parameters such as the code length N, the dimensionality K of applied polar code and the number of error vectors, $\mathcal{N}_E = 2^{N-K}$. However, the main causes of insecurity of the RN scheme against these attacks is short length and dimensionality of used (72,64) Hamming code and consequently low number of error vectors, $\mathcal{N}_E = 2^8$. In our proposed cryptosystem, due to applying $(1024, 832)$ nonsystematic polar code, the number of error vectors is $\mathcal{N}_E = 2^{192}$. This implies that our scheme has an acceptable resistance against chosen plaintext attacks and other well-known cryptanalytic attacks. Therefore, the parameters of used

code should be chosen in some way to keep the work factors of these attacks high enough and guarantee the security of the system.

VII. CONCLUSION AND FUTURE WORK

The issue of using non-systematic polar codes in the structure of symmetric key cryptosystem is addressed in this paper. The proposed scheme has some advantages in terms of efficiency and security in comparison with Rao-Nam secret key cryptosystem such as higher security level and shorter key length. In addition, by taking the advantage of combining security and channel coding processes, this scheme can be implemented with reasonable complexity which makes it suitable for secure high speed communications.

The proposed scheme is employed the non-systematic polar codes based on the following reasons: (1) By selecting K bit-channels from NR_0 good bit-channels, the generator matrix $G_{A(\mathcal{s})}$ is hidden and a large family of equivalent polar codes is constructed. This method helps the proposed cryptosystem to provide an efficient security and also to forbid an exhaustive attack. (2) By using the interesting structure of the generator matrix of polar codes, a secret information set $A(\mathcal{s})$ is stored instead of the generator matrix $G_{A(\mathcal{s})}$ which leads in reducing the key size. (3) By utilizing the non-systematic property of polar codes, the information bits do not appear explicitly as part of codeword. Also, a specific form of error vectors can be calculated from the random frozen vectors. These are resulted to provide the security and reduce the key length. (4) Polar codes have low complexity encoding/decoding algorithms which allow the proposed scheme to have easy encryption/decryption algorithms. Also, the method of constructing these codes is simple over BEC and can be performed with complexity $\mathcal{O}(N)$.

Simulation results depict that security and reliability of this scheme depend on some factors such as code length N, code rate R, erasure probability ϵ, secret information set $A(\mathcal{s})$ and cutoff rate R_0. In order to design a secure and reliable secret key scheme based on polar codes, these parameters should be selected in such a way that suitable trade-off between security and reliability is established.

Our future work is to apply the polar codes in the structure of McEliece public key cryptosystem. However, reducing the key length of McEliece cryptosystem based on polar codes is an open problem of this concept.

REFERENCES

[1] S. Lin, D. J. Costello, Error control coding: fundamentals and applications, Prentice-Hall, Upper Saddle River, NJ, 1983, 2nd edition, 2004.

[2] E. Arıkan, ''Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,'' *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051-3073, Jul. 2009.

[3] A. D. Wyner, ''The wire-tap channel,'' *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1367, Oct. 1975.

[4] A. Thangaraj, S. Dihidar, A. Calderbank, S. McLaughlin, J. M. Merolla, ''Applications of LDPC codes to the wiretap channel,'' *IEEE Trans. Inform. Theory*, vol. 53, no. 8, pp. 2933-2945, Aug. 2007.

[5] H. Mahdavifar, A. Vardy, ''Achieving the secrecy capacity of wiretap channels using polar Codes,'' *IEEE Trans. on Information Theory*, pp. 6428-6443, vol. 57, issue. 10, Oct. 2011.

[6] R. J. McEliece, ''A public-key cryptosystem based on algebraic coding Theory,'' DNS Progress Rep., Jet Propulsion Labaratory, pp. 114-116, CA, Pasadena, 1978.

[7] T. R. N. Rao, K. H. Nam, ''Private-Key Algebraic-Code Encryption,'' *IEEE Trans. on Information Theory*, vol. 35, no. 4, pp. 829-833, 1987.

[8] D. J. Bernstein, J. Buchmann, E. Dahmen, Post-Quantum Cryptography, Springer, 2008.

[9] E. R. Berlekamp, R. J. McEliece, H. C. A. van Tilborg, ''On the Inherent Intractability of Certain Coding Problems,'' *IEEE Trans. on Information Theory*, vol. 24, no. 5, pp. 384-386, May 1978.

[10] T. Johansson, F. Jonsson, ''On the complexity of some cryptographic problems based on the general decoding problem,'' *IEEE Trans. on Information Theory*, vol. 48, Issue. 10, pp. 2669-2678, Oct. 2002.

[11] T.R.N. Rao, ''Cryptosystems using algebraic codes'', *in Proc. Int. Conf. on Computer Systems and Signal Processing*, Bangalore, India, Dec. 1984.

[12] T. R. N. Rao, K. H. Nam, "Private-key algebraic cryptosysterns," *in Advances in Cryptology - CRYPTO'86*, New York: Springer-Verlag, pp. 35-48, 1986.

[13] R. Struik, J. van Tilburg, ''The Rao-Nam scheme is insecure against a chosen-plaintext attack,'' *in Advances in Cryptology - CRYPTO'87*, New York: Springer-Verlag, pp. 445-457, 1987.

[14] R. Struik, ''On the Rao-Nam scheme using nonlinear codes,'' *in Proc. IEEE Int. Symp. on Information Theory*, p. 174, 1991.

[15] F. M. R. Alencar, A. M. P. Leo, R. M. Campello de Souza, ''Private-key burst correcting code encryption,'' *in Proc. IEEE Int. Symp. on Information Theory*, pp. 227, Jan. 1993,.

[16] H. M. Sun, S. P. Shieh, ''Cryptanalysis of private-key encryption schemes based on burst-error-correcting codes,'' *in Proc. 3rd ACM Conf. on Computer and Communications Security*, pp. 153-156, New Delhi, 14-16 March 1996.

[17] R. M. Campello de Souza, J. Campello de Souza, ''Array codes for private-key encryption,'' *IEEE Electronics Letters*, vol. 30, no. 17, pp. 1394-1396, 1994.

[18] A. Al Jabri, ''Security of private-key encryption based on array codes,'' *IEEE Electronics Letters*, pp. 2226-2227, vol. 32, no. 24. 1996.

[19] H. M. Sun, ''Private key cryptosystem based on burst error correcting codes,'' *IEEE Electronics Letters*, vol. 33, no. 24, pp. 2035-2036, 1997.

[20] A. I. Barbero, O. Ytrehus, ''Modifications of the Rao-Nam Cryptosystem,'' in Proc. *Int. Conf. on Coding Theory, Cryptography and Related Areas*, pp. 1-13, , Guanajuato, Mexico, April 1998.

[21] A. Payandeh, M. Ahmadian, M. R. Aref, ''Adaptive secure channel coding based on punctured turbo codes,'' *IEE Proceeding Communications*, vol. 153, no. 2, pp. 313-316, April 2006.

[22] A. Payandeh, M. Ahmadian, M. R. Aref, ''An Adaptive Secure Channel Coding Scheme for Data Transmission over LEO Satellite Channels,'' *Scientica Iranica*, vol. 13, no. 4, pp. 373-378, Oct. 2006.

[23] A. A. SobhiAfshar, T. Eghlidos, M. R. Aref, ''Efficient secure channel coding based on quasi cyclic-low density parity check codes,'' *IET Communications Journals*, vol. 3, pp. 279-292, 2009.

[24] R. Hooshmand, T. Eghlidos, M.R. Aref, 'Improving the Rao-Nam Secret Key Cryptosystem Using Regular EDF-QC-LDPC Codes', *ISeCure Journal*, vol. 4, no. 1, pp. 3-14, Jan. 2012.

[25] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, M. Skoglund, ''Nested polar codes for wiretap and relay channels,'' *IEEE Communications Letters*, vol. 14, no. 8, pp. 752-754, Aug. 2010.

[26] E. Hof, S. Shamai, ''Secrecy-achieving polar-coding,'' *in Proc. IEEE Information Theory Workshop*, pp. 1-5, Sep. 2010.

[27] R. Hooshmand, M. K. Shooshtari, M.R. Aref, ''Secret Key Cryptosystem based on Polar Codes over Binary Erasure Channel,'' *in Proc. 10th Int. ISC conf. on Information Security &*

cryptology (ISCISC2013), pp. 21-22, Yazd University, Iran, 29-30 Aug. 2013.

[28] N. Goela, S. B. Korada, M. Gastpar, ''On LP decoding of polar codes,'' *in Proc. IEEE Information Theory Workshop*, pp. 1-5, 2010.

[29] E. Arıkan, ''A performance comparison of polar codes and Reed-Muller codes,'' *IEEE Communications Letters*, vol. 12, pp. 447–449, June 2008.

[30] E. Arıkan, ''Systematic Polar Coding,'' *IEEE Communications Letters*, vol. 15, no. 8, pp. 860-862, Aug. 2011.

[31] Robert H. Morelos-Zaragoza, The Art of Error Correcting Coding, Second Edition, John Wiley & Sons, 2006.

[32] S. B. Korada, Polar Codes for Channel and Source Coding, Ph.D. dissertation, EPFL, Lausanne, Switzerland, 2009.

[33] S. B. Korada, A. Montanari, E. Telatar and R. Urbanke , "An emprical scaling law for polar codes,'' *in Proc. IEEE Int. Symp. on Information Theory*, pp.884-888, 2010.

[34] A. Goli, S. H. Hassani, R. Urbanke, ''Universal Bounds on the Scaling Behavior of Polar codes,'' *in Proc. IEEE Int. Symp. on Information Theory*, pp. 1957-888, 2012.

[35] J. Meijers, J. V. Tilburg, ''Extended Majority Voting and Private-Key Algebraic Code Encryptions,'' *in Proc. ASIACRYPT'91*, vol. 739, pp. 288-298, Fujiyoshida, Japan, Nov. 1991.

[36] T. R. N. Rao, "On Struik-Tilburg cryptoanalysis of Rao-Nam scheme," *in Advances in Cryptology - CRYPTO'87*, pp. 458-461, New York: Springer-Verlag, 1987.

[37] J. van Tilburg, Security-Analysis of a Class of Cryptosystems Based on Linear Error-Correcting Codes, Ph.D. dissertation, Tech. Univ. Eindhoven, Eindhoven, The Netherlands, 1994.

[38] CCSDS 131.0-B-1, 'TM synchronization and channel coding', *Recommendation for Space Data System Standards*, CCSDS, Washington, DC, Aug. 2011, Blue Book,

[39] H. M. Sun, T. Hwang, "Key Generation of Algebraic-Code Cryptosystems,'' *Computers and Mathematics with Applications*, vol. 27, no. 2, pp. 99-106, 1994.

[40] H. M. Sun, T. Hwang, ''One Double-One Matrices and Double-Zero Matrices,'' *Linear and Multilinear Algebra*, vol. 31, pp. 47-55, 1992.