# One Class of Ideally Secret Autonomous Symmetric Ciphering Systems Based on Wiretap Polar Codes

Milan Milosavljević * , Jelica Radomirović , Tomislav Unkašević , Boško Božilović

Posted Date: 13 August 2024

doi: 10.20944/preprints202408.0756.v1

Keywords: cryptography; ideal secrecy; polar coding; wire-tap channel; key equivocation; privacy amplification

 doi:10.20944/preprints202408.0756.v1



*Article*

# One Class of Ideally Secret Autonomous Symmetric Ciphering Systems Based on Wiretap Polar Codes

**Milan Milosavljević [1], Jelica Radomirović [1,2,*] and Tomislav Unkašević [1] and Boško Božilović [1]**

[1] Vlatacom Institute of High Technology, Milutina Milankovica 5, 11070 Belgrade, Serbia; tomislav.unkasevic@vlatacom.com, milan.milosavljevic@vlatacom.com, bosko@vlatacom.com
[2] School of Electrical Engineering, Belgrade University, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia;
* Correspondence: jelica.radomirovic@vlatacom.com

**Abstract:** This paper introduces a class of symmetric ciphering systems with a finite secret key, which provides ideal secrecy, autonomy in key generation and distribution, and robustness against the probabilistic structure of messages (Ideally Secret Autonomous Robust – ISAR system). The ISAR system is based on wiretap polar codes constructed over an artificial wiretap channel with a maximum secrecy capacity of 0.5. The system autonomously maintains a minimum level of key equivocation by continuously refreshing secret keys without needing additional key generation and distribution infrastructure. Thus, this class of systems generalizes Shannon's ideal and strongly ideal ciphering systems. Additionally, it can transform any stream ciphering system with a finite secret key of known length into an ISAR system without the necessity of knowing and/or changing its algorithm. Therefore, this class of system strongly supports privacy, a critical requirement for contemporary security systems. The ISAR system's reliance on wiretap polar coding for strong secrecy ensures resistance to passive attacks with known plaintext messages. Furthermore, resistance to passive attacks on generated keys follows directly from ideal secrecy and autonomy. The results presented offer an efficient methodology for synthesizing this class of systems with predetermined security margins and a complexity of the order of $n \log n$, where $n$ is the block length of the applied polar code.

**Keywords:** cryptography; ideal secrecy; polar coding; wire-tap channel; key equivocation; privacy amplification

**MSC:** 94A60

## 1. Introduction

The notion of unconditional secrecy was introduced by Claude Shannon in [1], to describe an encryption system resistant to any passive attack on messages based on ciphertext. However, he showed in the same paper that the necessary condition for unconditional secrecy is that the secret key length must be at least as long as the message, which is impractical for most applications. To address this, various approaches have been developed to create secure ciphers with shorter keys. Shannon himself suggested so-called ideal and strongly ideal cipher systems, where an adversary is left with many equally probable decryption options, rather than a unique solution.

In this paper, we address two main questions:

1. Is it possible to develop a general concept for constructing ideal cipher systems with a predetermined minimum value of key equivocation, independent of the telecommunication environment, the probabilistic structure of the message, and without additional infrastructure for generating and distributing secret keys?
2. Is it possible to apply this solution to any existing symmetric stream ciphering system without modifying it or knowing the encryption key generation algorithm, except for the length of the applied secret key?

If feasible, such a system would possess the following properties:

 doi:10.20944/preprints202408.0756.v1



- *Ideal secrecy* - Guaranteed minimum value of key equivocation, regardless of the length of the ciphertext used by the adversary.
- *Autonomy* - The autonomous ability to maintain a given minimum value of key equivocation by continuously refreshing keys without additional infrastructure for key generation and distribution.
- *Robustness* - Retains guaranteed properties regardless of the probabilistic structure of the messages.

We refer to this system as the ISAR (Ideally Secure Autonomous Robust) ciphering system. This paper presents a class of ISAR systems based on wiretap polar codes.

### *1.1. Related Works*

Since key equivocation directly depends on the entropy of messages, the first group of works on ideal ciphering systems focuses on preprocessing messages before encryption, to reduce redundancies (or equivalently to increase entropy). In this way, the minimum key equivocation available to a system attacker could be controlled. Compression and randomization techniques are standard for this approach and have a long tradition in cryptography, [1–5]. The homophonic coding technique, which transforms a sequence of message symbols into a uniquely decodable sequence where all symbols have the same frequency, also belongs to this approach, [2,6–8]. The effectiveness of this approach depends on knowledge of the probabilistic properties of messages and reliable estimates of the statistical properties of the message source being encrypted.

Another approach, not much different from the first, is entropic security, [4,9–11]. The main difference lies in the security metrics used as a criterion. The goal of this metric is that any function of the original message is unattainable to passive adversaries. In the limiting case, when the so-called leakage parameter is equal to zero, this criterion aligns with Shannon's definition of a perfect ciphering system. However, then all the obtained results for the length of the secret key become practically unusable or give lengths equal to the length of the messages, reducing these procedures to the Vernam cipher.

The honey cipher, introduced in [12], is similar to the concept of ideal cipher. The main goal of this approach is that an adversary is left with many highly probable hypotheses about secret keys or messages. In [13] authors combine the ideas of honey ciphers and entropic security to create, practically implementable short-key ciphering system. The dependence of the efficiency of these systems on the knowledge of the probabilistic properties of the source being encrypted is even more critical than previous approaches. This is a consequence of the very idea of the system that, in the process of brute force attacks, hypothesis messages are generated that are difficult to distinguish from the true ones. However, the future development of these systems will likely enhance understanding of the relationship between the probabilistic properties of the source and the minimum length of secret keys sufficient to prevent brute force attacks by an unbounded adversary, which brings us back to Shannon's original ideas about the meaning of non-zero equivocation values of secret keys.

The final group of methods, relevant for the construction of ideally secret systems, comes from the broad research domain of combining the methods of error correcting codes, especially wiretap codes and cryptography, [6,14–19]. In [6], a system based on wiretap codes was proposed, which proved to improve the performance of an arbitrary stream cyphering system in the regime of short messages, but whose key equivocation drops to zero for sufficiently long messages. In the paper [20], the authors integrated wiretap polar coding in encryption schemes based on learning with errors problem and showed that with appropriate refresh of secret key procedures, they achieve non-zero equivocation of keys. In paper [21], the connection between ideal secrecy and wiretap coding approach is clarified. The same authors in [22] proposed an encryption system over the MIMO wiretap channel, which for infinite lattice input alphabet guarantees perfect security, while for finite constellations guarantees ideal secrecy with high probability. In [23] authors proposed a polar coding scheme which achieves maximal secrecy capacity for a secret key of arbitrary non-zero rate, shared between transmitter and legitimate receiver. In [24] an encrypted secure polar coding scheme for

 doi:10.20944/preprints202408.0756.v1



general two-way wiretap channel is presented. To achieve strong security and reliability criteria, without any key pre-sharing, it is necessary to apply a complex cooperative jumming strategy.

Based on the presented analysis of the relevant published results, we can conclude the following:

- Wiretap coding provides a promising environment for implementing cryptographic systems of increased security.
- Works in which ideal secrecy is proven are always related to a specific ciphering system, such as in [20], or is limited to a particular physical model of the wiretap channel, as in [22] and [24].

Based on this analysis that includes all four approaches, we conclude that there is no example of an ISAR system in the available literature.

### *1.2. Paper Organization*

Section 2 provides a basic conceptual and theoretical basis for understanding ISAR systems, particularly from the domain of wiretap channels, polar coding, privacy amplification techniques and Shannon's notion of ideal and strongly ideal systems.

Section 3 describes the system architecture and security properties of the ISAR ciphering system. First, it is shown that the system is equivalent to a wiretap model, whose main channel is error-free, while the wiretap channel is equivalent to an embedded symmetric stream cyphering system. In Theorem 2, it is proved how wiretap coding should be constructed in order to achieve strong secrecy and reliability. Theorem 4 gives the capacity of the wiretap channel, which turns out to decrease linearly with the length of the secret key. Theorem 5 gives equivocation of secret key, while Theorem 6 gives secrecy capacity of proposed ISAR as a function of polar code length and security margins for equivocation and privacy amplification. Then it was shown that for an arbitrary cipher system to be ideal with the same lower bound of key equivocation as the ISAR system, its secret key must be $t$-times greater than that of ISAR, where $t$ is proportional to message length. Thus, the superiority of the ISAR system increases with the length of the messages. The chapter concludes with a demonstration of how any symmetric stream ciphering system can be transformed into ISAR.

Section 4 provides security analysis of ISAR ciphering system, including its resistance to passive attacks both on secret keys and messages.

Section 5 summarizes the practical aspects of the implementation and application of the ISAR system in contemporary information and communication infrastructure. It is shown that the complexity of the system is $O(n \log n)$, where n is the length of the polar code. For point-to-point protection, the initial exchange of the secret key and the hash function seed is performed only once, regardless of the number of sessions or communication disruptions.

Section 6 concludes the paper with a summary of findings and suggestions for future research directions.

### *1.3. Notations*

We define integer interval $[a, b]$ as the integer set between $a$ and $b$. We denote $X, Y, Z, \ldots$ random variables taking values in alphabets $X, Y, Z, \ldots$ and their realization is denoted by $x, y, z, \ldots$ respectively. Also, we denote a $n$ size vector $X^n = (X_1, X_2, \ldots, X_n)$ and denote $X_a^b = (X_a, X_{a+1}, \ldots, X_b)$. Further, for any index set $A \subseteq [1, n]$, we define $X_A = \{X_i\}_{i \in A}$. $H(\cdot)$ denotes entropy and $I(\cdot)$ denotes mutual information.

## 2. Preliminaries

In order to make it easier to understand the operation of the proposed system, in this section we will present basic concepts from the domain of wiretap channel model, polar coding and privacy amplification, since these three elements are its basic building blocks. In the final part of the chapter, the definition of an ideal autonomous ciphering system is given, as a kind of generalization of Shannon's notions of ideal and strongly ideal ciphering systems, [25].

### *2.1. Wiretap Channel Model*

 doi:10.20944/preprints202408.0756.v1



The wiretap model consists of two channels $W_m: X \to Y$ and $W_e: X \to Z$, the main and wiretap channels, respectively, first time introduced by Wyner [14], see Figure 1. $M$ denotes the $k$-bit message that Alice wants to send to Bob. With the help of random bits, the encoder maps $M$ into sequence $X$ of $n$-bit channel symbols. This sequence is sent on the main and wiretap channels, giving the corresponding channel outputs $Y$ and $Z$. On Bob's side, the decoder maps $Y$ into estimate $\widehat{M}$ of the original message.

If the main channel is superior to the wiretap channel, we say that the wiretap channel is degraded relative to the main channel. Formally, for each wiretap channel $W_e$ degraded with respect to the main channel $W_m$ can be find a third channel $W_*: Y \to Z$ , such that $W_e$ can be represented as a cascade of channels $W_m$ and $W_*$. In other words, the degraded channel always has a lower capacity compared to the main channel.

The basic purpose of wiretap coding is the design of encoders and decoders that provide for legitimate users (Alice, Bob) at the same time reliability and security when the message length $k = |M|$ tends to infinity. Reliability is usually measured in terms of the probability of error in recovering the message sent by Alice at Bob's side

$$\lim_{k\to\infty} Pr\{\widehat{M} \neq M\} = 0 \ . \tag{1}$$

As a rule, security is measured in terms of mutual information between the message M and Eve's observation Z, giving the two most frequently analyzed criteria, so called week security criterion

$$\lim_{k\to\infty} \frac{I(M;Z)}{k} = 0 \ , \tag{2}$$

and strong security criterion

$$\lim_{k\to\infty} I(M;Z) = 0 \ . \tag{3}$$

As Maurer first pointed out [26,27], Eve can disclose $k^{1-\varepsilon}, \varepsilon \in (0,1)$ message bits, while (2) is still true, which is obviously unacceptable. Therefore, satisfying the strong security criterion (3) is the generally accepted gold standard of the cryptographic community. Note that there are other secrecy measures, such as advantage and semantic security, which are shown to be equivalent to the strong security criterion [15].

The largest transmission rate at which both reliability and secrecy conditions are simultaneously satisfied is commonly referred to as the secrecy capacity $C_s$. In case the main and wiretap channels are binary symmetric channels (BSC) and the wiretap channel is degraded with respect to the main channel, $C_s$ was given by [28]

$$C_s = C(W_m) - C(W_e) = H(X|Z) - H(X|Y) \ , \tag{4}$$

where the random variable $X$ at the input to the channel is uniform over $X$, while $C(W_m)$ and $C(W_e)$ denotes the capacities of the main and wiretap channels.

### *2.2. Wiretap Polar Coding*

Polar codes, introduced by Erdal Arikan [29,30], are a class of error-correcting codes that achieve the capacity of binary-input symmetric memoryless (BSM) discrete channels. The key idea behind polar coding is to transform a set of $n = 2^b$ communication channels into a set of polarized channels, where a subset of channels become nearly perfect (noiseless) and the others become completely noisy. If with $W$ we denote BSM channel $\langle\{0,1\}, Y, W\rangle$, the Bhattacharyya parameter of $W$ is

$$Z(W) \triangleq \sum_{y\in Y} \sqrt{W(y|0)W(y|1)} \underset{k\to\infty}{=} 0 \ . \tag{5}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 August 2024 doi:10.20944/preprints202408.0756.v1



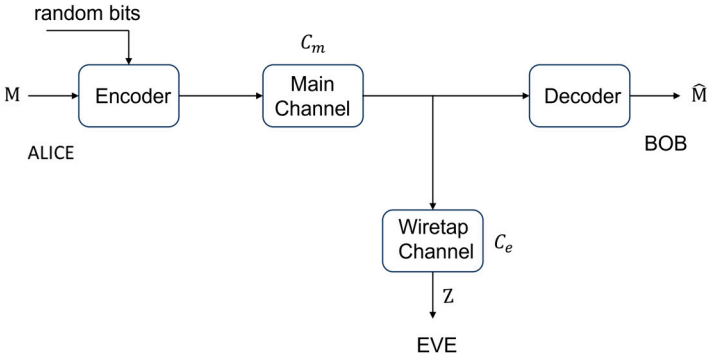


**Figure 1.** Generic Wiretap Channel Model.

Value (5) is equal to the upper bound probability of error Maximum Likelihood (ML) decoding on a single use of the channel. It is shown that $Z(W)$ takes values from the interval [0,1]. Channels with small $Z(W)$ values are almost noiseless, while channels with $Z(W)$ values close to 1 are almost pure-noise channels [30]. The essence of wiretap polar coding is the transformation of n instances of $W_m$ and $W_e$ into n-input channels $W_m^{(i)}$ and $W_e^{(i)}, i \in [1,n]$, which are polarized either as good or bad channels depending on their Bhattacharyya parameters $Z(W_m^{(i)})$ and $Z(W_e^{(i)})$ and/or symmetric capacity $C(W_m^{(i)})$, $C(W_e^{(i)})$. In order to achieve strong secrecy [17,18] proposed the following definition of good and bad polarized channels

$$P_n(W_e, \delta_n) = \left\{ i \in [1,n] : C(W_e^{(i)}) \leq \delta_n \right\} \qquad \delta_n \text{ - poor bit channels for Eve} \tag{6}$$

$$G_n(W_m, \beta) = \left\{ i \in [1,n] : Z(W_m^{(i)}) < 2^{-n^\beta}/n \right\} \qquad \text{good for Bob} \tag{7}$$

$$B_n(W_m, \beta) = \left\{ i \in [1,n] : Z(W_m^{(i)}) \geq 2^{-n^\beta}/n \right\} \qquad \text{bad for Bob} \tag{8}$$

using security function $\delta_n$ and a parameter $\beta \in (0,1/2)$. For fixed $\beta$, we can define next partition of $[1,n]$

$$R = [1,n] \setminus P_n(W_e, \delta_n). \tag{9}$$

$$A = P_n(W_e, \delta_n) \cap G_n(W_m, \beta) \tag{10}$$

$$B = P_n(W_e, \delta_n) \setminus G_n(W_m, \beta). \tag{11}$$

Next, let's partition the set R into two subsets R1 and R2 (R= R1 ∪ R2)

$$R1 = R \cap B_n(W_m, \beta) \tag{12}$$

$$R2 = R \cap G_n(W_m, \beta). \tag{13}$$

This obtained subsets of polarized channels, and their meaning in terms of transmission quality in the main and wiretap channels, are shown in Table 1.

**Table 1.** Partition of the index of polarized channels $W_m^{(i)}$ and $W_e^{(i)}, i \in [1,n]$, according to (6) – (13), which ensures strong secrecy.

| | **not $\delta_n$ – poor bit channels for Eve** | | **$\delta_n$ – poor bit channels for Eve** |
|---|---|---|---|
| **good for Bob** | $R$ | $R2$ - random bits | $A$ – message bits |
| **bad for Bob** | | $R1$ – random bits | $B$ – frozen bits (zeros) |

A polar wiretap encoder performs a 1-1 transformation of the original input vector $V_1^n$ into an $n$-dimensional code vector $X_1^n$,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 August 2024 doi:10.20944/preprints202408.0756.v1



$$X_1^n = V_1^n \cdot G_n \quad , \tag{14}$$

where $G_n$ is Arikan's generator matrix, given by

$$G_n = P_n \cdot F^{\otimes b} \quad , \tag{15}$$

$P_n$ is the $n \times n$ bit-reversal permutation matrix defined in [30], while $F^{\otimes b}$ is the b-fold tensor product of $F \triangleq \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ with itself. The input vector $V$ has the following structure

$$V_1^n = [V_R, V_A, V_B] \quad , \tag{16}$$

The corresponding components of the input vector $V$, are set to the following values

$$V_R = e, \quad V_A = M, V_B = 0, \quad , \tag{17}$$

where $e$ is random vector, selected by Alice uniformly at random from {0, 1}. Vector of frozen bits $V_B$ is set to 0 vector of dimension $|V_B| = n - |V_B^C| = n - |V_R \cup V_A|$, where $V_B^C$ is the complement of $V_B$ in $[1, n]$. As it was mentioned in [30], in the case of symmetric channels, any choice of frozen values is as good as other. From now on, we will assume that it is

$$k = |M|, \quad m = |R| \quad , \tag{18}$$

For the described coding scheme, in [17], see Proposition 16., is proven to hold

$$I(M_k; Z_1^n) \leq |P_n(W_e, \delta_n)|\delta_n \quad , \tag{19}$$

From here follows the conclusion that by choosing the security function $\delta_n$, we can bound mutual information leaked to Eve, for any message distribution. Specifically, according to [17], see Theorem 17., for any security function such that

$$\delta_n = o(1/n) \quad , \tag{20}$$

described wiretap coding scheme guarantees strong security.

As for achieving reliability conditions, since it is in the general case

$$R1 = R \cap B_n(W_m, \beta) \neq \emptyset \quad , \tag{21}$$

it is necessary to apply involving chaining construction [18] over multiple blocks to ensure reliability. Namely, since non frozen channels in strong security coding scheme are given by

$$A \cup R = G_n(W_m, \beta) \cup R1 \quad , \tag{22}$$

it is necessary to ensure reliable transmission in the channels indexed in the index set $R1$. This is exactly what chaining construction achieves by taking a subset of indices from set $A$, i.e., $H \subset A$ such that $|H| = |R1|$. Random bits that are placed in $H$ in the $j$-th block are used in $H$ in the $(j+1)$-th block for $j = 1, 2, 3, \ldots$. In the first block, random bits are used that were distributed to Alice and Bob before starting communication. In this way, the SC decoder on Bob's side in each block can successfully decode the contents in the channels with indices from (22), and therefore $V_A = M$.

Regarding the obtained secrecy capacity, the answer is provided by Theorem 1, from [17]:

**Theorem 1.** *For any security function $\delta_n$, and constants $\beta, c_1, c_2$ that satisfies condition*

$$c_1 2^{-n^\beta} \leq \delta_n \leq 1 - c_2, \quad \beta \in (0, 1/2), \quad c_1, c_2 > 0 \tag{23}$$

*the rate $R_n$ of the corresponding strong-security coding scheme approaches the secrecy capacity $C_s$ given by (4), namely*

$$\lim_{n\to\infty} R_n = \lim_{n\to\infty} \frac{|A|}{n} = C(W_m) - C(W_e) = C_s \quad . \tag{24}$$

*2.3. Privacy Amplification (PA)*

 doi:10.20944/preprints202408.0756.v1



Privacy amplification is a technique used in cryptography to enhance the security of a shared secret key between two parties. The primary goal is to convert a partially secure or somewhat compromised key into a highly secure key, even in the presence of an eavesdropper who may have some knowledge about the original key [19,31]. In terms of practical application, hash functions called universal families have a special role. [32].

**Definition 1.** *Family of* $G$ *functions* $g: A \to B$, *where* $A$ *and* $B$ *are two final sets, is* $2-$*universal, or in short* $-$ *universal, if it holds*

$$\forall x_1, x_2 \in A\ ,\ x_1 \neq x_2 \Longrightarrow P_G\{G(x_1) = G(x_2)\} \leq \frac{1}{|B|}\ . \tag{25}$$

*where* $G$ *is a random variable that represents a random choice of a afunction* $g \in G$ *uniformly at random in* $G$.

An example of a frequently used class of universal hash functions, see for example [33–35], is given by

$$H_M = \{h_M(x) = M \cdot x,\ x \in GF(2)^k, M \in GF(2)^{k\times r}\}\ . \tag{26}$$

In applications, a random choice of a hash function from a given set is implemented using a seeded pseudorandom generator. To emphasize this fact, we explicitly introduce seed notation. Therefore, a hash function from the class $H_M$, is denoted as

$$h_M(x, K_h)\ . \tag{27}$$

where $K_h$ denotes seed of adequate length. If $K_h$ is available only to legitimate users, the hash functions are typically known as cryptographic hash functions.

**Definition 2.** *The difference between the dimensions of the domain and codomain of a given class of hash functions will be referred to as its compression rate* $\Delta R$.

Compression rate of class $H_M$, defined in (26) is $\Delta R = r - k$.

*2.4. Ideal and Ideal Autonomous Cipher System*

Consider a general symmetric cipher system $(X, Z, K)$, where $X$ is the input plaintext, $Z$ is the ciphertext, and $K$ is the secret key.

According to Shannon, an ideal ciphering system is one in which the key equivocation, i.e., $H(K|\mathrm{Z}_1^n)$ remains non-zero even as the number of ciphertext symbols $n$ approaches infinity, [25]. This means that the uncertainty about the key does not reduce to zero, regardless of how much ciphertext is available to an attacker. Important properties of ideal encryption systems are:

- Non-zero Key Equivocation: The key remains partially unknown, no matter the length of the intercepted ciphertext.
- Security over Time: The security of the system does not degrade with the amount of data encrypted and transmitted.
- Practical Key Length: The key length can be shorter than the message length, unlike a one-time pad, but must be sufficient to maintain key equivocation.

A strongly ideal system has key equivocation constant at entropy of secret key, i.e $H(K|\mathrm{Z}_1^n) = H(K)$. This implies an even higher level of security, ensuring that the key's uncertainty remains unchanged from its apriory value, regardless of the volume of ciphertext available. Key Characteristics:

- Constant Key Equivocation: The amount of information about the key that remains unknown does not diminish with increasing ciphertext.
- Enhanced Security: Offers superior protection against extensive ciphertext analysis, maintaining key secrecy over an indefinite amount of encrypted data.

 doi:10.20944/preprints202408.0756.v1



- Robust Design: Typically requires more sophisticated cryptographic techniques to ensure that key equivocation remains constant.

By distinguishing between these two concepts, Shannon's theories provide a framework for evaluating and designing cryptographic systems based on the desired level of security and practical constraints.

Starting from these founding Shannon definitions, we will define a new term - ideal autonomous ciphering system.

**Definition 3.** *An ideally secret autonomous robust (ISAR) ciphering system is one which can maintain minimal key equivocation at predefined value* $\Delta_K$*, i.e.*

$$H(K|Z_1^n) \geq \Delta_K \quad . \tag{27}$$

*If several different secret keys are used during operation of one cipher system,* $K_1, K_2, \ldots K_t$, *which correspond to ciphertexts* $Z_1, Z_2, \ldots Z_t$, *we say that the system is ideal and autonomous if it is able to maintain a predetermined minimum value of equivocation for each of the applied keys, i.e.*

$$H(K_i|Z_i) \geq \Delta_K, \forall i \in [1,t] \quad . \tag{28}$$

**Remark 1.** *In Shannon's definition of an ideal system, it is important that the key equivocation is not reduced to zero with an unlimited increase of the available ciphertext. On the other hand, a strongly ideal cipher system requires constant maintenance of key equivocation at the entropy level of the secret key. An ideal autonomous cipher system is in a certain sense between these two extremes. It maintains the minimum value of key equivocation at a predetermined value, which in general belongs to the open interval* $(0, H(K))$.

**Remark 2.** *Since* $H(K|Z_1^n)$ *is a non-increasing function of the number of observed ciphertext symbols n, the minimum in (27) is always reached at the maximum available n.*

### 3. System Architecture and Security Properties of ISAR Cyphering System

Let the protected communication between the legitimate parties Alice and Bob take place in the successive exchange of messages $M$, which were previously divided into a series of blocks of length $k$. At their disposal is a symmetric stream cyphering system (GPSN, $K$) with short secret key $K$. As is known [1], such systems cannot provide strong security, and practical secrecy is measured by the amount of spent computer resources of the adversary (Eve) in arriving either to the message (partial system cracking) or to the secret key (total system cracking). Block $E$ denotes a polar coder, which performs a 1-1 transformation of messages $M$ into an $n$-dimensional codeword vector $X$, while block $E^{-1}$ performs an inverse transformation of vector $X$ into message $M$. Alice and Bob have local sources of randomness (denoted by RS on Alice's side), as well as Privacy Amplification (PA) blocks, which, based on the input random string obtained in block $i-1$, generate a shorter random string that serves as a new secret key $K(i)$ for encrypting the vector $X$ in block $i$. We will assume that the system uses a cryptographic hash function $h_M(x, K_h)$ from the class of universal hash functions (25), as well as that the seed $K_h$ was previously exchanged between Alice and Bob. In the initial block, $K(1) = S$, i.e., the initial secret key of the given stream cipher system (GPSN, $S$).

**Remark 3.** *Additionally, we will assume that the cipher system (GPSN,* $K$*) is semi-injective with respect to* $K$*, i.e., that the knowledge of ciphers and messages uniquely determines* $K$*, i.e., holds* $H(K|X,Z) = 0$.

A system conceived in this way can also be viewed as a kind of wiretap model. Legitimate users Alice and Bob are communicating over an equivalent noiseless main channel $\langle M, X, W_m \rangle = \langle \{0,1\}, \{0,1\}, I \rangle$, where $I$ is identity matrix, see Figure 2. An eavesdropper Eve is wiretapping over an equivalent wiretap channel $\langle M, Z, W_e \rangle = \langle \{0,1\}, \{0,1\}, W_e \rangle$, where $W_e$ is an $|M| \times |Z|$ matrix $W_e(z|m)$ being the probability of receiving $z \in Z$ given that $m \in M$ was sent, see Figure 4. In order to be able to apply the results presented in chapter 2, we have to prove that the equivalent wiretap channel is BSC.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 August 2024 doi:10.20944/preprints202408.0756.v1



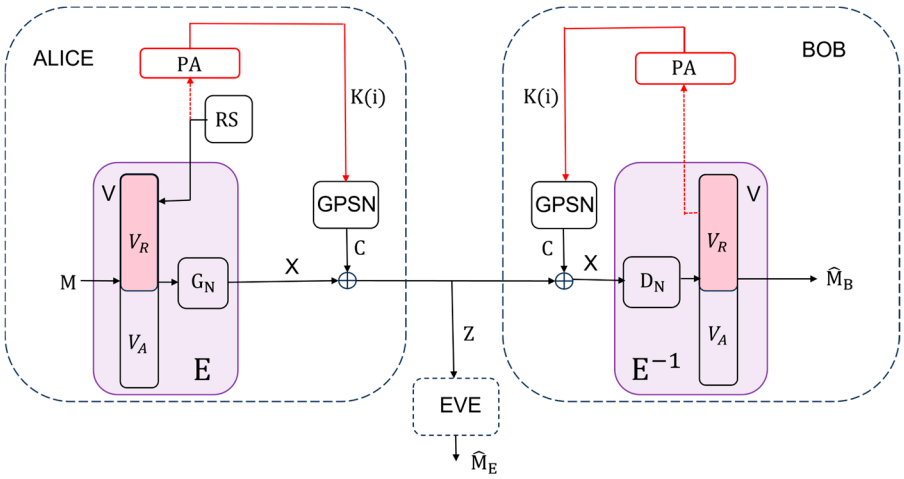


**Figure 2.** ISAR ciphering system for secure communication without explicit secret key exchange.

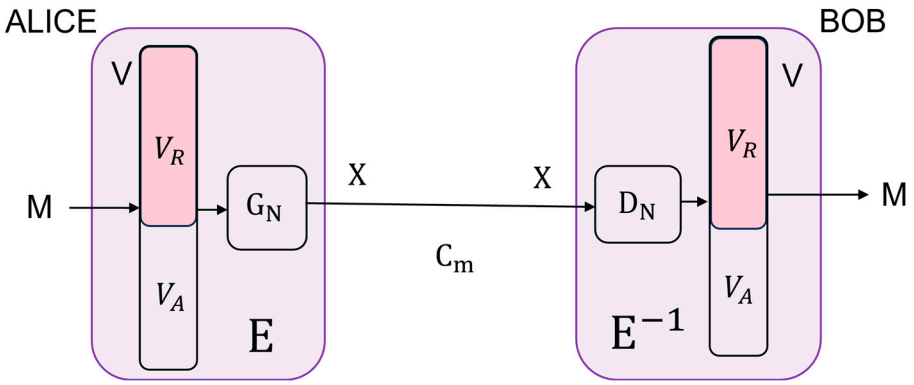


**Figure 3.** Equivalent main channel

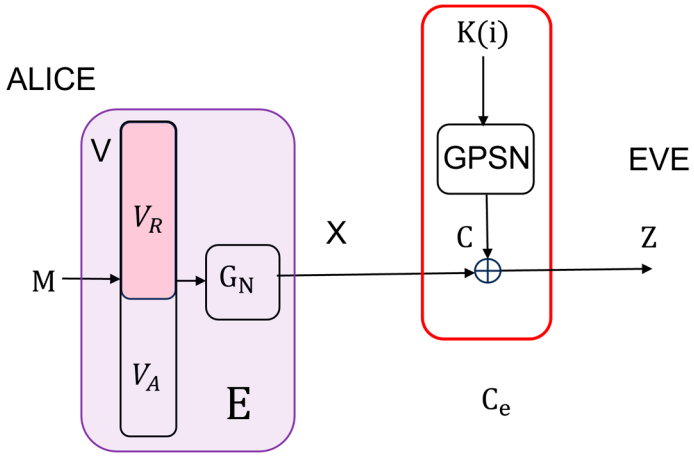


**Figure 4.** Equivalent wiretap channel.

**Lemma 1.** *Equivalent wiretap channel* $\langle M, Z, W_e \rangle$ *is a BSC.*

**Proof.** Equivalent wiretap channel $\langle M, Z, W_e \rangle$ is a BSC.
According to the proposed scheme, see Figure 4,

$$Z = C \oplus X. \tag{29}$$

Since $C$ is independent of $X$, (3.1) je equivalent to BSC ($\sigma$), where crossover probability data is $\sigma = Pr\{C = 1\}$. □

We can now formulate the main result of this part of the work.

 doi:10.20944/preprints202408.0756.v1



**Theorem 2**. *Let $\delta_n$, be an arbitrary security function that satisfies condition (1). Let the index sets $A$ and $R$ be given by $A = P_n(W_e, \delta_n)$, $R = [1,n] \setminus P_n(W_e, \delta_n)$, while $\mathbf{e}$ is random vector, selected by Alice uniformly at random from {0, 1}. If in the ISAR system, the input vector is structured as follows*

$$V_1^n = [V_R, V_A], \quad V_R = e, \quad V_A = M \quad . \tag{30}$$

*then it satisfies both reliability and strong security criteria, precisely*

$$Pr\{M \neq \widehat{M_B}\} = 0, \tag{31}$$

$$\lim_{|M|\to\infty} I(M;Z) = 0. \tag{32}$$

**Proof.** Bearing in mind that the equivalent main channel for ISAR system is noiseless, the wiretap coding scheme for strong secrecy is determined by the sets of polarized channels indices (6) – (13), which now have the following values

$$P_n(W_e, \delta_n) = \left\{ i \in [1,n] : C(W_e^{(i)}) \leq \delta_n \right\} \tag{33}$$

$$G_n(W_m, \beta) = [1,n] \tag{34}$$

$$B_n(W_m, \beta) = \emptyset \tag{35}$$

$$R = [1,n] \setminus P_n(W_e, \delta_n) \tag{36}$$

$$A = P_n(W_e, \delta_n) \cap G_n(W_m, \beta) = P_n(W_e, \delta_n) \tag{37}$$

$$B = P_n(W_e, \delta_n) \setminus G_n(W_m, \beta) = P_n(W_e, \delta_n) \setminus [1,n] = \emptyset \tag{38}$$

$$R1 = R \cap B_n(W_m, \beta) = [1,n] \setminus P_n(W_e, \delta_n) \cap \emptyset = \emptyset \tag{39}$$

$$R2 = R \cap G_n(W_m, \beta) = R \cap [1,n] = R. \tag{40}$$

According to (39) the problematic set of indices $R1$ is an empty set, therefore it is not necessary to apply the chaining scheme. By structuring the input vector $V$ according to (30), where the sets of indices $R$ and $A$ are given by (36) and (37) respectively, we conclude that the proposed polar coding scheme is merely an instantiation of the general polar wiretap coding scheme from section 2.2, which, under the assumptions of Theorem 2, ensures both reliability and strong secrecy. Since the main channel is noiseless, decoding on Bob's side is not performed using the SC decoder, but rather by a simple inverse operation with respect to the encoding, that is

$$XG_n^{-1} = (VG_n)G_n^{-1} = (VG_n)G_n = V \tag{41}$$

given that the Arikan transform matrix $G_n$ is its own inverse over Galois field $GF_2$. Therefore, the reliability condition expressed by (31), is actually deterministically satisfied. This completes the proof. □

**Theorem 3.** *For any security function $\delta_n$, and constants $\beta, c_1, c_2$ that satisfies condition (22) the rate of the coding scheme of proposed system from Figure 2, approaches the secrecy capacity, namely*

$$\lim_{n\to\infty} R_n = \lim_{n\to\infty} \frac{|A|}{n} = 1 - C(W_e). \tag{42}$$

**Proof.** The proof follows directly from Theorem 1 [17], and the fact that the capacity of main channel is $C(W_m) = 1$. □

**Remark 4.** *Eve's optimal strategy is to attempt to decode the wiretapped $Z$, using the Successive Cancellation (SC) decoder, after receiving it [30]. The average block error probability at Eve's side can be lower bounded, applying Lemma 2.9 of [36] by Korada*

 doi:10.20944/preprints202408.0756.v1



$$P_e(A) \geq \max_{i \in A} \frac{1}{2}\left(1 - \sqrt{1 - Z(W_n^i)^2}\right) \tag{43}$$

*where $A$ is information set. Considering that Eve does not know frozen bits on her side, the information set includes all indices, i.e., $A = [1, n]$, and the polar code applied is of rate 1, see the similar argumentation in [37]. The maximum value of $Z(W_n^i)$ is very close to 1, having in mind that $A$ includes bad channels as well. Therefore, according to (43) it follows*

$$P_e(A) \geq \frac{1}{2} - \delta, \qquad \delta = \frac{1}{2}\sqrt{1 - \max_{i \in A}\{Z(W_n^i)^2\}} \tag{44}$$

*where $\delta$ is small. From this, we conclude that Eve's optimal decoding strategy of using the SC decoder results in the maximum decoding error, preventing her from obtaining both the message $M$ and the purely random sequence $e$.*

It is evident that the key properties of the proposed system depend on the capacity of the wiretap channel $C(W_e)$. The following theorem determines the value of this quantity, depending on the length of the polar code $n$ and the length of the secret key $k_C$.

**Theorem 4.** *The capacity of Eve's channel in the proposed system from Figure 2 is given by*

$$C(W_e) = 1 - \frac{k_C}{n} \tag{45}$$

*where $k_C = |K|$ is the length of the secret key of the given symmetric stream cyphering system (GPSN, $K$), while $n$ is the length of the polar code.*

**Proof.** For a discrete memoryless symmetric channel $W_e$ with input $X$ and output $Z$, channel capacity is defined as

$$C(W_e) = \max_{p(x)} I(X; Z), \tag{46}$$

where the maximum is taken over all possible input distributions $p(x)$, [38]. Further, we have

$$I(X_1^n; Z_1^n) = H(X_1^n) - H(X_1^n | Z_1^n). \tag{47}$$

The input to the wiretap channel $W_e$ is simultaneously the input (message) to the symmetric stream ciphering system (GPSN, $K$). Message equivocation is equal to key equivocation

$$H(X_1^n | Z_1^n) = H(K | Z_1^n) \tag{48}$$

for every semi-injective symmetric ciphering system, see Theorem 2, [39]. On the other hand, it well known that (see for example [25])

$$H(K | Z_1^n) = H(K) + H(X_1^n) - H(Z_1^n), \tag{49}$$

which, by substituting into (48), and then into (47) gives

$$I(X_1^n; Z_1^n) = H(Z_1^n) - H(K). \tag{50}$$

One of the primary goals in the designing any cipher system is for ciphertext to appear to be totally random for as long as possible, i.e., that

$$H(Z_1^n) \approx n \tag{51}$$

this holds for larger values of $n$. The assumption (51) is referred to by Massey in [40] as "total randomness" assumption, and it is shown to be valid as long as

$$n \leq n_u \tag{52}$$

where $n_u$ is the unicity distance of given cipher system, [1]. Considering (51), and the fact that $H(K) = k_C$, since secret keys are chosen as purely random sequences, from (50) we obtain $I(X_1^n; \mathrm{Z}_1^n) = n - k_C$, or normalized per bit

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 August 2024 doi:10.20944/preprints202408.0756.v1



$$C(W_e) = \max_{p(x)} I(X;Z) = \frac{1}{n}(n - k_C) = 1 - \frac{k_C}{n}, \tag{53}$$

which had to be proven. □

***Remark 5.*** *Given Theorems 3 and 4, the secrecy capacity of the proposed system is*

$$C_s = 1 - C(W_e) = \frac{k_C}{n}, \tag{54}$$

*from this, the fundamental impact of the secret key length is clearly evident, see Figure 5. For* $k_C = 0$*, system does not provide any security, while for* $k_C = n$*,* $C_s = 1$*. It means that in latter case we can choose all n bits of polar code for secure transmitting of n message bits.*

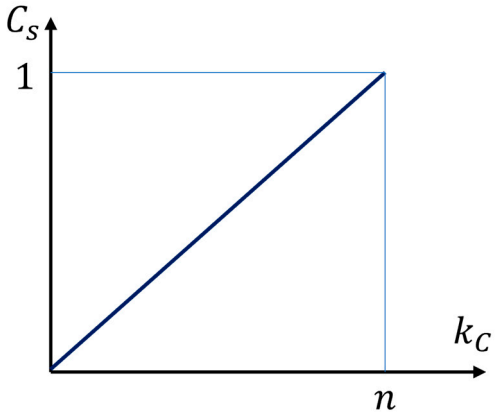


**Figure 5.** The linear dependence of the security capacity of the proposed system on the secret key length.

In order to examine the properties of ideally secret autonomous cipher systems, we need the following theorem which provides a lower bound on the equivocation of the system's secret keys.

**Theorem 5.** *The equivocation of the secret keys* $K$*, GPSN(K), when the ciphertext is known, satisfies the inequality*

$$H(K|Z_1^n) \geq H(K) + H(V_R) - n, \tag{55}$$

*where n is the length of the polar code,* $H(K)$ *is the entropy of the secret key, and* $V_R = e$ *is a purely random vector of dimension* $|V_R| = |R| = |[1,n] \setminus P_n(W_e, \delta_n)|$.

**Proof.** Generally, according to [25], for any cipher system with a secret key $K$, input $X_1^n$, and ciphertext(output) $Z_1^n$, it holds $H(K|Z_1^n) = H(K) + H(X_1^n) - H(Z_1^n)$. Based on (13), the output vector $X_1^n$ can be written in the form

$$X_1^n = V_R G_R \oplus V_A G_A = e_1^{n-k} G_R \oplus m_1^k G_A \tag{56}$$

where $G_R$ and $G_A$ are submatrices of polar code generation matrix $G$, consisting of corresponding rows in $G$.

The ranks of the matrices $G_R$ and $G_A$ are $n-k$ and $k$, respectively, because the generator matrix $G$ has full rank $n$. Therefore $e_1^{n-k} G_R$ and $m_1^k G_A$ can be uniquely represented by a set of basis vectors of dimensions $n-k$ and $k$. These basis vectors are some of the column vectors of the matrices $G_R$ and $G_A$ respectively. We will denote these sets of column indices as $\mathcal{T}_R$ and $\mathcal{T}_A$. Then there exists a one-to-one correspondence between $e_1^{n-k}$ and $X_{\mathcal{T}_R}$, and between $m_1^k$ and $X_{\mathcal{T}_A}$. Hence

$$H(V_R G_R) = H(V_R), \tag{57}$$

$$H(V_A G_A) = H(V_A). \tag{58}$$

 doi:10.20944/preprints202408.0756.v1



Based on the data processing properties of entropy [38] , we conclude that in the PA procedure there is a limitation, $nC_s = k_c \leq H(V_R) = n - k = n - nC_s$, from which it follows that the secrecy capacity in the ISAR system must be $C_s < \frac{1}{2}$. Thus, it holds that

$$H(X_1^n) = V_R G_R \oplus V_A G_A \geq max\{H(V_R), H(V_A)\} = H(V_R) \tag{59}$$

given the condition $C_s < \frac{1}{2}$ and the fact that $V_R$ is purely random vector with maximum entropy equal to $H(V_R) = |V_R| = n - k$. Since it is always true that

$$H(Z_1^n) \leq n, \tag{60}$$

we finally get

$$H(K|Z_1^n) = H(K) + H(X_1^n) - H(Z_1^n) \geq H(K) + H(V_R) - n \tag{61}$$

which had to be proven. □

To ensure the system possesses autonomy in generating secret keys, we must restrict the secrecy capacity, as some of the bits of the polar code need to be used for generating and distributing secret keys. Additionally, for the system to be ideal and autonomous, a further reduction in secrecy capacity is necessary to maintain the desired minimum level of key equivocation. These facts are summarized in Theorem 6 below.

**Theorem 6.** *The proposed system is ideal and autonomous, with secrecy capacity*

$$C_s = \frac{1}{2} - \frac{1}{2n}(\Delta_K + \Delta R), \tag{62}$$

*where* $\Delta_K \in \left(0, \frac{n}{2}\right)$ *is a given minimum value of key equivocation, and* $\Delta R$ *is the compression rate of the applied class of universal hash functions.*

**Proof.** To prove that the proposed system is ideal, it is necessary to show that $H(K|Z) > 0$, kada $|Z| \to \infty$. Since $|Z| = |M|$, this is equivalent to the condition $H(K|Z) > 0$, as $M \to \infty$. Consider the general case where Alice sends Bob a message $M$ of arbitrary length $|M| = n_M$. The message $M$ will be divided into blocks of length $k = |\mathrm{A}| = |P_n(W_e, \delta_n)|$. The total number of blocks will be $t = \left\lceil \frac{n_M}{n} \right\rceil$, where $n$ is the length of the polar code, so $M = [m_1|m_2|\ldots|m_t]$. The last block, if it is not of length $k$, can be padded with arbitrary content. According to the proposed coding scheme that provides strong security, $n - k = |R| = n - |A| = n - |P_n(W_e, \delta_n)|$. Based on (41), in addition to the transmitted message $m_i$, Bob decodes without error a purely random vector $e_i$, of length $n - k$, which was written on the indices from the set $R$ given in (36) during the encoding process on Alice's side. Thus, Alice and Bob, after decoding, possess identical random sequences $\{e_1, e_2, \ldots, e_t\}$, in each of the $t$ transmitted blocks of the polar code. Eve's optimal strategy is to decode her received signal at the output of the wiretap channel using the SC decoder. Eve knows the parameters of the applied polar code, such as the length $n$, and the index sets A and R, but does not know the values of the bits at those positions. This situation is equivalent to the SC decoder operating in the mode of unknown so-called frozen bits. As noted in Remark 4, Eve's error in optimal decoding is close to the maximum, so it can be said that Eve's information about the sequence $\{e_1, e_2, \ldots, e_t\}$ is close to zero. By applying PA to the common random sequences $\{e_i\}$, Alice and Bob can further reduce Eve's residual information about these sequences. In each current block, a secret key of length $k_C$ is generated for the next block using some chosen PA algorithm

$$K_i = PA(e_{i-1}),\ e_0 = S,\ |K_i| = k_C,\ \ i = 1,2,\ldots,t. \tag{63}$$

Given that $\Delta R$ is the compression rate of the class of universal hash functions applied in the PA process, it follows that

$$|K_i| = |e_{i-1}| - \Delta R\ ,\qquad i = 1,2,\ldots,t. \tag{64}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 August 2024 doi:10.20944/preprints202408.0756.v1



Considering Theorem 2 and (64), and assuming the ergodicity of the local randomness source, i.e., $H(e_{i-1}) = n - k,\ i = 1,2,\dots,t$, it can be obtained from (53)

$$k_C = n - k - \Delta R = n - nC_s - \Delta R, \qquad i = 1,2,\dots,t. \tag{65}$$

For the system to be ideal, it is sufficient to ensure that the right-hand side of (61) equals a predetermined minimum value of key equivocation $\Delta_K > 0$, in each block, i.e.

$$H(K_i) + H(V_{R\,i}) - n = \Delta_K, \quad i = 1,2,\dots,t. \tag{66}$$

that is

$$k_C + n - k - n = k_C - k = k_C - nC_s = \Delta_K. \tag{67}$$

By substituting (65) into (67) it is obtained that

$$C_s = \frac{1}{2} - \frac{1}{2n}(\Delta_K + \Delta R) \tag{68}$$

which had to be proven. □

In Figure 6 and Figure 7, the dependence of the secrecy capacity $C_s$ on the normalized minimum value of key equivocation $\frac{\Delta_K}{n}$ and the normalized compression rate $\frac{\Delta R}{n}$ of the hash function applied in the PA process is shown.

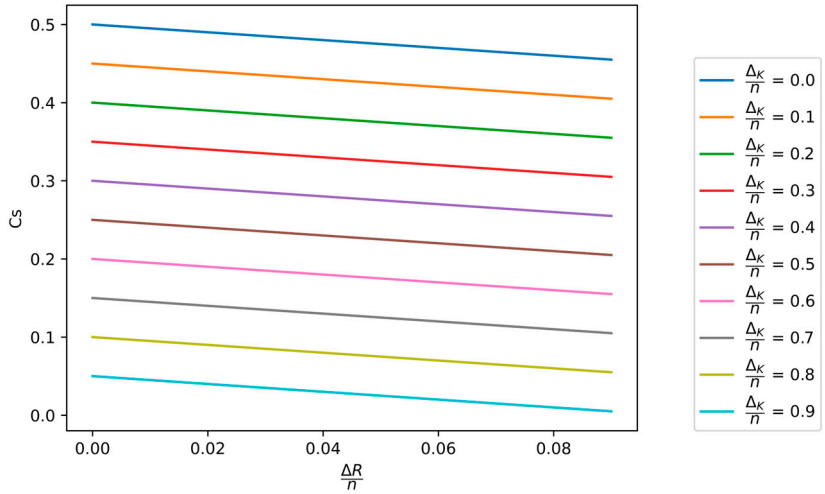


**Figure 6.** The dependence of the secrecy capacity $C_s$ on the normalized compression rate $\frac{\Delta R}{n}$ of the hash function applied in the PA process, for different values of the normalized minimum key equivocation $\frac{\Delta_K}{n}$.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 August 2024 doi:10.20944/preprints202408.0756.v1



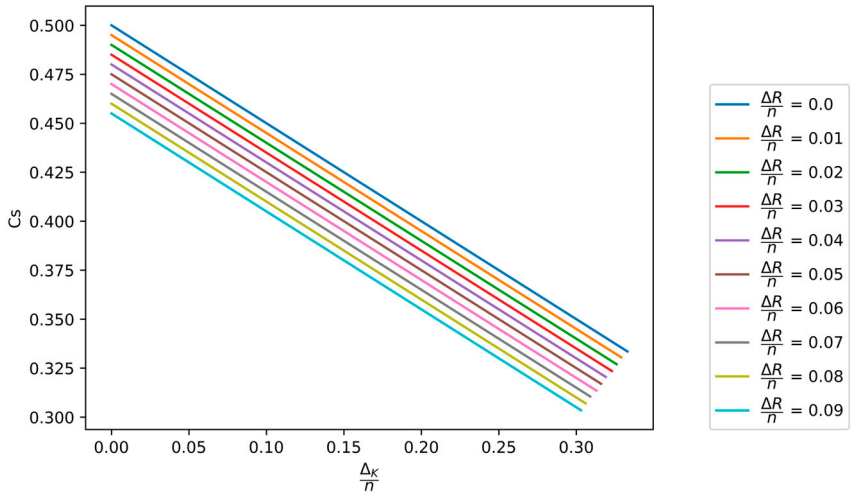


**Figure 7.** The dependence of the secrecy capacity $C_s$ on the normalized minimum key equivocation $\frac{\Delta_K}{n}$, for different values of the normalized compression rate $\frac{\Delta R}{n}$ of the hash function applied in the PA process.

Note that if the security margins $\Delta_K$ and $\Delta R$ are fixed constants, then based on Theorem 6, we conclude that $\lim_{n\to\infty} C_s = \frac{1}{2}$ holds, which is the maximum possible value of the secret capacity of ISAR system.

The presented theoretical offer an effective methodology for synthesizing this class of systems with predetermined security margins. Table 2 outlines the algorithms for configuring the ISAR system and for enciphering/deciphering messages of arbitrary length. Based on the initial values of parameters $\Delta_K$ and $\Delta R$, the secrecy capacity $C_s$ and the length of the secret key $k_C$ are determined. If these parameters are acceptable, the synthesis of the corresponding wiretap polar code proceeds. Identifying the index sets $A = P_n(W_e, \delta_n)$ and $R = [1, n] \setminus P_n(W_e, \delta_n)$, requires the polarization of the wiretap channel. For practical purposes, it is simpler to replace this channel with an equivalent $BSC(\varepsilon)$ channel of the same capacity. Consequently, the crossover probability must satisfy the condition

$$\varepsilon = h_b^{-1}\left(\frac{k_C}{n}\right) \tag{69}$$

where

$$h_b(a) = -a\,log\,a - (1-a)\,log\,1-a), \quad 0 < a < 1, \;\; h_b(0) = h_b(1) = 0 \tag{70}$$

is the binary entropy function, and $h_b^{-1}$ is its inversion. The polarization procedure under SC decoding can be performed using various methods, such as Monte-Carlo [30], Density Evolution (DE) [41,42], DE with Gaussian Approximation (GA) [43], among others. In the ISAR system, where the wiretap channel is designed by the system developer, all essential parameters for the wiretap channel are much more accessible and accurately estimated compared to a scenarios involving real wiretap channels. For instance, when determining the Bhattacharyya parameters $Z(W_e^{(i)})$ using the Monte-Carlo method for polarization, generating an ensemble of wiretap channel output samples is manageable. This process is facilitated by the precise knowledge of the crossover probability (69), allowing for accurate assessment of the results.

 doi:10.20944/preprints202408.0756.v1



**Table 2.** ISAR system setup and enciphering and deciphering algorithm

**System setup**

- Choose $n, \Delta K, \Delta R, \delta_n$.
- Calculate $C_s$ and $k_C = n \cdot C_s$ based on (3.34).
- Form the corresponding equivalent $BSC(\varepsilon)$ wiretap channel,

  $\varepsilon = h_b^{-1}\left(\frac{k_C}{n}\right)$ and perform polarization.
- Determine the set of indices $P_n(W_e, \delta_n) = \{i \epsilon [1, n], C(W_e(i)) \leq \delta_n\}$
- Determine the set of indices $A = P_n(W_e, \delta_n)$ and $R = [1: n] \setminus P_n(W_e, \delta_n)$
- Agree on initial secret key $S$, and seed of hash function $K_h$

**Enciphering**

- Split message $M$ into $t$ blocks so that $t = \left\lceil \frac{|M|}{N} \right\rceil$, $M = [m_1|m_2 \ldots |m_n]$ with padding of the last block with random bits if necessary
- Creation of ciphertext, $z_i$

  For $i = 1, \ldots, t$

  $e_i$ // generate random sequence of length $|e_i| = |A|$

  If $i = 1$ then

  $K = S$

  else

  $K_i = h_M(e_i - 1, K_h)$

  $c_i = \text{GPSN}(K_i, n)$ // generate streem ciphering sequence

  $V_R = e_i$

  $V_A = m_i$

  $x_i = [V_R, V_A] \cdot G_n$ // polar coding

  $z_i = c_i \oplus x_i$ // encription

**Deciphering**

- $M_d = \{ 0 \}$ // empty set at the beginning of the algorithm
- For $i = 1, \ldots, t$

  If $i = 1$ then

  $K = S$

  else

  $K_i = h_M(e_i - 1, K_h)$

  $c_i = \text{GPSN}(K_i, n)$

  $[V_R, V_A] = x_i \cdot G_n$ // deciphering

  $e_i = V_R$ // deciphered random sequence

  $m_i = V_A$ // deciphered message

  $M_d = [M_d|m_i]$ // completly deciphered message

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 August 2024 doi:10.20944/preprints202408.0756.v1



**Remark 6.** *Let's compare the proposed ISAR system with a classic ideal cipher system that has the same minimal key equivocation value for the same length of observed ciphertext. To ensure a fair comparison, we will assume that the entropy of the input messages is identical in both systems. After encrypting blocks of messages, and assuming the blocks are independent, the ISAR system has a total equivocation of all applied keys equal to*

$$H(K_1, \dots, K_t|Z_1, \dots, Z_t) = \sum_{i=1}^{t} H(K_i|Z_i) = tH(K) + tH(V_R) - nt \geq t\Delta_k \ , \tag{71}$$

*since the formation of each block in the polar code is independent of the previously formed blocks, a classic ideal cipher system with secret key* $K_{class}$ *under the same conditions and for the same length of ciphertext, has an equivocation given by*

$$H(K_{class}|Z_1^{nt}) = H(K_{class}) + tH(V_R) - nt \geq \ t\Delta_k. \tag{72}$$

*Equating (71) and (72) yields the condition that*

$$H(K_{class}) = \ tH(K), \tag{73}$$

*that is,*

$$|K_{class}| = t|K|, \tag{74}$$

*assumed that all secret keys in both systems have maximum entropy.*

To ensure that a classic encryption system achieves the same lower bound on key equivocation as the ISAR system for a given ciphertext length, it would need a secret key t-times longer than the key length of the GPSN($K$), see Figure 8. Additionally, $K_{class}$ must be pre-distributed to legitimate parties. The length of this key scales linearly with the message length $|M|$, specifically with $t = \left\lceil \frac{|M|}{n} \right\rceil$. In contrast, the proposed ideal autonomous system requires only a fixed-length secret key K, independent of the message length, and necessitates only an initial exchange of the key values $K_1 = S$. Therefore, the benefits of the ISAR system become more significant as the length of the messages increases.

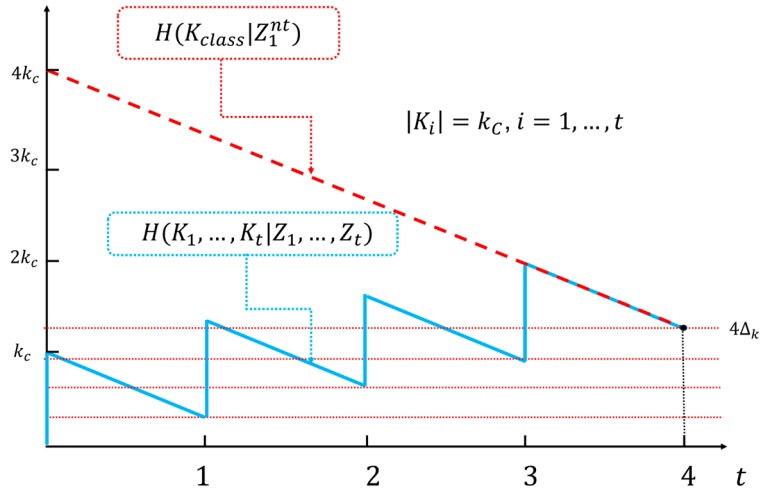


**Figure 8.** An example of the relationship between the key equivocation of the ISAR system (blue line) and the classic cipher system (red line) as a function of the number of encrypted blocks $t$, for a given minimum value $\Delta_K$, provided that after 4 blocks both systems have the same minimum key equivocation value. The identical slopes of both equivocations are a consequence of the same probabilistic properties of the source being encrypted. Note that in that case, the classic system must have a secret key 4 times longer than the ISAR System.

**Remark 7.** *The ISAR system can be obtained by a suitable transformation of any given stream ciphering system GPSN(K) with a known secret key length* $|K| = k_c$. *The procedure includes the following steps:*

 doi:10.20944/preprints202408.0756.v1



- *Based on the expression (62) for the given value of the length of the secret key $k_C$, we can directly obtain the required length of the polar code, i.e.,*

$$n = 2k_C + \Delta_K + \Delta R. \tag{75}$$

*Since the length of the polar code must be a power of 2, it is necessary to correct (75) to the value*

$$\tilde{n} = 2^{\lceil log_2 n \rceil}. \tag{76}$$

- *The corrected value of the length of the polar code block (76) allows the eventual correction of the total security margin to the new value*

$$\tilde{\Delta}_K + \tilde{\Delta}R = \tilde{n} - 2k_C \tag{77}$$

*bearing in mind the limitation*

$$\tilde{\Delta}_K \leq k_C. \tag{78}$$

- *Since all elements for ISAR system setup are available in this step, i.e., $\tilde{n}, \tilde{\Delta}_K, \tilde{\Delta}R, \delta_{\tilde{n}}$, further operation of the system takes place according to the algorithm from Table 2.*

## 4. Security Analysis of ISAR Ciphering System

In the following analysis, we will focus on the examination of passive attacks on the system. This implies that the attacker has access to the output of the wiretap channel and all information related to the architecture and operation of the system, except for the initial value $K_1 = S$ and the seed of the chosen hash function $K_h$. Based on the primary objective of the attack, we can distinguish three types: message attacks, attacks on locally generated random sequences, and attacks on the secret key.

### *4.1. Message Attacks*

In this type of attack, Eve attempts to obtain the message $m_i$ in each block based on the observation of the wiretap channel output $Z_i$.

According to Theorem 2, the proposed system satisfies the strong security criterion (32), which means that asymptotically, with an increase in the codeword length $n$, Eve can only acquire a negligible amount of information about the messages $m_i$.

### *4.2. Attacks on Locally Generated Random Sequences*

In this type of attack, Eve first tries to obtain one of the random sequences $\{e_1, e_2, \dots, e_t\}$ locally generated by Alice based on the observation of the wiretap channel outputs $\{Z_1, \dots, Z_t\}$. In the second step, using some obtained value $e_{i^*}, i^* \in [1, t-1]$, Alice could potentially generate the correct secret key for the next block $i^* + 1$. Knowing the secret key for block $i^* + 1$, Eve could successfully decode all subsequent blocks, thereby obtaining all messages in the sequence $\{m_{i^*+1}, m_{i^*+2}, \dots, m_t\}$. However, this scenario is not feasible. The first step cannot be realized since, according to Remark 3.2, the average block error probability on Eve's side is close to maximal (see (44)). The second step is not possible without knowing the seed of the chosen hash function $K_h$. As the success of this attack depends on the success of both steps, this type of attack is practically unfeasible because the overall difficulty of the attack is equal to the product of the difficulties of both steps.

### *4.3. Attacks on the Secret Key GPSN(K)*

A far more powerful cryptanalytic attack than an attack on individual messages is an attack on the secret key of the GPSN. If Eve were to obtain the secret key $K_{i^*}$ in block $i^*$ where $i^* \in [1, t-1]$, she would be able to access message $m_{i^*}$, the random sequence $e_{i^*}$, and consequently the secret key $K_{i^*+1}$ for the next block. This would allow her to repeat the same process for each subsequent block, thereby gaining access to all subsequent transmitted messages $\{m_{i^*+1}, m_{i^*+2}, \dots, m_t\}$, random sequenced $\{e_{i^*+1}, e_{i^*+2}, \dots, e_t\}$ and generated secret keys $\{K_{i^*+1}, K_{i^*+2}, \dots, K_t\}$. From the system designer's perspective, it is crucial to prevent this scenario. The information-theoretic quantity that quantifies the likelihood of such an attack is the equivocation of the secret keys for a given wiretap

 doi:10.20944/preprints202408.0756.v1



channel output. As the ISAR system is ideally secure, its secret key equivocation, according to (66), never falls below the value $\Delta_k$, which is the security parameter in the synthesis process of ISAR. Therefore, the system designer can render this attack unsuccessful with any chosen margin of security. It is important to note that increasing the margin of security leads to a reduction in the secrecy capacity of ISAR (see Theorem 6 and Figure 7).

**Example** To understand the order of magnitude of the difficulty of executing this attack, let us consider a typical example of ISAR with parameters (expressed in bits) $n = 4096$ , $\Delta_k = 100$, $\Delta R = 40$. According to (62), $C_s = 0.4829$, and the key length is $k_c = nC_s = 1978$ bits. Let's assume that Eve's optimal strategy for each block of the polar wiretap code reaches the lower bound of equivocation of 100 bits of the key. This means that Eve cannot resolve this uncertainty in any way. Recall that in a brute force attack examining all $2^{100}$ possible keys, all decrypted messages would be equally likely and valid as potential final solutions.

The above example shows that for the polar code length n of order $2^b$, the key length of GPSN($K$) is of order $2^{b-1}$. These lengths are not common for commercial stream ciphering systems. However, this does not mean that such systems cannot be relatively easily upgraded to the required key lengths. As an example, we cite a generic model of a pseudo-random generator described in [44]. If the address and selection sequences are chosen so that they are the outputs of two multiple linear shift registers, then the lengths of the equivalent secret key of this pseudo-random generator can easily be set in ranges of order $2^{b-1}$.

## 5. Practical Aspects

### *5.1. Complexity of ISAR*

Computational complexity of the ISAR system includes the total complexity of polar coding and decoding, as well as the complexity of the PA block.

Polar codes are attractive in practice due to their relatively low complexity compared to other coding schemes, especially when considering the powerful error-correcting capabilities they offer. The complexity of encoding a polar code of block size n is $O(n \log n)$. This efficient complexity is due to the structured way in which the polar transform combines inputs, leveraging the recursive nature of the polar code construction, and can be performed using the Fast Fourier Transform (FFT) approach.

In general, decoding complexity depends on the applied decoding algorithm. However, in the ISAR system decoding is a deterministic procedure identical to encoding since Arikan's generator matrix is involutory. Therefore, the complexity of decoding is identical to the complexity of encoding, i.e., $O(n \log n)$, [30].

The PA block was implemented using random binary matrices of dimension $n \times k_c$, which have the Toeplitz structure. It is known that this class of hash functions belongs to the universal family of hash functions, see [45,46]. The Toeplitz hash functions are particularly suitable for typical values for $n$ in polar coding (order of $2^{10} - 2^{16}$). Namely, the PA block can be efficiently implemented with a complexity of $O(n \log n)$, instead of $O(n^2)$, in the case of hash functions in the form of random binary matrices without the Toeplitz structure.

Consequently, we can say that the complexity of the ISAR system, without taking into account the complexity of GPSN($K$), is of the order of $O(n \log n)$, which indicates that the implementation of the proposed algorithm does not require special memory and computational resources. This makes it very attractive for practical implementation and use.

### *5.2. Integration of ISAR Ciphering System in Contemporary Information and Communication Infrastructure*

When applying ISAR in modern information and communication infrastructure, you should keep in mind the two most important advantages of this system:

- Guaranteed security margins in terms of key equivocation (ideality)
- Independence of the length of the keys from the length of the messages.

 doi:10.20944/preprints202408.0756.v1



The ideal position of the ISAR system is within the framework of permanent point-to-point protection of large-capacity information flows. That's when ideality and the independence of the length of the keys from the length of the messages come into play. As already mentioned, the start of communication requires the exchange of the secret key of GPSN(K) and the seed of the applied crypto hash function. It is interesting to note that disruption of protected communication or its regular termination does not require additional exchange of these values for subsequent communication. Namely, the decoded random sequences of the last block of polar code can be memorized and used to generate the initial key for subsequent communication. This raised the autonomy of the ISAR system to an even higher level.

## 6. Conclusions

In this paper, we introduced a class of symmetric ciphering systems termed ISAR (Ideally Secret Autonomous Robust) systems. These systems are designed to provide ideal secrecy, autonomy in key generation and distribution, and robustness to the probabilistic structure of messages. These systems address the longstanding challenge in cryptography of balancing security with practicality, particularly by eliminating the need for lengthy secret keys and additional key distribution infrastructure. By ensuring a predetermined minimum value of key equivocation and continuous key refreshing, ISAR systems provide robust security against passive attacks on both keys and messages.

Our work demonstrates that ISAR systems can be applied to any existing symmetric stream ciphering system without requiring changes to the original encryption algorithm, thus offering a versatile and efficient solution for enhancing the security of current cryptographic practices. This transformation greatly supports privacy, a critical requirement for modern security systems.

Overall, the ISAR system represents an advancement in cryptographic security, offering an efficient methodology for creating ciphering systems with predetermined security margins. Future research directions include exploring further optimizations in the ISAR architecture, extending its applicability to a broader range of cryptographic scenarios, and investigating additional techniques for enhancing its resistance to more sophisticated attack vectors. The continued development of ISAR systems holds the promise of ensuring stronger privacy and security measures in an increasingly digital world.

**Author Contributions:** Conceptualization, M.M.; methodology, M.M., T.U., J.R.; software, J.R.; validation, M.M., J.R., B.B.; formal analysis, M.M., J.R., T.U.; investigation, M.M., J.R.; resources, B.B.; data curation, J.R., B.B.; writing—original draft preparation, M.M., J.R.; writing—review and editing, M.M., J.R.; visualization, M.M., J.R.; supervision, M.M., B.B.; project administration, J.R.; funding acquisition, B.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research is funded by the Vlatacom Institute of High Technologies under project #164 EEG_Keys.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Shannon, C.E. Communication theory of secrecy systems. *Bell Syst. Tech. J.* **1949**, *28*, 656–715.
2. Gunther, C. G. A universal algorithm for homophonic coding. In *Workshop on the Theory and Application of Cryptographic Techniques*, Springer: Berlin, Heidelberg, 1988, pp. 405-414.
3. Massey, J. Some Applications of Source Coding in Cryptography. *European Transactions on Telecommunications* **1994**, *5*, 421-429.
4. Ryabko, B. Unconditionally secure short key ciphers based on data compression and randomization. *Des. Codes Cryptogr.* **2023**, *91*, 2201–2212.
5. Ryabko, B. Ya. A simply realizable ideal cryptographic system. *Problems of Information Transmission* **2000**, *36*, 84-89.
6. Oggier, F.; Mihaljević, M.J. An information-theoretic security evaluation of a class of randomized encryption schemes. *IEEE Trans. Inf. Forensics Sec.* **2014**, *9*, 158–168.
7. Ryabko, B.; Fionov, A. Efficient homophonic coding. *IEEE Trans. Inform. Theory* **1999**, *45*, 2083–2091.
8. Agrikola, T.; Couteau, G.; Ishai, Y.; Jarecki, S.; Sahai, A. On pseudorandom encodings. In *Theory of Cryptography Conference*, Springer, Cham, 2020, pp. 639-669.

9. Russell, A.; Wang, H. How to fool an unbounded adversary with a short key. *IEEE Trans. Inf. Theory* **2006**, *52*, 1130–1140.
10. Dodis, Y.; Smith, A. Entropic security and the encryption of high entropy messages. In: *Theory of Cryptography Conference*, Springer, Berlin, Heidelberg, 2005, pp. 556-577.
11. Ryabko, B. Unconditionally Secure Ciphers with a Short Key for a Source with Unknown Statistics. *Entropy* **2023**, *25*, 1406.
12. Juels, A.; Ristenpart, T. Honey encryption: Security beyond the brute-force bound. In *Advances in Cryptology-EUROCRYPT 2014*, Springer, Berlin, Heidelberg, 2014, pp. 293-310.
13. Li, X.; Tang, Q.; Zhang, Z. Fooling an Unbounded Adversary with a Short Key, Repeatedly: The Honey Encryption Perspective. In *2nd Conference on Information-Theoretic Cryptography*, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021, pp. 23:1-23:21.
14. Wyner, A.D. The wire-tap channel. *Bell System Tech. J.* **1975**, *54*, 1355–1387.
15. Bellare, M.; Tessaro, S.; Vardy, A. Semantic security for the wiretap channel. In International cryptology conference (CRYPTO). Springer, 2012, pp. 294–311.
16. Harrison, W.K.; Almeida, J.; Bloch, M.R.; McLaughlin, S.W.; Barros, J. Coding for secrecy: An overview of error-control coding techniques for physical-layer security. *IEEE Signal Processing Magazine* **2013,** *30*, 41-50.
17. Mahdavifar, H.; Vardy, A. Achieving the secrecy capacity of wiretap channels using polar codes. *IEEE Transactions on Information Theory* **2011**, *57*, 6428–6443.
18. Sasoglu, E.; Vardy, A. A new polar coding scheme for strong security on wiretap channels. In Information Theory Proceedings (ISIT), Istanbul, Turkey 2013, pp. 1117–1121.
19. Bloch, M.; Barros, J. *Physical-layer security: From information theory to security engineering*; Cambridge University Press: Cambridge, UK, 2011.
20. A Rajagopalan, A Thangaraj, S Agrawal, Wiretap polar codes in encryption schemes based on learning with errors problem. 2018 IEEE International Symposium on Information Theory (ISIT), Vail, CO, USA 2018, 1146-1150.
21. Liu, S.; Hong, Y.; Viterbo, E. Unshared secret key cryptography: Achieving Shannon's ideal secrecy and perfect secrecy. 2014 IEEE Information Theory Workshop (ITW 2014), Hobart, TAS, Australia, 2014, 636-640.
22. Liu, S.; et al. Unshared Secret Key Cryptography. *IEEE Transactions on Wireless Communications* **2014**, *13*, 6670-6683.
23. Wang, H.; Tao, X.; Li N.; Han, Z. Polar Coding for the Wiretap Channel with Shared Key. *IEEE Transactions on Information Forensics and Security* **2018**, *13*, 1351-1360.
24. Zhao, Y.; Xu, S.; Chi, H. Encrypted Secure Polar Coding Scheme for General Two-way Wiretap Channel. *IET Information Security* **2019,** *13*, 393-403.
25. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
26. Maurer, U.M. The strong secret key rate of discrete random triples. In *Communication and Cryptography — Two Sides of One Tapestry*, Blahut R.E. et al. (Eds.), Boston: Kluwer Academic, 1994; Volume 276, pp.271-285.
27. Maurer, U.M.; Wolf, S. Information-theoretic key agreement: From weak to strong secrecy for free. *Lect. Notes Computer Science* **2000,** *1807*, 351–368.
28. Leung-Yan-Cheong, S. On a special class of wire-tap channels. *IEEE Trans. Inform. Theory* **1977**, *23*, 625–627.
29. Arikan, E. A performance comparison of polar codes and Reed-Muller codes. *IEEE Comm. Letters* **2008,** *12*, pp. 447-449.
30. Arikan, E. Channel polarization: A method for constructing capacity achieving codes for symmetric binary-input memoryless channels. *IEEE Trans. Inform. Theory* **2009,** *55*, pp. 3051-3073.
31. Bennett, C.H.; Brassard, G.; Crepeau, C.; Maurer, U.M. Generalized privacy amplification. *IEEE Trans. Inf. Theory* **1995**, *41*, 1915–1923.
32. Carter, J. L.; Wegman, M. N. Universal classes of hash functions. *Journal of Computer and System Sciences* **1979**, *18*, 143–154.
33. Galis, M.; Milosavljević, M.; Jevremović, A.; Banjac, Z.; Makarov, A.; Radomirović, J. Secret-key agreement by asynchronous EEG over authenticated public channels. *Entropy* **2021**, *23*, 1327.
34. Radomirović, J.; Milosavljević, M.; Kovačević, B.; Jovanović, M. Privacy amplification strategies in sequential secret key distillation protocols based on machine learning. *Symmetry* **2022**, *14*, 2028.
35. Radomirović, J.; Milosavljević, M.; Banjac, Z.; Jovanović, M. Secret Key Distillation with Speech Input and Deep Neural Network-Controlled Privacy Amplification. *Mathematics* **2023**, *11*, 1524.
36. Korada, S. B. Polar Codes for Channel and Source Coding, Ph.D. Thesis, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland, 2009.
37. Kim, Y.S.; Kim, J.H.; Kim, S.H. A Secure Information Transmission Scheme With a Secret Key Based on Polar Coding. *IEEE Communications Letters* **2014**, *18(6)*, 937-940.
38. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed., NJ: John Wiley & Sons, Hoboken, 2006.
39. Biondi, F.; Given-Wilson, T.; Legay, A. Attainable unconditional security for shared-key cryptosystems. *Information Sciences* **2016**, *369*, 80-99.

 doi:10.20944/preprints202408.0756.v1




40. Massey, J. L. Applied Digital Information Theory II, Lecture notes. Available online: https://www.isiweb.ee.ethz.ch/archive/massey_scr/ (accessed on 31 July 2024).
41. Tal, I.; Vardy, A. How to construct polar codes. *IEEE Transactions on Information Theory* **2013**, *59*, 6562–6582.
42. Mori, R.; Tanaka, T. Performance of polar codes with the construction using density evolution. *IEEE Communications Letters* **2009** *13*, 519–521.
43. Trifonov, P.; Efficient design and decoding of polar codes. *IEEE transactions on communications* **2012**, *60*, 3221–3227.
44. Unkašević, T.; Banjac, Z.; Milosavljević, M. A Generic Model of the Pseudo-Random Generator Based on Permutations Suitable for Security Solutions in Computationally-Constrained Environments. *Sensors* **2019**, *19*, 5322.
45. Tsurumaru, T.; Hayashi, M. Dual universality of hash functions and its applications to quantum cryptography. *IEEE Trans. Inf. Theory* **2013**, *59*, 4700–4717.
46. Hayashi, M.; Tsurumaru, T. More efficient privacy amplification with less random seeds via dual universal hash function. *IEEE Trans. Inf. Theory* **2016**, *62*, 2213–2232.