entropy

MDPI

Article

Strong Secrecy on a Class of Degraded Broadcast Channels Using Polar Codes

Jaume del Olmo Alos* **and Javier Rodríguez Fonollosa**

Departament de Teoria del Senyal i Communications (TSC), Universitat Politècnica de Catalunya, Barcelona 08034, Spain; javier.fonollosa@upc.edu

* Correspondence: jaume.del.olmo@upc.edu

Received: 15 May 2018; Accepted: 12 June 2018; Published: 15 June 2018

Abstract: Asymptotic secrecy-capacity achieving polar coding schemes are proposed for the memoryless degraded broadcast channel under different reliability and secrecy requirements: layered decoding or layered secrecy. In these settings, the transmitter wishes to send multiple messages to a set of legitimate receivers keeping them masked from a set of eavesdroppers. The layered decoding structure requires receivers with better channel quality to reliably decode more messages, while the layered secrecy structure requires eavesdroppers with worse channel quality to be kept ignorant of more messages. Practical constructions for the proposed polar coding schemes are discussed and their performance evaluated by means of simulations.

Keywords: polar codes; information-theoretic security; degraded broadcast channels; strong secrecy

1. Introduction

Information-theoretic security over noisy channels was introduced by Wyner in [1], which characterized the (secrecy-)capacity of the degraded wiretap channel. Later, Csiszár and Körner in [2] generalized Wyner's results to the general wiretap channel. In these settings, one transmitter wishes to reliably send one message to a legitimate receiver, while keeping it secret from an eavesdropper, where secrecy is defined based on a condition on some information-theoretic measure that is fully quantifiable. One of these measures is the information leakage, defined as the mutual information $I(W;Z^n)$ between a uniformly-distributed random message W and the channel observations Z^n at the eavesdropper, n being the number of uses of the channel. Based on this measure, the most common secrecy conditions required to be satisfied by channel codes are the weak secrecy, which requires $\lim_{n\to\infty} \frac{1}{n} I(W;Z^n) = 0$, and the strong secrecy, requiring $\lim_{n\to\infty} I(W;Z^n) = 0$. Although the second notion of security is stronger, surprisingly, both secrecy conditions result in the same secrecy-capacity region [3].

In the last decade, information-theoretic security has been extended to a large variety of contexts, and this paper focuses on two different classes of discrete memoryless Degraded Broadcast Channels (DBC) surveyed in [4]: (a) with Non-Layered Decoding and Layered Secrecy (DBC-NLD-LS) and (b) with Layered Decoding and Non-Layered Secrecy (DBC-LD-NLS). In these models, the transmitter wishes to send a set of messages through the DBC, and each message must be reliably decoded by a particular set of receivers and kept masked from a particular set of eavesdroppers. The degradedness condition of the channel implies that individual channels can be ordered based on the quality of their received signals. The layered decoding structure requires receivers with better channel quality to reliably decode more messages, while the layered secrecy requires eavesdroppers with worse channel quality to be kept ignorant of more messages.

The capacity region of these models was first characterized in [4–6]. However, the achievable schemes used by these works rely on random coding arguments that are nonconstructive in practice.

In this sense, the purpose of this paper is to provide coding schemes based on polar codes, which were originally proposed by Arikan [7] to achieve the capacity of binary-input, symmetric, point-to-point channels under Successive Cancellation (SC) decoding. Capacity achieving polar codes for the binary symmetric degraded wiretap channel were introduced in [8,9], satisfying the weak and the strong secrecy condition, respectively. Recently, polar coding has been extended to the general wiretap channel in [10–13]. Indeed, [12,13] generalize their results providing polar coding schemes for the broadcast channel with confidential messages, and [11] also proposes polar coding strategies to achieve the best-known inner bounds on the secrecy-capacity region of some multi-user settings.

Although recent literature has proven the existence of different secrecy-capacity achieving polar coding schemes for multi-user scenarios (for instance, see [11–18]), polar codes for the two models on which this paper is focused have, as far as we know, not been analyzed yet. As mentioned in [4], these settings capture practical scenarios in wireless systems, in which channels can be ordered based on the quality of the received signals (for example, Gaussian channels are degraded). Hence, the ultimate goal of this work is not only to prove the existence of two asymptotic secrecy-capacity achieving polar coding schemes for these models under the strong secrecy condition, but also to discuss their practical construction and evaluate their performance for a finite blocklength by means of simulations.

1.1. Relation to Prior Work

A good overview of the similarities and differences between the polar codes proposed in [10–13] for the general wiretap channel can be found in [13] (Figure 1). The polar coding schemes proposed in this paper are based mainly on those introduced by [13] because of the following reasons:

- To provide strong secrecy. Despite both weak and strong secrecy conditions resulting in the same secrecy-capacity region, the weak secrecy requirement in practical applications can result in important system vulnerabilities [19] (Section 3.3).
- To provide polar coding schemes that are implementable in practice. Notice in [13] (Figure 1) that the coding scheme presented in [10] relies on a construction for which no efficient code is presently known. Moreover, the polar coding scheme in [12] relies on the existence, through averaging, of certain deterministic mappings for the encoding/decoding process.

As in [13], our polar coding schemes are totally explicit. However, to provide strong secrecy and reliability simultaneously, the transmitter and the legitimate receivers need to share a secret key of negligible size in terms of rate, and the distribution induced by the encoder must be close in terms of statistical distance to the original one considered for the code construction. Moreover, we adapt the deterministic SC encoder of [20] to our channel models, and we show that it can perform well in practice. As concluded in [20], this deterministic SC encoder will avoid the need to draw large sequences according to specific distributions at the encoder, which can be useful in communication systems requiring low complexity at the transmitter.

In [13] (Remark 3), the authors highlight the connection between polar code constructions and random binning proofs that allows them to apply their designs to different problems in network information theory. Nevertheless, in our polar coding schemes, the chaining construction used in [13] is not needed because of the degradedness condition of the channels, and consequently, we can introduce small changes in the design in order to make our proposed coding schemes more practical. In this sense, we assume that a source of common randomness is accessible to all parties, which allows the transmitter to send secret information in just one block of size n by only using a secret key with negligible size in terms of rate. Despite this common randomness being available to the eavesdroppers, no information will be leaked about the messages. Moreover, if we consider a communication system requiring transmissions over several blocks of size n, the same realization of this source of common randomness can be used at each block without compromising the strong secrecy condition.

1.2. Overview of Novel Contributions

The main novelties of this paper can be summarized as follows:

1. Scenario. This paper focuses on two different models of the DBC with an arbitrary number of legitimate receivers and an arbitrary number of eavesdroppers for which polar codes have not yet been proposed. These two models arise very commonly in wireless communications.
2. Existence of the polar coding schemes. We prove the existence for sufficiently large n of two secrecy-capacity achieving polar coding schemes under the strong secrecy condition.
3. Practical implementation. We provide polar codes that are implementable in real communication systems, and we discuss further how to construct them in practice. As far as we know, although the construction of polar codes has been covered in a large number of references (for instance, see [21–23]), they only focus on polar code constructions under reliability constraints.
4. Performance evaluation. Simulations results are provided in order to evaluate the reliability and secrecy performance of the polar coding schemes. The performance is evaluated according to different design parameters of the practical code construction. As far as we know, this paper is the first to evaluate the secrecy performance in terms of the strong secrecy, which is done by upper-bounding the information leakage at the eavesdroppers.

1.3. Notation

Through this paper, let $[n] = \{1, \dots, n\}$ for $n \in \mathbb{Z}^+$, a^n denote a row vector $(a(1), \dots, a(n))$. We write $a^{1:j}$ for $j \in [n]$ to denote the subvector $(a(1), \dots, a(j))$. Let $\mathcal{A} \subset [n]$, then we write $a[\mathcal{A}]$ to denote the sequence $\{a(j)\}_{j \in \mathcal{A}}$, and we use $\mathcal{A}^{\mathrm{C}}$ to denote the set complement with respect to the universal set $[n]$, that is $\mathcal{A}^{\mathrm{C}} = [n] \setminus \mathcal{A}$. If $\mathcal{A}$ denotes an event, then $\mathcal{A}^{\mathrm{C}}$ also denotes its complement. We use ln to denote the natural logarithm, whereas log denotes the logarithm base two. Let X be a random variable taking values in $\mathcal{X}$, and let q_x and p_x be two different distributions with support $\mathcal{X}$, then $\mathbb{D}(q_x, p_x)$ and $\mathbb{V}(q_x, p_x)$ denote the Kullback-Leibler divergence and the total variation distance, respectively. Finally, $h_2(p)$ denotes the binary entropy function, i.e., $h_2(p) = -p \log p - (1-p)\log(1-p)$, and we define the indicator function $\mathbb{1}\{u\}$ such that it equals one if the predicate u is true and zero otherwise.

1.4. Organization

The remainder of this paper is organized as follows. In Section 2, the channel models DBC-NLD-LS and DBC-LD-NLS are introduced formally, and their secrecy-capacity regions are characterized. In Section 3, the fundamentals theorems of polar codes are revisited. In Sections 4 and 5, two polar coding schemes are proposed for the DBC-NLD-LS and DBC-LD-NLS, respectively, and we prove that both are asymptotic secrecy-capacity achieving. In Section 6, practical polar code constructions are discussed for both models, and the performances of the polar codes are evaluated by means of simulations. Finally, the concluding remarks are presented in Section 7.

2. System Model and Secrecy-Capacity Region

Formally, a DBC $(\mathcal{X}, p_{Y_K \dots Y_1 Z_M \dots Z_1|X}, \mathcal{Y}_K \times \cdots \times \mathcal{Y}_1 \times \mathcal{Z}_M \times \cdots \times \mathcal{Z}_1)$ with K legitimate receivers and M eavesdroppers is characterized by the probability transition function $p_{Y_K \dots Y_1 Z_M \dots Z_1|X}$, where $X \in \mathcal{X}$ denotes the channel input, $Y_k \in \mathcal{Y}_k$ denotes the channel output corresponding to the legitimate receiver $k \in [1, K]$ and $Z_m \in \mathcal{Z}_m$ denotes the channel output corresponding to the eavesdropper $m \in [1, M]$. The broadcast channel is assumed to gradually degrade in such a way that each legitimate receiver has a better channel than any eavesdropper, that is:

$$X - Y_K - \cdots - Y_1 - Z_M - \cdots - Z_1 \tag{1}$$

forms a Markov chain. Although we consider physically degradation, the polar coding schemes proposed in this paper are also suitable for stochastically degraded channels (see Remark 2).

2.1. Degraded Broadcast Channel with Non-Layered Decoding and Layered Secrecy

In this model (see Figure 1), the transmitter wishes to send M messages $\{W_m\}_{m=1}^{M}$ to the K legitimate receivers. The non-layered decoding structure requires the legitimate receiver $k \in [1, K]$ to reliably decode all M messages, and the layered secrecy structure requires the eavesdropper $m \in [1, M]$ to be kept ignorant about messages $\{W_i\}_{i=m}^{M}$. Consider a $(\lceil 2^{nR_1} \rceil, \ldots, \lceil 2^{nR_M} \rceil, n)$ code for the DBC-NLD-LS, where $W_m \in [\lceil 2^{nR_m} \rceil]$ for any $m \in [1, M]$. The reliability condition to be satisfied by this code is measured in terms of the average probability of error at each legitimate receiver and is given by:

$$\lim_{n \to \infty} \mathbb{P}\left[(\hat{W}_1, \ldots, \hat{W}_M) \neq (W_1, \ldots, W_M)\right] = 0, \quad \text{for any legitimate receiver } k \in [1, K]. \tag{2}$$

On the other hand, the strong secrecy condition to be satisfied by the code is measured in terms of the information leakage at each eavesdropper and is given by:

$$\lim_{n \to \infty} I(W_m, W_{m+1}, \ldots, W_M; Z_m^n) = 0, \quad \text{for the eavesdropper } m \in [1, M]. \tag{3}$$

A tuple of rates $(R_1, \ldots, R_M) \in \mathbb{R}_+^M$ is achievable for the DBC-NLD-LS if there exists a sequence of $(\lceil 2^{nR_1} \rceil, \ldots, \lceil 2^{nR_M} \rceil, n)$ codes satisfying Equations (2) and (3).

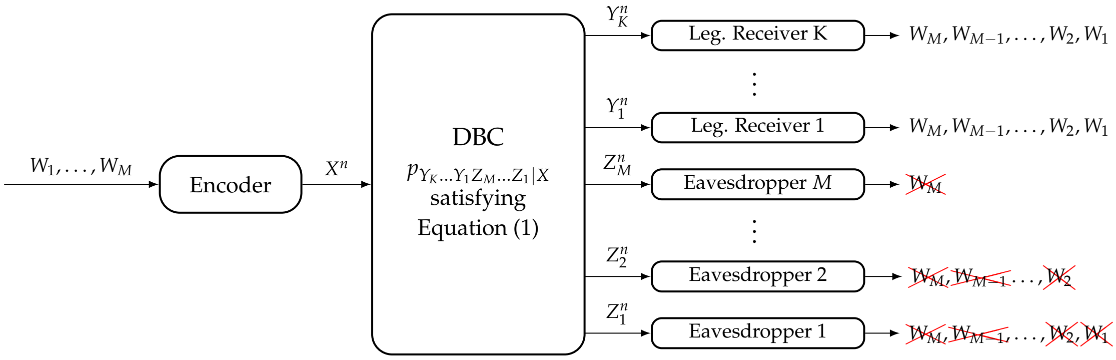
$W_1, \ldots, W_M$
Encoder
X^n
DBC
$p_{Y_K \ldots Y_1 Z_M \ldots Z_1 | X}$ satisfying Equation (1)
Y_K^n
Leg. Receiver K
$W_M, W_{M-1}, \ldots, W_2, W_1$
Y_1^n
Leg. Receiver 1
$W_M, W_{M-1}, \ldots, W_2, W_1$
Z_M^n
Eavesdropper M
~~W_M~~
Z_2^n
Eavesdropper 2
~~$W_M, W_{M-1}, \ldots, W_2$~~
Z_1^n
Eavesdropper 1
~~$W_M, W_{M-1}, \ldots, W_2, W_1$~~

Figure 1. DBC with Non-Layered Decoding and Layered Secrecy (DBC-NLD-LS).

Proposition 1 (Adapted from [4,5]). *The achievable region of the DBC-NLD-LS is the union of all M-tuples of rates $(R_1, \ldots, R_M) \in \mathbb{R}_+^M$ satisfying the following inequalities,*

$$\sum_{i=m}^{M} R_i \leq I(X; Y_1) - I(X; Z_m), \qquad m = 1, \ldots, M,$$

where the union is taken over all distributions p_X.

The proof for the case of only one legitimate receiver in the context of the fading wiretap channel is provided in [5], where the information-theoretic achievable scheme is based on embedded coding, stochastic encoding and rate sharing. Due to the degradedness condition of Equation (1), by applying the data processing inequality and Fano's inequality, an achievable scheme ensuring the reliability condition in Equation (2) for the legitimate Receiver 1 will satisfy it for any legitimate receiver $k \in [2, K]$.

Corollary 1. *The achievable subregion of the DBC-NLD-LS without considering rate sharing is a K-orthotope defined by the closure of all K-tuples of rates $(R_1, \ldots, R_M) \in \mathbb{R}_+^M$ satisfying:*

$$\begin{aligned} R_m &\leq I(X; Z_{m+1}) - I(X; Z_m), \qquad m = 1, \ldots, M-1, \\ R_M &\leq I(X; Y_1) - I(X; Z_M). \end{aligned}$$

2.2. Degraded Broadcast Channel with Layered Decoding and Non-Layered Secrecy

In this model (see Figure 2), the transmitter wishes to send K messages $\{W_\ell\}_{\ell=1}^{K}$ to the K legitimate receivers. The layered decoding structure requires the legitimate receiver $k \in [1, K]$ to reliably decode the messages $\{W_\ell\}_{\ell=1}^{k}$, and the non-layered secrecy structure requires the eavesdropper $m \in [1, M]$ to be kept ignorant of all K messages. Consider a $(\lceil 2^{nR_1} \rceil, \ldots, \lceil 2^{nR_K} \rceil, n)$ code for the DBC-LD-NLS, where $W_\ell \in [\lceil 2^{nR_\ell} \rceil]$ for any $\ell \in [1, K]$. The reliability condition to be satisfied by this code is:

$$\lim_{n\to\infty} \mathbb{P}\left[(\hat{W}_1, \ldots, \hat{W}_{k-1}, \hat{W}_k) \neq (W_1, \ldots, W_{k-1}, W_k)\right] = 0, \quad \text{for the legitimate receiver } k \in [1, K], \tag{4}$$

and the strong secrecy condition is given by:

$$\lim_{n\to\infty} I(W_1, \ldots, W_K; Z_m^n) = 0, \quad \text{for any eavesdropper } m \in [1, M]. \tag{5}$$

A tuple of rates $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ is achievable for the DBC-LD-NLS if there exists a sequence of $(\lceil 2^{nR_1} \rceil, \ldots, \lceil 2^{nR_K} \rceil, n)$ codes such that they satisfy Equations (4) and (5).

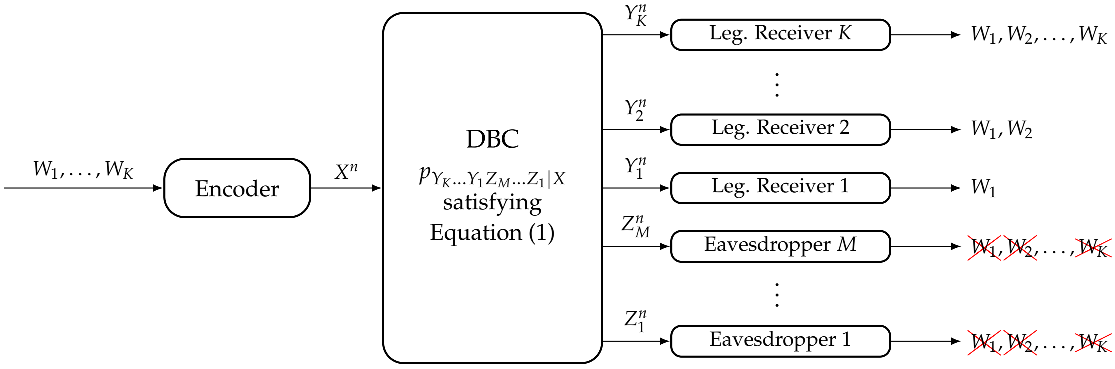

Figure 2. DBC with Layered Decoding and Non-Layered Secrecy (DBC-LD-NLS).

Proposition 2 (Adapted from [4,6]). *The achievable region of the DBC-LD-NLS is the union of all K-tuples of rates* $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ *satisfying the following inequalities,*

$$\sum_{\ell=1}^{k} R_\ell \leq \sum_{\ell=1}^{k} I(V_\ell; Y_\ell | V_{\ell-1}) - I(V_k, Z_M), \qquad k = 1, \ldots, K,$$

where $V_0 \triangleq \varnothing$ *and* $V_K \triangleq X$*, and the union is taken over all distributions* $p_{V_1 \ldots V_K}$ *such that* $V_1 - V_2 - \cdots - V_K$ *forms a Markov chain.*

The proof for the case of only one eavesdropper is provided in [6], where the information-theoretic achievable scheme is based on superposition coding, stochastic encoding and rate sharing. Due to the degradedness condition of Equation (1), note that any achievable scheme ensuring the strong secrecy condition in Equation (5) for the eavesdropper M will also satisfy it for any eavesdropper $m \in [1, M-1]$.

Corollary 2. *The achievable subregion of the DBC-LD-NLS without considering rate sharing is a K-orthotope defined by the closure of all K-tuples of rates* $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ *satisfying:*

$$R_\ell \leq I(V_\ell; Y_\ell | V_{\ell-1}) - I(V_\ell; Z_M | V_{\ell-1}), \qquad \ell = 1, \ldots, K.$$

3. Review of Polar Codes

Let $(\mathcal{X} \times \mathcal{Y}, p_{XY})$ be a Discrete Memoryless Source (DMS), where $X \in \{0,1\}$ (see Endnote [24]—which refers to References [25,26]) and $Y \in \mathcal{Y}$. The polar transform over the n-sequence X^n, n being any power of two, is defined as $U^n \triangleq X^n G_n$, where $G_n \triangleq \left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]^{\otimes n}$ is the source polarization matrix [27]. Since $G_n = G_n^{-1}$, then $X^n = U^n G_n$.

The polarization theorem for source coding with side information [27] (Theorem 1) states that the polar transform extracts the randomness of X^n in the sense that, as $n \to \infty$, the set of indices $j \in [n]$ can be divided practically into two disjoint sets, namely $\mathcal{H}_{X|Y}^{(n)}$ and $\mathcal{L}_{X|Y}^{(n)}$, such that $U(j)$ for $j \in \mathcal{H}_{X|Y}^{(n)}$ is practically independent of $(U^{1:j-1}, Y^n)$ and uniformly distributed, i.e., $H(U(j)|U^{1:j-1}, Y^n) \to 1$, and $U(j)$ for $j \in \mathcal{L}_{X|Y}^{(n)}$ is almost determined by $(U^{1:j-1}, Y^n)$, i.e., $H(U(j)|U^{1:j-1}, Y^n) \to 0$. Formally, let:

$$\mathcal{H}_{X|Y}^{(n)} \triangleq \{j \in [n] : H(U(j)|U^{1:j-1}, Y^n) \geq 1 - \delta_n\},$$
$$\mathcal{L}_{X|Y}^{(n)} \triangleq \{j \in [n] : H(U(j)|U^{1:j-1}, Y^n) \leq \delta_n\},$$

where $\delta_n \triangleq 2^{-n^\beta}$ for some $\beta \in (0, \frac{1}{2})$. Then, by [27] (Theorem 1), we have $\lim_{n\to\infty} \frac{1}{n}|\mathcal{H}_{X|Y}^{(n)}| = H(X|Y)$ and $\lim_{n\to\infty} \frac{1}{n}|\mathcal{L}_{X|Y}^{(n)}| = 1 - H(X|Y)$, which imply that $\lim_{n\to\infty} \frac{1}{n}|(\mathcal{H}_{X|Y}^{(n)})^{\mathrm{C}} \cap (\mathcal{L}_{X|Y}^{(n)})^{\mathrm{C}}| = 0$, i.e., the number of elements that have not been polarized is asymptotically negligible in terms of rate. Furthermore, [27] (Theorem 2) states that given $U[(\mathcal{L}_{X|Y}^{(n)})^{\mathrm{C}}]$ and Y^n, $U[\mathcal{L}_{X|Y}^{(n)}]$ can be reconstructed using SC decoding with error probability in $O(n\delta_n)$. Alternatively, the previous sets can be defined based on the Bhattacharyya parameters $\{Z(U(j)|U^{1:j-1}, Y^n)\}_{j=1}^n$ because both parameters polarize simultaneously [27] (Proposition 2). It is worth mentioning that both the entropy terms and the Bhattacharyya parameters required to define these sets can be obtained deterministically from p_{XY} and the algebraic properties of G_n [21–23].

Similarly to $\mathcal{H}_{X|Y}^{(n)}$ and $\mathcal{L}_{X|Y}^{(n)}$, the sets $\mathcal{H}_{X}^{(n)}$ and $\mathcal{L}_{X}^{(n)}$ can be defined by considering that observations Y^n are absent. A discrete memoryless channel $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ with some arbitrary p_X can be seen as a DMS $(\mathcal{X} \times \mathcal{Y}, p_X p_{Y|X})$. In channel polar coding, first, we define $\mathcal{H}_{X|Y}^{(n)}$, $\mathcal{L}_{X|Y}^{(n)}$, $\mathcal{H}_{X}^{(n)}$ and $\mathcal{L}_{X}^{(n)}$ from the target distribution $p_X p_{Y|X}$ (polar construction). Then, based on the previous sets, the encoder somehow constructs $\tilde{U}^n$ and applies the inverse polar transform $\tilde{X}^n = \tilde{U}^n G_n$, with distribution $\tilde{q}_{X^n}$ (since the polar-based encoder will construct random variables that must approach the target distribution of the DMS, throughout this paper, we use a tilde above the random variables to emphasize this purpose). Afterwards, the transmitter sends $\tilde{X}^n$ over the channel, which induces $\tilde{Y}^n \sim \tilde{q}_{Y^n}$. If $\mathbb{V}(\tilde{q}_{X^n Y^n}, p_{X^n Y^n}) \to 0$, then the receiver can reliably reconstruct $\tilde{U}[\mathcal{L}_{X|Y}^{(n)}]$ from $\tilde{Y}^n$ and $\tilde{U}[(\mathcal{L}_{X|Y}^{(n)})^{\mathrm{C}}]$ by using SC decoding [28].

To conclude this part, the following lemma provides a useful property of polar codes for the DBC.

Lemma 1 (Subset property, adapted from [14] (Lemma 4))**.** *Let (X, Y_2, Y_1) be random variables such that $X - Y_2 - Y_1$ forms a Markov chain. Then, the following property holds for the polar transform $U^n = X^n G_n$,*

$$H(U(j)|U^{1:j-1}) \geq H(U(j)|U^{1:j-1}, Y_1^n) \geq H(U(j)|U^{1:j-1}, Y_2^n) \quad \forall j \in [n], \quad \textit{which implies}$$
$$\mathcal{L}_{X}^{(n)} \subseteq \mathcal{L}_{X|Y_1}^{(n)} \subseteq \mathcal{L}_{X|Y_2}^{(n)}, \quad \textit{and} \quad \mathcal{H}_{X|Y_2}^{(n)} \subseteq \mathcal{H}_{X|Y_1}^{(n)} \subseteq \mathcal{H}_{X}^{(n)}.$$

Remark 1. *The subset property also holds if the sets are defined based on the Bhattacharyya parameters because, under the previous Markov chain condition, $Z(U(j)|U^{1:j-1}) \geq Z(U(j)|U^{1:j-1}, Y_1^n) \geq Z(U(j)|U^{1:j-1}, Y_2^n)$.*

Remark 2. *According to [14] (Lemma 4), the subset property also holds if the channels are stochastically degraded. Therefore, since the construction of the polar codes proposed in the following sections is based basically on Lemma 1, the polar coding schemes are suitable for physically- and stochastically-degraded channels.*

4. Polar Coding Scheme For the DBC-NLD-LS

The polar coding scheme provided in this section is designed to achieve the supremum of the achievable rates given in Corollary 1 (secrecy-capacity without rate sharing). Thus, consider the DMS $\big(\mathcal{X} \times \mathcal{Y}_K \times \cdots \times \mathcal{Y}_1 \times \mathcal{Z}_M \times \cdots \times \mathcal{Z}_1, p_{XY_K \ldots Y_1 Z_M \ldots Z_1}\big)$ that represents the input and output random variables involved in the achievable subregion of Corollary 1, where $\mathcal{X} = \{0,1\}$. Let $(X^n, Y_K^n, \ldots, Y_1^n, Z_M^n, \ldots, Z_1^n)$ be an i.i.d. n-sequence of this source. We define the polar transform $U^n \triangleq X^n G_n$, whose distribution is $p_{U^n}(u^n) = p_{X^n}(u^n G_n)$ (due to the invertibility of G_n), and we write:

$$p_{U^n}(u^n) \triangleq \prod_{j=1}^{n} p_{U(j)|U^{1:j-1}}\big(u(j)\big|u^{1:j-1}\big). \tag{6}$$

4.1. Polar Code Construction

Let $\delta_n \triangleq 2^{-n^\beta}$, where $\beta \in (0, \frac{1}{2})$. Based on $p_{XY_K \ldots Y_1 Z_M \ldots Z_1}$, we define:

$$\mathcal{H}_X^{(n)} \triangleq \big\{j \in [n] : H\big(U(j)\big|U^{1:j-1}\big) \geq 1 - \delta_n\big\}, \tag{7}$$

$$\mathcal{L}_X^{(n)} \triangleq \big\{j \in [n] : H\big(U(j)\big|U^{1:j-1}\big) \leq \delta_n\big\}, \tag{8}$$

$$\mathcal{L}_{X|Y_k}^{(n)} \triangleq \big\{j \in [n] : H\big(U(j)\big|U^{1:j-1}, Y_k^n\big) \leq \delta_n\big\}, \quad k = 1, \ldots, K, \tag{9}$$

$$\mathcal{H}_{X|Y_k}^{(n)} \triangleq \big\{j \in [n] : H\big(U(j)\big|U^{1:j-1}, Y_k^n\big) \geq 1 - \delta_n\big\}, \quad k = 1, \ldots, K, \tag{10}$$

$$\mathcal{H}_{X|Z_m}^{(n)} \triangleq \big\{j \in [n] : H\big(U(j)\big|U^{1:j-1}, Z_m^n\big) \geq 1 - \delta_n\big\}, \quad m = 1, \ldots, M. \tag{11}$$

Then, based on the previous sets, we define the following partition of the universal set $[n]$,

$$\mathcal{I}_M^{(n)} \triangleq \mathcal{H}_{X|Z_M}^{(n)} \cap \big(\mathcal{H}_{X|Y_1}^{(n)}\big)^{\mathrm{C}}, \tag{12}$$

$$\mathcal{I}_m^{(n)} \triangleq \mathcal{H}_{X|Z_m}^{(n)} \cap \big(\mathcal{H}_{X|Z_{m+1}}^{(n)}\big)^{\mathrm{C}}, \quad m = 1, \ldots, M-1, \tag{13}$$

$$\mathcal{F}^{(n)} \triangleq \mathcal{H}_{X|Y_1}^{(n)}, \tag{14}$$

$$\mathcal{C}^{(n)} \triangleq \mathcal{H}_X^{(n)} \cap \big(\mathcal{H}_{X|Z_1}^{(n)}\big)^{\mathrm{C}}, \tag{15}$$

$$\mathcal{T}^{(n)} \triangleq \big(\mathcal{H}_X^{(n)}\big)^{\mathrm{C}}, \tag{16}$$

which is graphically represented in Figure 3. Roughly speaking, in order to ensure reliability and strong secrecy, the distribution of $\tilde{U}^n$ after the encoding process must be close in terms of statistical distance to the distribution given in Equation (6) corresponding to the original DMS. Hence, the elements $U(j)$ such that $j \in \mathcal{H}_X^{(n)}$ will be suitable for storing uniformly-distributed random sequences. On the other hand, $U[\mathcal{T}^{(n)}]$ will not, and the elements $U(j)$ such that $j \in \mathcal{T}^{(n)}$ will be constructed somehow from $U^{1:j-1}$ and the distribution $p_{U(j)|U^{1:j-1}}$. The set $\mathcal{I}_m^{(n)}$ ($m \in [1, M]$) belongs to $\mathcal{H}_{X|Z_m}^{(n)}$, and by Lemma 1, we have $\mathcal{H}_{X|Z_m}^{(n)} \subseteq \mathcal{H}_{X|Z_{m'}}^{(n)}$ for any $m' < m$. Thus, $U[\mathcal{I}_m^{(n)}]$ will be suitable for storing information to be secured from Eavesdroppers 1–m. Since $\mathcal{C}^{(n)} \subseteq \big(\mathcal{H}_{X|Z_m}^{(n)}\big)^{\mathrm{C}}$ for any $m \in [1, M]$, the sequence $U\big[\mathcal{C}^{(n)}\big]$ cannot contain information to be secured from any eavesdropper, and it will be used to store the local randomness [8] required to confuse the eavesdroppers (the local randomness in polar codes plays the same role as the stochastic encoding used in [1,2]). According to [27] (Theorem 2), the legitimate Receiver 1 will be able to reliably infer $U[\mathcal{L}_{X|Y_1}^{(n)}]$ given Y_1^n and $U[(\mathcal{L}_{X|Y_1}^{(n)})^{\mathrm{C}}]$. Hence, if the polar coding scheme somehow make the entries $U(j)$ such that j belongs to $\mathcal{F}^{(n)}$ and $(\mathcal{H}_{X|Y_1}^{(n)})^{\mathrm{C}} \cap (\mathcal{L}_{X|Y_1}^{(n)})^{\mathrm{C}}$ (hatched areas in Figure 3) available to the legitimate Receiver 1, this receiver will be able to reliably infer the entire sequence U^n. In this sense, $U[\mathcal{F}^{(n)}]$ will be used to store the uniformly-distributed random sequence provided by a source of common randomness that will be available to all parties.

Since $\mathcal{F}^{(n)} \subseteq \mathcal{H}^{(n)}_{X|Z_m}$ for any $m \in [1, M]$, the knowledge of $U[\mathcal{F}^{(n)}]$ of the eavesdroppers will not compromise the strong secrecy condition. On the other hand, $U[(\mathcal{H}^{(n)}_{X|Y_1})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{X|Y_1})^{\mathrm{C}}]$ will contain secret information or elements that cannot be known directly by all the eavesdroppers. Therefore, the transmitter somehow will secretly send it to the legitimate receivers. Nevertheless, as will be seen, this additional transmission will incur an asymptotically negligible rate penalty. Finally, by Lemma 1, we have $(\mathcal{L}^{(n)}_{X|Y_1})^{\mathrm{C}} \supseteq (\mathcal{L}^{(n)}_{X|Y_k})^{\mathrm{C}}$ for any $k > 1$. Hence, given $U[(\mathcal{L}^{(n)}_{X|Y_1})^{\mathrm{C}}]$, all the legitimate receivers will be able to reliably infer the entire sequence U^n from their own channel observations.

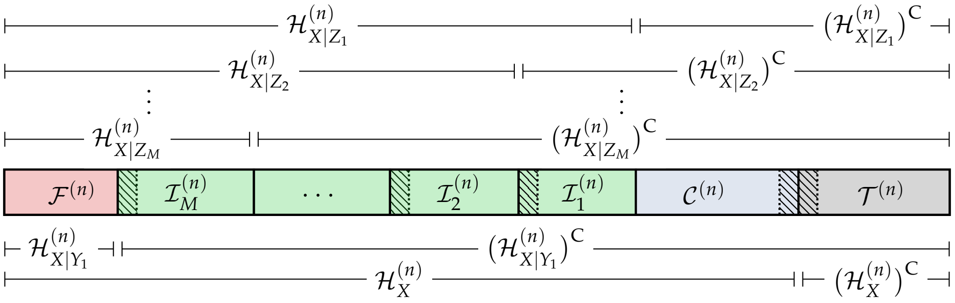
$\mathcal{H}^{(n)}_{X|Z_1}$
$(\mathcal{H}^{(n)}_{X|Z_1})^{\mathrm{C}}$
$\mathcal{H}^{(n)}_{X|Z_2}$
$(\mathcal{H}^{(n)}_{X|Z_2})^{\mathrm{C}}$
$\mathcal{H}^{(n)}_{X|Z_M}$
$(\mathcal{H}^{(n)}_{X|Z_M})^{\mathrm{C}}$
$\mathcal{F}^{(n)}$
$\mathcal{I}^{(n)}_M$
$\cdots$
$\mathcal{I}^{(n)}_2$
$\mathcal{I}^{(n)}_1$
$\mathcal{C}^{(n)}$
$\mathcal{T}^{(n)}$
$\mathcal{H}^{(n)}_{X|Y_1}$
$(\mathcal{H}^{(n)}_{X|Y_1})^{\mathrm{C}}$
$\mathcal{H}^{(n)}_{X}$
$(\mathcal{H}^{(n)}_{X})^{\mathrm{C}}$

Figure 3. Polar code construction for DBC-NLD-LS. The hatched area represents those indices $j \in (\mathcal{H}^{(n)}_{X|Y_1})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{X|Y_1})^{\mathrm{C}}$, which can belong to the sets $\mathcal{I}^{(n)}_m$ ($m \in [1, M]$), $\mathcal{C}^{(n)}$, $\mathcal{F}^{(n)}$ or $\mathcal{T}^{(n)}$.

Remark 3. *The goal of the polar code construction is to obtain the entropy terms* $\{H(U(j)|U^{1:j-1})\}_{j=1}^{n}$, $\{H(U(j)|U^{1:j-1}, Y_1^n)\}_{j=1}^{n}$ *and* $\{H(U(j)|U^{1:j-1}, Z_m^n)\}_{j=1}^{n}$ *for all* $m \in [1, M]$ *required to define the sets in Equations* (7)–(11) *and, consequently, to obtain the partition of* $[n]$ *given in Equations* (12)–(16)*. In Section* 6*, we discuss further how to construct polar codes under both reliability and secrecy constraints.*

4.2. Polar Encoding

The polarization-based encoder aims to construct the sequence $\tilde{U}^n$ and, consequently, $\tilde{X}^n = \tilde{U}^n G_n$. Let W_m for all $m \in [1, M]$ and C be uniformly-distributed random vectors of size $|\mathcal{I}^{(n)}_m|$ and $|\mathcal{C}^{(n)}|$, respectively, where C represents the local randomness required to confuse the eavesdroppers, and recall that W_m represents the message m that is intended for all legitimate receivers. Let F be a given uniformly-distributed random $|\mathcal{F}^{(n)}|$-sequence, which represents the source of common randomness that is available to all parties. The encoder constructs the sequence $\tilde{u}^n$ as follows. Consider the realizations w_m for all $m \in [1, M]$, c and f, whose elements have been indexed by the set of indices $\mathcal{I}^{(n)}_m$, $\mathcal{C}^{(n)}$ and $\mathcal{F}^{(n)}$, respectively. The encoder draws $\tilde{u}^n$ from the distribution:

$$\tilde{q}_{U(j)|U^{1:j-1}}(\tilde{u}(j)|\tilde{u}^{1:j-1}) \triangleq \begin{cases} \mathbb{1}\{\tilde{u}(j) = w_m(j)\} & \text{if } j \in \mathcal{I}^{(n)}_m, \;\; m = 1, \ldots, M, \\ \mathbb{1}\{\tilde{u}(j) = c(j)\} & \text{if } j \in \mathcal{C}^{(n)}, \\ \mathbb{1}\{\tilde{u}(j) = f(j)\} & \text{if } j \in \mathcal{F}^{(n)}, \\ p_{U(j)|U^{1:j-1}}(\tilde{u}(j)|\tilde{u}^{1:j-1}) & \text{if } j \in (\mathcal{H}^{(n)}_X)^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_X)^{\mathrm{C}}, \\ \mathbb{1}\{\tilde{u}(j) = \xi^{(j)}(\tilde{u}^{1:j-1})\} & \text{if } j \in \mathcal{L}^{(n)}_X, \end{cases} \tag{17}$$

where:

$$\xi^{(j)}(\tilde{u}^{1:j-1}) \triangleq \arg\max_{u \in \mathcal{X}} p_{U(j)|U^{1:j-1}}(u|\tilde{u}^{1:j-1}), \tag{18}$$

$p_{U(j)|U^{1:j-1}}$ being the distribution induced by the original DMS. Note that $\mathcal{T}^{(n)} = ((\mathcal{H}^{(n)}_X)^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_X)^{\mathrm{C}}) \cup \mathcal{L}^{(n)}_X$, and according to Equation (17), $\tilde{U}[\mathcal{L}^{(n)}_X]$ is constructed deterministically by adapting the SC encoding algorithm in [20], while $\tilde{U}[(\mathcal{H}^{(n)}_X)^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_X)^{\mathrm{C}}]$ is constructed randomly. By [27] (Theorem 1), we have that the amount of randomness for SC encoding will be asymptotically negligible in terms of rate. Then, the encoder computes $\tilde{X}^n = \tilde{U}^n G_n$ and transmits it over the DBC, inducing $(\tilde{Y}_K, \ldots, \tilde{Y}_1, \tilde{Z}_M, \ldots, \tilde{Z}_1)$.

Finally, besides the sequence $\tilde{X}^n$, the encoder outputs the following additional secret sequence,

$$\Phi \triangleq \tilde{U}\Big[\big(\mathcal{H}_{X|Y_1}^{(n)}\big)^{\mathrm{C}} \cap \big(\mathcal{L}_{X|Y_1}^{(n)}\big)^{\mathrm{C}}\Big]. \tag{19}$$

This sequence Φ must be additionally transmitted to all legitimate receivers keeping it masked from the eavesdroppers. To do so, the transmitter can perform a modulo-two addition between Φ and a uniformly-distributed secret key that is privately shared with the legitimate receivers and somehow additionally send it to them. Nevertheless, by [27] (Theorem 1), we know that this additional transmission is asymptotically negligible in terms of rate.

Remark 4. *The additional secret sequence Φ can be divided into two parts: $\tilde{U}[\mathcal{H}_X^{(n)} \cap (\mathcal{H}_{X|Y_1}^{(n)})^{\mathrm{C}} \cap (\mathcal{L}_{X|Y_1}^{(n)})^{\mathrm{C}}]$, which will be uniformly distributed according to Equation (17), and the remaining part that will not. The transmitter could make the uniformly-distributed part available to the legitimate receivers by using a chaining structure as the one presented in [9]. However, such a scheme requires the transmission to take place over several blocks of size n. Moreover, it requires having a large memory capacity on either the transmitter or the legitimate receivers, which can make the polar coding scheme unpractical in communication systems.*

4.3. Polar Decoding

Before the decoding process, consider that the realization of the source of common randomness F is available to all parties and the sequence Φ has been successfully received by the legitimate receivers.

The legitimate receiver $k \in [1, K]$ forms an estimate $\hat{U}^n$ of the sequence $\tilde{U}^n$ as follows. Given that Φ and F are available, notice that it knows $\tilde{U}[(\mathcal{L}_{X|Y_1}^{(n)})^{\mathrm{C}}]$. Moreover, by Lemma 1, $(\mathcal{L}_{X|Y_1}^{(n)})^{\mathrm{C}} \supseteq (\mathcal{L}_{X|Y_k}^{(n)})^{\mathrm{C}}$ for any $k > 1$. Thus, the k-th legitimate receiver performs SC decoding for source coding with side information [27] to construct $\tilde{U}^n$ from $\tilde{U}[(\mathcal{L}_{X|Y_1}^{(n)})^{\mathrm{C}}]$ and its channel output observations $\tilde{Y}_k$. In Section 4.5.3, we show formally that the reliability condition in Equation (2) is satisfied at each legitimate receiver $k \in [1, K]$.

4.4. Information Leakage

Besides the observations $\tilde{Z}_m^n$, the eavesdropper $m \in [1, M]$ has access to the common randomness $F = \tilde{U}[\mathcal{F}^{(n)}]$. Thus, the information about the messages $\{W_i\}_{i=m}^M$ leaked to this eavesdropper is:

$$I(W_m, \ldots, W_M; F, \tilde{Z}_m^n) = I\big(\tilde{U}\big[\cup_{i=m}^M \mathcal{I}_i^{(n)}\big]; \tilde{U}\big[\mathcal{F}^{(n)}\big], \tilde{Z}_m^n\big). \tag{20}$$

In Section 4.5.4, we prove that $(W_m, W_{m+1}, \ldots, W_M)$ is asymptotically statistically independent of $(F, \tilde{Z}_m^n)$.

4.5. Performance of the Polar Coding Scheme

The analysis of the polar coding scheme described previously leads to the following theorem.

Theorem 1. *Consider an arbitrary DBC $\big(\mathcal{X}, p_{Y_K \cdots Y_1 Z_M \ldots Z_1|X}, \mathcal{Y}_K \times \cdots \times \mathcal{Y}_1 \times \mathcal{Z}_M \times \cdots \times \mathcal{Z}_1\big)$ such that $\mathcal{X} \in \{0, 1\}$ and $p_{Y_K \ldots Y_1 Z_M \ldots Z_1|X}$ satisfies the Markov chain condition $X - Y_K - \cdots - Y_1 - Z_M - \cdots - Z_1$. The polar coding scheme described in Sections 4.1–4.4 achieves any rate tuple of the region defined in Corollary 1, satisfying the reliability and strong secrecy conditions given in Equations (2) and (3), respectively.*

Corollary 3. *Since $\tilde{U}[\mathcal{I}_m^{(n)}]$ for some $m \in [1, M]$ can contain information to be secured from Eavesdroppers 1–m, the polar coding scheme described in Sections 4.1–4.4 can achieve the entire region considering rate sharing of Proposition 1 by storing part of any message $W_{m'}$ such that $m' < m$ into $\tilde{U}[\mathcal{I}_m^{(n)}]$ instead of part of W_m.*

Corollary 4. *If we consider a communication scenario requiring transmissions over several blocks of size n, the same realization of the source of common randomness F that is known by all parties could be used at each block, and the reliability and the strong secrecy conditions would still be ensured.*

The proof of Theorem 1 follows in four steps with similar reasoning as in [13] and is provided in Sections 4.5.1–4.5.4. The proof of Corollary 3 is immediate, and the proof of Corollary 4 is provided in Section 4.5.5.

4.5.1. Transmission Rates

In this step, we prove that the polar coding scheme approaches the corner point of the subregion defined in Corollary 1. For any $m \in [1, M-1]$, the rate R_m corresponding to the message W_m satisfies:

$$\begin{aligned} \lim_{n\to\infty} R_m = \lim_{n\to\infty} \frac{1}{n}\left|\mathcal{I}_m^{(n)}\right| &\overset{(a)}{=} \lim_{n\to\infty} \frac{1}{n}\left|\mathcal{H}_{X|Z_m}^{(n)} \cap \left(\mathcal{H}_{X|Z_{m+1}}^{(n)}\right)^{\mathrm{C}}\right| \\ &\overset{(b)}{=} \lim_{n\to\infty} \frac{1}{n}\left(\left|\mathcal{H}_{X|Z_m}^{(n)}\right| - \left|\mathcal{H}_{X|Z_{m+1}}^{(n)}\right|\right) \\ &\overset{(c)}{=} I(X;Z_{m+1}) - I(X;Z_m), \end{aligned}$$

where (a) follows from the definition of the set $\mathcal{I}_m^{(n)}$ in Equation (13), (b) holds because, by Lemma 1, $\mathcal{H}_{X|Z_m}^{(n)} \supseteq \mathcal{H}_{X|Z_{m+1}}^{(n)}$, and (c) follows from [27] (Theorem 1). Similarly, according to Equation (12), we obtain:

$$\lim_{n\to\infty} R_M = \lim_{n\to\infty} \frac{1}{n}\left|\mathcal{I}_M^{(n)}\right| = \lim_{n\to\infty} \frac{1}{n}\left|\mathcal{H}_{X|Z_M}^{(n)} \cap \left(\mathcal{H}_{X|Y_1}^{(n)}\right)^{\mathrm{C}}\right| = I(X;Y_1) - I(X;Z_M).$$

4.5.2. Distribution of the DMS after the Polar Encoding

Let $\tilde{q}_{U^n}$ be the distribution of $\tilde{U}^n$ after the encoding in Section 4.2. The following lemma shows that $\tilde{q}_{U^n}$ and the distribution p_{U^n} in Equation (6) of the original DMS are nearly statistically indistinguishable for sufficiently large n and, consequently, so are the overall distributions $\tilde{q}_{XY_K\ldots Y_1Z_M\ldots Z_1}$ and $p_{XY_K\ldots Y_1Z_M\ldots Z_1}$.

Lemma 2. *Let $\delta_n = 2^{-n^\beta}$ for some $\beta \in (0, \frac{1}{2})$. Then,*

$$\mathbb{V}(\tilde{q}_{U^n}, p_{U^n}) \leq \delta_{\textit{nld-ls}}^{(n)},$$

$$\mathbb{V}(\tilde{q}_{X^nY_K^n\ldots Y_1^nZ_M^n\ldots Z_1^n}, p_{XY_K\ldots Y_1Z_M\ldots Z_1}) = \mathbb{V}(\tilde{q}_{U^n}, p_{U^n}) \leq \delta_{\textit{nld-ls}}^{(n)},$$

where $\delta_{\textit{nld-ls}}^{(n)} \triangleq n\sqrt{4\sqrt{n\delta_n \ln 2}(2n - \log(2\sqrt{n\delta_n \ln 2})) + \delta_n} + \sqrt{2n\delta_n \ln 2}.$

Proof. See Appendix A, setting $L = 1$. □

Remark 5. *The first term of $\delta_{\textit{nld-ls}}^{(n)}$ bounds the impact on the total variation distance of using the deterministic SC encoding in Equation (18) for the entries $\tilde{U}[\mathcal{L}_X^{(n)}]$, while the second term bounds the impact of storing uniformly-distributed random sequences (messages, local randomness and common randomness) into the entries $\tilde{U}[\mathcal{H}_X^{(n)}]$.*

As will be seen in the following subsections, an encoding process satisfying Lemma 2 is crucial for the reliability and the secrecy performance of the polar code.

4.5.3. Reliability Performance

Consider the probability of incorrectly decoding all messages $\{W_m\}_{m=1}^{M}$ at the legitimate receiver $k \in [1, K]$. Let $\tilde{q}_{X^n Y_k^n}$ and $p_{X^n Y_k^n}$ be the marginal distributions of $\tilde{q}_{X^n Y_K^n \dots Y_1^n Z_M^n \dots Z_1^n}$ and $p_{X^n Y_K^n \dots Y_1^n Z_M^n \dots Z_1^n}$, respectively. Consider an optimal coupling [29] (Proposition 4.7) between $\tilde{q}_{X^n Y_k^n}$ and $p_{X^n Y_k^n}$ such that:

$$\mathbb{P}\big[\mathcal{E}_{X^n Y_k^n}\big] = \mathbb{V}(\tilde{q}_{X^n Y_k^n}, p_{X^n Y_k^n}),$$

where $\mathcal{E}_{X^n Y_k^n} \triangleq \{(\tilde{X}^n, \tilde{Y}_k^n) \neq (X^n, Y_k^n)\}$ or, equivalently, $\mathcal{E}_{X^n Y_k^n} \triangleq \{(\tilde{U}^n, \tilde{Y}_k^n) \neq (U^n, Y_k^n)\}$ because of the invertibility of G_n. Thus, for the legitimate receiver $k \in [1, K]$, we obtain:

$$\begin{aligned}
\mathbb{P}\Big[(\hat{W}_1, \dots \hat{W}_M) \neq (W_1, \dots, W_M)\Big] &\leq \mathbb{P}\big[\hat{U}^n \neq \tilde{U}^n\big] \\
&= \mathbb{P}\big[\hat{U}^n \neq \tilde{U}^n \,|\, \mathcal{E}^{\mathrm{C}}_{X^n Y_k^n}\big]\mathbb{P}\big[\mathcal{E}^{\mathrm{C}}_{X^n Y_k^n}\big] + \mathbb{P}\big[\hat{U}^n \neq \tilde{U}^n \,|\, \mathcal{E}_{X^n Y_k^n}\big]\mathbb{P}\big[\mathcal{E}_{X^n Y_k^n}\big] \\
&\leq \mathbb{P}\big[\hat{U}^n \neq \tilde{U}^n \,|\, \mathcal{E}^{\mathrm{C}}_{X^n Y_k^n}\big] + \mathbb{P}\big[\mathcal{E}_{X^n Y_k^n}\big] \\
&\overset{(a)}{\leq} \textstyle\sum_{j \in \mathcal{L}^{(n)}_{X|Y_1}} Z(U(j) | U^{1:j-1}, Y_k^n) + \mathbb{P}\big[\mathcal{E}_{X^n Y_k^n}\big] \\
&\overset{(b)}{\leq} n\sqrt{\delta_n} + \mathbb{P}\big[\mathcal{E}_{X^n Y_k^n}\big] \\
&\overset{(c)}{\leq} n\sqrt{\delta_n} + \delta^{(n)}_{\text{nld-ls}},
\end{aligned} \tag{21}$$

where (a) holds by [27] (Theorem 2) because $\tilde{U}[(\mathcal{L}^{(n)}_{X|Y_1})^{\mathrm{C}}]$ is available to all receivers, (b) holds by Lemma 1, that is, $Z(U(j)|U^{1:j-1}, Y_k^n) \leq Z(U(j)|U^{1:j-1}, Y_1^n)$ for any $k > 1$, and by the definition of $\mathcal{L}^{(n)}_{X|Y_1}$ in Equation (9) and [27] (Proposition 2), that is $Z(U(j)|U^{1:j-1}, Y_1^n) \leq (H(U(j)|U^{1:j-1}, Y_1^n))^{1/2}$, and (c) holds by the optimal coupling and Lemma 2 because $\mathbb{V}(\tilde{q}_{X^n Y_k^n}, p_{X^n Y_k^n}) \leq \mathbb{V}(\tilde{q}_{X^n Y_K^n \dots Y_1^n Z_M^n \dots Z_1^n}, p_{X^n Y_K^n \dots Y_1^n Z_M^n \dots Z_1^n})$. Therefore, the polar coding scheme satisfies the reliability condition given in Equation (2).

4.5.4. Secrecy Performance

Consider the information leakage at the eavesdropper $m \in [1, M]$ given in Equation (20). We obtain:

$$\begin{aligned}
I(W_m, \dots, W_M; F, \tilde{Z}_m^n) &= H(\tilde{U}[\cup_{i=m}^{M} \mathcal{I}_i^{(n)}]) + H(\tilde{U}[\mathcal{F}^{(n)}] | \tilde{Z}_m^n) - H(\tilde{U}[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] | \tilde{Z}_m^n) \\
&\leq \textstyle\sum_{i=m}^{M} |\mathcal{I}_i^{(n)}| + |\mathcal{F}^{(n)}| - H(\tilde{U}[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] | \tilde{Z}_m^n).
\end{aligned} \tag{22}$$

Now, we provide a lower-bound for the conditional entropy term of Equation (22). First, for large enough n,

$$\begin{aligned}
&\Big| H(\tilde{U}[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] | \tilde{Z}_m^n) - H(U[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] | Z_m^n) \Big| \\
&\overset{(a)}{\leq} |H(\tilde{Z}_m^n) - H(Z_m^n)| + \Big| H(\tilde{U}[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}], \tilde{Z}_m^n) - H(U[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}], Z_m^n) \Big| \\
&\overset{(b)}{\leq} \mathbb{V}(\tilde{q}_{Z_m^n}, p_{Z_m^n}) \log \tfrac{2^n}{\mathbb{V}(\tilde{q}_{Z_m^n}, p_{Z_m^n})} \\
&\quad + \mathbb{V}(\tilde{q}_{U[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] Z_m^n}, p_{U[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] Z_m^n}) \log \tfrac{2^{(n + |(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}|)}}{\mathbb{V}(\tilde{q}_{U[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] Z_m^n}, p_{U[(\cup_{i=m}^{M} \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] Z_m^n})} \\
&\overset{(c)}{\leq} 3n\delta^{(n)}_{\text{nld-ls}} - 2\delta^{(n)}_{\text{nld-ls}} \log \delta^{(n)}_{\text{nld-ls}},
\end{aligned} \tag{23}$$

where (a) holds by the chain rule of entropy and the triangle inequality, (b) holds by [30] (Lemma 2.9) and (c) holds because the function $x \mapsto x \log x$ is decreasing for $x > 0$ small enough and by

Lemma 2 because $\mathbb{V}(\tilde{q}_{Z_m^n}, p_{Z_m^n}) \leq \mathbb{V}(\tilde{q}_{X^n Y_K^n \ldots Y_1^n Z_M^n \ldots Z_1^n}, p_{X^n Y_K^n \ldots Y_1^n Z_M^n \ldots Z_1^n})$, as well as by the invertibility of G_n, $\mathbb{V}(\tilde{q}_{U[(\cup_{i=m}^M \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] Z_m^n}, p_{U[(\cup_{i=m}^M \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] Z_m^n}) \leq \mathbb{V}(\tilde{q}_{X^n Y_K^n \ldots Y_1^n Z_M^n \ldots Z_1^n}, p_{X^n Y_K^n \ldots Y_1^n Z_M^n \ldots Z_1^n})$. Hence, we have:

$$\begin{aligned} H(\tilde{U}[(\cup_{i=m}^M \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] | \tilde{Z}_m^n) &\geq H(U[(\cup_{i=m}^M \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}] | Z_m^n) - (3n\delta_{\text{nld-ls}}^{(n)} - 2\delta_{\text{nld-ls}}^{(n)} \log \delta_{\text{nld-ls}}^{(n)}) \\ &\overset{(a)}{\geq} \textstyle\sum_{j \in (\cup_{i=m}^M \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)}} H(U(j) | U^{1:j-1}, Z_m^n) - (3n\delta_{\text{nld-ls}}^{(n)} - 2\delta_{\text{nld-ls}}^{(n)} \log \delta_{\text{nld-ls}}^{(n)}) \\ &\overset{(b)}{\geq} \left(\textstyle\sum_{i=m}^M |\mathcal{I}_i^{(n)}| + |\mathcal{F}^{(n)}| \right) (1 - \delta_n) - (3n\delta_{\text{nld-ls}}^{(n)} - 2\delta_{\text{nld-ls}}^{(n)} \log \delta_{\text{nld-ls}}^{(n)}), \end{aligned} \tag{24}$$

where (a) holds because conditioning does not increase the entropy and (b) holds because, according to Equations (12)–(14) and Lemma 1, $(\cup_{i=m}^M \mathcal{I}_i^{(n)}) \cup \mathcal{F}^{(n)} \subseteq \mathcal{H}_{X|Z_m}^{(n)}$, as well as by the definition of $\mathcal{H}_{X|Z_m}^{(n)}$ in Equation (11).

Finally, by substituting Equation (24) into Equation (22), for n sufficiently large, we obtain:

$$I(W_m, \ldots, W_M; F, \tilde{Z}_m^n) \leq n\delta_n + 3n\delta_{\text{nld-ls}}^{(n)} - 2\delta_{\text{nld-ls}}^{(n)} \log \delta_{\text{nld-ls}}^{(n)}, \tag{25}$$

Hence, the polar code satisfies the strong secrecy condition in Equation (3), and the proof of Theorem 1 is concluded.

4.5.5. Reuse of the Source of Common Randomness

Consider that the transmission takes place over B blocks of size n. We use the subscript $b \in [1, B]$ between parentheses to denote random variables associated with the block b. From Lemma 2, we have $\mathbb{V}(\tilde{q}_{U_{(b)}^n}, p_{U^n}) \leq \delta_{\text{nld-ls}}^{(n)}$ for any $b \in [1, B]$ because we use the same encoding of Equation (17) at each block. Hence, by the union bound, the polar code satisfies the reliability condition given in Equation (2) because:

$$\mathbb{P}\left[\cup_{b=1}^B \{ \hat{U}_{(b)}^n \neq \tilde{U}_{(b)}^n \} \right] \leq \sum_{b=1}^B \mathbb{P}[\hat{U}_{(b)}^n \neq \tilde{U}_{(b)}^n] \leq B(n\sqrt{\delta_n} + \delta_{\text{nld-ls}}^{(n)}),$$

where the last inequality follows from the fact that, since F and $\Phi_{(b)}$ are perfectly known, $\mathbb{P}[\hat{U}_{(b)}^n \neq \tilde{U}_{(b)}^n]$ only depends on the decoding at block b and, consequently, can be bounded as in Equation (21).

With a slight abuse of notation, let $W_{m:M,(b_1:b_2)}$, where $1 \leq b_1 \leq b_2 \leq B$, denote the messages $\{(W_{m,(b)}, \ldots, W_{M,(b)})\}_{b=b_1}^{b_2}$. It remains to show that $W_{m:M,(1:B)}$ is asymptotically statistically independent of $(F, \tilde{Z}_{m,(1:B)}^n)$. Since F is reused at each block, we have to consider the dependencies between the random variables of different blocks that are involved in the secrecy analysis. According to these dependencies, which are represented in the Bayesian graph of Figure 4, we obtain:

$$\begin{aligned} I(W_{m:M,(1:B)}; \tilde{Z}_{m,(1:B)}^n, F) &\overset{(a)}{=} I(W_{m:M,(1:B)}; \tilde{Z}_{m,(1:B)}^n | F) \\ &= \sum_{b=0}^{B-1} I(W_{m:M,(1:B)}; \tilde{Z}_{m,(b+1)}^n | F, \tilde{Z}_{m,(1:b)}^n) \\ &\overset{(b)}{\leq} B\big(n\delta_n + 3n\delta_{\text{nld-ls}}^{(n)} - 2\delta_{\text{nld-ls}}^{(n)} \log \delta_{\text{nld-ls}}^{(n)}\big), \end{aligned}$$

where (a) follows from the independence between $W_{m:M,(1:B)}$ and F, and (b) holds because:

$$\begin{aligned}
&I(W_{m:M,(1:B)}; \tilde{Z}^n_{m,(b+1)} | F, \tilde{Z}^n_{m,(1:b)}) \\
&\quad = I(W_{m:M,(1:b+1)}; \tilde{Z}^n_{m,(b+1)} | F, \tilde{Z}^n_{m,(1:b)}) + I(W_{m:M,(b+2:B)}; \tilde{Z}^n_{m,(b+1)} | F, \tilde{Z}^n_{m,(1:b)}, W_{m:M,(1:b+1)}) \\
&\quad \leq I(W_{m:M,(1:b+1)}, F, \tilde{Z}^n_{m,(1:b)}; \tilde{Z}^n_{m,(b+1)}) + I(W_{m:M,(b+2:B)}; \tilde{Z}^n_{m,(1:b+1)}, F, W_{m:M,(1:b+1)}) \\
&\quad \overset{(a)}{=} I(W_{m:M,(1:b+1)}, F, \tilde{Z}^n_{m,(1:b)}; \tilde{Z}^n_{m,(b+1)}) \\
&\quad \leq I(W_{m:M,(b+1)}, F; \tilde{Z}^n_{m,(b+1)}) + I(W_{m:M,(1:b)}, \tilde{Z}^n_{m,(1:b)}; \tilde{Z}^n_{m,(b+1)} | W_{m:M,(b+1)}, F) \\
&\quad \overset{(b)}{\leq} \left(n\delta_n + 3n\delta^{(n)}_{\text{nld-ls}} - 2\delta^{(n)}_{\text{nld-ls}} \log \delta^{(n)}_{\text{nld-ls}}\right) + I(W_{m:M,(1:b)}, \tilde{Z}^n_{m,(1:b)}; W_{m:M,(b+1)}, \tilde{Z}^n_{m,(b+1)} | F) \\
&\quad \overset{(c)}{=} n\delta_n + 3n\delta^{(n)}_{\text{nld-ls}} - 2\delta^{(n)}_{\text{nld-ls}} \log \delta^{(n)}_{\text{nld-ls}},
\end{aligned}$$

where (a) holds because the messages at blocks $b+2$–B are independent of F and all the random variables of the previous blocks, (b) follows from Equation (25) and (c) holds by applying d-separation [31] over the graph of Figure 4 because $(W_{m:M,(1:b)}, \tilde{Z}^n_{m,(1:b)}) \leftarrow F \rightarrow (W_{m:M,(b+1)}, \tilde{Z}^n_{m,(b+1)})$ forms a common cause and, consequently, $(W_{m:M,(1:b)}, \tilde{Z}^n_{m,(1:b)})$ and $(W_{m:M,(b+1)}, \tilde{Z}^n_{m,(b+1)})$ are independent given F.

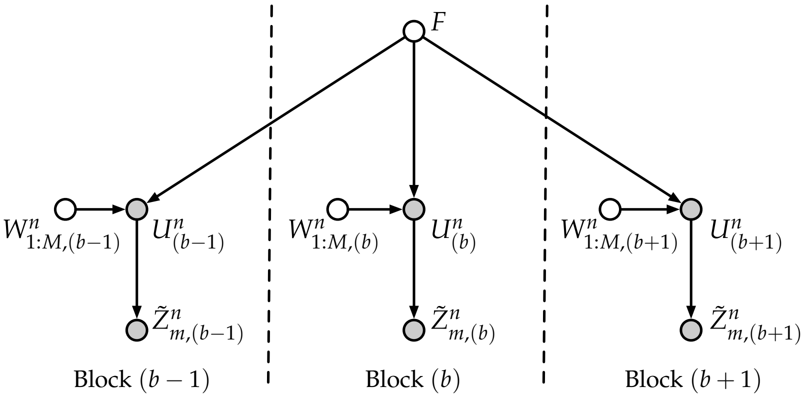

Figure 4. Bayesian graph plotting the dependencies between the random variables of different blocks that are involved in the secrecy analysis when we consider a transmission over several blocks of size n.

5. Polar Coding Scheme for the DBC-LD-NLS

The polar coding scheme provided in this section is designed to achieve the supremum of the achievable rates given in Corollary 2 (secrecy-capacity without rate sharing). In this model, there are K input random variables $\{V_\ell\}_{\ell=1}^{K}$ (where $V_K \triangleq X$), each one corresponding to a different superposition layer. Consider the DMS $\left(\mathcal{V}_1 \times \cdots \times \mathcal{V}_K \times \mathcal{Y}_K \times \cdots \times \mathcal{Y}_1 \times \mathcal{Z}_M \times \cdots \times \mathcal{Z}_1, p_{V_1 \ldots V_K Y_K \ldots Y_1 Z_M \ldots Z_1}\right)$ that represents the input and output random variables involved in the achievable subregion of Corollary 2, where $\mathcal{V}_\ell = \{0,1\}$ for any $\ell \in [1,K]$. Let $(V_1^n, \ldots, V_K^n, Y_K^n, \ldots, Y_1^n, Z_M^n, \ldots, Z_1^n)$ be an i.i.d. n-sequence of this source. Then, we define the K polar transforms $U_\ell^n \triangleq V_\ell^n G_n$, where $\ell \in [1,K]$. Since $V_1 - V_2 - \cdots - V_K$ and, consequently, $U_1 - U_2 - \cdots - U_K$ (by the invertibility of G_n) form a Markov chain, the joint distribution of $(U_1^n, \ldots, U_K^n)$ satisfies"

$$p_{U_1^n \ldots U_K^n}(u_1^n, \ldots, u_K^n) \triangleq \prod_{\ell=1}^{K} \prod_{j=1}^{n} p_{U_\ell(j) | U_\ell^{1:j-1} V_{\ell-1}^n}\left(u_\ell(j) \middle| u_\ell^{1:j-1}, u_{\ell-1}^n G_n\right). \tag{26}$$

5.1. Polar Code Construction

Based on $p_{V_1 \dots V_K Y_K \dots Y_1 Z_M \dots Z_1}$, the construction is carried out similarly at each superposition layer. Consider the polar construction at layer $\ell \in [1, K]$. Let $\delta_n \triangleq 2^{-n^{\beta}}$, where $\beta \in (0, \frac{1}{2})$. For the polar transform $U_\ell^n = V_\ell^n G_n$ associated with the ℓ-th layer, we define the sets:

$$\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}} \triangleq \{j \in [n] : H(U_\ell(j)|U_\ell^{1:j-1}, V_{\ell-1}^n) \geq 1 - \delta_n\}, \tag{27}$$

$$\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}} \triangleq \{j \in [n] : H(U_\ell(j)|U_\ell^{1:j-1}, V_{\ell-1}^n) \leq \delta_n\}, \tag{28}$$

$$\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_k} \triangleq \{j \in [n] : H(U_\ell(j)|U_\ell^{1:j-1}, V_{\ell-1}^n, Y_k^n) \leq \delta_n\}, \quad k = \ell, \dots, K, \tag{29}$$

$$\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Y_k} \triangleq \{j \in [n] : H(U_\ell(j)|U_\ell^{1:j-1}, V_{\ell-1}^n, Y_k^n) \geq 1 - \delta_n\}, \quad k = \ell, \dots, K, \tag{30}$$

$$\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_m} \triangleq \{j \in [n] : H(U_\ell(j)|U_\ell^{1:j-1}, V_{\ell-1}^n, Z_m^n) \geq 1 - \delta_n\}, \quad m = 1, \dots, M, \tag{31}$$

where we recall that $V_0 = \varnothing$ when $\ell = 1$ and $V_K \triangleq X$ when $\ell = K$. At each layer $\ell \in [1, K]$, based on these previous sets, we define the following partition of the universal set $[n]$,

$$\mathcal{I}_\ell^{(n)} \triangleq \mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M} \cap \left(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell}\right)^{\mathrm{C}}, \tag{32}$$

$$\mathcal{F}_\ell^{(n)} \triangleq \mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell}, \tag{33}$$

$$\mathcal{C}_\ell^{(n)} \triangleq \mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}} \cap \left(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M}\right)^{\mathrm{C}}, \tag{34}$$

$$\mathcal{T}_\ell^{(n)} \triangleq \left(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}}\right)^{\mathrm{C}}, \tag{35}$$

which is graphically represented in Figure 5. The way we define this partition at the ℓ-th layer follows similar reasoning as the one to define the partition in Section 4.1 for the DBC-NLD-LS. In this sense, $U_\ell[\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}}]$ will be suitable for storing uniformly-distributed random sequences. Otherwise, $U_\ell[\mathcal{T}_\ell^{(n)}]$ will not and $U_\ell(j)$ such that $j \in \mathcal{T}_\ell^{(n)}$ will be constructed somehow from $(U_\ell^{1:j-1}, V_{\ell-1})$ and the distribution $p_{U_\ell(j)|U_\ell^{1:j-1}V_{\ell-1}^n}$. Now, $U_\ell[\mathcal{I}_\ell^{(n)}]$ will be suitable for storing information to be secured from all eavesdroppers because $\mathcal{I}_\ell^{(n)}$ belongs to $\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M}$, and by Lemma 1, $\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M} \subseteq \mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_{m'}}$ for any $m' \in [1, M-1]$. Since $\mathcal{C}_\ell^{(n)} \subseteq (\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M})^{\mathrm{C}}$, $U[\mathcal{C}_\ell^{(n)}]$ will be used to store the local randomness required to confuse all eavesdroppers about the secret information carried on this layer. According to [27] (Theorem 2), the legitimate receiver $k \in [1, K]$ will be able to reliably infer $U_\ell[\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_k}]$ given Y_k^n and $U_\ell[(\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_k})^{\mathrm{C}}]$. By Lemma 1, we have $(\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell})^{\mathrm{C}} \supseteq (\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_k})^{\mathrm{C}}$ for any $\ell < k$. Therefore, given $U_\ell[(\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell})^{\mathrm{C}}]$, the legitimate receivers ℓ–K will be able to reliably reconstruct U_ℓ^n from its own channel observations. In this sense, $U_\ell[\mathcal{F}_\ell^{(n)}]$ will be used to store the random sequence provided by the source of common randomness. Since $\mathcal{F}_\ell^{(n)} \subseteq \mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M}$, the strong secrecy condition will not be compromised. On the other hand, $U[(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell})^{\mathrm{C}}]$ (hatched areas in Figure 5) will contain secret information or elements that cannot be known directly by the eavesdroppers. Therefore, the transmitter somehow will make those elements available to the legitimate receivers ℓ–K keeping them masked from all eavesdroppers by incurring an asymptotically-negligible rate penalty.

As mentioned in Remark 3, the goal of the polar construction is to obtain the entropy terms associated with the sets in Equations (27)–(31) and then define the partition of $[n]$ given in Equations (32)–(35).

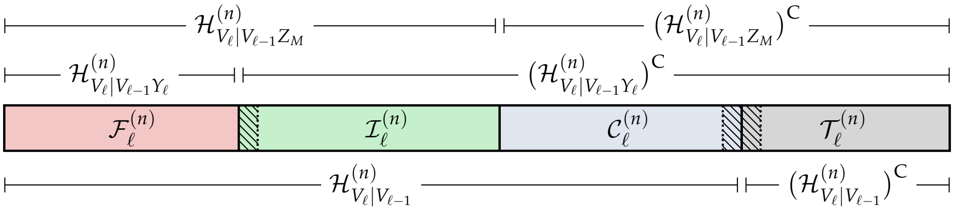
$\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M}$
$(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M})^{\mathrm{C}}$
$\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell}$
$(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell})^{\mathrm{C}}$
$\mathcal{F}^{(n)}_\ell$
$\mathcal{I}^{(n)}_\ell$
$\mathcal{C}^{(n)}_\ell$
$\mathcal{T}^{(n)}_\ell$
$\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}}$
$(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}})^{\mathrm{C}}$

Figure 5. Polar code construction for the DBC-LD-NLS at the ℓ-th layer. The hatched area represents those indices $j \in (\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell})^{\mathrm{C}}$, which can belong to the sets $\mathcal{I}^{(n)}_\ell$, $\mathcal{C}^{(n)}_\ell$ or $\mathcal{T}^{(n)}_\ell$.

5.2. Polar Encoding

The superposition-based polar encoder will consist of K encoding blocks operating sequentially at each superposition layer, the block at layer $\ell \in [1, K]$ being responsible for the construction of $\tilde{U}^n_\ell$. In order to construct $\tilde{U}^n_\ell$ for some $\ell \in [2, K]$, the encoder block needs $\tilde{V}^n_{\ell-1} = \tilde{U}^n_{\ell-1} G_n$, which have been constructed previously by the encoding block operating at the $(\ell - 1)$-th layer.

Consider the encoding procedure at layer $\ell \in [1, K]$. Let W_ℓ and C_ℓ be uniformly-distributed random vectors of size $|\mathcal{I}^{(n)}_\ell|$ and $|\mathcal{C}^{(n)}_\ell|$, respectively, where W_ℓ represents the message intended for receivers ℓ–K and C_ℓ the local randomness required at the ℓ-th layer to confuse all eavesdroppers about this message. Let F_ℓ be a given uniformly-distributed random $|\mathcal{F}^{(n)}_\ell|$-sequence, which represents the source of common randomness that is available to all parties. The ℓ-th encoding block constructs the sequence $\tilde{u}^n_\ell$ as follows. Given the realizations w_ℓ, c_ℓ and f_ℓ, whose elements have been indexed by the set of indices $\mathcal{I}^{(n)}_\ell$, $\mathcal{C}^{(n)}_\ell$ and $\mathcal{F}^{(n)}_\ell$, respectively, and given $\tilde{v}^n_{\ell-1} = \tilde{u}^n_{\ell-1} G_n$ provided by the previous encoding block (recall that $\tilde{v}^n_0 \triangleq \varnothing$ at the first layer), the ℓ-th encoding block draws $\tilde{u}^n_\ell$ from:

$$\tilde{q}_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}\big(\tilde{u}_\ell(j)|\tilde{u}_\ell^{1:j-1}, \tilde{v}^n_{\ell-1}\big)$$
$$\triangleq \begin{cases} \mathbb{1}\{\tilde{u}_\ell(j) = w_\ell(j)\} & \text{if } j \in \mathcal{I}^{(n)}_\ell, \\ \mathbb{1}\{\tilde{u}_\ell(j) = c_\ell(j)\} & \text{if } j \in \mathcal{C}^{(n)}_\ell, \\ \mathbb{1}\{\tilde{u}_\ell(j) = f_\ell(j)\} & \text{if } j \in \mathcal{F}^{(n)}_\ell, \\ p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}\big(\tilde{u}_\ell(j)|\tilde{u}_\ell^{1:j-1}, \tilde{v}^n_{\ell-1}\big) & \text{if } j \in \big(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}}\big)^{\mathrm{C}} \cap \big(\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}}\big)^{\mathrm{C}}, \\ \mathbb{1}\{\tilde{u}_\ell(j) = \xi^{(j)}_\ell(\tilde{u}_\ell^{1:j-1}, \tilde{v}^n_{\ell-1})\} & \text{if } j \in \mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}}, \end{cases} \tag{36}$$

where:

$$\xi^{(j)}_\ell\big(\tilde{u}_\ell^{1:j-1}, \tilde{v}^n_{\ell-1}\big) \triangleq \arg\max_{u \in \mathcal{V}_\ell} p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}\big(u|\tilde{u}_\ell^{1:j-1}, \tilde{v}^n_{\ell-1}\big), \tag{37}$$

$p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}$ being the distribution induced by the original DMS. Notice that $\mathcal{T}^{(n)}_\ell = ((\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}})^{\mathrm{C}}) \cup \mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}}$, and similarly to the previous model, $\tilde{U}[\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}}]$ is constructed in a deterministic way by adapting the SC encoding algorithm in [20]; and $\tilde{U}[(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}})^{\mathrm{C}}]$ is constructed randomly. By [27] (Theorem 1), the rate of the amount of randomness for SC encoding will be asymptotically negligible. After constructing $\tilde{U}^n_\ell$, the ℓ-th encoding block computes the sequence $\tilde{V}^n_\ell = \tilde{U}^n_\ell G_n$ and delivers it to the next encoding block. If $\ell = K$, then $\tilde{V}^n_K \triangleq \tilde{X}^n$, and the encoder transmits it over the DBC, which induces the channel outputs $(\tilde{Y}^n_K, \ldots, \tilde{Y}^n_1, \tilde{Z}^n_M, \ldots, \tilde{Z}^n_1)$.

Finally, besides the sequence $\tilde{X}^n$, the encoder outputs the following additional secret sequences,

$$\Phi_\ell \triangleq \tilde{U}_\ell\Big[\big(\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell}\big)^{\mathrm{C}} \cap \big(\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell}\big)^{\mathrm{C}}\Big], \quad \ell = 1, \ldots, K, \tag{38}$$

The sequence Φ_ℓ corresponding to the layer $\ell \in [1, K]$ must be additionally transmitted to the legitimate receivers ℓ–K keeping it masked from the eavesdroppers. To do so, the transmitter can perform a modulo-two addition between $\{\Phi_\ell\}_{\ell=1}^{K}$ and a uniformly-distributed secret key privately shared with the legitimate receivers and somehow additionally send it to them. If $K \ll n$, by [27] (Theorem 1), we have that the overall rate required to transmit these additional secret sequences is asymptotically negligible, i.e., $\lim_{n\to\infty} \sum_{\ell=1}^{K} \frac{|\Phi_\ell|}{n} = 0$. As for the previous model, the uniformly-distributed part of any Φ_ℓ could be made available to the corresponding legitimate receivers by using a chaining structure as in [9]. However, this approach will present the same disadvantages as those mentioned in Remark 4.

5.3. Polar Decoding

Consider that the realizations of $\{F_\ell\}_{\ell=1}^{K}$ are available to all parties, and the sequences $\{\Phi_\ell\}_{\ell=1}^{K}$ have been successfully received by the corresponding legitimate receivers before the decoding process.

Consider the decoding at the legitimate receiver $k \in [1, K]$. This receiver forms the estimates $\{\hat{U}_\ell^n\}_{\ell=1}^{k}$ of the sequences $\{\tilde{U}_\ell^n\}_{\ell=1}^{k}$ in a successive manner from $\hat{U}_1^n$-$\hat{U}_k^n$, and the procedure to estimate $\tilde{U}_\ell^n$ for some $\ell \in [1, k]$ is as follows. First, given that Φ_ℓ and F_ℓ are available, the receiver knows $\tilde{U}_\ell[(\mathcal{L}_{V_\ell|V_{\ell-1}Y_\ell}^{(n)})^{\mathrm{C}}]$. Moreover, by Lemma 1, $(\mathcal{L}_{V_\ell|V_{\ell-1}Y_k}^{(n)})^{\mathrm{C}} \subseteq (\mathcal{L}_{V_\ell|V_{\ell-1}Y_\ell}^{(n)})^{\mathrm{C}}$ for any $\ell < k$. Thus, given $\tilde{U}_\ell[(\mathcal{L}_{V_\ell|V_{\ell-1}Y_\ell}^{(n)})^{\mathrm{C}}]$, the k-th legitimate receiver performs SC decoding for source coding with side information [27] to construct $\hat{U}_\ell[\mathcal{L}_{V_\ell|V_{\ell-1}Y_\ell}^{(n)}]$ from $\tilde{Y}_k^n$, and from $\hat{V}_{\ell-1}^n = \hat{U}_{\ell-1}^n G_n$ estimated previously. In Section 5.5.3, we show formally that the polar coding scheme satisfies the reliability condition in Equation (4).

5.4. Information Leakage

Besides the observations $\tilde{Z}_m^n$, the eavesdropper $m \in [1, M]$ has access to the common randomness $\{F_\ell\}_{\ell=1}^{K}$. Therefore, the information about all messages leaked to the m-th eavesdropper is:

$$I(W_1, \ldots, W_K; F_1, \ldots, F_K, \tilde{Z}_m^n) = I(\tilde{U}_1[\mathcal{I}_1^{(n)}], \ldots, \tilde{U}_K[\mathcal{I}_K^{(n)}]; \tilde{U}_1[\mathcal{F}_1^{(n)}], \ldots, \tilde{U}_K[\mathcal{F}_K^{(n)}], \tilde{Z}_m^n). \quad (39)$$

In Section 5.5.4, we prove that $(W_1, \ldots, W_K)$ is asymptotically statistically independent of $(F_1, \ldots, F_K, \tilde{Z}_m^n)$.

5.5. Performance of the Polar Coding Scheme

The analysis of the polar coding scheme leads to the following theorem.

Theorem 2. *Consider an arbitrary DBC* $(\mathcal{X}, p_{Y_K \cdots Y_1 Z_M \cdots Z_1|X}, \mathcal{Y}_K \times \cdots \times \mathcal{Y}_1 \times \mathcal{Z}_M \times \cdots \times \mathcal{Z}_1)$ *such that* $\mathcal{X} \in \{0, 1\}$ *and* $p_{Y_K \cdots Y_1 Z_M \cdots Z_1|X}$ *satisfies the Markov chain condition* $X - Y_K - \cdots - Y_1 - Z_M - \cdots - Z_1$. *The polar coding scheme described in Sections 5.1–5.4 achieves any rate tuple of the achievable region defined in Corollary 2, satisfying the reliability and strong secrecy conditions in Equations (4) and (5), respectively.*

Corollary 5. *Since* $\tilde{U}_\ell[\mathcal{I}_\ell^{(n)}]$ *for some* $\ell \in [1, K]$ *can contain any information to be reliably decoded by the legitimate receivers* ℓ*–*K*, the coding scheme in Sections 5.1–5.4 can achieve the entire region considering the rate sharing of Proposition 2 by storing part of any message* $W_{\ell'}$ *such that* $\ell' > \ell$ *into* $\tilde{U}_\ell[\mathcal{I}_\ell^{(n)}]$ *instead of part of* W_ℓ.

Corollary 6. *If we consider a communication scenario requiring transmissions over several blocks of size* n*, the same realization of the source of common randomness* $(F_1, \ldots, F_K)$ *that is known by all parties could be used at each block, and the reliability and the strong secrecy conditions would still be ensured.*

As in Theorem 1, the proof of Theorem 2 follows in four steps and is provided in Sections 4.5.1–4.5.4. The proof of Corollary 5 is immediate. The proof of Corollary 6 is omitted because it follows similar reasoning as in Corollary 4. Despite that in this model, we have different superposition layers, the dependencies between the random variables at different blocks have the same structure of those graphically represented in Figure 4.

5.5.1. Transmission Rates

We prove that the polar coding scheme approaches the corner point of the subregion defined in Corollary 2. For any $\ell \in [1, K]$, the transmission rate R_ℓ corresponding to the message W_ℓ satisfies:

$$\begin{aligned} \lim_{n\to\infty} R_\ell &= \lim_{n\to\infty} \tfrac{1}{n}\big|\mathcal{I}_\ell^{(n)}\big| \overset{(a)}{=} \lim_{n\to\infty} \tfrac{1}{n}\Big|\mathcal{H}_{V_\ell|V_{\ell-1}Z_M}^{(n)} \cap \big(\mathcal{H}_{V_\ell|V_{\ell-1}Y_\ell}^{(n)}\big)^{\mathrm{C}}\Big| \\ &\overset{(b)}{=} \lim_{n\to\infty} \tfrac{1}{n}\big|\mathcal{H}_{V_\ell|V_{\ell-1}Z_M}^{(n)}\big| - \lim_{n\to\infty} \tfrac{1}{n}\big|\mathcal{H}_{V_\ell|V_{\ell-1}Y_\ell}^{(n)}\big| \\ &\overset{(c)}{=} I(V_\ell; Y_\ell|V_{\ell-1}) - I(V_\ell; Z_M|V_{\ell-1}), \end{aligned} \tag{40}$$

where (a) follows from the definition of the set $\mathcal{I}_\ell^{(n)}$ in Equation (32), (b) holds because, by Lemma 1, $\mathcal{H}_{V_\ell|V_{\ell-1}Z_M}^{(n)} \supseteq \mathcal{H}_{V_\ell|V_{\ell-1}Y_\ell}^{(n)}$, and (c) holds by [27] (Theorem 1).

5.5.2. Distribution of the DMS after the Polar Encoding

Let $\tilde{q}_{U_1^n\ldots U_K^n}$ be the distribution of $(\tilde{U}_1^n, \ldots, \tilde{U}_K^n)$ after the encoding in Section 5.2. The following lemma shows that $\tilde{q}_{U_1^n\ldots U_K^n}$ and $p_{U_1^n\ldots U_K^n}$ of the DMS are nearly statistically indistinguishable for sufficiently large n and, consequently, so are the overall distributions $\tilde{q}_{V_1^n\ldots V_K^n Y_K^n\ldots Y_1^n Z_M^n\ldots Z_1^n}$ and $p_{V_1^n\ldots V_K^n Y_K^n\ldots Y_1^n Z_M^n\ldots Z_1^n}$.

Lemma 3. *Let $\delta_n = 2^{-n^\beta}$ for some $\beta \in (0, \frac{1}{2})$. Then,*

$$\mathbb{V}(\tilde{q}_{U_1^n\ldots U_K^n}, p_{U_1^n\ldots U_K^n}) \le \delta_{\text{ld-nls}}^{(n)},$$
$$\mathbb{V}(\tilde{q}_{V_1^n\ldots V_K^n Y_K^n\ldots Y_1^n Z_M^n\ldots Z_1^n}, p_{V_1^n\ldots V_K^n Y_K^n\ldots Y_1^n Z_M^n\ldots Z_1^n}) = \mathbb{V}(\tilde{q}_{U_1^n\ldots U_K^n}, p_{U_1^n\ldots U_K^n}) \le \delta_{\text{ld-nls}}^{(n)},$$

where $\delta_{\text{ld-nls}}^{(n)} \triangleq Kn\sqrt{4\sqrt{n\delta_n \ln 2}\big(2n - \log\big(2\sqrt{n\delta_n \ln 2}\big)\big) + \delta_n} + \sqrt{K2n\delta_n \ln 2}.$

Proof. See Appendix A setting $L = K$. □

Remark 6. *The first term of $\delta_{\text{ld-nls}}^{(n)}$ bounds the impact on the total variation distance of using the deterministic SC encoding in Equation (37) for $\tilde{U}_\ell\big[\mathcal{L}_{V_\ell|V_{\ell-1}}^{(n)}\big]$ at each layer $\ell \in [1, K]$. The second term bounds the impact of storing uniformly-distributed random sequences that are independent of $\tilde{V}_{\ell-1}^n$ into $\tilde{U}_\ell\big[\mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}\big]$.*

5.5.3. Reliability Performance

Consider the probability of incorrectly decoding $\{W_\ell\}_{\ell=1}^k$ at the legitimate receiver $k \in [1, K]$. Let $\tilde{q}_{V_\ell^n Y_k^n}$ and $p_{V_\ell^n Y_k^n}$ for any $\ell \le k$ be marginals of $\tilde{q}_{V_1^n\ldots V_K^n Y_K^n\ldots Y_1^n Z_M^n\ldots Z_1^n}$ and $p_{V_1^n\ldots V_K^n Y_K^n\ldots Y_1^n Z_M^n\ldots Z_1^n}$, respectively. Consider an optimal coupling [29] (Proposition 4.7) between $\tilde{q}_{V_\ell^n Y_k^n}$ and $p_{V_\ell^n Y_k^n}$ such that:

$$\mathbb{P}\big[\mathcal{E}_{V_\ell^n Y_k^n}\big] = \mathbb{V}(\tilde{q}_{V_\ell^n Y_k^n}, p_{V_\ell^n Y_k^n}),$$

where $\mathcal{E}_{V_\ell^n Y_k^n} \triangleq \{(\tilde{V}_\ell^n, \tilde{Y}_k^n) \neq (V_\ell^n, Y_k^n)\}$ or, equivalently, $\mathcal{E}_{V_\ell^n Y_k^n} \triangleq \{(\tilde{U}_\ell^n, \tilde{Y}_k^n) \neq (U_\ell^n, Y_k^n)\}$ due to the invertibility of G_n. Furthermore, for all $\ell \in [1, k]$, we define the error events $\mathcal{E}_{\hat{V}_\ell^n} \triangleq \{\hat{V}_\ell^n \neq \tilde{V}_\ell^n\}$ or, equivalently, $\mathcal{E}_{\hat{V}_\ell^n} \triangleq \{\hat{U}_\ell^n \neq \tilde{U}_\ell^n\}$; and we define $\mathcal{E}_{\hat{V}_0^n} \triangleq \emptyset$. Hence, for any $\ell \in [1, k]$, the average probability of incorrectly decoding the message W_ℓ at the k-th receiver can be upper-bounded as:

$$\begin{aligned}
\mathbb{P}[\hat{W}_\ell \neq W_\ell] \quad &\leq \mathbb{P}\big[\hat{U}_\ell^n \neq \tilde{U}_\ell^n\big] \\
&= \mathbb{P}\big[\hat{U}_\ell^n \neq \tilde{U}_\ell^n \big| \mathcal{E}^{\mathrm{C}}_{V_\ell^n Y_k^n} \cap \mathcal{E}^{\mathrm{C}}_{\hat{V}_{\ell-1}^n}\big]\mathbb{P}\big[\mathcal{E}^{\mathrm{C}}_{V_\ell^n Y_k^n} \cap \mathcal{E}^{\mathrm{C}}_{\hat{V}_{\ell-1}^n}\big] \\
&\quad +\mathbb{P}\big[\hat{U}_\ell^n \neq \tilde{U}_\ell^n \big| \mathcal{E}_{V_\ell^n Y_k^n} \cup \mathcal{E}_{\hat{V}_{\ell-1}^n}\big]\mathbb{P}\big[\mathcal{E}_{V_\ell^n Y_k^n} \cup \mathcal{E}_{\hat{V}_{\ell-1}^n}\big] \\
&\leq \mathbb{P}\big[\hat{U}_\ell^n \neq \tilde{U}_\ell^n \big| \mathcal{E}^{\mathrm{C}}_{V_\ell^n Y_k^n} \cap \mathcal{E}^{\mathrm{C}}_{\hat{V}_{\ell-1}^n}\big] + \mathbb{P}\big[\mathcal{E}_{V_\ell^n Y_k^n}\big] + \mathbb{P}\big[\mathcal{E}_{\hat{V}_{\ell-1}^n}\big] \\
&\overset{(a)}{\leq} \textstyle\sum_{j \in \mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell}} Z\big(U_\ell(j)\big|U_\ell^{1:j-1}, V_{\ell-1}^n, Y_k^n\big) + \mathbb{P}\big[\mathcal{E}_{V_\ell^n Y_k^n}\big] + \mathbb{P}\big[\mathcal{E}_{\hat{V}_{\ell-1}^n}\big] \\
&\overset{(b)}{\leq} n\sqrt{\delta_n} + \mathbb{P}\big[\mathcal{E}_{V_\ell^n Y_k^n}\big] + \mathbb{P}\big[\mathcal{E}_{\hat{V}_{\ell-1}^n}\big] \\
&\overset{(c)}{\leq} n\sqrt{\delta_n} + \delta^{(n)}_{\text{ld-nls}} + \mathbb{P}\big[\mathcal{E}_{\hat{V}_{\ell-1}^n}\big]
\end{aligned} \tag{41}$$

where (a) holds by [27] (Theorem 2) because $\tilde{U}_\ell[(\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_\ell})^{\mathrm{C}}]$ for any $\ell \leq k$ is available to the k-th receiver, (b) holds by Lemma 1, by the definition of the set $\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}Y_1}$ in Equation (29) and by applying [27] (Proposition 2) and (c) holds by the optimal coupling and Lemma 3 because $\mathbb{V}(\tilde{q}_{V_\ell^n Y_k^n}, p_{V_\ell^n Y_k^n}) \leq \mathbb{V}(\tilde{q}_{V_1^n \ldots V_K^n Y_K^n \ldots Y_1^n Z_M^n \ldots Z_1^n}, p_{V_1^n \ldots V_K^n Y_K^n \ldots Y_1^n Z_M^n \ldots Z_1^n})$. Thus, by induction, we obtain:

$$\mathbb{P}\big[(\hat{W}_1, \ldots \hat{W}_k) \neq (W_1, \ldots, W_k)\big] \leq \sum_{\ell=1}^{k} \mathbb{P}[\hat{U}_\ell \neq \tilde{U}_\ell] \leq \frac{k(k+1)}{2}\big(n\sqrt{\delta_n} + \delta^{(n)}_{\text{ld-nls}}\big). \tag{42}$$

Consequently, if $K \ll n$, the polar coding scheme satisfies the reliability condition in Equation (4).

5.5.4. Secrecy Performance

Consider the leakage at the eavesdropper $m \in [1, M]$ given in Equation (39). As in Equation (22), we obtain:

$$I(W_1, \ldots, W_K; F_1, \ldots, F_K, \tilde{Z}_m^n) \leq \sum_{\ell=1}^{K} \big|\mathcal{I}_\ell^{(n)} \cup \mathcal{F}_\ell^{(n)}\big| - H\big(\tilde{U}_1\big[\mathcal{I}_1^{(n)} \cup \mathcal{F}_1^{(n)}\big], \ldots, \tilde{U}_K\big[\mathcal{I}_K^{(n)} \cup \mathcal{F}_K^{(n)}\big]\big|\tilde{Z}_m^n\big). \tag{43}$$

Following similar reasoning as in Equation (23), for n large enough, we have:

$$\begin{aligned}
&\Big|H\big(\tilde{U}_1\big[\mathcal{I}_1^{(n)} \cup \mathcal{F}_1^{(n)}\big], \ldots, \tilde{U}_K\big[\mathcal{I}_K^{(n)} \cup \mathcal{F}_K^{(n)}\big]\big|\tilde{Z}_m^n\big) - H\big(U_1\big[\mathcal{I}_1^{(n)} \cup \mathcal{F}_1^{(n)}\big], \ldots, U_K\big[\mathcal{I}_K^{(n)} \cup \mathcal{F}_K^{(n)}\big]\big|Z_m^n\big)\Big| \\
&\overset{(a)}{\leq} \mathbb{V}(\tilde{q}_{Z_m^n}, p_{Z_m^n}) \log \tfrac{2^n}{\mathbb{V}(\tilde{q}_{Z_m^n}, p_{Z_m^n})} + \mathbb{V}^{\dagger} \log \tfrac{2^{(n + \sum_{\ell=1}^{K} |\mathcal{I}_\ell^{(n)} \cup \mathcal{F}_\ell^{(n)}|)}}{\mathbb{V}^{\dagger}} \\
&\overset{(b)}{\leq} (K+2)n\delta^{(n)}_{\text{ld-nls}} - 2\delta^{(n)}_{\text{ld-nls}} \log \delta^{(n)}_{\text{ld-nls}},
\end{aligned} \tag{44}$$

where (a) holds by defining $\mathbb{V}^{\dagger} \triangleq \mathbb{V}(\tilde{q}_{U_1[\mathcal{I}_1^{(n)} \cup \mathcal{F}_1^{(n)}], \ldots, U_K[\mathcal{I}_K^{(n)} \cup \mathcal{F}_K^{(n)}] Z_m^n}, p_{U_1[\mathcal{I}_1^{(n)} \cup \mathcal{F}_1^{(n)}], \ldots, U_K[\mathcal{I}_K^{(n)} \cup \mathcal{F}_K^{(n)}] Z_m^n})$ and [30] (Lemma 2.9) and (b) follows from Lemma 2 by using similar reasoning as in Equation (23) and because the function $x \mapsto x \log x$ is decreasing for $x > 0$ small enough. Hence, we obtain:

$$\begin{aligned}
&H\big(\tilde{U}_1\big[\mathcal{I}_1^{(n)} \cup \mathcal{F}_1^{(n)}\big], \ldots, \tilde{U}_K\big[\mathcal{I}_K^{(n)} \cup \mathcal{F}_K^{(n)}\big]\big|\tilde{Z}_m^n\big) \\
&\quad \geq H\big(U_1\big[\mathcal{I}_1^{(n)} \cup \mathcal{F}_1^{(n)}\big], \ldots, U_K\big[\mathcal{I}_K^{(n)} \cup \mathcal{F}_K^{(n)}\big]\big|Z_m^n\big) - \big((K+2)n\delta^{(n)}_{\text{ld-nls}} - 2\delta^{(n)}_{\text{ld-nls}} \log \delta^{(n)}_{\text{ld-nls}}\big) \\
&\quad \overset{(a)}{\geq} \textstyle\sum_{\ell=1}^{K} \sum_{j \in \mathcal{I}_\ell^{(n)} \cup \mathcal{F}_\ell^{(n)}} H\big(U_\ell(j)\big|U_\ell^{1:j-1}, V_{\ell-1}^n, Z_m^n\big) - \big((K+2)n\delta^{(n)}_{\text{ld-nls}} - 2\delta^{(n)}_{\text{ld-nls}} \log \delta^{(n)}_{\text{ld-nls}}\big) \\
&\quad \overset{(b)}{\geq} \textstyle\sum_{\ell=1}^{K} \big|\mathcal{I}_\ell^{(n)} \cup \mathcal{F}_\ell^{(n)}\big| (1 - 2\delta_n) - \big((K+2)n\delta^{(n)}_{\text{ld-nls}} - 2\delta^{(n)}_{\text{ld-nls}} \log \delta^{(n)}_{\text{ld-nls}}\big),
\end{aligned} \tag{45}$$

where (a) holds because conditioning does not increase the entropy and because $U_1^n - \cdots - U_{K-1}^n - U_K^n$ forms a Markov chain and the invertibility of G_n and (b) holds because, according to Equations (32) and (33), $\mathcal{I}_\ell^{(n)} \cup \mathcal{F}_\ell^{(n)} \subseteq \mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M}$ for all $\ell \in [1, K]$, because by Lemma 1, we have $\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_M} \subseteq \mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_m}$ for any $m \in [1, M-1]$, and by the definition of the set $\mathcal{H}^{(n)}_{V_\ell|V_{\ell-1}Z_m}$ given in Equation (31).

Finally, by substituting Equation (45) into Equation (43), we obtain:

$$I(W_1,\dots,W_K;F_1,\dots,F_K,\tilde{Z}_m^n) \leq n\delta_n + (K+2)n\delta_{\text{ld-nls}}^{(n)} - 2\delta_{\text{ld-nls}}^{(n)}\log\delta_{\text{ld-nls}}^{(n)}, \tag{46}$$

Hence, if $K \ll n$, the polar code satisfies the secrecy condition in Equation (5), and the proof is concluded.

6. Polar Construction and Performance Evaluation

In this section, we discuss further how to construct the polar codes for the DBC-NLD-LS and DBC-LD-NLS proposed in Sections 4 and 5, respectively. Moreover, we evaluate the reliability and the secrecy performance of both polar coding schemes according to different parameters involved in the polar code construction. Although the construction of polar codes has been covered in a large number of references (see, for instance, [21–23]), they only focus on polar codes under reliability constraints.

For the DBC-NLD-LS, we consider the Binary Erasure Broadcast Channel (BE-BC), where each individual channel of the DBC is a Binary Erasure Channel (BEC). For this model, we propose a construction of the polar code that is based on the Bhattacharyya parameters instead of the corresponding entropy terms. The reason is that, for the BE-BC, the Bhattacharyya parameters associated with the sets in Equations (7)–(11) can be computed exactly [7] (Proposition 5). Then, we evaluate the reliability and the secrecy performance of the code, and we focus on how different parameters involved in the proposed polar code construction impact its performance.

On the other hand, for the DBC-LD-NLS, we consider the Binary Symmetric Broadcast Channel (BS-BC), where each individual channel is a Binary Symmetric Channel (BSC). From [7] (Proposition 5), we know that the method to compute the exact values of the Bhattacharyya parameters for a BEC provides an upper-bound on the Bhattacharyya parameters of the BSC. Although this method can be useful to construct polar codes under reliability constraints [21–23], it fails when the code must guarantee some secrecy condition based on the information leakage. Indeed, in order to upper-bound the information leakage in Equation (39), according to Equation (45), notice that we need a lower-bound on the entropy terms (or Bhattacharyya parameters). Hence, for this model, we focus more on proposing a new polar code construction that is based directly on the entropy terms associated with the sets in Equations (27)–(31).

Throughout this section, as in [7], we say that a channel or a conditional distribution $p_{Y|X}(y|x)$ with $x \in \mathcal{X} \triangleq \{0,1\}$ and $y \in \mathcal{Y} \triangleq \{0,\dots,|\mathcal{Y}|-1\}$ is symmetric if the columns of the probability transition matrix $\mathbf{P}_{Y|X} \triangleq \begin{bmatrix} p_{Y|X}(0|0) & \cdots & p_{Y|X}(|\mathcal{Y}|-1|0) \\ p_{Y|X}(0|1) & \cdots & p_{Y|X}(|\mathcal{Y}|-1|1) \end{bmatrix}$ can be grouped into sub-matrices such that for each sub-matrix, each row is a permutation of each other row and each column is a permutation of each other column. Therefore, the individual channels of both BE-BC and the BS-BC are symmetric.

Due to the symmetry of BE-BC, we will see that the distribution induced by the encoding described in Section 4.2 for the DBC-NLD-LS will approach exactly the optimum distribution of the original DMS used in the polar code construction. Consequently, the performance of the polar code will depend only on the parameters involved in the construction. On the other hand, despite the symmetry of the BS-BC, due to its superposition-based structure, the encoding described in Section 5.2 for the DBC-NLD-LS only approaches the target distribution asymptotically. Hence, this encoding will impact the reliability and secrecy performance of the polar code when we consider a finite blocklength.

6.1. DBC-NLD-LS

For this model, we consider BE-BC with two legitimate receivers ($K = 2$) and two eavesdroppers ($M = 2$). Therefore, each individual channel is a BEC with $\mathcal{X} \triangleq \{0,1\}$ and $\mathcal{Y}_k = \mathcal{Z}_m \triangleq \{0,1,E\}$, E being the erasure symbol and $k,m \in \{1,2\}$. The individual channels are defined simply by their erasure probability, which is denoted by ϵ_{Y_k} for the corresponding legitimate receiver k ($\mathbb{P}[Y_k = E] = \epsilon_{Y_k}$) and ϵ_{Z_m} for the eavesdropper m ($\mathbb{P}[Z_m = E] = \epsilon_{Z_m}$). Due to the degradedness condition of the broadcast channel given in Equation (1), we have $\epsilon_{Y_2} < \epsilon_{Y_1} < \epsilon_{Z_2} < \epsilon_{Z_1}$. By properly applying [19] (Proposition 3.2), it is easy to shown that the secrecy-capacity achieving distribution $p^{\star}_{X}$ for this model is the uniform, i.e., $p^{\star}_{X}(x) = \frac{1}{2}\ \forall x \in \{0,1\}$. For the simulations, we consider a BE-BC such that $\epsilon_{Y_2} = 0.01$, $\epsilon_{Y_1} = 0.04$, $\epsilon_{Z_2} = 0.2$ and $\epsilon_{Z_1} = 0.35$. According to Corollary 1 and since $p^{\star}_{X}(x)$ is uniform, we obtain that the capacity without considering rate sharing is $R^{\star}_1 = 0.15$ and $R^{\star}_2 = 0.16$.

6.1.1. Practical Polar Code Construction

Given the blocklength n and the distribution $p^{\star}_{XY_2Y_1Z_2Z_1} = p^{\star}_{X} p_{Y_2Y_1Z_2Z_1|X}$, the goal of the polar code construction is to obtain the partition of the universal set $[n]$ defined in Equations (12)–(16) and graphically represented in Figure 3. Hence, we need to define first the required sets of Equations (7)–(11), which means having to compute the entropy terms $\{H(U(j)|U^{1:j-1})\}_{j=1}^{n}$, $\{H(U(j)|U^{1:j-1}, Y_1^n)\}_{j=1}^{n}$ and $\{H(U(j)|U^{1:j-1}, Z_m^n)\}_{j=1}^{n}$ $\forall m \in \{1,2\}$ associated with the polar transform $U^n = X^n G_n$. Alternatively, as mentioned in Section 3, we can define the sets in Equations (7)–(11) from the corresponding Bhattacharyya parameters. Indeed, since each individual channel is a BEC, by [7] (Proposition 5), we can compute with very low complexity the exact values of $\{Z(U(j)|U^{1:j-1})\}_{j=1}^{n}$, $\{Z(U(j)|U^{1:j-1}, Y_1^n)\}_{j=1}^{n}$ and $\{Z(U(j)|U^{1:j-1}, Z_m^n)\}_{j=1}^{n}$ $\forall m \in \{1,2\}$. To do so, we use the recursive algorithm [22] (PCC-0) adapted to the BEC, which, for instance, will obtain $\{Z(U(j)|U^{1:j-1}, Y_1^n)\}_{j=1}^{n}$ from the initial value $Z(X|Y_1) = \epsilon_{Y_1}$ (the entire code in MATLAB used for this section is provided as Supplementary Material—see Endnote [32]). Regarding $\{Z(U(j)|U^{1:j-1})\}_{j=1}^{n}$, since $p^{\star}_{X}$ is uniform, it is clear that $Z(U(j)|U^{1:j-1}) = H(U(j)|U^{1:j-1}) = 1$ for all $j \in [n]$, which means $\mathcal{H}_X^{(n)} = [n]$. Consequently, the set $\mathcal{T}^{(n)} = \emptyset$, and according to Equation (17), neither random, nor deterministic SC encoding will be needed.

In order to compare the performance of the polar coding scheme according to different parameters and to provide more flexibility in the design, instead of using only δ_n to define the sets in Equations (7)–(11), we introduce the pair $(\delta_n^{(\mathrm{r})}, \delta_n^{(\mathrm{s})})$, where $\delta_n^{(\mathrm{r})} \triangleq 2^{-n^{\beta^{(\mathrm{r})}}}$ and $\delta_n^{(\mathrm{s})} \triangleq 2^{-n^{\beta^{(\mathrm{s})}}}$ for some $\beta^{(\mathrm{r})}, \beta^{(\mathrm{s})} \in (0, \frac{1}{2})$. Let $R'_1 \in [0, R^{\star}_1]$ and $R'_2 \in [0, R^{\star}_2]$ denote the target rates that the polar coding scheme must approach. We obtain the partition defined in Equations (12)–(16) as follows. First, we define $(\mathcal{H}_{X|Y_1}^{(n)})^{\mathrm{C}} \triangleq \{j \in [n] : H\left(U(j)\,|U^{1:j-1}, Y_1^n\right) \leq 1 - \delta_n^{(\mathrm{s})}\}$, where one can notice that we have used $\delta_n^{(\mathrm{s})}$. Then, we choose $\mathcal{I}_2^{(n)}$ by taking the $\lceil nR'_2 \rceil$ indices $j \in (\mathcal{H}_{X|Y_1}^{(n)})^{\mathrm{C}}$ that correspond to the highest Bhattacharyya parameters $\{Z(U(j)|U^{1:j-1}, Z_2^n)\}_{j=1}^{n}$ for Eavesdropper 2. Second, we choose $\mathcal{I}_1^{(n)}$ by taking the $\lceil nR'_1 \rceil$ indices $j \in (\mathcal{H}_{X|Y_1}^{(n)})^{\mathrm{C}} \setminus \mathcal{I}_2^{(n)}$ that correspond to the highest Bhattacharyya parameters $\{Z(U(j)|U^{1:j-1}, Z_1^n)\}_{j=1}^{n}$ for Eavesdropper 1. Finally, we obtain $\mathcal{C}^{(n)} = (\mathcal{H}_{X|Y_1}^{(n)})^{\mathrm{C}} \setminus (\mathcal{I}_1^{(n)} \cup \mathcal{I}_2^{(n)})$ and $\mathcal{F}^{(n)} = \mathcal{H}_{X|Y_1}^{(n)}$. Furthermore, in order to evaluate the reliability performance of the code, we define $\mathcal{L}_{X|Y_1}^{(n)} \triangleq \{j \in [n] : H\left(U(j)\,|U^{1:j-1}, Y_1^n\right) \leq \delta_n^{(\mathrm{r})}\}$, where one can notice that we have used $\delta_n^{(\mathrm{r})}$. Since the additional secret sequence Φ corresponds to those entries belonging to $(\mathcal{H}_{X|Y_1}^{(n)})^{\mathrm{C}} \cap (\mathcal{L}_{X|Y_1}^{(n)})^{\mathrm{C}}$, its length will depend on $(\delta_n^{(\mathrm{r})}, \delta_n^{(\mathrm{s})})$. According to the polar code construction proposed in this section, notice that $\delta_n^{(\mathrm{s})}$ must be small enough to guarantee that $|(\mathcal{H}_{X|Y_1}^{(n)})^{\mathrm{C}}| \geq R'_1 + R'_2$.

6.1.2. Performance Evaluation

First, notice that the encoding of Section 4.2 will induce a distribution $\tilde{q}_{X^nY_2^nY_1^nZ_2^nZ_1^n} = p^{\star}_{X^nY_2^nY_1^nZ_2^nZ_1^n}$ because $\mathcal{T}^{(n)} = \emptyset$ (we do not use SC encoding), and the encoder will store uniformly-distributed sequences into the entries $U(j)$ that satisfy $H(U(j)|U^{1:j-1}) = 1$ for all $j \in \mathcal{H}_X^{(n)} = [n]$. Hence, $\mathbb{V}(\tilde{q}_{X^nY_2^nY_1^nZ_2^nZ_1^n}, p^{\star}_{X^nY_2^nY_1^nZ_2^nZ_1^n}) = 0$, and the performance will only depend on the code construction.

To evaluate the reliability performance, we obtain an upper-bound $P_{\mathrm{b}}^{\mathrm{ub}(1)}$ on the average bit error probability at the legitimate Receiver 1. Since $\mathbb{V}(\tilde{q}_{X^n Y_2^n Y_1^n Z_2^n Z_1^n}, p^{\star}_{X^n Y_2^n Y_1^n Z_2^n Z_1^n}) = 0$, from Equation (21), we have:

$$P_{\mathrm{b}}^{\mathrm{ub}(1)} \triangleq \frac{1}{|\mathcal{L}^{(n)}_{X|Y_1}|} \sum_{j \in \mathcal{L}^{(n)}_{X|Y_1}} Z(U(j)|U^{1:j-1}, Y_1^n). \tag{47}$$

Due to the degradedness condition of the BE-BC and, consequently, by Lemma 1, the average bit error probability at the legitimate Receiver 2 will be always less than the one at the legitimate Receiver 1. Since the legitimate receivers must estimate the entries belonging to $\mathcal{L}^{(n)}_{X|Y_1}$ regardless of $(\mathcal{H}^{(n)}_{X|Y_1})^{\mathrm{C}}$ and the target rates (R_1', R_2'), the reliability performance only depends on the pair $(n, \delta_n^{(\mathrm{r})})$.

In order to evaluate the secrecy performance, we compute an upper-bound on the information leakage $I(W_1, W_2; F, \tilde{Z}_1^n)$ and an upper-bound on the information leakage $I(W_2; F, \tilde{Z}_2^n)$. Since $\mathbb{V}(\tilde{q}_{X^n Y_2^n Y_1^n Z_2^n Z_1^n}, p^{\star}_{X^n Y_2^n Y_1^n Z_2^n Z_1^n}) = 0$, from Equations (22) and (24), we obtain:

$$I^{\mathrm{ub}}(W_1, W_2; F, \tilde{Z}_1^n) \triangleq \sum_{i=1}^{2} |\mathcal{I}_i^{(n)}| + |\mathcal{F}^{(n)}| - \sum_{j \in \mathcal{I}_1^{(n)} \cup \mathcal{I}_2^{(n)} \cup \mathcal{F}^{(n)}} Z(U(j)|U^{1:j-1}, Z_1^n)^2, \tag{48}$$

$$I^{\mathrm{ub}}(W_2; F, \tilde{Z}_2^n) \triangleq |\mathcal{I}_2^{(n)}| + |\mathcal{F}^{(n)}| - \sum_{j \in \mathcal{I}_2^{(n)} \cup \mathcal{F}^{(n)}} Z(U(j)|U^{1:j-1}, Z_2^n)^2, \tag{49}$$

where we have used [27] (Proposition 2) to express the information leakage in terms of the Bhattacharyya parameters because $H(U(j)|U^{1:j-1}, Z_m^n) \geq Z(U(j)|U^{1:j-1}, Z_m^n)^2$. According to the proposed polar code construction, the secrecy performance will depend on $(n, \delta_n^{(\mathrm{s})})$ and the rates (R_1', R_2'), but not on $\delta_n^{(\mathrm{r})}$.

Additionally, we evaluate the rate of the additional sequence Φ simply by computing:

$$\frac{1}{n}|\Phi| = \frac{1}{n}\left|(\mathcal{H}^{(n)}_{X|Y_1})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{X|Y_1})^{\mathrm{C}}\right|, \tag{50}$$

which will depend on the triple $(n, \delta_n^{(\mathrm{r})}, \delta_n^{(\mathrm{s})})$, but not on (R_1', R_2').

Let ρ_{R} be the normalized target rate in which the polar coding scheme operates, that is $\rho_{\mathrm{R}} \triangleq \frac{R_1'}{R_1^{\star}} = \frac{R_2'}{R_2^{\star}}$. In Figure 6A,B, we evaluate the upper-bounds on the information leakage defined in Equations (48) and (49), respectively, as a function of the blocklength n for different values of ρ_{R}. To do so, we set $\beta^{(\mathrm{r})} = 0.16$ and $\beta^{(\mathrm{s})} = 0.30$, which defines a particular pair $(\delta_n^{(\mathrm{r})}, \delta_n^{(\mathrm{s})})$ for each value of n (recall that $\delta_n^{(\mathrm{r})}$ does not impact on the secrecy performance of the polar code). As we proved in Section 4.5.4, for large enough n, the secrecy performance improves as n increases. Moreover, to achieve a particular secrecy performance level, the polar code will require a larger blocklength n as the rates approach the capacity. This happens because, given $(n, \delta_n^{(\mathrm{s})})$ and, consequently, $(\mathcal{H}^{(n)}_{X|Y_1})^{\mathrm{C}}$, the parameter ρ_{R} only determines the amount of indices that will belong to $\mathcal{I}_1^{(n)} \cup \mathcal{I}_2^{(n)} \subseteq (\mathcal{H}^{(n)}_{X|Y_1})^{\mathrm{C}}$. Since, by construction, we take those indices corresponding to the highest Bhattacharyya parameters associated with the eavesdroppers, taking more elements always increases the corresponding leakage. For rates approaching the capacity and small values of n, notice that we obtain a secrecy performance that is getting worse as n increases (for instance, for $\rho_{\mathrm{R}} = 0.94$, we obtain that the information leakage is increasing from $n = 2^9$ to $n = 2^{12}$). This behavior is mainly explained because the elements of U^n have not been polarized enough for small values of n. Consequently, for a given value of $\beta^{(\mathrm{s})}$, not all the Bhattacharyya parameters associated with the eavesdroppers corresponding to the sets $\mathcal{I}_1^{(n)}$ and $\mathcal{I}_2^{(n)}$ are sufficiently close to one. Since, for a given ρ_{R}, the cardinality of $\mathcal{I}_1^{(n)}$ and $\mathcal{I}_2^{(n)}$ increases with n, then the information leakage can increase with n when n is not large enough. Moreover, since operating at lower rates means taking a fewer number of indices in $\mathcal{I}_1^{(n)}$ and $\mathcal{I}_2^{(n)}$, but taking those that are closest to one, this behavior appears only for large values of ρ_R.

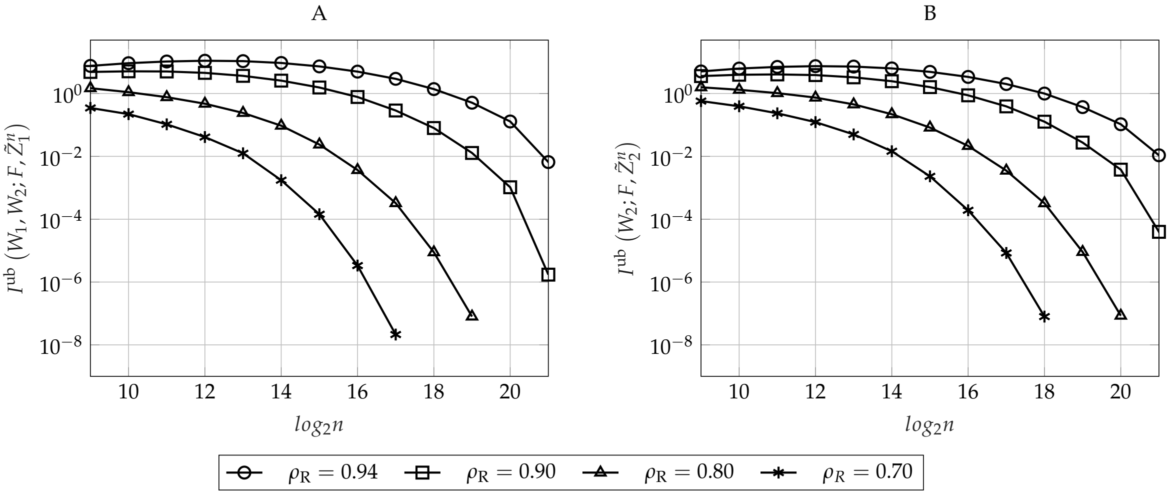

Figure 6. Secrecy performance of the polar coding scheme for DBC-NLD-LS over BE-BC as a function of the blocklength n and the normalized target rate ρ_R when we set $\beta^{(\mathrm{r})} = 0.16$ and $\beta^{(\mathrm{s})} = 0.30$. (**A**) Upper-bound on the information about (W_1, W_2) leaked to Eavesdropper 1 defined as in Equation (48). (**B**) Upper-bound on the information about W_2 leaked to Eavesdropper 2 defined as in Equation (49).

The impact of $\delta_n^{(\mathrm{s})}$ on the secrecy performance is graphically represented in Figure 7A,B, where the former plots the upper-bound defined in Equation (48) and the latter the upper-bound in Equation (49) as a function of the blocklength n for different values of $\beta^{(\mathrm{s})}$. Now, we set $\beta^{(\mathrm{r})} = 0.16$ and $\rho_\mathrm{R} = 0.90$. As can be seen in Figure 7, the secrecy performance improves as the value of $\beta^{(\mathrm{s})}$ increases (or equivalently, as $\delta_n^{(\mathrm{s})}$ decreases). This behavior is as expected because notice that $\delta_n^{(\mathrm{s})}$ defines the value of the highest Bhattacharyya parameter $Z\big(U(j)\big|U^{1:j-1}, Y_1^n\big)$ that will belong to $\big(\mathcal{H}_{X|Y_1}^{(n)}\big)^{\mathrm{C}}$, that is the set containing the possible candidates for $\mathcal{I}_1^{(n)} \cup \mathcal{I}_2^{(n)}$. Since the polar construction chooses the indices that will belong to $\mathcal{I}_1^{(n)}$ and $\mathcal{I}_2^{(n)}$ by taking the ones corresponding to the highest Bhattacharyya parameters associated with the eavesdroppers and since, by Lemma 1, $Z\big(U(j)\big|U^{1:j-1}, Z_1^n\big) \geq Z\big(U(j)\big|U^{1:j-1}, Z_2^n\big) \geq Z\big(U(j)\big|U^{1:j-1}, Y_1^n\big)$ for any $j \in [n]$, the sums in Equations (48) and (49) over the indices $j \in \mathcal{I}_1^{(n)} \cup \mathcal{I}_2^{(n)}$ will be larger as $\beta^{(\mathrm{s})}$ increases (as $\delta_n^{(\mathrm{s})}$ decreases), while their cardinality remains the same for a given ρ_R. Furthermore, notice that $\delta_n^{(\mathrm{s})}$ also defines $\mathcal{F}^{(n)} = \mathcal{H}_{X|Y_1}^{(n)} = \{j \in [n] : Z\big(U(j)\big|U^{1:j-1}, Y_1^n\big) > 1 - \delta_n^{(\mathrm{s})}\}$. Thus, the larger is the value of $\beta^{(\mathrm{s})}$ (the lower is $\delta_n^{(\mathrm{s})}$), the smaller is the cardinality of $\mathcal{F}^{(n)}$ and the higher are the Bhattacharyya parameters associated with the eavesdroppers that belong to this set.

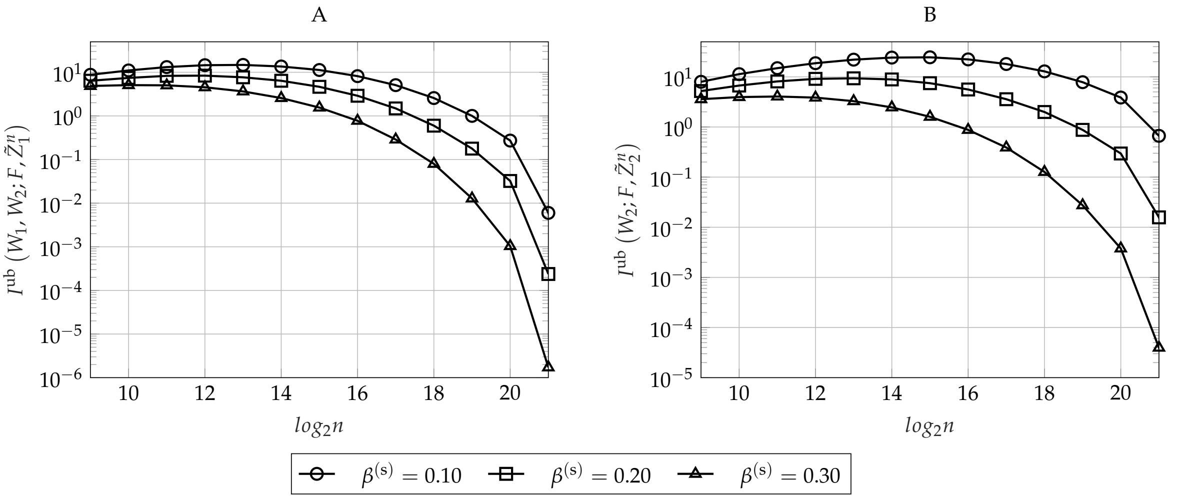

Figure 7. Secrecy performance of the polar coding scheme for DBC-NLD-LS over BE-BC as a function of n and $\beta^{(\text{s})}$, which defines $\delta_n^{(\text{s})}$ for each n, when we set $\beta^{(\text{r})} = 0.16$ and $\rho_{\text{R}} = 0.90$. (**A**) Upper-bound on the information about (W_1, W_2) leaked to Eavesdropper 1 defined as in Equation (48). (**B**) Upper-bound on the information about W_2 leaked to Eavesdropper 2 defined as in Equation (49).

Figure 8 plots the upper-bound on the average bit error probability at the legitimate Receiver 1 defined in Equation (47) as a function of the blocklength n for different values of $\beta^{(\text{r})}$ (which defines a particular $\delta_n^{(\text{r})}$ for each n). For this figure, we set $\beta^{(\text{s})} = 0.30$ and $\rho_{\text{R}} = 0.90$. As can be seen in Figure 8, the higher is the value of $\beta^{(\text{r})}$ (the smaller is the value of $\delta_n^{(\text{r})}$), the better is the reliability performance of the polar code. This is because $\delta_n^{(\text{r})}$ defines the higher Bhattacharyya parameter associated with the legitimate Receiver 1 whose corresponding index will belong to the set $\mathcal{L}_{X|Y_1}^{(n)}$ (recall that this set contains the indices of those entries that the legitimate receivers have to estimate). Hence, it is clear that the upper-bound in Equation (47) is decreasing as $\delta_n^{(\text{r})}$ decreases (as $\beta^{(\text{r})}$ increases). Moreover, as we have proven in Section 4.5.3, we can see that the reliability performance is always improving as n increases.

Finally, how the values of the pair $(\beta^{(\text{r})}, \beta^{(\text{s})})$, or equivalently, the values of $(\delta_n^{(\text{r})}, \delta_n^{(\text{s})})$, impact the rate of the additional secret sequence Φ given in Equation (50) is represented graphically in Figure 9. In Figure 9A, we set $\rho_{\text{R}} = 0.90$ and $\beta^{(\text{r})} = 0.16$, and we represent the rate of Φ as a function of the blocklength n for different values of $\beta^{(\text{s})}$. Otherwise, in Figure 9B, we evaluate the rate of Φ as a function of n for different values of $\beta^{(\text{r})}$ when $\rho_{\text{R}} = 0.90$ and $\beta^{(\text{s})} = 0.30$. As mentioned in Section 4.2, this rate tends to be negligible for sufficiently large n. Moreover, according to the polar code construction proposed previously, for a fixed n, the cardinality of the set $\big(\mathcal{H}_{X|Y_1}^{(n)}\big)^{\text{C}} \cap \big(\mathcal{L}_{X|Y_1}^{(n)}\big)^{\text{C}}$ will be higher for larger values of $(\beta^{(\text{r})}, \beta^{(\text{s})})$, or equivalently, smaller values of $(\delta_n^{(\text{r})}, \delta_n^{(\text{s})})$. Therefore, as can be seen in Figure 9, it is clear that higher values of $(\beta^{(\text{r})}, \beta^{(\text{s})})$ mean also higher rate of the additional secret sequence.

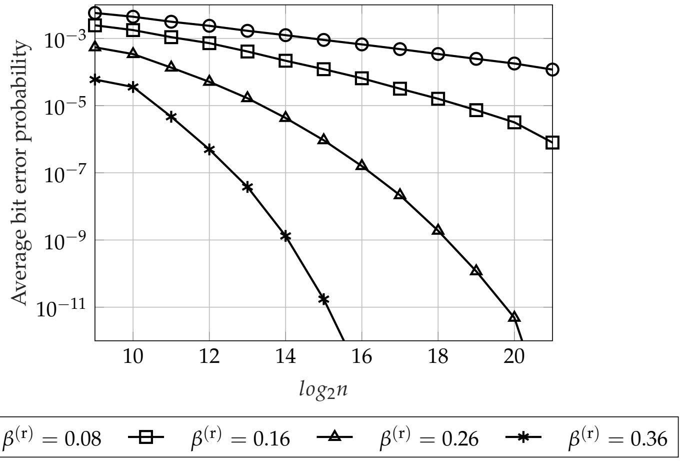

Figure 8. Reliability performance of the polar coding scheme for DBC-NLD-LS over BE-BC as a function of n and $\beta^{(r)}$, which defines $\delta_n^{(r)}$ for each n, when we set $\beta^{(s)} = 0.30$ and $\rho_R = 0.90$. That is, the bound $P_b^{ub(1)}$ on the average bit error probability at the legitimate Receiver 1 is defined as in Equation (47).

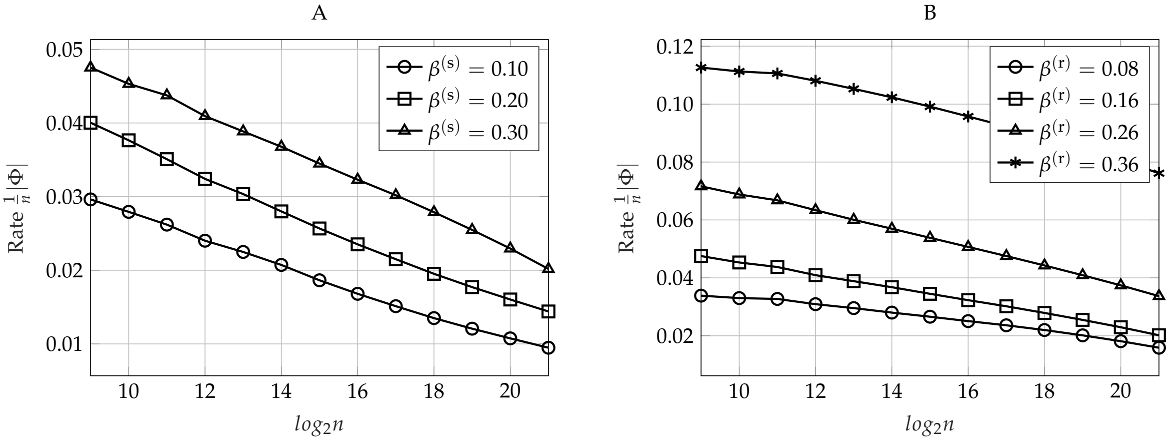

Figure 9. Rate of the additional secret sequence Φ computed as in Equation (50) for DBC-NLD-LS over BE-BC as a function of the blocklength n for different values of $(\beta^{(r)}, \beta^{(s)})$, which defines $(\delta_n^{(r)}, \delta_n^{(s)})$ for each n. (**A**) Rate of Φ for different values of $\beta^{(s)}$ when $\beta^{(r)} = 0.16$ and $\rho_R = 0.90$. (**B**) Rate of Φ for different values of $\beta^{(r)}$ when $\beta^{(s)} = 0.30$ and $\rho_R = 0.90$.

In conclusion, Figures 6–9 show that, for a particular value of the blocklength n, there is a trade-off between the reliability or the secrecy performance of the polar code and the length of the additional secret sequence Φ, which can be controlled by the value of $\beta^{(r)}$ or $\beta^{(s)}$, respectively, in the polar code construction. Moreover, for sufficiently large n, the performance of the polar coding scheme always is improving as n increases. Indeed, these figures show that we can transmit at rates very close to the capacity, providing good reliability and secrecy performance levels.

6.2. DBC-LD-NLS

For this model, we consider BS-BC with two legitimate receivers ($K = 2$) and two eavesdroppers ($M = 2$). Hence, each individual channel is a BSC where $\mathcal{X} = \mathcal{Y}_k = \mathcal{Z}_m = \{0,1\}$, and $k, m \in \{1,2\}$. The individual channels are defined simply by their crossover probability, which is denoted by α_{Y_k} for the corresponding legitimate receiver k ($\mathbb{P}[Y_k = 0|X = 1] = \mathbb{P}[Y_k = 1|X = 0] = \alpha_{Y_k}$) and α_{Z_m} for the corresponding eavesdropper m ($\mathbb{P}[Z_m = 0|X = 1] = \mathbb{P}[Z_m = 0|X = 1] = \alpha_{Z_m}$). Due to the degradedness condition of the broadcast channel given in Equation (1), we have $\alpha_{Y_2} < \alpha_{Y_1} < \alpha_{Z_2} < \alpha_{Z_1}$. Due to the symmetry of the channel, it is easy to prove by using similar reasoning as in [33] (Ex. 15.6.5) and by properly applying [19] (Proposition 3.2) that the secrecy-capacity achieving distribution $p^\star_{VX}$ satisfies $p^\star_V(v) = p^\star_X(x) = \frac{1}{2}\ \forall v, x \in \{0,1\}$, and consequently, $p^\star_{X|V}$ is symmetric. Thus, the distribution $p^\star_{X|V}$ can be characterized simply by the crossover probability $\alpha_{X|V} \triangleq p^\star_{X|V}(0|1) = p^\star_{X|V}(1|0)$, where $\alpha_{X|V} \in [0, \frac{1}{2}]$. Indeed, the overall rate in Proposition 2 is maximized when $\alpha_{X|V} = \frac{1}{2}$, which implies that $R_1 = 0$. Then, by taking $\alpha_{X|V} < \frac{1}{2}$, we can transfer part of the rate associated with the message W_2 to the rate R_1, $R_2 = 0$ and R_1 being maximum if $\alpha_{X|V} = 0$. For the simulations, we consider a BS-BC with $\alpha_{Y_2} = 0.01$, $\alpha_{Y_1} = 0.04$, $\alpha_{Z_2} = 0.2$ and $\alpha_{Z_1} = 0.35$. We set $\alpha_{X|V} = 0.1084$, which corresponds to the distribution that maximizes $\ln(R_1) + \ln(R_2)$ for this particular channel (proportional fair allocation). Thus, according to Corollary 2, the maximum achievable rates are $R_1^\star = 0.2507$ and $R_2^\star = 0.3254$.

6.2.1. Practical Polar Code Construction

Given the blocklength n and the distribution $p^\star_{VXY_2Y_1Z_2Z_1} = p^\star_{VX} p_{Y_2Y_1Z_2Z_1|X}$, the goal of the polar code construction is to obtain the partition of the universal set $[n]$ defined in Equations (32)–(35) and graphically represented in Figure 5. Hence, we need to define first the sets in Equations (27)–(31), which means having to compute the entropy terms $\{H(U_1(j)|U_1^{1:j-1})\}_{j=1}^n$, $\{H(U_1(j)|U_1^{1:j-1}, Y_1^n)\}_{j=1}^n$ and $\{H(U_1(j)|U_1^{1:j-1}, Z_2^n)\}_{j=1}^n$ associated with the polar transform $U_1^n = V^n G_n$ for the first superposition layer and $\{H(U_2(j)|U_2^{1:j-1}, V^n)\}_{j=1}^n$, $\{H(U_2(j)|U_2^{1:j-1}, V^n, Y_2^n)\}_{j=1}^n$ and $\{H(U_2(j)|U_2^{1:j-1}, V^n, Z_2^n)\}_{j=1}^n$ associated with the polar transform $U_2^n = X^n G_n$ for the second layer. In the following, we propose an adaptation of the Monte Carlo method [22] (PCC-1), which is based on the butterfly algorithm described in [7] for SC decoding, to directly estimate these entropy terms.

Monte-Carlo method to estimate the entropy terms. First, consider the entropy terms associated with to the first layer. As for the previous model, since $p^\star_V(v) = \frac{1}{2}$, we have $H(U_1(j)|U_1^{1:j-1}) = 1$ for all $j \in [n]$. In order to compute $\{H(U_1(j)|U_1^{1:j-1}, Y_k^n)\}_{j=1}^n$ and $\{H(U_1(j)|U_1^{1:j-1}, Z_m^n)\}_{j=1}^n$ for some $k, m \in \{1,2\}$, we run the Monte Carlo simulation as follows. First, due to the symmetry of the channel and the symmetry of $p^\star_{X|V}$, as in [22] (PCC-1), we can set $v^n = u_1^n = 0^n$ at each iteration. For the realization $\tau \in [1, N_\tau]$, N_τ being the number of realizations, we randomly generate $y_k^{n(\tau)}$ and $z_m^{n(\tau)}$ from $p^\star_{Y_k^n|V^n}$ and $p^\star_{Z_m^n|V^n}$, respectively (by abuse of notation, we use (τ) in any sequence $a^{n(\tau)}$ to emphasize that it is generated at the iteration $\tau \in [1, N_\tau]$). Next, we obtain the log-likelihood ratios $\{L^{(\tau)}_{Y_k|V}(j)\}_{j=1}^n$ and $\{L^{(\tau)}_{Z_m|V}(j)\}_{j=1}^n$ by using the algorithm [22] (PCC-1). For instance, consider $\{L^{(\tau)}_{Y_k|V}(j)\}_{j=1}^n$. From the initial values $\{p^\star_{Y_k|V}(y_k^{(\tau)}(j)|0)/p^\star_{Y_k|V}(y_k^{(\tau)}(j)|1)\}_{j=1}^n$, the algorithm recursively computes:

$$L^{(\tau)}_{Y_k|V}(j) \triangleq \ln \frac{p^\star_{Y_k^n U_1^{1:j-1}|U_1(j)}(y_k^{n(\tau)}, 0^{j-1}|0)}{p^\star_{Y_k^n U_1^{1:j-1}|U_1(j)}(y_k^{n(\tau)}, 0^{j-1}|1)} \overset{(a)}{=} \frac{p^\star_{U_1(j)|U_1^{1:j-1}Y_k^n}(0|0^{j-1}, y_k^{n(\tau)})}{1 - p^\star_{U_1(j)|U_1^{1:j-1}Y_k^n}(0|0^{j-1}, y_k^{n(\tau)})},$$

for all $j \in [n]$, where (a) follows from the fact that $p^\star_{U_1(j)}(0) = p^\star_{U_1(j)}(1) = \frac{1}{2}$ because $H(U_1(j)|U_1^{1:j-1}) = 1$ for all $j \in [n]$. Hence, we can obtain $p^\star_{U_1(j)|U_1^{1:j-1}Y_k^n}(0|0^{j-1}, y_k^{n(\tau)})$ from $L^{(\tau)}_{Y_k|V}(j)$, and since:

$$H(U_1(j)|U_1^{1:j-1}, Y_k^n) = \mathbb{E}_{U_1^{1:j-1}Y_k^n}\left[h_2\left(p^\star_{U_1(j)|U_1^{1:j-1}Y_k^n}(0|u_1^{1:j-1}, y_k^n)\right)\right],$$

after N_τ realizations, we can estimate $H(U_1(j)|U_1^{1:j-1}, Y_k^n)$ by computing the empirical mean, that is,

$$H(U_1(j)|U_1^{1:j-1}, Y_k^n) \approx \frac{1}{N_\mathrm{r}} \sum_{\tau=1}^{N_\tau} h_2\left(p^\star_{U_1(j)|U_1^{1:j-1}Y_k^n}(0|0^{j-1}, y_k^{n(\tau)})\right).$$

Now, consider the Monte Carlo method to estimate $\{H(U_2(j)|U_2^{1:j-1}, V^n)\}_{j=1}^n$, $\{H(U_2(j)|U_2^{1:j-1}, V^n, Y_k^n)\}_{j=1}^n$ and $\{H(U_2(j)|U_2^{1:j-1}, V^n, Z_m^n)\}_{j=1}^n$ for any $k, m \in \{1, 2\}$ associated with the second layer. To obtain $\{H(U_2(j)|U_2^{1:j-1}, V^n)\}_{j=1}^n$, we can see X and V as the input and output random variables, respectively, of a symmetric channel with distribution $p^\star_{V|X}$. Now, although $p^\star_X$ is uniform and, consequently, $H(U_2(j)|U_2^{1:j-1}) = 1$ for all $j \in [n]$, notice that $\mathcal{H}^{(n)}_{X|V} \neq [n]$ and $\mathcal{T}^{(n)}_1 \neq \emptyset$ because $\mathcal{H}^{(n)}_{X|V}$ and its complementary set depend on $p^\star_{X|V}$. On the other hand, to obtain $\{H(U_2(j)|U_2^{1:j-1}, V^n, Y_k^n)\}_{j=1}^n$ or $\{H(U_2(j)|U_2^{1:j-1}, V^n, Z_m^n)\}_{j=1}^n$, we can see (V, Y_k) or (V, Z_m) as the output of a symmetric channel with distribution $p^\star_{VY_k|X}$ or $p^\star_{VZ_m|X}$, respectively, where notice that $p^\star_{VY_k|X} = p^\star_{V|X} p^\star_{Y_k|X}$ and $p^\star_{VZ_m|X} = p^\star_{V|X} p^\star_{Z_m|X}$ because $V - X - Y_k - Z_m$ forms a Markov chain. Hence, due to the symmetry of the previous distributions, we can set $x^n = u_2^n = 0^n$ at each iteration. Then, for the realization $\tau \in [1, N_\tau]$, we draw $v^{n(\tau)}$, $y_k^{n(\tau)}$ and $z_m^{n(\tau)}$ from the distributions $p^\star_{V^n|X^n}$, $p_{Y_k^n|X^n}$ and $p_{Z_m^n|X^n}$, respectively. Next, we obtain the log-likelihood ratios $\{L^{(\tau)}_{V|X}(j)\}_{j=1}^n$, $\{L^{(\tau)}_{VY_k|X}(j)\}_{j=1}^n$ and $\{L^{(\tau)}_{VZ_m|X}(j)\}_{j=1}^n$ by using [22] (PCC-1). Since $H(U_2(j)|U_2^{1:j-1}) = 1$ for all $j \in [n]$, we have $p^\star_{U_2(j)}(u) = \frac{1}{2}$ for all $u \in \{0, 1\}$, and we can compute $p^\star_{U_2(j)|U_2^{1:j-1}V^n}(0|0^{j-1}, v^{n(\tau)})$, $p^\star_{U_2(j)|U_2^{1:j-1}V^nY_k^n}(0|0^{j-1}, v^{n(\tau)}, y_k^{n(\tau)})$ and $p^\star_{U_2(j)|U_2^{1:j-1}V^nZ_m^n}(0|0^{j-1}, v^{n(\tau)}, z_m^{n(\tau)})$ from the corresponding log-likelihood ratios. Finally, after N_τ realizations, we can estimate the corresponding entropy terms by computing the empirical mean.

Partition of the universal set [n]. In order to provide more flexibility on the design, now we introduce $(\delta_n^{(1,\mathrm{r})}, \delta_n^{(1,\mathrm{s})})$ for the first layer, where $\delta_n^{(1,\mathrm{r})} \triangleq 2^{-n^{\beta^{(1,\mathrm{r})}}}$ and $\delta_n^{(1,\mathrm{s})} \triangleq 2^{-n^{\beta^{(1,\mathrm{s})}}}$ for some $\beta^{(1,\mathrm{r})}, \beta^{(1,\mathrm{s})} \in (0, \frac{1}{2})$. For the second layer, we introduce $(\delta_n^{(2,\mathrm{r})}, \delta_n^{(2,\mathrm{s})})$ and $(\delta_n^{(2,\mathrm{L})}, \delta_n^{(2,\mathrm{H})})$, where $\delta_n^{(2,\mathrm{r})} \triangleq 2^{-n^{\beta^{(2,\mathrm{r})}}}$, $\delta_n^{(2,\mathrm{s})} \triangleq 2^{-n^{\beta^{(2,\mathrm{s})}}}$, $\delta_n^{(2,\mathrm{L})} \triangleq 2^{-n^{\beta^{(2,\mathrm{L})}}}$ and $\delta_n^{(2,\mathrm{H})} \triangleq 2^{-n^{\beta^{(2,\mathrm{H})}}}$ for some $\beta^{(2,\mathrm{r})}, \beta^{(2,\mathrm{s})}, \beta^{(2,\mathrm{L})}, \beta^{(2,\mathrm{H})} \in (0, \frac{1}{2})$.

Consider the partition of $[n]$ for the first layer ($\ell = 1$ in Equations (32)–(35)). As mentioned previously, since $p^\star_V(v) = \frac{1}{2}$, we have $\mathcal{H}^{(n)}_V = [n]$ and $\mathcal{T}^{(n)}_1 = \emptyset$. Let $R'_1 \in [0, R^\star_1]$ denote the target rate corresponding to the message W_1 that the polar coding scheme must approach. We obtain the partition in Equations (32)–(35) as follows. First, we define $(\mathcal{H}^{(n)}_{V|Y_1})^{\mathrm{C}} \triangleq \{j \in [n] : H(U_1(j)|U_1^{1:j-1}, Y_1^n) \leq 1 - \delta_n^{(1,\mathrm{s})}\}$. Then, we choose $\mathcal{I}^{(n)}_1$ by taking the $\lceil nR'_1 \rceil$ indices $j \in (\mathcal{H}^{(n)}_{V|Y_1})^{\mathrm{C}}$ that correspond to the highest entropy terms $\{H(U_1(j)|U_1^{1:j-1}, Z_2^n)\}_{j=1}^n$ associated with Eavesdropper 2. Notice that $\delta_n^{(1,\mathrm{s})}$ must guarantee $|(\mathcal{H}^{(n)}_{V|Y_1})^{\mathrm{C}}| \leq R'_1$. Finally, we obtain $\mathcal{C}^{(n)}_1 = (\mathcal{H}^{(n)}_{V|Y_1})^{\mathrm{C}} \setminus \mathcal{I}^{(n)}_1$ and $\mathcal{F}^{(n)}_1 = \mathcal{H}^{(n)}_{V|Y_1}$. Furthermore, in order to evaluate the reliability performance, we define $\mathcal{L}^{(n)}_{V|Y_1} \triangleq \{j \in [n] : H(U_1(j)|U_1^{1:j-1}, Y_1^n) \leq \delta_n^{(1,\mathrm{r})}\}$.

Consider the partition of $[n]$ for the second layer ($\ell = 2$ in Equations (32)–(35)). Since $\mathcal{H}^{(n)}_{X|V} \neq [n]$ and $\mathcal{T}^{(n)}_1 \neq \emptyset$, we define $\mathcal{H}^{(n)}_{X|V} \triangleq \{j \in [n] : H(U_2(j)|U_2^{1:j-1}, V^n) \geq 1 - \delta_n^{(2,\mathrm{H})}\}$ and $\mathcal{L}^{(n)}_{X|V} \triangleq \{j \in [n] : H(U_2(j)|U_2^{1:j-1}, V^n) \leq \delta_n^{(2,\mathrm{L})}\}$, where we have used $\delta_n^{(2,\mathrm{H})}$ and $\delta_n^{(2,\mathrm{L})}$, respectively. Let $R'_2 \in [0, R^\star_2]$ denote the target rate corresponding to W_2. We define $(\mathcal{H}^{(n)}_{X|VY_2})^{\mathrm{C}} \triangleq \{j \in \mathcal{H}^{(n)}_{X|V} : H(U_2(j)|U_2^{1:j-1}, V^n, Y_2^n) \leq 1 - \delta_n^{(2,\mathrm{s})}\}$. Then, we choose $\mathcal{I}^{(n)}_2$ by taking the $\lceil nR'_2 \rceil$ indices $j \in (\mathcal{H}^{(n)}_{X|VY_2})^{\mathrm{C}}$ that correspond to the highest entropy terms $\{H(U_2(j)|U_2^{1:j-1}, V^n, Z_2^n)\}_{j=1}^n$ associated with Eavesdropper 2. Thus, notice that $\delta_n^{(2,\mathrm{H})}$ and $\delta_n^{(2,\mathrm{s})}$ must guarantee $|\mathcal{H}^{(n)}_{X|V}| \geq |(\mathcal{H}^{(n)}_{X|VY_2})^{\mathrm{C}}| \geq R'_2$. Then, we obtain $\mathcal{C}^{(n)}_2 = (\mathcal{H}^{(n)}_{X|VY_2})^{\mathrm{C}} \setminus \mathcal{I}^{(n)}_2$ and $\mathcal{F}^{(n)}_2 = \mathcal{H}^{(n)}_{X|VY_2}$. Finally, in order to evaluate the reliability performance, we define $\mathcal{L}^{(n)}_{X|VY_2} \triangleq \{j \in [n] : H(U_2(j)|U_2^{1:j-1}, V^n, Y_2^n) \leq \delta_n^{(2,\mathrm{r})}\}$.

6.2.2. Performance Evaluation

First, notice that the encoding at the first layer induces a distribution $\tilde{q}_{V^n} = p_{V^n}$. For the second layer, the entries $U[\mathcal{H}^{(n)}_{X|V}]$ of the original DMS only are almost independent of V^n because $H(U_2(j)|U_2^{1:j-1}, V^n) \leq 1 - \delta_n^{(2,\mathrm{s})}$ for $j \in \mathcal{H}^{(n)}_{X|V}$. Nevertheless, the encoding will construct $\tilde{U}_2[\mathcal{H}^{(n)}_{X|V}]$ by storing uniformly-distributed sequences that are totally independent of V^n. On the other hand, since $\mathcal{L}^{(n)}_{X|V} \subseteq \mathcal{T}_2^{(n)} \neq \emptyset$, the encoder will use the deterministic SC encoding in Equation (37) to construct $\tilde{U}_2[\mathcal{L}^{(n)}_{X|V}]$. Therefore, according to Lemma 3 and Remark 6, we will have $\mathbb{V}(\tilde{q}_{V^n X^n Y_2^n Y_1^n Z_2^n Z_1^n}, p^{\star}_{V^n X^n Y_2^n Y_1^n Z_2^n Z_1^n}) \neq 0$ for finite n. Since, as seen in Section 5.5, this total variation distance impacts the performance, we obtain first an upper-bound $d_{\mathrm{TV}}^{\mathrm{ub}}$ on $\mathbb{V}(\tilde{q}_{V^n X^n Y_2^n Y_1^n Z_2^n Z_1^n}, p^{\star}_{V^n X^n Y_2^n Y_1^n Z_2^n Z_1^n})$, which is defined as:

$$d_{\mathrm{TV}}^{\mathrm{ub}} \triangleq d_{\mathrm{TV}}^{\mathrm{ub(L)}} + d_{\mathrm{TV}}^{\mathrm{ub(H)}},$$

where $d_{\mathrm{TV}}^{\mathrm{ub(L)}}$ will measure the impact of using the deterministic SC encoding in Equation (37) for the entries $\tilde{U}_2\big[\mathcal{L}^{(n)}_{X|V}\big]$, and $d_{\mathrm{TV}}^{\mathrm{ub(H)}}$ is the contribution on the total variation distance of storing uniformly-distributed random sequences into $\tilde{U}_2\big[\mathcal{H}^{(n)}_{X|V}\big]$ that are totally independent of V^n.

Consider $d_{\mathrm{TV}}^{\mathrm{ub(L)}}$, which corresponds to the analytic bound found in Lemma A2. For the simulations, we can use the Monte Carlo method to directly estimate Equation (A4) by computing the empirical mean,

$$d_{\mathrm{TV}}^{\mathrm{ub(L)}} \triangleq \frac{1}{N_{\tau'}} \sum_{\tau'=1}^{N_{\tau'}} \left[\sum_{j \in \mathcal{L}^{(n)}_{X|V}} \left(1 - p^{\star}_{U_2(j)|U_2^{1:j-1}V^n}\left(u_2^{*}(j) \middle| \check{u}_2^{1:j-1(\tau')}, \check{v}^{n(\tau')} \right) \right) \right], \tag{51}$$

where $(\check{v}^{n(\tau')}, \check{u}_2^{n(\tau')})$ must be drawn at each iteration $\tau' \in [1, N_{\tau'}]$ according to Equation (A2), $\mathcal{L}^{(n)}_{X|V}$ has been obtained previously in the polar code construction and, according to Equation (A4), $u_2^{*}(j) \triangleq \arg\max_{u \in \{0,1\}} p^{\star}_{U_2(j)|U_2^{1:j-1}V^n}(u|\check{u}_2^{1:j-1(\tau')}, \check{v}^{n(\tau')})$. Due to the symmetry of $p^{\star}_{V|X}$, the probabilities $p^{\star}_{U_2(j)|U_2^{1:j-1}V^n}$ can be obtained with low complexity using the butterfly algorithm described in [7].

Consider now $d_{\mathrm{TV}}^{\mathrm{ub(H)}}$, which corresponds to the analytic bound found in Lemma A1. We can compute exactly the Kullback-Leibler divergence as in Equation (A3) by using the corresponding entropy terms obtained in the polar code construction. Thus, by applying Pinsker's inequality, we have:

$$d_{\mathrm{TV}}^{\mathrm{ub(H)}} \triangleq \left(2 \ln 2 \sum_{j \in \mathcal{H}^{(n)}_{X|V}} \left(1 - H\left(U_2(j) \middle| U_2^{1:j-1}, V^n \right) \right) \right)^{1/2}. \tag{52}$$

According to the polar code construction, $|\mathcal{L}^{(n)}_{X|V}|$ and $|\mathcal{H}^{(n)}_{X|V}|$ will depend only on the values of $\delta_n^{(2,\mathrm{L})}$ and $\delta_n^{(2,\mathrm{H})}$, respectively, for a particular n. Hence, the value of $d_{\mathrm{TV}}^{\mathrm{ub}}$ can be controlled by adjusting $(\beta^{(2,\mathrm{L})}, \beta^{(2,\mathrm{H})})$. It is clear that higher values of $(\beta^{(2,\mathrm{L})}, \beta^{(2,\mathrm{H})})$ mean lower cardinalities of the sets $\mathcal{L}^{(n)}_{X|V}$ and $\mathcal{H}^{(n)}_{X|V}$ and, consequently, lower $d_{\mathrm{TV}}^{\mathrm{ub}}$. However, $|(\mathcal{H}^{(n)}_{X|V})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{X|V})^{\mathrm{C}}|$ increases with $(\beta^{(2,\mathrm{L})}, \beta^{(2,\mathrm{H})})$, and the encoder in Equation (36) requires more randomness to form $\tilde{U}_2[(\mathcal{H}^{(n)}_{X|V})^{\mathrm{C}} \cap (\mathcal{L}^{(n)}_{X|V})^{\mathrm{C}}]$.

To evaluate the reliability performance, we obtain the upper-bounds $P_{\mathrm{b}}^{\mathrm{ub}(1)}$ and $P_{\mathrm{b}}^{\mathrm{ub}(2)}$ on the average bit error probability at Receivers 1 and 2, respectively. From Equations (41) and (42) and by applying [27] (Proposition 2) to upper-bound the Bhattacharyya parameters from the entropy terms, we have:

$$P_{\mathrm{b}}^{\mathrm{ub}(1)} \triangleq d_{\mathrm{TV}}^{\mathrm{ub}} + \frac{1}{\left|\mathcal{L}_{V|Y_1}^{(n)}\right|} \sum_{j \in \mathcal{L}_{V|Y_1}^{(n)}} \sqrt{H\big(U_1(j)\big|U_1^{1:j-1}, Y_1^n\big)}, \tag{53}$$

$$P_{\mathrm{b}}^{\mathrm{ub}(2)} \triangleq 2d_{\mathrm{TV}}^{\mathrm{ub}} + \frac{2}{\left|\mathcal{L}_{V|Y_1}^{(n)}\right|} \sum_{j \in \mathcal{L}_{V|Y_1}^{(n)}} \sqrt{H\big(U_1(j)\big|U_1^{1:j-1}, Y_2^n\big)} + \frac{1}{\left|\mathcal{L}_{X|VY_2}^{(n)}\right|} \sum_{j \in \mathcal{L}_{X|VY_2}^{(n)}} \sqrt{H\big(U_2(j)\big|U_2^{1:j-1}, V^n, Y_2^n\big)}. \tag{54}$$

To evaluate the secrecy performance, we compute an upper-bound $I^{\mathrm{ub}}(W_1, W_2; F_1, F_2, \tilde{Z}_2^n)$ on the information leakage $I(W_1, W_2; F_1, F_2, \tilde{Z}_2^n)$ for Eavesdropper 2. From Equation (45) we obtain:

$$\begin{aligned} I^{\mathrm{ub}}(W_1, W_2; F_1, F_2, \tilde{Z}_2^n) \triangleq 4nd_{\mathrm{TV}}^{\mathrm{ub}} - 2d_{\mathrm{TV}}^{\mathrm{ub}} \log d_{\mathrm{TV}}^{\mathrm{ub}} + \sum_{\ell=1}^{2} \left|\mathcal{I}_\ell^{(n)} \cup \mathcal{F}_\ell^{(n)}\right| \\ - \sum_{j \in \mathcal{I}_1^{(n)} \cup \mathcal{F}_1^{(n)}} H\big(U_1(j)\big|U_1^{1:j-1}, Z_2^n\big) - \sum_{j \in \mathcal{I}_2^{(n)} \cup \mathcal{F}_2^{(n)}} H\big(U_2(j)\big|U_2^{1:j-1}, V^n, Z_2^n\big), \end{aligned} \tag{55}$$

Due to the degradedness condition of BS-BC and, consequently, by Lemma 1, the information leakage at Eavesdropper 1 will be always less than the one at Eavesdropper 2.

Finally, we evaluate the overall rate of the additional sequences $\{\Phi_1, \Phi_2\}$ by computing:

$$\frac{1}{n}\left(|\Phi_1| + |\Phi_2|\right) = \frac{1}{n}\left(\left|\left(\mathcal{H}_{V|Y_1}^{(n)}\right)^{\mathrm{C}} \cap \left(\mathcal{L}_{V|Y_1}^{(n)}\right)^{\mathrm{C}}\right| + \left|\left(\mathcal{H}_{X|VY_2}^{(n)}\right)^{\mathrm{C}} \cap \left(\mathcal{L}_{X|VY_2}^{(n)}\right)^{\mathrm{C}}\right|\right). \tag{56}$$

The performance of the polar coding scheme is graphically shown in Figure 10. As for the previous model, let ρ_{R} be the normalized target rate in which the polar coding scheme operates, that is $\rho_{\mathrm{R}} \triangleq \frac{R_1'}{R_1^\star} = \frac{R_2'}{R_2^\star}$. In Figure 10A, we evaluate the upper-bound $I_0^{\mathrm{ub}}(W_1, W_2; F_1, F_2, Z_2^n)$, which corresponds to the upper-bound on the information leakage defined in Equation (55) when we consider $d_{\mathrm{TV}}^{\mathrm{ub}} = 0$, as a function of the blocklength n for different values of ρ_{R}. For this plot, we set $\beta^{(1,\mathrm{s})} = 0.30$ and $\beta^{(2,\mathrm{s})} = 0.36$. Notice that $(\beta^{(1,\mathrm{r})}, \beta^{(2,\mathrm{r})})$ and $(\beta^{(2,\mathrm{L})}, \beta^{(2,\mathrm{H})})$ if we set $d_{\mathrm{TV}}^{\mathrm{ub}} = 0$ will not impact the information leakage. As we have proven in Section 5.5.4, the secrecy performance is improving as n increases. Moreover, to satisfy a particular secrecy performance level, the polar code will need higher values of n as the target rates approach the capacity.

In Figure 10B, we evaluate the upper-bounds $P_{\mathrm{b},0}^{\mathrm{ub}(1)}$ and $P_{\mathrm{b},0}^{\mathrm{ub}(2)}$, which correspond to the bounds on the average bit error probability at the legitimate Receivers 1 and 2, respectively, when we set $d_{\mathrm{TV}}^{\mathrm{ub}} = 0$, as a function of the blocklength n. For this plot, we set $\beta^{(1,\mathrm{r})} = \beta^{(2,\mathrm{r})} = 0.24$ and notice that the reliability performance will not depend on the values of $(\beta^{(1,\mathrm{s})}, \beta^{(2,\mathrm{s})})$ and ρ_{R}. If we set $d_{\mathrm{TV}}^{\mathrm{ub}} = 0$, then it is clear that it will not depend on $(\beta^{(2,\mathrm{L})}, \beta^{(2,\mathrm{H})})$ either. As shown theoretically in Section 5.5.3, the error probability becomes lower as the blocklength n increases.

Figure 10C plots the overall rate of the additional secret sequences computed as in Equation (56) when we set $\beta^{(1,\mathrm{r})} = \beta^{(2,\mathrm{r})} = 0.24$, $\beta^{(1,\mathrm{s})} = 0.30$ and $\beta^{(2,\mathrm{s})} = 0.36$. As mentioned in Section 5.2, we can see that this rate tends to be negligible for n sufficiently large.

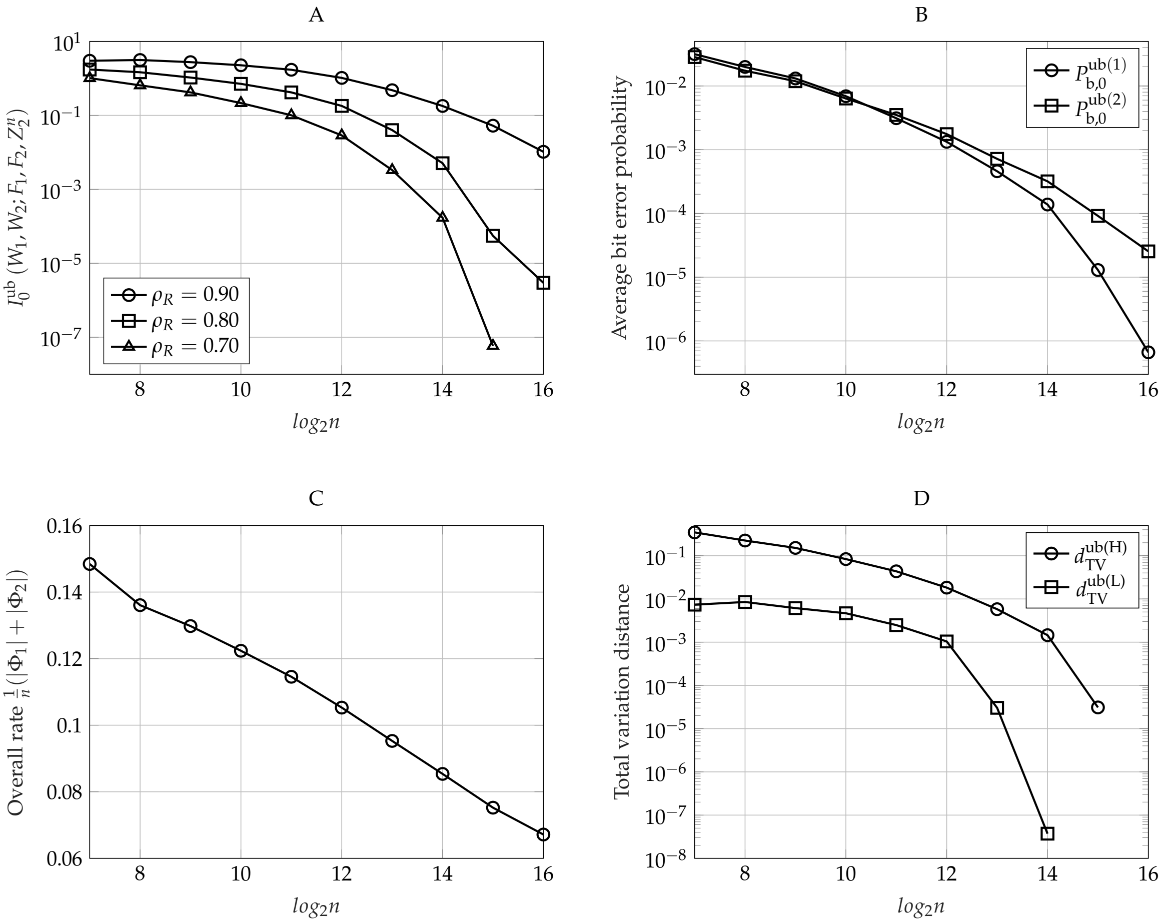

Figure 10. Performance of the polar coding scheme for DBC-LD-NLS over BS-BC as a function of the blocklength n when $\beta^{(1,\mathrm{r})} = \beta^{(2,\mathrm{r})} = 0.24$, $\beta^{(1,\mathrm{s})} = 0.30$, $\beta^{(2,\mathrm{s})} = 0.36$ and $\beta^{(2,\mathrm{H})} = \beta^{(2,\mathrm{H})} = 0.36$. (**A**) Upper-bound on the information about (W_1, W_2) leaked to Eavesdropper 2 defined as in Equation (55) for different normalized target rates ρ_{R} when we set $d_{\mathrm{TV}}^{\mathrm{ub}} = 0$. (**B**) Upper-bounds on the average error probability at legitimate Receivers 1 and 2 defined as in Equations (53) and (54), respectively, when $d_{\mathrm{TV}}^{\mathrm{ub}} = 0$. (**C**) Overall rate of the sequences $\{\Phi_1, \Phi_2\}$ computed as in Equation (56). (**D**) terms $d_{\mathrm{TV}}^{\mathrm{ub(H)}}$ and $d_{\mathrm{TV}}^{\mathrm{ub(L)}}$ that contribute to the bound on the total variation distance $d_{\mathrm{TV}}^{\mathrm{ub}}$ defined as in Equations (51) and (52), respectively.

Finally, Figure 10D plots the upper-bounds $d_{\mathrm{TV}}^{\mathrm{ub(L)}}$ and $d_{\mathrm{TV}}^{\mathrm{ub(H)}}$ defined in Equations (51) and (52), respectively, when we set $\beta^{(2,\mathrm{L})} = \beta^{(2,\mathrm{H})} = 0.36$. As we have proven theoretically in Lemma 3, notice that the total variation distance decays with the blocklength n. Precisely, notice that $d_{\mathrm{TV}}^{\mathrm{ub(L)}}$ is lower than $d_{\mathrm{TV}}^{\mathrm{ub(H)}}$, and therefore, the bound on the total variation distance is practically governed by $d_{\mathrm{TV}}^{\mathrm{ub(H)}}$ ($d_{\mathrm{TV}}^{\mathrm{ub}} \approx d_{\mathrm{TV}}^{\mathrm{ub(H)}}$). This happens because although we can compute exactly the Kullback–Leibler divergence as in Equation (A3) from the entropy terms estimated in the polar code construction, Pinsker's inequality to obtain $d_{\mathrm{TV}}^{\mathrm{ub(H)}}$ as in Equation (52) can be too loose for n not sufficiently large. Consider the impact of $d_{\mathrm{TV}}^{\mathrm{ub}}$ on the reliability performance of the code. The average error probability bounds in Equations (53) and (54) are modeled as the sum of two terms, one depending directly on $d_{\mathrm{TV}}^{\mathrm{ub}}$ and the other depending on the polar construction (which has been plotted in Figure 10B). Since $d_{\mathrm{TV}}^{\mathrm{ub}(H)}$ is too loose, what we obtain is that the reliability performance of the code will be governed practically by the bound $d_{\mathrm{TV}}^{\mathrm{ub}}$ for small values of the blocklength n. Now, consider the impact of $d_{\mathrm{TV}}^{\mathrm{ub}}$ on the secrecy performance of the code. The bound on the information leakage in Equation (55) is modeled as the sum of two terms, one also depending only on the polar code construction (which has been plotted in Figure 10A) and the other depending on $d_{\mathrm{TV}}^{\mathrm{ub}}$. However, in this situation, $d_{\mathrm{TV}}^{\mathrm{ub}}$ impacts the information leakage approximately as $n \cdot d_{\mathrm{TV}}^{\mathrm{ub}}$, which means that this term will totally govern the secrecy performance. Recall that this term follows from Equation (44), which bounds the impact of the encoding in Equation (36) on the conditional entropy term of the information leakage as a function of the total variation distance. Hence, we can conclude that this bound, which follows from applying [30] (Lemma 2.9), can be too loose for n not sufficiently large.

7. Conclusions

We have described two polar coding schemes for two different models over the degraded broadcast channel: DBC-NLD-LS and DBC-LD-NLS. For both models, we have proven that the proposed polar coding schemes are asymptotically secrecy-capacity achieving, providing reliability and strong secrecy simultaneously. Then, we have discussed how to construct these polar codes in practice, and we have evaluated their performance for a finite blocklength by means of simulations. Although several polar code constructions methods have been proposed in the literature, this paper, as far as we know, is the first to discuss practical constructions when the polar code must satisfy both reliability and secrecy constraints. In addition, we have evaluated the secrecy performance of the polar code in terms of the strong secrecy performance, which has been possible by obtaining an upper-bound on the corresponding information leakage at the eavesdroppers. Indeed, we have shown that the proposed polar coding schemes can perform well in practice for a finite blocklength.

The criteria we have chosen for designing the polar codes are: to provide reliability and strong secrecy in one block of size n by using only a secret key that is negligible in terms of rate and to minimize the amount of random decisions for the SC encoding. For the first purpose, we have introduced the source of common randomness, and we have avoided the use of the chaining construction given in [9] (which is possible due to the degraded nature of the broadcast channel); for the second one, we have adapted the deterministic SC encoding given in [20]. These two types of randomness have different implications on the practical design: while the common randomness is uniformly distributed and can be provided by the communication system, the randomness for SC encoding is not and must be drawn by the encoder. In communication scenarios requiring several transmissions of size n, we have shown that one realization of the common randomness can be reused without worsening the performance.

Despite the good performance of the polar coding schemes, some issues still persist. How to avoid the transmissions of the additional secret sequences is a problem that remains open. Despite the length of the required secret key being asymptotically negligible in terms of rate, these additional transmissions can be problematic in practical scenarios. As pointed out in Remark 4, one can adopt the chaining construction in [9] to further reduce the length of these sequences, but this requires the transmission to take place over several blocks of size n and a very large memory capacity at the transmitter or receiver side. Furthermore, despite the rate of the amount of randomness required for SC encoding being negligible, how to replace the random decisions entirely by deterministic ones is a problem that still remains unsolved. Another problem that remains open is how to avoid the use of the common randomness, which allows keyless secret communication over a single block of size n (keyless in the sense that the rate of the required secret key is negligible). Finally, to design polar codes based on the proposed performance evaluation, it seems necessary to find tighter upper-bounds on the total variation distance between the distribution induced by the encoder and the original distribution used in the code construction, particularly for the term that models the impact of storing uniformly-distributed sequences. Also, for the secrecy performance, it would be interesting to find a tighter upper-bound to evaluate the impact of the total variation distance on the information leakage.

Lastly, it is worth mentioning that having to know the statistics of the eavesdropper channels for the polar code construction may seem problematic. Nevertheless, for the polar code construction, one can consider virtual eavesdroppers with some target channel qualities. For DBC-LD-NLS, we can design a polar code according to the statistics of this virtual eavesdropper, and due to the degradedness condition of the channel, this code will perform well if the real eavesdroppers have worse channel quality (worst-case design). On the other hand, for the DBC-NLD-LS, one can simply consider different levels of secrecy depending on different target channel qualities. Depending on the channel quality of the real eavesdropper with respect to the virtual ones considered for the design, the polar coding scheme will provide a particular secrecy performance level.

Supplementary Materials: The MATLAB code used in this paper for Section 6 is available at http://www.mdpi.com/1099-4300/20/6/467/?.

Author Contributions: Conceptualization, J.d.O.A. and J.R.F. Formal analysis, J.d.O.A. Funding acquisition, J.R.F. Investigation, J.d.O.A. and J.R.F. Methodology, J.d.O.A. and J.R.F. Software, J.d.O.A. Supervision, J.R.F. Validation, J.R.F. Writing, original draft, J.d.O.A.

Funding: This work is supported by the "Ministerio de Ciencia, Innovación y Universidades" and the "Agencia Estatal de Investigación" of the Spanish Government, ERDF funds (TEC2013-41315-R, TEC2015-69648-REDC, TEC2016-75067-C4-2-R) and the Catalan Government (2017 SGR 578 AGAUR).

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

Abbreviations

The following abbreviations are used in this manuscript:

DBC	Degraded Broadcast Channel
DBC-NLD-LS	Degraded Broadcast Channel with Non-Layered Decoding and Layered Secrecy
DBC-LD-NLS	Degraded Broadcast Channel with Layered Decoding and Non-Layered Secrecy
SC	Successive Cancellation
DMS	Discrete Memoryless Source
BEC	Binary Erasure Channel
BSC	Binary Symmetric Channel
BE-BC	Binary Erasure Broadcast Channel
BS-BC	Binary Symmetric Broadcast Channel

Appendix A. Proof of Lemmas 2 and 3

Consider a DMS $(\mathcal{V}_1 \times \cdots \times \mathcal{V}_L \times \mathcal{Y}_K \times \cdots \times \mathcal{Y}_1 \times \mathcal{Z}_M \times \cdots \times \mathcal{Z}_1, p_{V_1 \ldots V_L Y_K \ldots Y_1 Z_M \ldots Z_1})$, the joint distribution of which satisfies the Markov chain condition $V_1 - \cdots - V_L - Y_K - \cdots - Y_1 - Z_M - \cdots - Z_1$. Consider an i.i.d. n-sequence $(V_1^n, \ldots, V_L^n, Y_K^n, \ldots, Y_1^n, Z_M^n, \ldots, Z_1^n)$ of this DMS, n being any power of two. We define the polar transforms $(U_1^n, \ldots, U_L^n)$, where $U_\ell^n \triangleq V_\ell^n G_n$ for each $\ell \in [1, L]$, with joint distribution $p_{U_1^n \ldots U_L^n}$. Then, define $\mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}$ and $\mathcal{L}_{V_\ell|V_{\ell-1}}^{(n)}$ as in Equations (27) and (28), where $V_0 = U_0 \triangleq \varnothing$. Let $V_L \triangleq X$; if $L \triangleq 1$, notice that this DMS is the one considered for the code construction of DBC-NLD-LS. Otherwise, if $L \triangleq K$, it is the one considered for DBC-LD-NLS.

Now, consider the polar encoding procedures described for both models in Sections 4.2 and 5.2. Let $\tilde{q}_{U_1^n \ldots U_L^n}$ be the joint distribution of $(\tilde{U}_1^n, \ldots, \tilde{U}_L^n)$ after the encoding. For both models, we have:

$$\tilde{q}_{U_1^n \ldots U_L^n}(\tilde{u}_1^n, \ldots, \tilde{u}_L^n) = \prod_{\ell=1}^{L} \prod_{j=1}^{n} \tilde{q}_{U_\ell(j)|U_\ell^{1:j-1} V_{\ell-1}^n}\big(\tilde{u}_\ell(j) \big| \tilde{u}_\ell^{1:j-1}, \tilde{u}_{\ell-1}^n G_n\big),$$

where, for all $\ell \in [1, L]$,

$$\tilde{q}_{U_\ell(j)|U_\ell^{1:j-1} V_{\ell-1}^n}\big(\tilde{u}_\ell(j) \big| \tilde{u}_\ell^{1:j-1}, \tilde{v}_{\ell-1}^n\big) = \begin{cases} \frac{1}{2} & \text{if } j \in \mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}, \\ p_{U_\ell(j)|U_\ell^{1:j-1} V_{\ell-1}^n}\big(\tilde{u}_\ell(j) \big| \tilde{u}_\ell^{1:j-1}, \tilde{v}_{\ell-1}^n\big) & \text{if } j \in \big(\mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}\big)^{\mathrm{C}} \cap \big(\mathcal{L}_{V_\ell|V_{\ell-1}}^{(n)}\big)^{\mathrm{C}}, \\ \mathbb{1}\big\{\tilde{u}_\ell(j) = \xi^{(j)}\big(\tilde{u}_\ell^{1:j-1}, \tilde{v}_{\ell-1}^n\big)\big\} & \text{if } j \in \mathcal{L}_{V_\ell|V_{\ell-1}}^{(n)}, \end{cases} \tag{A1}$$

$p_{U_\ell(j)|U_\ell^{1:j-1} V_{\ell-1}^n}$ being the distribution induced by the original DMS and $\xi^{(j)}$ being the deterministic arg max function given in Equation (18) for DBC-NLD-LS or given in Equation (37) for DBC-LD-NLS.

Additionally, consider another encoding process that constructs $(\check{U}_1^n, \ldots, \check{U}_L^n)$ by omitting the use of the deterministic arg max function, but samples $\check{U}_1(j)$ from the distribution:

$$\check{q}_{U_\ell(j)|U_\ell^{1:j-1}V_{\ell-1}^n}(\check{u}_\ell(j)|\check{u}_\ell^{1:j-1}, \check{v}_{\ell-1}^n) = \begin{cases} \frac{1}{2} & \text{if } j \in \mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}, \\ p_{U_\ell(j)|U_\ell^{1:j-1}V_{\ell-1}^n}(\check{u}_\ell(j)|\check{u}_\ell^{1:j-1}, \check{v}_{\ell-1}^n) & \text{if } j \in (\mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)})^{\mathrm{C}}. \end{cases} \tag{A2}$$

First, the following lemma shows that the joint distributions $p_{U_1^n \ldots U_L^n}$ and $\check{q}_{U_1^n \ldots U_L^n}$ are nearly statistically indistinguishable for sufficiently large n.

Lemma A1. *Let $\delta_n = 2^{-n^\beta}$ for some $\beta \in (0, \frac{1}{2})$, and define $\delta_n^{(1)} \triangleq \sqrt{2n\delta_n \ln 2}$. Then,*

$$\mathbb{V}(\check{q}_{U_1^n \ldots U_L^n}, p_{U_1^n \ldots U_L^n}) \leq \sqrt{L}\delta_n^{(1)}.$$

Proof. The Kullback-Leibler distance between $p_{U_1^n \ldots U_L^n}$ and $\check{q}_{U_1^n \ldots U_L^n}$ is:

$$\begin{aligned} \mathbb{D}(p_{U_1^n \ldots U_L^n} \| \check{q}_{U_1^n \ldots U_L^n}) &\overset{(a)}{=} \textstyle\sum_{\ell=1}^{L} \sum_{j=1}^{n} \mathbb{E}_{p_{U_\ell^{1:j-1}V_{\ell-1}^n}} \left[\mathbb{D}\left(p_{U_\ell(j)|U_\ell^{1:j-1}V_{\ell-1}^n} \middle\| \check{q}_{U_\ell(j)|U_\ell^{1:j-1}V_{\ell-1}^n} \right) \right] \\ &\overset{(b)}{=} \textstyle\sum_{\ell=1}^{L} \sum_{j \in \mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}} \left(1 - H\left(U_\ell(j) \middle| U_\ell^{1:j-1}, V_{\ell-1}^n \right) \right) \\ &\overset{(c)}{\leq} L\delta_n |\mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}|, \end{aligned} \tag{A3}$$

where (a) holds by the chain rule, the invertibility of G_n and the fact that $U_1^n - U_2^n - \cdots - U_L$ (and $\check{U}_1^n - \check{U}_2^n - \cdots - \check{U}_L$) forms a Markov chain, (b) follows from Equation (A2) and by applying [14] (Lemma 10), and (c) holds by the definition of $\mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}$ in Equation (27). Finally, since $|\mathcal{H}_{V_\ell|V_{\ell-1}}^{(n)}| \leq n$ and by using Pinsker's inequality, we obtain $\mathbb{V}(\check{q}_{U_1^n \ldots U_L^n}, p_{U_1^n \ldots U_L^n}) \leq \sqrt{2Ln\delta_n \ln 2}$. □

Now, we show that $\check{q}_{U_1^n \ldots U_L^n}$ and $\tilde{q}_{U_1^n \ldots U_L^n}$ are nearly indistinguishable for n large enough.

Lemma A2. *Let $\delta_n = 2^{-n^\beta}$ for some $\beta \in (0, \frac{1}{2})$. Then,*

$$\mathbb{V}(\tilde{q}_{U_1^n \ldots U_L^n}, \check{q}_{U_1^n \ldots U_L^n}) \leq \delta_n^{(2)},$$

where $\delta_n^{(2)} \triangleq Ln\sqrt{2\sqrt{2}\delta_n^{(1)}(2n - \log\sqrt{2}\delta_n^{(1)}) + \delta_n}$ and $\delta_n^{(1)}$ defined as in Lemma A1.

Proof. The proof follows similar reasoning as the one for [20] (Lemma 2). Hence, define a coupling [29] for $(\check{U}_1^n, \ldots, \check{U}_L^n)$ and $(\tilde{U}_1^n, \ldots, \tilde{U}_L^n)$ such that $\check{U}_\ell[(\mathcal{L}_{V_\ell|V_{\ell-1}}^{(n)})^{\mathrm{C}}] = \tilde{U}_\ell[(\mathcal{L}_{V_\ell|V_{\ell-1}}^{(n)})^{\mathrm{C}}]$. Thus, we have:

$$\begin{aligned} \mathbb{V}(\tilde{q}_{U_1^n \ldots U_L^n}, \check{q}_{U_1^n \ldots U_L^n}) &\overset{(a)}{\leq} \mathbb{P}\left[(\tilde{U}_1^n, \ldots, \tilde{U}_L^n) \neq (\check{U}_1^n, \ldots, \check{U}_L^n) \right] \\ &\overset{(b)}{\leq} \textstyle\sum_{\ell=1}^{L} \mathbb{P}\left[\tilde{U}_\ell^n \neq \check{U}_\ell^n \middle| \tilde{V}_{\ell-1}^n = \check{V}_{\ell-1}^n \right] \\ &\overset{(c)}{\leq} \textstyle\sum_{\ell=1}^{L} \sum_{j=1}^{n} \mathbb{P}\left[\tilde{U}_\ell(j) \neq \check{U}_\ell(j) \middle| \tilde{U}_\ell^{1:j-1} = \check{U}_\ell^{1:j-1}, \tilde{V}_{\ell-1}^n = \check{V}_{\ell-1}^n \right] \\ &\overset{(d)}{=} \textstyle\sum_{\ell=1}^{L} \sum_{j \in \mathcal{L}_{V_\ell|V_{\ell-1}}^{(n)}} \mathbb{E}_{(\check{U}_\ell^{1:j-1}, \check{V}_{\ell-1}^n)} \left[\left(1 - p_{U_\ell(j)|U_\ell^{1:j-1}V_{\ell-1}^n}\left(u_\ell^*(j) \middle| \check{U}_\ell^{1:j-1}, \check{V}_{\ell-1}^n \right) \right) \right], \end{aligned} \tag{A4}$$

where (a) follows from the coupling lemma [29] (Proposition 4.7), (b) holds by the union bound, the invertibility of G_n and the fact that $\tilde{U}_1^n - \tilde{U}_2^n - \cdots - \tilde{U}_L$ (and $\check{U}_1^n - \check{U}_2^n - \cdots - \check{U}_L$) forms a Markov chain, (c) also holds by the union bound and (d) follows from

Equations (A1) and (A2) given that $\check{U}_\ell[(\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}})^{\mathrm{C}}] = \tilde{U}_\ell[(\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}})^{\mathrm{C}}]$ and from defining $u^*_\ell(j) \triangleq \arg\max_{u\in\{0,1\}} p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}(u|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1})$.

Next, for any $\ell \in [1, L]$ and $j \in [n]$, for sufficiently large n, we have:

$$\begin{aligned}
&\left|H(U_\ell(j)|U_\ell^{1:j-1}, V^n_{\ell-1}) - H(U_\ell(j)|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1})\right| \\
&\overset{(a)}{\leq} \left|H(U_\ell^{1:j-1}, V^n_{\ell-1}) - H(\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1})\right| + \left|H(U_\ell^{1:j}, V^n_{\ell-1}) - H(U_\ell(j), \check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1})\right| \\
&\overset{(b)}{\leq} 2\mathbb{V}(\check{q}_{U_\ell^{1:j-1}U^n_{\ell-1}}, p_{U_\ell^{1:j-1}U^n_{\ell-1}}) \log \frac{2^n}{\mathbb{V}(\check{q}_{U_\ell^{1:j-1}U^n_{\ell-1}}, p_{U_\ell^{1:j-1}U^n_{\ell-1}})} \\
&\overset{(c)}{\leq} 2\sqrt{2}\delta_n^{(1)}(2n - \log\sqrt{2}\delta_n^{(1)}),
\end{aligned} \tag{A5}$$

where (a) holds by the chain rule of entropy and the triangle inequality, (b) follows from applying [30] (Lemma 2.9), the invertibility of G_n and because $\mathbb{V}(p_{U_\ell(j)|U_\ell^{1:j-1}U^n_{\ell-1}}\check{q}_{U_\ell^{1:j-1}U^n_{\ell-1}}, p_{U_\ell^{1:j}U^n_{\ell-1}}) = \mathbb{V}(\check{q}_{U_\ell^{1:j-1}U^n_{\ell-1}}, p_{U_\ell^{1:j-1}U^n_{\ell-1}})$, and (c) holds because $\mathbb{V}(\check{q}_{U_\ell^{1:j-1}U^n_{\ell-1}}, p_{U_\ell^{1:j-1}U^n_{\ell-1}}) \leq \mathbb{V}(\check{q}_{U^n_{\ell-1}U^n_\ell}, p_{U^n_{\ell-1}U^n_\ell}) \leq \sqrt{2}\delta_n^{(1)}$ (by using Lemma A1 and taking $L \triangleq 2$) and because the function $x \mapsto x\log x$ is monotonically decreasing for $x > 0$ small enough.

Thus, for any $\ell \in [1, L]$ and $j \in \mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}}$, we have:

$$\begin{aligned}
&2\sqrt{2}\delta_n^{(1)}(2n - \log\sqrt{2}\delta_n^{(1)}) + \delta_n \\
&\overset{(a)}{\geq} 2\sqrt{2}\delta_n^{(1)}(2n - \log\sqrt{2}\delta_n^{(1)}) + H\Big(U_\ell(j)|U_\ell^{1:j-1}, V^n_{\ell-1}\Big) \\
&\overset{(b)}{\geq} H\Big(U_\ell(j)|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1}\Big) \\
&= \mathbb{E}_{(\check{U}_\ell^{1:j-1}, \check{U}^n_{\ell-1})}\left[h_2\left(p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}\left(u^\star_\ell(j)\middle|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1}\right)\right)\right] \\
&\geq \mathbb{E}_{(\check{U}_\ell^{1:j-1}, \check{U}^n_{\ell-1})}\left[-\left(1 - p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}\left(u^\star_\ell(j)\middle|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1}\right)\right)\right. \\
&\quad \left.\cdot \log\left(1 - p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}\left(u^\star_\ell(j)\middle|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1}\right)\right)\right] \\
&\overset{(c)}{\geq} \mathbb{E}_{(\check{U}_\ell^{1:j-1}, \check{U}^n_{\ell-1})}\left[\left(1 - p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}\left(u^\star_\ell(j)\middle|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1}\right)\right)^2\right] \\
&\overset{(d)}{\geq} \left(\mathbb{E}_{(\check{U}_\ell^{1:j-1}, \check{U}^n_{\ell-1})}\left[\left(1 - p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}\left(u^\star_\ell(j)\middle|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1}\right)\right)\right]\right)^2,
\end{aligned} \tag{A6}$$

where (a) holds because, by definition, $H\Big(U_\ell(j)|U_\ell^{1:j-1}, V^n_{\ell-1}\Big) \leq \delta_n$ if $j \in \mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}}$, (b) holds by Equation (A5), (c) holds because $p_{U_\ell(j)|U_\ell^{1:j-1}V^n_{\ell-1}}(u^\star_\ell(j)|\check{U}_\ell^{1:j-1}, \check{V}^n_{\ell-1}) \geq 1/2$ and $\log(x) < -x$ if $x \in [0, 1/2)$ and (d) follows from Jensen's inequality.

Finally, by combining Equations (A4) and (A6) and because $|\mathcal{L}^{(n)}_{V_\ell|V_{\ell-1}}| \leq n$, we have $\mathbb{V}(\tilde{q}_{U_1^n \ldots U_L^n}, \check{q}_{U_1^n \ldots U_L^n}) \leq Ln\sqrt{2\sqrt{2}\delta_n^{(1)}(2n - \log\sqrt{2}\delta_n^{(1)}) + \delta_n}$. □

Hence, by Lemma A1, Lemma A2 and by applying the triangle inequality, we obtain:

$$\begin{aligned}
\mathbb{V}(\tilde{q}_{U_1^n \ldots U_L^n}, p_{U_1^n \ldots U_L^n}) &\leq \mathbb{V}(\tilde{q}_{U_1^n \ldots U_L^n}, \check{q}_{U_1^n \ldots U_L^n}) + \mathbb{V}(\check{q}_{U_1^n \ldots U_L^n}, p_{U_1^n \ldots U_L^n}) \\
&\leq Ln\sqrt{2\sqrt{2}\delta_n^{(1)}(2n - \log\sqrt{2}\delta_n^{(1)}) + \delta_n} + \sqrt{L}\delta_n^{(1)}.
\end{aligned} \tag{A7}$$

Consequently, since $\tilde{q}_{Y_K^n \dots Y_1^n Z_M^n \dots Z_1^n | V_1^n \dots V_L^n} = p_{Y_K^n \dots Y_1^n Z_M^n \dots Z_1^n | V_1^n \dots V_L^n}$ and the invertibility of G_n, we obtain $\mathbb{V}(\tilde{q}_{V_1^n \dots V_L^n Y_K^n \dots Y_1^n Z_M^n \dots Z_1^n}, p_{V_1^n \dots V_L^n Y_K^n \dots Y_1^n Z_M^n \dots Z_1^n}) = \mathbb{V}(\tilde{q}_{U_1^n \dots U_L^n}, p_{U_1^n \dots U_L^n})$, and this concludes the proof.

References and Notes

1. Wyner, A. The wire-tap channel. *Bell Syst. Tech. J.* **1975**, *54*, 1355–1387, doi:10.1002/j.1538-7305.1975.tb02040.x. [CrossRef]
2. Csiszár, I.; Körner, J. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory* **1978**, *24*, 339–348, doi:10.1109/TIT.1978.1055892. [CrossRef]
3. Maurer, U.; Wolf, S. Information-theoretic key agreement: From weak to strong secrecy for free. In Advances in Cryptology—EUROCRYPT 2000; Springer: Berlin/Heidelberg, Germany, 2000; pp. 351–368.
4. Zou, S.; Liang, Y.; Lai, L.; Poor, H.; Shamai, S. Broadcast networks with layered decoding and layered secrecy: Theory and applications. *Proc. IEEE* **2015**, *103*, 1841–1856, doi:10.1109/JPROC.2015.2458338. [CrossRef]
5. Liang, Y.; Lai, L.; Poor, H.V.; Shamai, S. A broadcast approach for fading wiretap channels. *IEEE Trans. Inf. Theory* **2014**, *60*, 842–858, doi:10.1109/TIT.2013.2293756. [CrossRef]
6. Ekrem, E.; Ulukus, S. Secrecy capacity of a class of broadcast channels with an eavesdropper. *EURASIP J. Wirel. Commun. Netw.* **2009**, *2009*, doi:10.1155/2009/824235. [CrossRef]
7. Arikan, E. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *IEEE Trans. Inf. Theory* **2009**, *55*, 3051–3073. [CrossRef]
8. Mahdavifar, H.; Vardy, A. Achieving the secrecy capacity of wiretap channels using polar codes. *IEEE Trans. Inf. Theory* **2011**, *57*, 6428–6443, doi:10.1109/TIT.2011.2162275. [CrossRef]
9. Şaşoğlu, E.; Vardy, A. A new polar coding scheme for strong security on wiretap channels. In Proceedings of the IEEE International Symposium on Information Theory Proceedings (ISIT), Istanbul, Turkey, 7–12 July 2013; pp. 1117–1121, doi:10.1109/ISIT.2013.6620400. [CrossRef]
10. Renes, J.M.; Renner, R.; Sutter, D. Efficient one-way secret key agreement and private channel coding via polarization. In *Advances in Cryptology-ASIACRYPT*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 194–213.
11. Wei, Y.; Ulukus, S. Polar coding for the general wiretap channel with extensions to multiuser scenarios. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 278–291. doi:10.1109/JSAC.2015.2504275. [CrossRef]
12. Cihad Gulcu, T.; Barg, A. Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component. *arXiv* **2014**, arXiv:1410.3422.
13. Chou, R.A.; Bloch, M.R. Polar coding for the broadcast channel with confidential messages: A random binning analogy. *IEEE Trans. Inf. Theory* **2016**, *62*, 2410–2429, doi:10.1109/TIT.2016.2539145. [CrossRef]
14. Goela, N.; Abbe, E.; Gastpar, M. Polar codes for broadcast channels. *IEEE Trans. Inf. Theory* **2015**, *61*, 758–782, doi:10.1109/TIT.2014.2378172. [CrossRef]
15. Chou, R.A.; Bloch, M.R.; Abbe, E. Polar coding for secret-key generation. *IEEE Trans. Inf. Theory* **2015**, *61*, 6213–6237, doi:10.1109/TIT.2015.2471179. [CrossRef]
16. Wang, L.; Sasoglu, E. Polar coding for interference networks. In Proceedings of the 2014 IEEE International Symposium on Information Theory, Honolulu, HI, USA, 29 June–4 July 2014; pp. 311–315, doi:10.1109/ISIT.2014.6874845. [CrossRef]
17. Chou, R.A.; Yener, A. Polar coding for the multiple access wiretap channel via rate-splitting and cooperative jamming. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016; pp. 983–987, doi:10.1109/ISIT.2016.7541446. [CrossRef]
18. Hirche, C.; Morgan, C.; Wilde, M.M. Polar codes in network quantum information theory. *IEEE Trans. Inf. Theory* **2016**, *62*, 915–924, doi:10.1109/TIT.2016.2514319. [CrossRef]
19. Bloch, M.; Barros, J. *Physical-Layer Security: From Information Theory to Security Engineering*; Cambridge University Press: Cambridge, UK, 2011.
20. Chou, R.A.; Bloch, M.R. Using deterministic decisions for low-entropy bits in the encoding and decoding of polar codes. In Proceedings of the 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL, USA, 29 September–2 October 2015; pp. 1380–1385, doi:10.1109/ALLERTON.2015.7447169. [CrossRef]
21. Tal, I.; Vardy, A. How to construct polar codes. *IEEE Trans. Inf. Theory* **2013**, *59*, 6562–6582. [CrossRef]

22. Vangala, H.; Viterbo, E.; Hong, Y. A comparative study of polar code constructions for the AWGN channel. *arXiv* **2015**, arXiv:1501.02473.
23. Honda, J.; Yamamoto, H. Polar coding without alphabet extension for asymmetric models. *IEEE Trans. Inf. Theory* **2013**, *59*, 7829–7838, doi:10.1109/TIT.2013.2282305. [CrossRef]
24. Throughout this paper, we assume binary polarization. An extension to q-ary alphabets is possible [25,26].
25. Karzand, M.; Telatar, E. Polar codes for q-ary source coding. In Proceedings of the 2010 IEEE International Symposium on Information Theory, Austin, TX, USA, 12–18 June 2010; pp. 909–912, doi:10.1109/ISIT.2010.5513555. [CrossRef]
26. Şasoğlu, E.; Telatar, E.; Arikan, E. Polarization for arbitrary discrete memoryless channels. In Proceedings of the IEEE Information Theory Workshop, Sicily, Italy, 11–16 October 2009; pp. 144–148.
27. Arikan, E. Source polarization. In Proceedings of the 2010 IEEE International Symposium on Information Theory, Austin, TX, USA, 12–18 June 2010; pp. 899–903.
28. Korada, S.B.; Urbanke, R.L. Polar codes are optimal for lossy source coding. *IEEE Trans. Inf. Theory* **2010**, *56*, 1751–1768. [CrossRef]
29. Levin, D.A.; Peres, Y.; Wilmer, E.L. *Markov Chains and Mixing Times*; American Mathematical Society: Providence, RI, USA, 2009.
30. Csiszar, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Cambridge University Press: Cambridge, UK, 2011.
31. Pearl, J. *Causality*; Cambridge University Press: Cambridge, UK, 2009.
32. Most of the code in MATLAB is adapted from https://ecse.monash.edu/staff/eviterbo/polarcodes.html.
33. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2012.