Polar Coding for the Broadcast Channel with Confidential Messages and Constrained Randomization

Rémi A. Chou, Matthieu R. Bloch

Abstract

We develop a low-complexity polar coding scheme for the discrete memoryless broadcast channel with confidential messages under strong secrecy and randomness constraints. Our scheme extends previous work by using an optimal rate of uniform randomness in the stochastic encoder, and avoiding assumptions regarding the symmetry or degraded nature of the channels. The price paid for these extensions is that the encoder and decoders are required to share a secret seed of negligible size and to increase the block length through chaining. We also highlight a close conceptual connection between the proposed polar coding scheme and a random binning proof of the secrecy capacity region.

I. INTRODUCTION

With the renewed interest for information-theoretic security, there have been several attempts to develop low-complexity coding schemes achieving the fundamental secrecy limits of the wiretap channel models. In particular, explicit coding schemes based on low-density parity-check codes [1]–[3], polar codes [4]–[7], and invertible extractors [8], [9] have been successfully developed for special cases of Wyner’s model [10], in which the channels are at least required to be symmetric. The recently introduced chaining techniques for polar codes provide, however, a convenient way to construct explicit low-complexity coding schemes for a variety of information-theoretic channel models [11], [12] without any restrictions on the channels.

In this paper, we develop a low-complexity polar coding scheme for the broadcast channel with confidential messages [13]. Rather than view randomness as a free resource, which could be used to
simulate random numbers at arbitrary rate with no cost, we adopt the point of view put forward in [14], [15], in which any randomness used for stochastic encoding must be explicitly accounted for. In particular, our proposed polar coding scheme exploits the optimal rate of randomness identified in [14] and relies on polar codes for channel prefixing.

Results closely related to the present work have been independently developed in [16], [17]. However, these works do not consider randomness as a resource and assume that channel prefixing can be performed through other means; in addition [17] only focuses on weak secrecy. When specialized to Wyner’s wiretap model, our scheme also resembles that in [6], but with a number of notable distinctions. Specifically, while no pre-shared secret seed is required in [6], the coding scheme therein relies on a two-layer construction for which no efficient code construction is presently known [6, Section 3.3]. In contrast, our coding scheme requires a pre-shared secret seed, but at the benefit of only using a single layer of polarization.

The remaining of the paper is organized as follows. Section II formally introduces the notation and the model under investigation. Section III develops a random binning proof of the results in [14], which serves as a guideline for the design of the polar coding scheme. Section IV describes the proposed polar coding scheme in details, while Section V provides its detailed analysis. Section VI offers some concluding remarks.

II. BROADCAST CHANNEL WITH CONFIDENTIAL MESSAGES AND CONSTRAINED RANDOMIZATION

A. Notation

We define the integer interval $[a, b]$, as the set of integers between $a$ and $b$. For $n \in \mathbb{N}$ and $N \triangleq 2^n$, we let $G_n \triangleq \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \otimes n$ be the source polarization transform defined in [18]. We note the components of a vector, $X^{1:N}$, of size $N$, with superscripts, i.e., $X^{1:N} \triangleq (X^1, X^2, \ldots, X^N)$. When the context makes clear that we are dealing with vectors, we write $X^N$ in place of $X^{1:N}$. We note the variational distance and the divergence, respectively, between two distributions. Finally, we note the indicator function $1\cdot \{\omega\}$, which is equal to 1 if the predicate $\omega$ is true and 0 otherwise.

B. Channel model and capacity region

We consider the problem of secure communication over a discrete memoryless broadcast channel $(\mathcal{X}, p_{YZ|X}, \mathcal{Y}, \mathcal{Z})$ illustrated in Figure 1. The marginal probabilities $p_{Y|X}$ and $p_{Z|X}$ define two Discrete Memoryless Channels (DMCs) $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ and $(\mathcal{X}, p_{Z|X}, \mathcal{Z})$, which we refer to as Bob’s channel and Eve’s channel, respectively.
Fig. 1. Communication over a broadcast channel with confidential messages. \( O \) is a common message that must be reconstructed by both Bob and Eve. \( S \) is a confidential message that must be reconstructed by Bob and kept secret from Eve. \( M \) is a private message that must be reconstructed by Bob, but may neither be secret nor reconstructed by Eve. \( R \) represents an additional randomization sequence used at the encoder.

\begin{definition}
A \( (2^{NR_O}, 2^{NR_M}, 2^{NR_S}, 2^{NR_R}, N) \) code \( C_N \) for the broadcast channel consists of

\begin{itemize}
  \item a common message set \( \mathcal{O} \triangleq [1, 2^{NR_O}] \)
  \item a private message set \( \mathcal{M} \triangleq [1, 2^{NR_M}] \)
  \item a confidential message set \( \mathcal{S} \triangleq [1, 2^{NR_S}] \)
  \item a randomization sequence set \( \mathcal{R} \triangleq [1, 2^{NR_R}] \)
  \item an encoding function \( f : \mathcal{O} \times \mathcal{M} \times \mathcal{S} \times \mathcal{R} \to X^N \), which maps the messages \((o, m, s)\) and the randomness \(r\) to a codeword \(x^N\)
  \item a decoding function \( g : Y^N \to \mathcal{O} \times \mathcal{M} \times \mathcal{S} \), which maps each observation of Bob’s channel \(y^N\) to the messages \((\hat{o}, \hat{m}, \hat{s})\)
  \item a decoding function \( h : Z^N \to \mathcal{O} \), which maps each observation of Eve’s channel \(z^N\) to the message \(\hat{o}\)
\end{itemize}

For uniformly distributed \( O, M, S, \) and \( R \), the performance of a \( (2^{NR_O}, 2^{NR_M}, 2^{NR_S}, 2^{NR_R}, N) \) code \( C_N \) for the broadcast channel is measured in terms of its probability of error

\[ P_e(C_N) \triangleq \mathbb{P} \left[ (\hat{O}, \hat{M}, \hat{S}) \neq (O, M, S) \text{ or } \hat{O} \neq O \right], \]

and its leakage of information about the confidential message to Eve

\[ L_e(C_N) \triangleq I(S; Z^N). \]
Definition 2. A rate tuple $(R_O, R_M, R_S, R_R)$ is achievable for the broadcast channel if there exists a sequence of $(2^{NR_O}, 2^{NR_M}, 2^{NR_S}, 2^{NR_R}, N)$ codes $\{C_N\}_{N \geq 1}$ such that
\[
\lim_{N \to \infty} P_e(C_N) = 0, \text{(reliability condition)}
\]
\[
\lim_{N \to \infty} L_s(C_N) = 0. \text{(strong secrecy)}
\]
The achievable region $\mathcal{R}_{BCC}$ is defined as the closure of the set of all achievable rate quadruples.

The exact characterization of $\mathcal{R}_{BCC}$ was obtained in [14].

Theorem 1 ([14]). $\mathcal{R}_{BCC}$ is the closed convex set consisting of the quadruples $(R_O, R_M, R_S, R_R)$ for which there exist auxiliary random variables $(U, V)$ such that $U - V - X - (Y, Z)$ and
\[
R_O \leq \min[I(U; Y), I(U; Z)],
\]
\[
R_O + R_M + R_S \leq I(V; Y | U) + \min[I(U; Y), I(U; Z)],
\]
\[
R_S \leq I(V; Y | U) - I(V; Z | U),
\]
\[
R_M + R_R \geq I(X; Z | U),
\]
\[
R_R \geq I(X; Z | V).
\]

The main contribution of the present work is to develop a polar coding scheme achieving the rates in $\mathcal{R}_{BCC}$.

III. A BINNING APPROACH TO CODE DESIGN: FROM RANDOM BINNING TO POLAR BINNING

In this section, we argue that our construction of polar codes for the broadcast channel with confidential messages is essentially the constructive counterpart of a random binning proof of the region $\mathcal{R}_{BCC}$. While random coding is often the natural tool to address channel coding problems, random binning is already found in [19] to establish the strong secrecy of the wiretap channel, and is the tool of choice in quantum information theory [20]; there has also been a renewed interest for random binning proofs in multi-user information theory, motivated in part by [21]. In Section III-A, we sketch a random binning proof of the characterization of $\mathcal{R}_{BCC}$ established in [14], which may be viewed as a refinement of the analysis in [21] to obtain a more precise characterization of the stochastic encoder. While the results we derive are not new, we use this alternative proof in Section III-B to obtain high-level insight into the construction of polar codes. The main benefit is to clearly highlight the crucial steps of the construction in Section IV and of its analysis in Section V. In particular, the rate conditions developed in the random binning proof of Section III-A directly translate into the definition of the polarization sets in Section III-B.
A. Information-theoretic random binning

Information-theoretic random binning proofs rely on the following well-known lemmas. We use the notation $\delta(N)$ to denote an unspecified positive function of $N$ that vanishes as $N$ goes to infinity.

**Lemma 1** (Source-coding with side information). Consider a Discrete Memoryless Source (DMS) $(\mathcal{X} \times \mathcal{Y}, p_{XY})$. For each $x^N \in \mathcal{X}^N$, assign an index $\Phi(x^N) \in [1, 2^{NR}]$ uniformly at random. If $R > H(X|Y)$, then $\exists N_0$ such that $\forall N \geq N_0$, there exists a deterministic function $g_N : [1, 2^{NR}] \times \mathcal{Y}^N \rightarrow \mathcal{X}^N$ such that

$$E_{\Phi}(V(p_{X^N X^N} p_{X^N g_N(Y^N)})) \leq \delta(N).$$

**Lemma 2** (Privacy amplification, channel intrinsic randomness, output statistics of random binning). Consider a DMS $(\mathcal{X} \times \mathcal{Z}, p_{XZ})$ and let $\epsilon > 0$. For each $x^N \in \mathcal{X}^N$, assign an index $\Psi(x^N) \in [1, 2^{NR}]$ uniformly at random. Denote by $q_M$ the uniform distribution on $[1, 2^{NR}]$. If $R < H(X|Z)$, then $\exists N_0$ such that $\forall N \geq N_0$

$$E_{\Psi}(V(p_{\Psi(x^N)Z^N} q_{M}p_{Z^N})) \leq \delta(N).$$

One may obtain more explicit results regarding the convergence to zero in Lemma 1 and Lemma 2, but we ignore this for brevity.

The principle of a random binning proof of Theorem 1 is to consider a DMS $(U \times V \times X \times Y \times Z, p_{UVXYZ})$ such that $U - V - X - Y Z$, and to assign two types of indices to source sequences by random binning. The first type identifies subset of sequences that play the roles of codebooks, while the second type labels sequences with indices that can be thought of as messages. As explained in the next paragraphs, the crux of the proof is to show that the binning can be “inverted,” so that the sources may be generated from independent choices of uniform codebooks and messages.

**Common message encoding.** We introduces two indices $\psi^U \in [1, 2^{N\rho_U}]$ and $o \in [1, 2^{N\rho_O}]$ by random binning on $u^N$ such that:

- $\rho_U > \max (H(U|Y), H(U|Z))$, so that Lemma 1 ensures that the knowledge of $\psi^U$ allows Bob and Eve to reconstruct the sequence $u^N$ with high probability knowing $y^N$ or $z^N$, respectively;
- $\rho_U + R_O < H(U)$, so that Lemma 2 ensures that the indices $\psi^U$ and $o$ are almost uniformly distributed and independent of each other.

The binning scheme induces a joint distribution $p_{U^N \Psi^U \Psi^O}$. To convert the binning scheme into a channel coding scheme, Alice operates as follows. Upon sampling indices $\tilde{\psi}^U \in [1, 2^{N\rho_U}]$ and $\tilde{o} \in [1, 2^{N\rho_O}]$
from independent uniform distributions, Alice stochastically encodes them into a sequence \( \tilde{u}^N \) drawn according to \( p_{U|\Psi V}(\tilde{u}^N | \tilde{\psi}^U, \tilde{\rho}) \). The choice of rates above guarantees that the joint distribution \( p_{U|\Psi V}(\tilde{u}^N | \tilde{\psi}^U, \tilde{\rho}) \) approximates the distribution \( p_{U|\Psi V} \) in variational distance, so that disclosing \( \tilde{\psi}^U \) allows Bob and Eve to decode the sequence \( \tilde{u}^N \).

**Secret and private message encoding.** Following the same approach, we introduce three indices \( \tilde{\psi}^V | U \in [1, 2^{N\rho_V | U}] \), \( \tilde{s} \in [1, 2^{N\rho_S}] \), and \( \tilde{m} \in [1, 2^{N\rho_M}] \) by random binning on \( v^N \) such that

- \( \rho_{V|U} > H(V|UY) \), to ensure that knowing \( \tilde{\psi}^V | U \) and \( u^N \), Bob may reconstruct the sequence \( x^N \);
- \( \rho_{V|U} + R_S + R_M < H(V|UZ) \), to ensure that the indices are almost uniformly distributed and independent of each other, as well as of the source sequences \( U^N \) and \( Z^N \).

The binning scheme induces a joint distribution \( p_{V^N U^N \Psi V | U S M} \). To obtain a channel coding scheme, Alice encodes the realizations of independent and uniformly distributed indices \( \tilde{\psi}^V | U \in [1, 2^{N\rho_V | U}] \), \( \tilde{s} \in [1, 2^{N\rho_S}] \), \( \tilde{m} \in [1, 2^{N\rho_M}] \), and the sequence \( \tilde{u}^N \), into a sequence \( \tilde{v}^N \) drawn according to the distribution \( p_{V^N | U^N \Psi V | U S M}(\tilde{v}^N | \tilde{u}^N, \tilde{\psi}^V | U, \tilde{s}, \tilde{m}) \). The resulting joint distribution is again a close approximation of \( p_{V^N U^N \Psi V | U S M} \), so that the scheme inherits the reliability and secrecy properties of the random binning scheme upon disclosing \( \tilde{\psi}^V | U \).

**Channel prefixing.** Finally, we introduce the indices \( \tilde{\psi}^X | V \in [1, 2^{N\rho_V | X}] \) and \( \tilde{r} \in [1, 2^{N\rho_R}] \) by random binning on \( x^N \) such that

- \( \rho_{X|V} < H(X|V) \) to ensure that \( \tilde{\psi}^X | V \) is independent of the source sequences \( X^N \) and \( Z^N \);
- \( \rho_{X|V} + R_R < H(X|V) \) to ensure that the indices are almost uniformly distributed and independent of each other, as well as of the source sequences \( V^N \).

The binning scheme induces a joint distribution \( p_{X^N V^N U^N \Psi X | V R} \). To obtain a channel prefixing scheme, Alice encodes the realizations of uniformly distributed indices \( \tilde{\psi}^X | V \) and \( \tilde{r} \), and the previously obtained \( \tilde{v}^N \) into a sequence \( \tilde{x}^N \) drawn according to \( p_{X^N | V^N \Psi X | V R}(\tilde{x}^N | \tilde{v}^N, \tilde{\psi}^X | V, \tilde{\rho}) \). The resulting joint distribution induced is once again a close approximation of \( p_{X^N V^N U^N \Psi X | V R} \).

**Chaining to de-randomize the codebooks.** The downside of the schemes described earlier is that they require sharing the indices \( \tilde{\psi}^U, \tilde{\psi}^V | U \), and \( \tilde{\psi}^X | V \), identifying the codebooks between Alice, Bob, and Eve; however, the rate cost may be amortized by reusing the same indices over sequences of \( k \) blocks. Specifically, the union bound shows that the average error probability over \( k \) blocks is at most \( k \) times that of an individual block, and a hybrid argument shows that the information leakage over \( k \) blocks is at most \( k \) times that of an individual block. Consequently, for \( k \) and \( N \) large enough, the impact on the transmission rates is negligible.
**Total amount of randomness.** The total amount of randomness required for encoding includes not only the explicit random numbers used for channel prefixing but also all the randomness required in the stochastic encoding to approximate the source distribution. One can show that the rate randomness specifically used in the stochastic encoding is negligible; we omit the proof of this result for random binning, but this is analyzed precisely for polar codes in Section V.

By combining all the rate constraints above and perform Fourier-Motzkin elimination, one recovers the rates in Theorem 1.

**B. Binning with polar codes**

The main observation to translate the analysis of Section III-A into a polar coding scheme is that Lemma 1 and Lemma 2 have the following counterparts in terms of source polarization.

**Lemma 3** (adapted from [18]). Consider a DMS $(X \times Y, p_{XY})$. For each $x^{1:N} \in \mathbb{F}_2^N$ polarized as $u^{1:N} = G_n x^{1:N}$, let $u^{1:N}[H_{X|Y}]$ denote the high entropy bits of $u^{1:N}$ in positions $H_{X|Y} \triangleq \{i \in [1,n] : H(U^i|U^{1:i-1}Y^N) > \delta_N\}$ and $\delta_N \triangleq 2^{-N^\beta}$ with $\beta \in [0, \frac{1}{2}]$. For every $i \in [1,N]$, sample $\tilde{u}^{1:N}$ from the distribution

$$
\tilde{p}_{U^i|U^{1:i-1}}(\tilde{u}^i|\tilde{u}^{1:i-1}) \triangleq \begin{cases} 1 \{\tilde{u}^i = u^i\} & \text{if } i \in H_{X|Y} \\ p_{U^i|U^{1:i-1}Y^N}(\tilde{u}^i|\tilde{u}^{1:i-1}y^N) & \text{if } i \in H^c_{X|Y}. \end{cases}
$$

and create $\tilde{x}^{1:N} = \tilde{u}^{1:N}G_n$. Then,

$$
\forall\left(p_{X^{1:N}X^{1:N},p_{X^{1:N}X^{1:N}}}\right) \leq \delta_N,
$$

and $\lim_{N \to \infty} |H_{X|Y}| = H(X|Y)$.

In other words, the high entropy bits in positions $H_{X|Y}$ play the same role as the random binning index in Lemma 1. However, note that the construction of $\tilde{x}^{1:N}$ in Lemma 3 is explicitly stochastic.

**Lemma 4** (adapted from [22]). Consider a DMS $(X \times Z, p_{XZ})$. For each $x^{1:N} \in \mathbb{F}_2^N$ polarized as $u^{1:N} = G_n x^{1:N}$, let $u^{1:N}[V_{X|Z}]$ denote the very high entropy bits of $u^{1:N}$ in positions $V_{X|Z} \triangleq \{i \in [1,n] : H(U^i|U^{1:i-1}Z^N) > 1 - \delta_N\}$ and $\delta_N \triangleq 2^{-N^\beta}$ with $\beta \in [0, \frac{1}{2}]$. Denote by $q_{U^{1:N}|V_{X|Z}}$ the uniform distribution of bits in positions $V_{X|Z}$. Then,

$$
\forall\left(p_{U^{1:N}|V_{X|Z}Z^{1:N},q_{U^{1:N}|V_{X|Z}}PZ^{1:N}}\right) \leq \delta_N
$$

and $\lim_{N \to \infty} |V_{X|Z}| = H(X|Z)$ by [22, Lemma 1].
The very high entropy bits in positions $\mathcal{V}_{X|Z}$ therefore play the same role as the random binning index in Lemma 2.

This suggests that any result obtained from random binning could also be derived using source polarization as a linear and low-complexity alternative; intuitively, information theoretic constraints resulting from Lemma 1 translate into the use of “high entropy” sets $\mathcal{H}$, while those resulting from Lemma 2 translate into the use of “very high entropy” sets $\mathcal{V}$. However, unlike the indices resulting from random binning, the high entropy and very high entropy sets may not necessarily be aligned, and the precise design of a polar coding scheme requires more care.

In the remainder of the paper, we consider a DMS $(U \times V \times \mathcal{X} \times \mathcal{Y} \times Z, p_{UV,XY,Z})$ such that $U - V - X - YZ$, $I(V; Y|U) - I(V; Z|U) > 0$, and $|\mathcal{X}| = |U| = |V| = 2$. The extension to larger alphabets is obtained following ideas in [23]. We also assume without loss of generality $I(U; Y) \leq I(U; Z)$, the case $I(U; Y) > I(U; Z)$ is treated similarly.

**Common message encoding.** Define the polar transform of $U^{1:N}$, as $A^{1:N} \triangleq U^{1:N} G_n$ and the associated sets

$$
\mathcal{H}_U \triangleq \{ i \in [1, N] : H(A^i|A^{1:i-1}) > \delta_N \},
$$

$$
\mathcal{V}_U \triangleq \{ i \in [1, N] : H(A^i|A^{1:i-1}) > 1 - \delta_N \},
$$

$$
\mathcal{H}_{U|Y} \triangleq \{ i \in [1, N] : H(A^i|A^{1:i-1}Y^{1:N}) > \delta_N \},
$$

$$
\mathcal{H}_{U|Z} \triangleq \{ i \in [1, N] : H(A^i|A^{1:i-1}Z^{1:N}) > \delta_N \}.
$$

If we could guarantee that $\mathcal{H}_{U|Z} \subseteq \mathcal{H}_{U|Y} \subseteq \mathcal{V}_U$, then we could directly mimic the information-theoretic random binning proof. We would use random bits in positions $\mathcal{H}_{U|Z}$ to identify the code, random bits in positions $\mathcal{V}_U \setminus \mathcal{H}_{U|Z}$ for the message, successive cancellation encoding to compute the bits in positions $\mathcal{V}_U^c$, and approximate the source distribution, and chaining to amortize the rate cost of the bits in positions $\mathcal{H}_{U|Z}$. Unfortunately, the inclusion $\mathcal{H}_{U|Y} \subseteq \mathcal{H}_{U|Z}$ is not true in general, and one must also use chaining as in [11] to “realign” the sets of indices. Furthermore, only the inclusions $\mathcal{H}_{U|Z} \subseteq \mathcal{H}_U$ and $\mathcal{H}_{U|Y} \subseteq \mathcal{H}_U$ are true in general, so that the bits in positions $\mathcal{H}_{U|Z} \cap \mathcal{V}_U^c$ and $\mathcal{H}_{U|Y} \cap \mathcal{V}_U^c$ must be transmitted separately. The precise coding scheme is detailed in Section IV-A.

**Secret and private messages encoding.** Define the polar transform of $V^{1:N}$ as $B^{1:N} \triangleq V^{1:N} G_n$ and the associated sets

$$
\mathcal{V}_{V|U} \triangleq \{ i \in [1, N] : H(B^i|B^{1:i-1}U^{1:N}) > 1 - \delta_N \},
$$

$$
\mathcal{V}_{V|UZ} \triangleq \{ i \in [1, N] : H(B^i|B^{1:i-1}U^{1:N}Z^{1:N}) > 1 - \delta_N \},
$$
\[\mathcal{H}_{V|UY} \triangleq \{ i \in [1, N] : H(B^{i}|B^{1:i-1}U^{1:N}Y^{1:N}) > \delta_N \}, \quad (7)\]
\[\mathcal{V}_{V|UY} \triangleq \{ i \in [1, N] : H(B^{i}|B^{1:i-1}U^{1:N}Y^{1:N}) > 1 - \delta_N \}, \quad (8)\]
\[\mathcal{M}_{UVZ} \triangleq \mathcal{V}_{V|U}\setminus\mathcal{V}_{V|UZ}. \quad (9)\]

If the inclusion \(\mathcal{H}_{V|UY} \subseteq \mathcal{V}_{V|UZ}\) were true, then we would place random bits identifying the codebook in positions \(\mathcal{H}_{V|UY}\), random bits describing the secret message in positions \(\mathcal{V}_{V|UZ}\setminus\mathcal{H}_{V|UY}\), random bits describing the private message in positions \(\mathcal{V}_{V|U}\setminus\mathcal{V}_{V|UZ}\), use successive cancellation encoding to compute the bits in positions \(\mathcal{V}_{V|U}\) and approximate the source distribution, and use chaining to amortize the rate cost of the bits in positions \(\mathcal{H}_{V|UY}\). This is unfortunately again not directly possible in general, and one needs to exploit chaining to realign the indices, and transmit the bits in positions \(\mathcal{H}_{V|UY} \cap \mathcal{V}_{V|U}\) separately and secretly to Bob. The precise coding scheme is detailed in Section IV-B.

**Channel prefixing.** Finally, define the polar transform of \(X^{1:N}\) as \(T^{1:N} \triangleq X^{1:N}G_n\) and the associated sets
\[\mathcal{V}_{X|V} \triangleq \{ i \in [1, N] : H(T^{i}|T^{1:i-1}V^{1:N}) > 1 - \delta_N \}, \quad (10)\]
\[\mathcal{V}_{X|VZ} \triangleq \{ i \in [1, N] : H(T^{i}|T^{1:i-1}V^{1:N}Z^{1:N}) > 1 - \delta_N \}. \quad (11)\]

One performs channel prefixing by placing random bits identifying the code in positions \(\mathcal{V}_{X|VZ}\), random bits describing the randomization sequence in positions \(\mathcal{V}_{X|V}\setminus\mathcal{V}_{X|VZ}\), and using successive cancellation encoding to compute the bits in positions \(\mathcal{V}_{X|V}\) and approximate the source distribution. Chaining is finally used to amortize the cost of randomness for describing the code. The precise coding scheme is detailed in Section IV-C.

### IV. POLAR CODING SCHEME

In this section, we describe the details of the polar coding scheme resulting from the discussion of the previous section. Recall that the joint probability distribution \(p_{UVXYZ}\) of the original source is fixed and defined as in Section III-B. As alluded to earlier, we perform the encoding over \(k\) blocks of size \(N\). We use the subscript \(i \in [1, k]\) to denote random variables associated to encoding Block \(i\). The chaining constructions corresponding to the encoding of the common, secret, and private messages, and randomization sequence, are described in Section IV-A, Section IV-B, and Section IV-C, respectively. Although each chaining is described independently, all messages should be encoded in every block before moving to the next. Specifically, in every block \(i \in [1, k-1]\), Alice successively encodes the common message, the secret and private messages, and performs channel prefixing, before she moves to the next block \(i+1\).
Fig. 2. Chaining for the encoding of the \( A_1^{1:N} \)'s, which corresponds to the encoding of the common messages.

### A. Common message encoding

In addition to the polarization sets defined in (1)-(4) we also define

\[
\mathcal{I}_{UY} \triangleq \mathcal{V}_U \setminus \mathcal{H}_{U|Y},
\]

\[
\mathcal{I}_{UZ} \triangleq \mathcal{V}_U \setminus \mathcal{H}_{U|Z},
\]

\[
\mathcal{A}_{UYZ} \triangleq \text{any subset of } \mathcal{I}_{UZ} \setminus \mathcal{I}_{UY} \text{ with size } |\mathcal{I}_{UZ} \setminus \mathcal{I}_{UY}|.
\]

Note that \( \mathcal{A}_{UYZ} \) exists since we have assumed \( I(U; Y) \leq I(U; Z) \). In fact,

\[
|\mathcal{I}_{UZ} \setminus \mathcal{I}_{UY}| - |\mathcal{I}_{UY} \setminus \mathcal{I}_{UZ}| = |\mathcal{I}_{UZ}| - |\mathcal{I}_{UY}| \geq 0.
\]

The encoding procedure with chaining is summarized in Figure 2.

In Block 1, the encoder forms \( \bar{O}_1^{1:N} \) as follows. Let \( O_1 \) be a vector of \( |\mathcal{I}_{UY}| \) uniformly distributed information bits that represents the common message to be reconstructed by Bob and Eve. Upon observing a realization \( o_1 \), the encoder samples \( \bar{a}_1^{1:N} \) from the distribution \( \bar{p}_{A_1^{1:N}} \) defined as

\[
\bar{p}_{A_1^{1:N}}(a_1^j | a_1^{j-1}) \triangleq \begin{cases} 
1 \left\{ a_1^j = a_1^j \right\} & \text{if } j \in \mathcal{I}_{UY} \\
1/2 & \text{if } j \in \mathcal{V}_U \setminus \mathcal{I}_{UY}, \\
\tilde{p}_{A_1^{1:N}}(a_1^j | a_1^{j-1}) & \text{if } j \in \mathcal{V}_U^c
\end{cases}
\]  
(12)
where the components of \( o_1 \) have been indexed by the set of indices \( \mathcal{I}_{UY} \) for convenience, so that \( O_1 \triangleq \tilde{A}_{1:N}^{1:1}[\mathcal{I}_{UY}] \). The random bits that identify the codebook and that are required to reconstruct \( \tilde{A}_{1:N}^{1:1} \) are \( \tilde{A}_{1:N}^{1:1}[\mathcal{H}_{U|Z}] \) for Eve and \( \tilde{A}_{1:N}^{1:1}[\mathcal{H}_{U|Y}] \) for Bob. Moreover, we note
\[
\Psi_{1}^{U} \triangleq \tilde{A}_{1:N}^{1:1}[\mathcal{V}_{U}\setminus\mathcal{I}_{UY}] = \tilde{A}_{1:N}^{1:1}[\mathcal{V}_{U} \cap \mathcal{H}_{U|Y}],
\]
\[
\Phi_{1}^{U} \triangleq \tilde{A}_{1:N}^{1:1}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \cap \mathcal{V}_{U}].
\]
Both \( \Psi_{1}^{U} \) and \( \Phi_{1}^{U} \) are publicly transmitted to both Bob and Eve. Note that, unlike in the random binning proof, the use of polarization forces us to distinguish the part \( \Psi_{1}^{U} \) that is nearly uniform from the part \( \Phi_{1}^{U} \) that is not. We show later that the rate cost of this additional transmission is negligible. We also write \( O_1 \triangleq [O_{1:1}, O_{1:2}] \), where \( O_{1:1} \triangleq \tilde{A}_{1:N}^{1:1}[\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}] \) and \( O_{1:2} \triangleq \tilde{A}_{1:N}^{1:1}[\mathcal{I}_{UY} \setminus \mathcal{I}_{UZ}] \). We will retransmit \( O_{1:2} \) in the next block following the same strategy as in [11]. Finally, we compute \( \tilde{O}_{1}^{1:1} \triangleq \tilde{A}_{1:N}^{1:1}G_{n} \).

In Block \( i \in [2, k - 1] \), the encoder forms \( \tilde{A}_{1:N}^{1:1} \) as follows. Let \( O_i \) be a vector of \( |\mathcal{I}_{UY}| \) uniformly distributed information bits representing the common message in that block. Upon observing the realizations \( o_i \) and \( o_{i-1} \), the encoder draws \( \tilde{a}_{1:N}^{1:1} \) from the distribution \( \tilde{p}_{A_{1:1}^{1:1}} \) defined as follows.
\[
\tilde{p}_{A_{1:1}^{1:1}}(a_{1:1}^{1:1-1} | a_{1:1}^{1:1-1}) \triangleq \begin{cases} 
1 \left\{ a_{1}^{1} = a_{i}^{1} \right\} & \text{if } j \in \mathcal{I}_{UY} \\
1 \left\{ a_{1}^{i} = a_{i-1,2}^{1} \right\} & \text{if } j \in \mathcal{A}_{UYZ} \\
1 \left\{ a_{1}^{i} = (\psi_{1}^{U})^{j} \right\} & \text{if } j \in \mathcal{V}_{U} \setminus (\mathcal{I}_{UY} \cup \mathcal{A}_{UYZ}) \\
p_{A_{1:1}^{1:1}}(a_{1:1}^{1:1-1} | a_{1:1}^{1:1-1}) & \text{if } j \in \mathcal{V}_{U} 
\end{cases}
\] (13)

where the components of \( o_i \), \( o_{i-1,2} \), and \( \psi_{1}^{U} \), have been indexed by the set of indices \( \mathcal{I}_{UY}, \mathcal{A}_{UYZ}, \) and \( \mathcal{V}_{U} \setminus (\mathcal{I}_{UY} \cup \mathcal{A}_{UYZ}) \), respectively. Consequently, note that
\[
O_i = \tilde{A}_{1:N}^{1:1}[\mathcal{I}_{UY}] \text{ and } O_{i-1,2} = \tilde{A}_{1:N}^{1:1}[\mathcal{A}_{UYZ}].
\]
The random bits that identify the codebook and that are required to reconstruct \( \tilde{A}_{1:N}^{1:1} \) are \( \tilde{A}_{1:N}^{1:1}[\mathcal{H}_{U|Y}] \) for Bob and \( \tilde{A}_{1:N}^{1:1}[\mathcal{H}_{U|Z}] \) for Eve. Parts of these bits depend on messages in previous blocks. For the others, we define
\[
\Psi_{1}^{U} \triangleq \tilde{A}_{1:N}^{1:1}[\mathcal{V}_{U} \setminus (\mathcal{I}_{UY} \cup \mathcal{A}_{UYZ})],
\]
\[
\Phi_{1}^{U} \triangleq \tilde{A}_{1:N}^{1:1}[(\mathcal{H}_{U|Y} \cup \mathcal{H}_{U|Z}) \setminus \mathcal{V}_{U}].
\]
Note that the bits in \( \Psi_{1}^{U} \) are reusing the bits in \( \Psi_{1}^{U} \); however, it is necessary to make the bits \( \Phi_{1}^{U} \) available to both Bob and Eve, to enable the reconstruction of \( O_i \). We show later that this entails a negligible rate
Finally, we write $O_i \triangleq [O_{i,1}, O_{i,2}]$, where $O_{i,1} \triangleq \widetilde{A}_i^{1:N}[\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}]$ and $O_{i,2} \triangleq \widetilde{A}_i^{1:N}[\mathcal{I}_{UY} \setminus \mathcal{I}_{UZ}]$, and we retransmit $O_{i,2}$ in the next block. We finally compute $\widetilde{U}_i^{1:N} \triangleq \widetilde{A}_i^{1:N} G_n$. 

Finally, the encoder forms $\widetilde{A}_k^{1:N}$ in Block $k$, as follows. Let $O_k$ be a vector of $[\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}]$ uniformly distributed bits representing the common message in that block. Given realizations $o_k$ and $o_{k-1}$, the encoder samples $\tilde{a}_k^{1:N}$ from the distribution $\overline{p}_{A_k^{1:N}}$ defined as follows.

$$\overline{p}_{A_k^{1:N}}(a_k^{1:j-1}) \triangleq \begin{cases} 1 \{a_k^j = a_k^j\} & \text{if } j \in \mathcal{I}_{UY} \cap \mathcal{I}_{UZ} \\ 1 \{a_k^j = a_k^{j-2}\} & \text{if } j \in \mathcal{A}_{UY} \\ 1 \{a_k^j = (\psi_1^U)^j\} & \text{if } j \in \mathcal{V}_U \setminus (\mathcal{A}_{UY} \cup (\mathcal{I}_{UY} \cap \mathcal{I}_{UZ})) \\ p_{A_k^{1:N}}(a_k^{1:j-1}) & \text{if } j \in \mathcal{V}_U \setminus \mathcal{I}_{UY} \cap \mathcal{I}_{UZ} \}, \quad (14) \end{cases}$$

where the components of $o_k, o_{k-1,2}$, and $\psi_1^U$ have been indexed by the set of indices $\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}, \mathcal{A}_{UY},$ and $\mathcal{V}_U \setminus (\mathcal{A}_{UY} \cup (\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}))$, respectively. Consequently,

$$O_k = \widetilde{A}_k^{1:N}[\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}], \quad O_{k-1,2} = \widetilde{A}_k^{1:N}[\mathcal{A}_{UY}].$$

The random bits that identify the codebook and that are required to reconstruct $\widetilde{A}_k^{1:N}$ are $\widetilde{A}_k^{1:N}[\mathcal{H}_{U|Y}]$ for Bob and $\widetilde{A}_k^{1:N}[\mathcal{H}_{U|Z}]$ for Eve. Parts of these bits depend on messages in previous blocks. For the others, we define

$$\Psi_k^U \triangleq \widetilde{A}_k^{1:N}[\mathcal{V}_U \setminus (\mathcal{A}_{UY} \cup (\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}))],$$

$$\Phi_k^U \triangleq \widetilde{A}_k^{1:N}[\mathcal{V}_U \setminus (\mathcal{A}_{UY} \cup (\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}))],$$

and note that $\Psi_k^U$ merely reuses the bits of $\Psi_1^U$. $\Phi_k^U$ is made available to both Bob and Eve to help them reconstruct $O_k$, but this incurs a negligible rate cost.

The public transmission of $(\Psi_1^U, \Phi_1^U)$ to perform the reconstruction of the common message is taken into account in the secrecy analysis in Section V.

**B. Secret and private message encoding**

In addition to the polarization set defined in (5)-(9), we also define

$$\mathcal{B}_V|UY \triangleq \text{a fixed subset of } \mathcal{V}_V|UZ \text{ with size } |\mathcal{V}_V|UY \setminus ((\mathcal{H}_{V|UY} \setminus \mathcal{V}_V|UY) \cap \mathcal{V}_V|U))|$$

$$\mathcal{M}_{UVZ} \triangleq \mathcal{V}_V|U \setminus \mathcal{V}_V|UZ.$$ 

The encoding procedure with chaining is summarized in Fig. 3.
In Block 1, the encoder forms $\tilde{V}_1^{1:N}$ as follows. Let $S_1$ be a vector of $|\mathcal{V}_{V\mid UZ}|$ uniformly distributed bits representing the secret message and let $M_1$ be a vector of $|\mathcal{M}_{UVZ}|$ uniformly distributed bits representing the private message to be reconstructed by Bob. Given a confidential message $s_1$, a private message $m_1$, and $\tilde{u}_1^{1:N}$ resulting from the encoding of the common message, the encoder samples $\tilde{b}_1^{1:N}$ from the distribution $\tilde{p}_{B_1^{1:N}}$ defined as follows.

$$
\tilde{p}_{B_1^{1:N}}(b_1^{1:j-1}u_1^{1:N} | s_1^{j-1}) \triangleq \begin{cases} 
1 \{ b_1 = s_1 \} & \text{if } j \in \mathcal{V}_{V\mid UZ} \\
1 \{ b_1 = m_1 \} & \text{if } j \in \mathcal{M}_{UVZ} \\
p_{B_1^{1:N}}(b_1^{1:j-1}u_1^{1:N}) & \text{if } j \in \mathcal{G}_{V\mid U} 
\end{cases}
$$

(15)

where the components of $s_1$ and $m_1$ have been indexed by the set of indices $\mathcal{V}_{V\mid UZ}$ and $\mathcal{M}_{UVZ}$, respectively. Consequently, note that $S_1 = \tilde{B}_1^{1:N}[\mathcal{V}_{V\mid UZ}]$ and $M_1 = \tilde{B}_1^{1:N}[\mathcal{M}_{UVZ}]$. The random bits that identify the codebook required for reconstruction are those in positions $\mathcal{H}_{V\mid UY}$, which we split as

$$
\Psi_1^{V\mid U} \triangleq \tilde{B}_1^{1:N}[\mathcal{V}_{V\mid UY} \cup ((\mathcal{H}_{V\mid UY} \setminus \mathcal{V}_{V\mid UY}) \cap \mathcal{V}_{V\mid U})],
$$

$$
\Phi_1^{V\mid U} \triangleq \tilde{B}_1^{1:N}[(\mathcal{H}_{V\mid UY} \setminus \mathcal{V}_{V\mid UY} \cap \mathcal{V}_{V\mid U})].
$$

Note that $\Psi_1^{V\mid U}$ is uniformly distributed but $\Phi_1^{V\mid U}$ is not. Consequently, we may reuse $\Psi_1^{V\mid U}$ in the next block but we cannot reuse $\Phi_1^{V\mid U}$. We instead share $\Phi_1^{V\mid U}$ secretly between Alice and Bob and we show later that this may be accomplished with negligible rate cost. Finally, define $\tilde{V}_1^{1:N} \triangleq \tilde{B}_1^{1:N}G_b$. 
In Block $i \in [2, k]$, the encoder forms $\tilde{V}_i^{1:N}$ as follows. Let $S_i$ be a vector of $|V_{V|UY}|$ uniformly distributed bits and $M_i$ be a vector of $|M_{UVZ}|$ uniformly distributed bits that represent the secret and private message in block $i$, respectively. Given a private message $m_i$, a confidential message $s_i$, $\psi_{i-1}^{V|U}$, and $\tilde{u}_i^{1:N}$ resulting from the encoding of the common message, the encoder draws $\tilde{b}_i^{1:N}$ from the distribution $\tilde{p}_{B_i^{1:N}}$ defined as follows.

$$\tilde{p}_{B_i^{1:N}}(b_i^{1:j-1}\tilde{u}_i^{1:N}) \triangleq \begin{cases} \frac{1}{2} & \text{if } j \in V_{V|UY} \\
\begin{cases} b_i^j = s_i^j \\
\begin{cases} b_i^j = \left(\psi_{i-1}^{V|U}\right)^j \\
\begin{cases} b_i^j = m_i^j \\
\begin{cases} b_i^j = \tilde{b}_i^j \end{cases} \\
\end{cases}
\end{cases}
\end{cases} & \text{if } j \in B_{V|UY} \\
\begin{cases} b_i^j = \tilde{b}_i^j \end{cases} & \text{if } j \in M_{UVZ} \\
\begin{cases} b_i^j = \tilde{m}_i^j \end{cases} & \text{if } j \in V_{V|U}
\end{cases}$$

(16)

where the components of $s_i$, $\psi_{i-1}^{V|U}$, and $m_i$ have been indexed by the set of indices $V_{V|UY}$, $B_{V|UY}$, and $M_{UVZ}$ respectively, so that $S_i = \tilde{B}_i^{1:N}[V_{V|UY}]$, $\psi_i^{V|U} = \tilde{B}_i^{1:N}[B_{V|UY}]$, and $M_i = \tilde{B}_i^{1:N}[M_{UVZ}]$. The random bits that identify the codebook required for reconstruction are those in positions $H_{V|UY}$, which we split as

$$\Psi_i^{V|U} \triangleq \tilde{B}_i^{1:N}[V_{V|UY} \cup (H_{V|UY} \setminus V_{V|UY}) \cap V_{V|U}],$$

$$\Phi_i^{V|U} \triangleq \tilde{B}_i^{1:N}[H_{V|UY} \setminus V_{V|UY} \cap V_{V|U}].$$

Again, $\Psi_i^{V|U}$ is uniformly distributed but $\Phi_i^{V|U}$ is not, so that we reuse $\Psi_i^{V|U}$ in the next block but we share $\Phi_i^{V|U}$ securely between Alice and Bob. We show later that the cost of sharing $\Phi_i^{V|U}$ is negligible.

In Block $k$, Alice securely shares $(\Psi_k^{V|U}, \Phi_{1:k}^{V|U})$ with Bob. Finally, define $\tilde{V}_i^{1:N} \triangleq \tilde{B}_i^{1:N} G_n$.

C. Channel prefixing

The channel prefixing procedure with chaining is illustrated in Fig. 4.

In Block 1, the encoder forms $\tilde{X}_1^{1:N}$ as follows. Let $R_1$ be a vector of $|V_{X|V} \setminus V_{X|VZ}|$ uniformly distributed bits representing the randomness required for channel prefixing. Given a randomization sequence $r_1$ and $\tilde{v}_1^{1:N}$ resulting from the encoding of secret and private messages, the encoder draws $\tilde{t}_1^{1:N}$ from the distribution $\tilde{p}_{T_1^{1:N}}$ defined as follows.

$$\tilde{p}_{T_1^{1:N}}(t_1^{1:j-1}\tilde{v}_1^{1:N}) \triangleq \begin{cases} \frac{1}{2} & \text{if } j \in V_{X|VZ} \\
\begin{cases} t_1^j = r_1^j \\
\begin{cases} t_1^j = \tilde{t}_1^j \end{cases} & \text{if } j \in V_{X|W} \setminus V_{X|VZ} \\
\begin{cases} t_1^j = \tilde{t}_1^j \end{cases} & \text{if } j \in V_{X|V}
\end{cases}
\end{cases}$$

(17)
where the components of \( r_1 \) have been indexed by the set of indices \( \mathcal{V}_X|V \setminus \mathcal{V}_X|VZ \), so that \( R_1 = \bar{T}_1^{1:N}[\mathcal{V}_X|V \setminus \mathcal{V}_X|VZ] \). The random bits that identify the codebook are those in position \( \mathcal{V}_X|VZ \), which we denote

\[
\Psi^X_{1|V} \triangleq \bar{T}_1^{1:N}[\mathcal{V}_X|VZ].
\]

Finally, compute \( \bar{X}_1^{1:N} \triangleq \bar{T}_1^{1:N}G_n \), which is transmitted over the channel \( W_{YZ|X} \). We note \( Y_1^{1:N}, Z_1^{1:N} \) the corresponding channel outputs.

In Block \( i \in [2, k] \), the encoder forms \( \bar{X}_i^{1:N} \) as follows. Let \( R_i \) be a vector of \( |\mathcal{V}_X|V \setminus \mathcal{V}_X|VZ| \) uniformly distributed bits representing the randomness required for channel prefixing in block \( i \). Given a randomization sequence \( r_i \) and \( \bar{v}_i^{1:N} \) resulting from the encoding of secret and private messages, the encoder draws \( \bar{r}_i^{1:N} \) from the distribution \( \bar{p}_{T_i^{1:N}} \) defined as follows.

\[
\bar{p}_{T_i^{1:N} \mid T_i^{1:j-1} V_i^{1:N}} (t_i^{1:j-1} \bar{v}_i^{1:N}) \triangleq \begin{cases} \frac{1}{2} \left\{ t_i^j = \bar{v}_i^{j-1} \right\} & \text{if } j \in \mathcal{V}_X|VZ \\ \frac{1}{2} \left\{ t_i^j = r_i^j \right\} & \text{if } j \in \mathcal{V}_X|V \setminus \mathcal{V}_X|VZ \\ p_{T_i^{1:j-1} V_i^{1:N} \mid T_i^{1:j-1} \bar{v}_i^{1:N}} (t_i^{1:j-1} \bar{v}_i^{1:N}) & \text{if } j \in \mathcal{V}_X|V} \end{cases}
\]

where the components of \( r_i \) have been indexed by the set of indices \( \mathcal{V}_X|V \setminus \mathcal{V}_X|VZ \), so that \( R_i = \bar{T}_i^{1:N}[\mathcal{V}_X|V \setminus \mathcal{V}_X|VZ] \). Note that the random bits describing the codebook are \( \Psi_i^{X|V} \triangleq \bar{T}_i^{1:N}[\mathcal{V}_X|VZ] \), and are reused from the previous block. Finally, define \( \bar{X}_i^{1:N} \triangleq \bar{T}_i^{1:N}G_n \) and transmit it over the channel \( W_{YZ|X} \). We note \( Y_i^{1:N}, Z_i^{1:N} \) the corresponding channel outputs.
D. Decoding

The decoding procedure is as follows.

Reconstruction of the common message by Bob. Bob forms the estimate $\tilde{A}_{1:k}^{1:N}$ of $\tilde{A}_{1:k}^{1:N}$ as follows. In Block 1, Bob knows $(\Psi^U, \Phi^U)$, which contains all the bits $\tilde{A}_{1}^{1:N}[H_U|Y]$ by construction. Bob runs the successive cancellation decoder for source coding with side information of [18] using $Y_1^{1:N}$ and $\tilde{A}_{1}^{1:N}[H_U|Y]$. In Block $i \in [2, k]$, Bob estimates $\tilde{A}_{i}^{1:N}[H_U|Y]$ with $(\Psi^U, \hat{A}_{i-1}^{1:N}[I_{UY}\setminus I_{UZ}], \Phi^U)$, and uses this estimate along with $Y_i^{1:N}$ to run the successive cancellation decoder for source coding with side information.

Reconstruction of the common message by Eve. Eve forms the estimate $\tilde{A}_{1:k}^{1:N}$ of $\tilde{A}_{1:k}^{1:N}$ starting from Block $k$ and going backwards as follows. In Block $k$, Eve knows $(\Psi^U, \Phi^U)$, which contains all the bits in $\tilde{A}_{k}^{1:N}[H_U|Z]$ by construction. Eve runs the successive cancellation decoder for source coding with side information using $Z_k^{1:N}$ and $\tilde{A}_{k}^{1:N}[H_U|Z]$. For $i \in [1, k - 1]$, Eve estimates $\tilde{A}_{k-i}^{1:N}[H_U|Z]$ with $(\Psi^U, \hat{A}_{k-i+1}^{1:N}[A_{UYZ}], \Phi^U)$, and uses this estimate along with $Z_k^{1:N}$ to run the successive cancellation decoder for source coding with side information.

Reconstruction of the private and confidential messages by Bob. Bob forms the estimate $\tilde{B}_{1:k}^{1:N}$ of $\tilde{B}_{1:k}^{1:N}$ as follows starting with Block $k$. In Block $k$, given $(\Psi^{V|U}, \Phi^{V|U}, Y_1^{1:N}, \hat{U}_k^{1:N})$, Bob estimates $\tilde{B}_k^{1:N}$ with the successive cancellation decoder for source coding with side information. From $\tilde{B}_k^{1:N}$, an estimate $\hat{\Psi}_{k-1}^{V|U} \triangleq \hat{B}_{k-1}^{1:N}[V|UY]$ of $\Psi_{k-1}^{V|U}$ is formed. For $i \in [1, k - 1]$, given $(\hat{\Psi}_{k-i}^{V|U}, \Phi_{k-i}, Y_{k-i}^{1:N}, \hat{U}_{k-i}^{1:N})$, Bob estimates $\tilde{B}_{k-i}^{1:N}$ with the successive cancellation decoder for source coding with side information. From $\tilde{B}_{k-i}^{1:N}$, an estimate of $\Psi_{k-i-1}^{V|U}$ is formed. Once all the estimates $\tilde{B}_{1:k}$ have been formed, Bob extracts the estimates $\tilde{S}_{1:k}$ and $\tilde{M}_{1:k}$ of $S_{1:k}$ and $M_{1:k}$, respectively.

V. ANALYSIS OF POLAR CODING SCHEME

We now analyze in details the characteristics and performances of the polar coding scheme described in Section IV. Specifically, we show the following.

Theorem 2. Consider a discrete memoryless broadcast channel $(\mathcal{X}, p_{Y,Z|X}, \mathcal{Y}, \mathcal{Z})$. The coding scheme of Section III, whose complexity is $O(N \log N)$ achieves the region $R_{BCC}$.

The result of Theorem 2, follows in four steps. First, we show that the polar coding scheme of Section IV approximates the statistics of the original DMS $(U \times V \times X \times Y \times Z, p_{UVXYZ})$ from which the polarization sets were defined. Second, we show that the various messages rates are indeed those in
Third, we show that the probability of decoding error vanishes with the block length. Finally, we show that the information leakage vanishes with the block length.

A. Approximation of original DMS statistics

Recall that the vectors \( \tilde{A}_i^{1:N} \), \( \tilde{B}_i^{1:N} \), \( \tilde{V}_i^{1:N} \), and \( \tilde{X}_i^{1:N} \), generated in block \( i \in [1, k] \) do not have the exact joint distribution of the vectors \( A_1^{1:N} \), \( B_1^{1:N} \), \( V_1^{1:N} \), and \( X_1^{1:N} \), induced by the source polarization of the original DMS \( (U \times V \times X \times Y \times Z, p_{UVXYZ}) \). However, the following lemmas show that the joint distributions are close to one another, which is crucial for the subsequent reliability and secrecy analysis.

**Lemma 5.** For \( i \in [1, k] \), we have

\[
\mathbb{D}(p_{U_1^{1:N}} || \tilde{p}_{U_1^{1:N}}) = \mathbb{D}(p_{A_1^{1:N}} || \tilde{p}_{A_1^{1:N}}) \leq N \delta_N.
\]

Hence, by Pinsker's inequality

\[
\mathbb{V}(p_{A_1^{1:N}} || \tilde{p}_{A_1^{1:N}}) \leq \delta_N^{(U)},
\]

where \( \delta_N^{(U)} \equiv \sqrt{2 \log 2 \sqrt{N \delta_N}} \).

**Proof:** See Appendix A. \( \blacksquare \)

**Lemma 6.** For \( i \in [1, k] \), we have

\[
\mathbb{D}(p_{V_1^{1:N} U_1^{1:N}} || \tilde{p}_{V_1^{1:N} U_1^{1:N}}) = \mathbb{D}(p_{B_1^{1:N} U_1^{1:N}} || \tilde{p}_{B_1^{1:N} U_1^{1:N}}) \leq 2N \delta_N.
\]

Hence, by Pinsker's inequality

\[
\mathbb{V}(p_{B_1^{1:N} U_1^{1:N}} || \tilde{p}_{B_1^{1:N} U_1^{1:N}}) \leq \delta_N^{(UV)},
\]

where \( \delta_N^{(UV)} \equiv 2 \sqrt{\log 2 \sqrt{N \delta_N}} \).

**Proof:** See Appendix B. \( \blacksquare \)

**Lemma 7.** For \( i \in [1, k] \), we have

\[
\mathbb{D}(p_{X_1^{1:N} V_1^{1:N}} || \tilde{p}_{X_1^{1:N} V_1^{1:N}}) = \mathbb{D}(p_{T_1^{1:N} V_1^{1:N}} || \tilde{p}_{T_1^{1:N} V_1^{1:N}}) \leq 3N \delta_N.
\]

Hence, by Pinsker's inequality

\[
\mathbb{V}(p_{T_1^{1:N} V_1^{1:N}} || \tilde{p}_{T_1^{1:N} V_1^{1:N}}) \leq \delta_N^{(XV)},
\]

where \( \delta_N^{(XV)} \equiv \sqrt{2 \log 2 \sqrt{3N \delta_N}} \).
Proof: See Appendix C.

Combining the three previous lemmas, we obtain the following.

Lemma 8. For $i \in [1, k]$, we have
\[
\mathcal{P}(p_{U_1, V_1, N^1, X_1, Y_1, N^1, Z_1, N}, \tilde{p}_{U_1, V_1, N^1, X_1, Y_1, N^1, Z_1, N}) \leq \delta^{(P)}_N.
\]
where $\delta^{(P)}_N \triangleq \sqrt{\frac{2 \log 2}{N \delta_N}}(2\sqrt{2} + \sqrt{3})$.

Proof: See Appendix D.

As noted in [24], upper-bounding the divergence with a chain rule is easier than directly upper-bounding the variational distance as in [25], [26].

B. Transmission rates

We now analyze the rate of common message, confidential message, private message, and randomization sequence, used at the encoder, as well as the different sum rates and the rate of additional information sent to Bob and Eve.

Common message rate. The overall rate $R_O$ of common information bits transmitted satisfies
\[
R_O = \frac{(k - 1)|\mathcal{I}_{UY}| + |\mathcal{I}_{UY} \cap \mathcal{I}_{UZ}|}{kN}
\]
\[
= \frac{|\mathcal{I}_{UY}|}{N} - \frac{|\mathcal{I}_{UY}\setminus\mathcal{I}_{UZ}|}{kN}
\]
\[
\geq \frac{|\mathcal{I}_{UY}|}{N} - \frac{|\mathcal{I}_{UY}|}{kN}
\]
\[
\xrightarrow{N \to \infty} I(Y; U) - \frac{I(Y; U)}{k}
\]
\[
\xrightarrow{k \to \infty} I(Y; U),
\]
where we have used [18]. Since we also have $R_O \leq \frac{|\mathcal{I}_{UY}|}{N} \xrightarrow{N \to \infty} I(Y; U)$, we conclude
\[
R_O \xrightarrow{N \to \infty, k \to \infty} I(Y; U).
\]

Confidential message rate. First, observe that
\[
|\Psi_1^{V|U}| = |\mathcal{V}_{V|UY} \cup ((\mathcal{H}_{V|UY} \setminus \mathcal{V}_{V|UY}) \cap \mathcal{V}_{V|U})|
\]
\[
\leq |\mathcal{V}_{V|UY}| + |\mathcal{H}_{V|UY} \setminus \mathcal{V}_{V|UY}|
\]
\[
= |\mathcal{V}_{V|UY}| + |\mathcal{H}_{V|UY}| - |\mathcal{V}_{V|UY}|
\]
\[
\leq |\mathcal{H}_{V|UY}|,\]
and $|\Psi_{1}^{V}| \geq |\nu_{V}|$. Hence, since $\lim_{N \to \infty} |\nu_{V}| / N = H(V|UY)$ by [22, Lemma 1] and $\lim_{N \to \infty} |\mathcal{H}_{V}| / N = H(V|UY)$ by [18], we have

$$\lim_{N \to \infty} \frac{|\Psi_{1}^{V}|}{N} = H(V|UY).$$

Then, the overall rate $R_{S}$ of secret information bits transmitted is

$$R_{S} = \frac{|\nu_{V}| + (k - 1)|\nu_{V}| - |\mathcal{B}_{V}|}{kN} \to \frac{|\nu_{V}| - |\mathcal{B}_{V}|}{N} + \frac{|\mathcal{B}_{V}|}{kN}$$

$$= \frac{|\nu_{V}| - |\Psi_{1}^{V}|}{N} + \frac{|\Psi_{1}^{V}|}{kN} \to \frac{|\nu_{V}| - |\Psi_{1}^{V}|}{N} + \frac{|\Psi_{1}^{V}|}{kN}$$

$$\to \frac{N \to \infty}{\to I(V; Y|U) - I(V; Z|U) + \frac{H(V|UY)}{k}}$$

Private message rate. The overall rate $R_{M}$ of private information bits transmitted is

$$R_{M} = \frac{k|\mathcal{M}_{UVZ}|}{kN} \to \frac{|\nu_{V}| - |\nu_{V}|}{N} \to \frac{|\nu_{V}| - |\nu_{V}|}{N}$$

where we have used [22, Lemma 1].

Randomization rate. The uniform random bits used in the stochastic encoder includes those of the randomization sequence for channel prefixing, as well as those required to identify the codebooks and run the successive cancellation encoding. Using [22, Lemma 1], we find that the rate required to identify the codebook for the common message is

$$\frac{|\nu_{V}| - |\nu_{V}|}{kN} \to \frac{|\nu_{V}|}{kN} \to \frac{H(U|Y)}{k} \to 0.$$

Similarly, the rate required to identify the codebook for the secret and private messages corresponds to
the rate of \( (\Psi_k^{V|U}, \Phi_k^{V|U}) \), which is transmitted to Bob to allow him to reconstruct \( \tilde{B}^{1:N}_i \),

\[
\frac{|(\Psi_k^{V|U}, \Phi_k^{V|U})|}{kN} = \frac{|\tilde{B}^{1:N}_k[\mathcal{H}_V|UY]|}{kN} \xrightarrow{N \to \infty} \frac{H(V|UY)}{k} \xrightarrow{k \to \infty} 0,
\]

where we have used [18].

The randomization sequence rate used in channel prefixing is

\[
\frac{|\mathcal{V}_X| + (k - 1)|\mathcal{V}_X \setminus \mathcal{V}_X \setminus \mathcal{V}_U|}{kN} = \frac{|\mathcal{V}_X \setminus \mathcal{V}_X \setminus \mathcal{V}_U|}{N} + \frac{|\mathcal{V}_X \setminus \mathcal{V}_U|}{kN} = \frac{|\mathcal{V}_X| - |\mathcal{V}_X \setminus \mathcal{V}_U|}{N} + \frac{|\mathcal{V}_X \setminus \mathcal{V}_U|}{kN} \xrightarrow{N \to \infty} I(X; Z|V) + \frac{H(X|VZ)}{k} \xrightarrow{k \to \infty} I(X; Z|V),
\]

where we have used [22, Lemma 1]. We now show that the rate of uniform bits required for successive cancellation encoding in (12), (13), (14), (15), (16), (17), (18) is negligible through a series of lemmas.

**Lemma 9.** For \( i \in [1, k] \), we have

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{j \in \mathcal{V}_U \cup \mathcal{V}} H(\tilde{A}_j^{1:j-1} | \tilde{A}_i^{1:j-1}) = 0.
\]

**Proof:** See Appendix E.

**Lemma 10.** For \( i \in [1, k] \), we have

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{j \in \mathcal{V}_U \cup \mathcal{V}} H(\tilde{B}_j^{1:j-1} | \tilde{B}_i^{1:j-1} \tilde{U}_i^{1:N}) = 0.
\]

**Proof:** See Appendix F.

**Lemma 11.** For \( i \in [1, k] \), we have

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{j \in \mathcal{V}_U \cup \mathcal{V}} H(\tilde{T}_j^{1:j-1} | \tilde{T}_i^{1:j-1} \tilde{V}_i^{1:N}) = 0.
\]
The proof of Lemma 11 is similar to that of Lemma 10 using Lemma 7 in place of Lemma 6.

Hence, the overall randomness rate $R_R$ used at the encoder is asymptotically

$$R_R \xrightarrow{N \to \infty, k \to \infty} I(X; Z|V).$$

**Sum rates.** The sum of the private message rate $R_M$ and the randomness rate $R_R$ is asymptotically

$$R_M + R_R \xrightarrow{N \to \infty, k \to \infty} I(V; Z|U) + I(X; Z|V)$$

$$(a) = H(Z|U) - H(Z|UV) + H(Z|V) - H(Z|XV)$$

$$= H(Z|U) - H(Z|XV)$$

$$(b) = H(Z|U) - H(Z|XU)$$

$$= I(X; Z|U),$$

where $(a)$ and $(b)$ hold by $U - V - X - Z$.

Moreover, the sum of the common message rate $R_O$, the private message rate $R_M$, and the confidential message rate $R_S$ is asymptotically

$$R_O + R_M + R_S \xrightarrow{N \to \infty, k \to \infty} I(Y; U) + I(V; Z|U) + I(V; Y|U) - I(V; Z|U)$$

$$= I(Y; U) + I(V; Y|U).$$

**Seed Rate.** The rate of the secret sequence that must be shared between the legitimate users to initialize the coding scheme is

$$\frac{|\Psi^V_k|+k|\Phi^V_1|}{kN}$$

$$= \frac{|\Psi^V_k|}{kN} + \frac{|\Phi^V_1|}{N}$$

$$\leq \frac{|\mathcal{H}_V|}{kN} + \frac{|\mathcal{H}_V| |\mathcal{X}_V|}{N}$$

$$= \frac{|\mathcal{H}_V|}{kN} + \frac{|\mathcal{H}_V| |\mathcal{X}_V|}{N}$$

$$= \frac{N}{k} \rightarrow_{N \to \infty} H(V|Y)$$

$$\rightarrow_{k \to \infty} 0,$$

where we have used [22, Lemma 1] and [18].

Moreover the rate of public communication from Alice to both Bob and Eve is

$$\frac{|\Psi^U_1|+|\Phi^U_{1,k}|}{kN}$$
\[ \begin{align*}
&\leq \frac{|\Psi_U|+k|\mathcal{H}_U\setminus V_U|}{kN} \\
&= \frac{|V_U \setminus I_{UY}|+k(|H_U|-|V_U|)}{kN} \\
&\leq \frac{|\mathcal{H}_U|Y|+k(|H_U|-|V_U|)}{kN} \\
&= \frac{|\mathcal{H}_U|Y|+|H_U|-|V_U|}{N} \\
&\xrightarrow{N\to\infty} \frac{H(U|Y)}{k} \\
&\xrightarrow{k\to\infty} 0.
\end{align*} \]

C. Average probability of error

We first show that Eve and Bob can reconstruct the common messages $O_{1:k}^{1:N}$ with small probability. For $i \in [1,k]$, consider an optimal coupling [25], [27] between $\hat{p}_{U_i^{1:N}Y_i^{1:N}}$ and $p_{U_i^{1:N}Y_i^{1:N}}$ such that

\[ P[\mathcal{E}_{U_i,Y_i}] = \forall(\hat{p}_{U_i^{1:N}Y_i^{1:N}}, p_{U_i^{1:N}Y_i^{1:N}}), \] where $\mathcal{E}_{U_i,Y_i} \triangleq \{ (\hat{U}_i^{1:N}, \hat{Y}_i^{1:N}) \neq (U_i^{1:N}, Y_i^{1:N}) \}$. Define also for $i \in [2,k]$, $E_i \triangleq \{ \tilde{A}_i^{1:N}[I_{UY}\setminus I_{UZ}] \neq \tilde{A}_i^{1:N-1}[I_{UY}\setminus I_{UZ}] \}$.

We have

\[ \begin{align*}
&\mathbb{P}[O_i \neq \hat{O}_i] \\
&= \mathbb{P}[\tilde{U}_i^{1:N} \neq \tilde{U}_i^{1:N}] \\
&= \mathbb{P}[\tilde{U}_i^{1:N} \neq \tilde{U}_i^{1:N} | \mathcal{E}_{U_i,Y_i} \cap \mathcal{E}_i^c] \mathbb{P}[\mathcal{E}_{U_i,Y_i} \cap \mathcal{E}_i^c] + \mathbb{P}[\tilde{U}_i^{1:N} \neq \tilde{U}_i^{1:N} | \mathcal{E}_{U_i,Y_i} \cup \mathcal{E}_i] \mathbb{P}[\mathcal{E}_{U_i,Y_i} \cup \mathcal{E}_i], \\
&\leq \mathbb{P}[\tilde{U}_i^{1:N} \neq \tilde{U}_i^{1:N} | \mathcal{E}_{U_i,Y_i} \cap \mathcal{E}_i^c] + \mathbb{P}[\mathcal{E}_{U_i,Y_i} \cup \mathcal{E}_i] \\
&\leq N\delta_N + \mathbb{P}[\mathcal{E}_{U_i,Y_i} \cup \mathcal{E}_i] \\
&\leq N\delta_N + \mathbb{P}[\mathcal{E}_i] \\
&\leq N\delta_N + \mathbb{P}[\mathcal{E}_i] \\
&\leq N\delta_N + \mathbb{P}[\tilde{U}_i^{1:N} \neq \tilde{U}_i^{1:N}] \\
&\leq (i-1)(N\delta_N + \mathbb{P}[\tilde{U}_i^{1:N} \neq \tilde{U}_i^{1:N}]) \\
&\leq i(N\delta_N + \mathbb{P}[\tilde{U}_1^{1:N} \neq \tilde{U}_1^{1:N}]),
\end{align*} \]

where (a) follows from the error probability of source coding with side information [18] and the union bound, (b) holds by the optimal coupling and Lemma 8, (c) holds by induction, (d) holds similarly to
the previous inequalities. We thus have by the union bound and (19)
\[
P[O_{1:k} \neq \hat{O}_{1:k}] \leq \sum_{i=1}^{k} P[O_i \neq \hat{O}_i] \leq \frac{k(k+1)}{2} (N\delta_N + \delta_N^{(P)}).
\]

We similarly obtain for Eve
\[
P[O_{1:k} \neq \hat{O}_{1:k}] \leq \frac{k(k+1)}{2} (N\delta_N + \delta_N^{(P)}).
\]

Next we show how Bob can recover the secret and private messages. Informally, the decoding process of the confidential and private messages \((M_{1:k}, S_{1:k})\) for Bob is as follows. Reconstruction starts with Block \(k\). Given \((\Psi_{k-1}^{V[U]}, \Phi_{k-1}^{V[U]}, Y_{k-1}^{1:N}, \hat{U}_{k-1}^{1:N})\), Bob can reconstruct \(\hat{V}_{k-1}^{1:N}\), from which \(\Psi_{k-1}^{V[U]}\) is deduced. Then, for \(i \in [1, k-1]\), given \((\Psi_{k-1}^{V[U]}, \Phi_{k-1}^{V[U]}, Y_{k-1}^{1:N}, \hat{U}_{k-1}^{1:N})\), Bob can reconstruct \(\hat{V}_{k-1}^{1:N}\), from which \(\Psi_{k-1}^{V[U]}\) is deduced. Finally, \(S_{1:k}\) can be recovered from \(\hat{V}_{1:k}^{1:N}\).

Formally, the analysis is as follows. For \(i \in [1, k]\), consider an optimal coupling [25], [27] between \(\hat{p}_{U_i^{1:N}V_i^{1:N}Y_i^{1:N}}\) and \(p_{U_i^{1:N}V_i^{1:N}Y_i^{1:N}}\) such that \(P[\mathcal{E}_{U_i,V_i,Y_i}] = \mathcal{V}(\hat{p}_{U_i^{1:N}V_i^{1:N}Y_i^{1:N}}, p_{U_i^{1:N}V_i^{1:N}Y_i^{1:N}})\), where \(\mathcal{E}_{U_i,V_i,Y_i} \triangleq \{(\hat{U}_{i}^{1:N}, \hat{V}_{i}^{1:N}, \hat{Y}_{i}^{1:N}) \neq (U_{1:N}^{i}, V_{1:N}^{i}, Y_{1:N}^{i})\}. Define also for \(i \in [1, k-1]\), \(\mathcal{E}_{\Psi_{i}^{V[U]} \neq \Psi_{i}^{V[U]}} \triangleq \hat{\mathcal{E}}_{U_{i}} \triangleq \{(\hat{U}_{i}^{1:N} \neq U_{i}^{1:N})\}, \text{ and } \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_{i}}} \triangleq \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_{i}}}.

For \(i \in [1, k-1]\), we have
\[
P[(M_i, S_i) \neq (\hat{M}_i, \hat{S}_i)] \overset{(a)}{=} P[\hat{V}_i \neq \hat{V}_i]
\]
\[
= P[\hat{V}_i \neq \hat{V}_i | \mathcal{E}_{\hat{U}_i, V_i, Y_i} \cap \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}] P[\mathcal{E}_{\hat{U}_i, V_i, Y_i} \cap \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}]
\]
\[
+ P[\hat{V}_i \neq \hat{V}_i | \mathcal{E}_{U_i, V_i, Y_i} \cup \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}] P[\mathcal{E}_{U_i, V_i, Y_i} \cup \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}]
\]
\[
\leq P[\hat{V}_i \neq \hat{V}_i | \mathcal{E}_{\hat{U}_i, V_i, Y_i} \cap \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}] + P[\mathcal{E}_{U_i, V_i, Y_i} \cup \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}]
\]
\[
\leq P[\hat{V}_i \neq \hat{V}_i | \mathcal{E}_{\hat{U}_i, V_i, Y_i} \cap \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}] + P[\mathcal{E}_{U_i, V_i, Y_i} | \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}]
\]
\[
\leq P[\hat{V}_i \neq \hat{V}_i | \mathcal{E}_{\hat{U}_i, V_i, Y_i} \cap \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}] + P[\mathcal{E}_{U_i, V_i, Y_i} | \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}] + P[\mathcal{E}_{U_i}]
\]
\[
\overset{(b)}{=} P[\hat{V}_i \neq \hat{V}_i | \mathcal{E}_{\hat{U}_i, V_i, Y_i} \cap \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}] + P[\mathcal{E}_{U_i, V_i, Y_i} | \mathcal{E}_{\Psi_{i}^{V[U]} \cup \mathcal{E}_{\hat{U}_i}}] + P[\hat{V}_{i+1} \neq \hat{V}_{i+1}] + P[\hat{U}_{i}^{1:N} \neq \hat{U}_{i}^{1:N}]
\]
\[
\overset{(c)}{=} N\delta_N + P[\mathcal{E}_{U_i, V_i, Y_i}] + P[\hat{V}_{i+1} \neq \hat{V}_{i+1}] + P[\hat{U}_{i}^{1:N} \neq \hat{U}_{i}^{1:N}]
\]
\[
\overset{(d)}{=} N\delta_N + \delta_N^{(P)} + P[\hat{V}_{i+1} \neq \hat{V}_{i+1}] + P[\hat{U}_{i}^{1:N} \neq \hat{U}_{i}^{1:N}]
\]
\[
\overset{(e)}{=} (i+1) \left(N\delta_N + \delta_N^{(P)}\right) + P[\hat{V}_{i+1} \neq \hat{V}_{i+1}]
\]
\[ f \leq (i + 1)(k - i) \left( N\delta_N + \delta_N^{(P)} \right) + P[\hat{V}_k \neq \tilde{V}_k] \]

\[ g \leq (i + 1)(k - i + 1) \left( N\delta_N + \delta_N^{(P)} \right) \]

where (a) holds because \( \hat{V}_i \) contains \( (M_i, S_i, \Psi^U_{i-1}) \) by construction, (b) holds because \( \hat{V}_{i+1} \) contains \( \Psi^U_i \) by construction, (c) follows from the error probability of lossless source coding with side information [18], (d) holds by the optimal coupling and Lemma 8, (e) holds by (19), (f) holds by induction, (g) is obtained similarly to the previous inequalities.

Hence,

\[ P[(M_{1:k}, S_{1:k}) \neq (\hat{M}_{1:k}, \hat{S}_{1:k})] \]
\[ \leq \sum_{i=1}^{k} P[(M_i, S_i) \neq (\hat{M}_i, \hat{S}_i)] \]
\[ \leq \sum_{i=1}^{k} (i + 1)(k - i + 1) \left( N\delta_N + \delta_N^{(P)} \right) \]
\[ = \left( \frac{k(k + 1)(k + 2)}{6} + k \right) \left( N\delta_N + \delta_N^{(P)} \right). \] (20)

D. Information leakage

The functional dependence graph for the coding scheme of Section III is given in Figure 5. For the secrecy analysis the following term must be upper bounded

\[ I(S_{1:k}; \Psi^U_1 \Phi^U_1 \Psi^U_{1:k}) \]

Note that we have introduced \( (\Psi^U_1, \Phi^U_1, \Psi^U_{1:k}) \), since these random variables have been made available to Eve. Recall that \( \Phi^U_1 \) is additional information transmitted to Bob and Eve to reconstruct the common messages \( O_{1:k} \). Recall also that \( \Psi^U_i \supset \Psi^U_i, i \in [2, k] \), as it is the randomness reused among all the blocks that allows the transmission of the common messages \( O_{1:k} \). We start by proving that secrecy holds for a given block \( i \in [2, k] \) in the following lemma.

**Lemma 12.** For \( i \in [2, k] \) and \( N \) large enough,

\[ I(S_i; Z_{1:N}^V \Phi^U_1 \Psi^U_i) \leq \delta_N^{(s)} \],

where \( \delta_N^{(s)} = \sqrt{2\log 2}\sqrt{N} \delta_N (1 + 6\sqrt{2} + 3\sqrt{3}) (N - \log_2 (\sqrt{2} \log 2 \sqrt{N} \delta_N (1 + 6\sqrt{2} + 3\sqrt{3}))) \).

**Proof:** See Appendix G. 

November 4, 2014 DRAFT
Recall that for channel prefixing in the encoding process we reuse some randomness $\Psi_1^{X|V}$ among all the blocks so that $\Psi_1^{X|V} = \Psi_{i-1}^{X|V}$, $i \in [2,k]$. We show in the following lemma that $\Psi_1^{X|V}$ is almost independent from $(Z_{1:N}^1, \Psi_{i-1}^{V|U}, S_i, \Phi_i, \Psi_i^U)$. This fact will be useful in the secrecy analysis of the overall scheme.

**Lemma 13.** For $i \in [2,k]$ and $N$ large enough,

$$I(\Psi_{1}^{X|V}; Z_{1:N}^1, \Psi_{i-1}^{V|U}, S_i, \Phi_i, \Psi_i^U) \leq \delta_N^{(s)},$$

where $\delta_N^{(s)}$ is defined as in Lemma 12.

**Proof:** See Appendix H.

Using Lemmas 12 and 13, we show in the following lemma a recurrence relation that will make the secrecy analysis over all blocks easier.
Lemma 14. Let $i \in [1, k - 1]$. Define $\tilde{L}_i \triangleq I(S_{1:k}; \Psi U_1 \Phi U_1 Z_{1:k}^i)$. We have

$$\tilde{L}_{i+1} - \tilde{L}_i \leq 3 \delta^{(s)} N,$$

Proof: See Appendix I.

We then have

$$\tilde{L}_1 = I(S_{1:k}; \Psi U_1 \Phi U_1 Z_{1:N}^1)$$

$$= I(S_1; \Psi U_1 \Phi U_1 Z_{1:N}^1) + I(S_{2:k}; \Psi U_1 \Phi U_1 Z_{1:N}^1 | S_1)$$

$$\leq \delta^{(s)} N + I(S_{2:k}; \Psi U_1 \Phi U_1 Z_{1:N}^1 | S_1)$$

$$\leq \delta^{(s)} N + I(S_{2:k}; \Psi U_1 \Phi U_1 Z_{1:N}^1 S_1)$$

$$\leq \delta^{(s)} N,$$

where (a) follows from Lemma 12, (b) follows from independence of $S_{2:k}$ and the random variables of Block 1.

Hence, strong secrecy follows from Lemma 14 by remarking that

$$I(S_{1:k}; \Psi U_1 \Phi U_1 Z_{1:k}^N) = \tilde{L}_1 + \sum_{i=1}^{k-1} (\tilde{L}_{i+1} - \tilde{L}_i)$$

$$\leq \delta^{(s)} N + (k - 1)(3 \delta^{(s)} N)$$

$$= (3k - 2) \delta^{(s)} N.$$

VI. CONCLUSION

Our proposed polar coding scheme for the broadcast channel with confidential messages and constrained randomization provides an explicit low-complexity scheme achieving the capacity region of [14]. Although the presence of auxiliary random variables and the need to re-align polarization sets through chaining introduces rather involved notation, the coding scheme is conceptually close to a binning proof of the capacity region, in which polarization is used in place of random binning. We believe that a systematic use of this connection will effectively allow one to translate any results proved with output statistics of random binning [21] into a polar coding scheme.

It is arguable whether the resulting schemes are truly practical, as the block length $N$ and the number of blocks $k$ are likely to be fairly large. In addition work remains to be done to circumvent the need for sharing random seeds between the transmitter and receivers.
APPENDIX A

PROOF OF LEMMA 5

Let \( i \in [2, k - 1] \). We have

\[
\begin{align*}
\mathbb{D}(p_U^{1:N} | | \tilde{p}_U^{1:N}) \\
&\overset{(a)}{=} \mathbb{D}(p_{A^{1:N}} | | \tilde{p}_{A^{1:N}}) \\
&\overset{(b)}{=} N \sum_{j=1}^N \mathbb{D}(p_{A^j | A^{1:j-1}} | | \tilde{p}_{A^j | A^{1:j-1}}) \\
&\overset{(c)}{=} \sum_{j \in V_U} \mathbb{D}(p_{A^j | A^{1:j-1}} | | \tilde{p}_{A^j | A^{1:j-1}}) \\
&\overset{(d)}{=} \sum_{j \in V_U} (1 - H(A^j | A^{1:j-1})) \\
&\overset{(e)}{\leq} |V_U| \delta_N \\
&\leq N \delta_N, \quad (21)
\end{align*}
\]

where \((a)\) holds by invertibility of \(G_n\), \((b)\) holds by the chain rule, \((c)\) holds by \((13)\), \((d)\) holds by \((13)\) and uniformity of \(O_i\) and \(O_{i-1,2}\), \((e)\) holds by definition of \(V_U\).

Similarly for \( i \in \{1, k\} \), using \((12)\) and \((14)\) we also have

\[
\mathbb{D}(p_U^{1:N} | | \tilde{p}_U^{1:N}) \leq N \delta_N. \quad (22)
\]

APPENDIX B

PROOF OF LEMMA 6

Let \( i \in [2, k] \). We have

\[
\begin{align*}
\mathbb{D}(p_{B^{1:N}} | U^{1:N} | | \tilde{p}_{B^{1:N}} | U^{1:N}) \\
&\overset{(a)}{=} \sum_{j=1}^N \mathbb{D}(p_{B^j | B^{1:j-1}U^{1:N}} | | \tilde{p}_{B^j | B^{1:j-1}U^{1:N}}) \\
&\overset{(b)}{=} \sum_{j \in V_{V|U}} \mathbb{D}(p_{B^j | B^{1:j-1}U^{1:N}} | | \tilde{p}_{B^j | B^{1:j-1}U^{1:N}}) \\
&\overset{(c)}{=} \sum_{j \in V_{V|U}} (1 - H(B^j | B^{1:j-1}U^{1:N})) \\
&\overset{(d)}{\leq} |V_{V|U}| \delta_N \\
&\leq N \delta_N, \quad (23)
\end{align*}
\]
where (a) holds by the chain rule, (b) holds by (16), (c) holds by (16) and uniformity of $\Psi_{i-1}^V$, $S_i$, and $M_i$, (d) holds by definition of $\mathcal{V}_{V|U}$.

Then,

\[
\mathbb{D}(p_{V_1^N|U_1^N}||\tilde{p}_{V_1^N|U_1^N})
\leq (a) \mathbb{D}(p_{B_1^N|U_1^N}||\tilde{p}_{B_1^N|U_1^N})
\leq (b) \mathbb{D}(p_{B_1^N|U_1^N}||\tilde{p}_{B_1^N|U_1^N}) + \mathbb{D}(p_{U_1^N}||\tilde{p}_{U_1^N})
\leq (c) 2N\delta_N,
\]

where (a) holds by invertibility of $G_n$, (b) holds by the chain rule, (c) holds by (23) and Lemma 5.

Similarly, using (15) and Lemma 5, we have

\[
\mathbb{D}(p_{V_1^N|U_1^N}||\tilde{p}_{V_1^N|U_1^N}) \leq 2N\delta_N.
\]

**Appendix C**

**Proof of Lemma 7**

Let $i \in [2, k]$. We have

\[
\mathbb{D}(p_{T_1^N|V_1^N}||\tilde{p}_{T_1^N|V_1^N})
\leq (a) \sum_{j=1}^N \mathbb{D}(p_{T_j|T^{1:j-1}V_1^N}||\tilde{p}_{T_j|T^{1:j-1}V_1^N})
\leq (b) \sum_{j \in \mathcal{V}_{X|V}} \mathbb{D}(p_{T_j|T^{1:j-1}V_1^N}||\tilde{p}_{T_j|T^{1:j-1}V_1^N})
\leq (c) \sum_{j \in \mathcal{V}_{X|V}} (1 - H(T_j|T^{1:j-1}V_1^N)))
\leq (d) |\mathcal{V}_{X|V}|\delta_N
\leq N\delta_N, \tag{24}
\]

where (a) holds by the chain rule, (b) holds by (18), (c) holds by (18) and uniformity of the bits in $\tilde{T}_i^1:N|\mathcal{V}_{X|V}$, (d) holds by definition of $\mathcal{V}_{X|V}$.
Similarly, using (17) and Lemma 6, we have

\[ \mathbb{D}(p_{X_1^N V_1^N} || p_{X_1^N V_1^N}) \]

\[ \overset{(a)}{=} \mathbb{D}(p_{T_1^N V_1^N} || p_{T_1^N V_1^N}) \]

\[ \overset{(b)}{=} \mathbb{D}(p_{T_1^N V_1^N} || p_{T_1^N V_1^N}) + \mathbb{D}(p_{V_1^N} || p_{V_1^N}) \]

\[ \overset{(c)}{\leq} 3N\delta_N, \]

where (a) holds by invertibility of \( G_n \), (b) holds by the chain rule, (c) holds by (24) and Lemma 6.

Similarly, using (17) and Lemma 6, we have

\[ \mathbb{D}(p_{X_1^N V_1^N} || p_{X_1^N V_1^N}) \leq 3N\delta_N. \]

**APPENDIX D**

**PROOF OF LEMMA 8**

We have

\[ \mathbb{V}(p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}) \]

\[ = \mathbb{V}(p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}) \]

\[ \overset{(a)}{=} \mathbb{V}(p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}) \]

\[ \overset{(b)}{=} \mathbb{V}(p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}) \]

\[ \overset{(c)}{=} \mathbb{V}(p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}) \]

\[ \overset{(d)}{\leq} \mathbb{V}(p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}) + \mathbb{V}(p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}) \]

\[ = \mathbb{V}(p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}) + \mathbb{V}(p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{U_1^N V_1^N X_1^N Y_1^N Z_1^N}) \]

\[ \overset{(e)}{\leq} \mathbb{V}(p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}) + \delta_{(UV)}^{(UV)} \]

\[ = \mathbb{V}(p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}) + \delta_{(UV)}^{(UV)} \]

\[ \overset{(f)}{\leq} \mathbb{V}(p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}, p_{X_1^N U_1^N V_1^N X_1^N Y_1^N Z_1^N}) + \delta_{(UV)}^{(UV)} \]

\[ \overset{(g)}{\leq} 2\delta_{(UV)}^{(UV)} + \delta_{(UV)}^{(UV)} \]
where (a) and (c) follow from the Markov condition $U \rightarrow V \rightarrow X \rightarrow (YZ)$ and $\tilde{U}_i^{1:N} \rightarrow \tilde{V}_i^{1:N} \rightarrow \tilde{X}_i^{1:N} \rightarrow (Y_i^{1:N}Z_i^{1:N})$, (b) follows from $p_{Y_i^{1:N}Z_i^{1:N}|X_i^{1:N}} = \tilde{p}_{Y_i^{1:N}Z_i^{1:N}|X_i^{1:N}}$ and [28, Lemma 17], (d) holds by the triangle inequality, (e) holds by Lemma 6, (f) hold by the triangle inequality, (g) holds by Lemmas 6 and 7.

**APPENDIX E**

**PROOF OF LEMMA 9**

We have for $i \in [1,k]$, for $j \in V_0^c$,

$$|H(\tilde{A}_j^i|\tilde{A}_i^{1:j-1}) - H(A_j^i|A_i^{1:j-1})|$$

$$\leq |H(\tilde{A}_i^j) - H(A_i^{1:j})| + |H(\tilde{A}_i^{1:j-1}) - H(A_i^{1:j-1})|$$

$$(a) \leq V(p_{A_i^{1:j}}, \tilde{p}_{A_i^{1:j}}) \log \frac{2^j}{V(p_{A_i^{1:j}}, \tilde{p}_{A_i^{1:j}})} + |H(\tilde{A}_i^{1:j-1}) - H(A_i^{1:j-1})|$$

$$(b) \leq \delta_N^{(U)} (N - \log_2 \delta_N^{(U)}) + |H(\tilde{A}_i^{1:j-1}) - H(A_i^{1:j-1})|$$

$$\leq 2\delta_N^{(U)} (N - \log_2 \delta_N^{(U)})$$

$$\triangleq \delta_N^{(A)} ,$$

where (a) holds by [29], (b) holds by Lemma 5 and because $x \mapsto x \log x$ is decreasing for $x > 0$ small enough.

Hence, we obtain

$$\sum_{j \in V_0^c} H(\tilde{A}_j^i|\tilde{A}_i^{1:j-1})$$

$$= \sum_{j \in H_U^c, j \in H_U \setminus V_U} \sum_{j \in H_U \setminus V_U} H(\tilde{A}_j^i|\tilde{A}_i^{1:j-1})$$

$$\leq |H_U \setminus V_U| + \sum_{j \in H_U} H(\tilde{A}_j^i|\tilde{A}_i^{1:j-1})$$

$$= |H_U| - |V_U| + \sum_{j \in H_U} H(\tilde{A}_j^i|\tilde{A}_i^{1:j-1})$$

$$\leq |H_U| - |V_U| + \sum_{j \in H_U} (H(A_j^i|A_i^{1:j-1}) + \delta_N^{(A)})$$

$$\leq |H_U| - |V_U| + |H_U| |(\delta_N + \delta_N^{(A)})$$

$$\leq |H_U| - |V_U| + N(\delta_N + \delta_N^{(A)})$$

and we obtain the result by [22, Lemma 1] and [18].
We have for $i \in [1, k]$, for $j \in \mathcal{V}_{V|U}$,

$$
\begin{align*}
&|H(\tilde{B}_i^j | \tilde{B}_i^{1:j-1} \tilde{U}_i^{1:N}) - H(B_i^j | B_i^{1:j-1} U_i^{1:N})| \\
&\leq |H(\tilde{B}_i^j | \tilde{U}_i^{1:N}) - H(B_i^{1:j} U_i^{1:N})| + |H(\tilde{B}_i^{1:j-1} \tilde{U}_i^{1:N}) - H(B_i^{1:j-1} U_i^{1:N})| \\
&\overset{(a)}{\leq} \mathbb{V}(p_{B_i^{1:j} U_i^{1:N}}, p_{\tilde{B}_i^{1:j} \tilde{U}_i^{1:N}}) \log \frac{2^{j+N}}{\mathbb{V}(p_{B_i^{1:j} U_i^{1:N}}, p_{\tilde{B}_i^{1:j} \tilde{U}_i^{1:N}})} + |H(\tilde{B}_i^{1:j-1} \tilde{U}_i^{1:N}) - H(B_i^{1:j-1} U_i^{1:N})| \\
&\overset{(b)}{\leq} \delta_N^{(UV)} \left(2N - \log_2 \delta_N^{(UV)}\right) + |H(\tilde{B}_i^{1:j-1} \tilde{U}_i^{1:N}) - H(B_i^{1:j-1} U_i^{1:N})| \\
&\leq 2\delta_N^{(UV)} \left(2N - \log_2 \delta_N^{(UV)}\right) \\
&\triangleq \delta_N^{(B)},
\end{align*}
$$

where (a) holds by [29], (b) holds by Lemma 6 and because $x \mapsto x \log x$ is decreasing for $x > 0$ small enough.

Then,

$$
\begin{align*}
\sum_{j \in \mathcal{V}_{V|U}} H(\tilde{B}_i^j | \tilde{B}_i^{1:j-1} \tilde{U}_i^{1:N}) \\
= \sum_{j \in \mathcal{H}_{V|U}} \sum_{j \in \mathcal{H}_{V|U} \setminus \mathcal{V}_{V|U}} H(\tilde{B}_i^j | \tilde{B}_i^{1:j-1} \tilde{U}_i^{1:N}) \\
\leq |\mathcal{H}_{V|U} \setminus \mathcal{V}_{V|U}| + \sum_{j \in \mathcal{H}_{V|U}} H(\tilde{B}_i^j | \tilde{B}_i^{1:j-1} \tilde{U}_i^{1:N}) \\
= |\mathcal{H}_{V|U} \setminus \mathcal{V}_{V|U}| \sum_{j \in \mathcal{H}_{V|U}} \left( H(B_i^j | B_i^{1:j-1} U_i^{1:N}) + \delta_N^{(B)} \right) \\
\leq |\mathcal{H}_{V|U} \setminus \mathcal{V}_{V|U}| + |\mathcal{H}_{V|U}| (\delta_N + \delta_N^{(B)}) \\
\leq |\mathcal{H}_{V|U} \setminus \mathcal{V}_{V|U}| + N(\delta_N + \delta_N^{(B)}),
\end{align*}
$$

and we obtain the result by [22, Lemma 1] and [18].
We have
\[
\mathbb{V}(p_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}}, \tilde{p}_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}})
\leq \mathbb{V}(p_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}}, p_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}}) + \mathbb{V}(p_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}}, \tilde{p}_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}})
\leq \sqrt{2 \log 2} \sqrt{I(B^{1:N} \mid \mathcal{V}_t[U\mid Z], U^{1:N} Z^{1:N})} + 2\delta_N^{(P)}
\leq \sqrt{2 \log 2} \sqrt{N \delta_N + 2\delta_N^{(P)}},
\] (25)
where (a) follows from the triangle inequality, (b) holds by Lemma 8, (c) holds by Pinsker’s inequality, (d) holds because using the fact that conditioning reduces entropy we have
\[
I(B^{1:N} \mid \mathcal{V}_t[U\mid Z], U^{1:N} Z^{1:N})
= H(B^{1:N} \mid \mathcal{V}_t[U\mid Z]) - H(B^{1:N} \mid \mathcal{V}_t[U\mid Z] \mid U^{1:N} Z^{1:N})
\leq |\mathcal{V}_t[U\mid Z]| - \sum_j H(B^j \mid B^{1:j-1} U^{1:N} Z^{1:N})
\leq |\mathcal{V}_t[U\mid Z]| + |\mathcal{V}_t[U\mid Z]|(\delta_N - 1)
\leq N \delta_N.
\]

We then obtain
\[
\mathbb{V}(\tilde{p}_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}}, \tilde{p}_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}})
\leq \mathbb{V}(\tilde{p}_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}}, p_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}}) + \mathbb{V}(p_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}}, \tilde{p}_{B^{1:N} \mid \mathcal{V}_t[U\mid Z]}^{U^{1:N} Z^{1:N}})
\leq \sqrt{2 \log 2} \sqrt{N \delta_N + 3\delta_N^{(P)}},
\] (26)
where (a) holds by the triangle inequality, (b) holds by Lemma 8, and (25).
Then, for $N$ large enough by [29],

$$ I(S_i \Psi_i^{V_{i}}; Z_{i}^{1:N} \Psi_i^{U_{i}}) $$

$$ \leq I(\tilde{B}_i^{1:N}[V_{i}U_{i}Z_i]; Z_{i}^{1:N} \tilde{U}_i^{1:N}) $$

$$ \leq \var(V(\tilde{B}_i^{1:N}[V_{i}U_{i}Z_i], \tilde{B}_i^{1:N}[V_{i}U_{i}Z_i] \tilde{U}_i^{1:N} Z_i^{1:N})) $$

$$ \times \frac{|V_{i}U_{i}Z_i|}{\var(V(\tilde{B}_i^{1:N}[V_{i}U_{i}Z_i], \tilde{B}_i^{1:N}[V_{i}U_{i}Z_i] \tilde{U}_i^{1:N} Z_i^{1:N}))} $$

$$ \leq \sqrt{2}\log 2(\sqrt{N} + (1 + 6)\sqrt{2} + 3\sqrt{3})(N - \log_2(\sqrt{2}\log 2(\sqrt{N} + (1 + 6)\sqrt{2} + 3\sqrt{3}))), $$

where we have used (26) and that $x \mapsto x \log x$ is decreasing for $x > 0$ small enough.

**APPENDIX H**

**PROOF OF LEMMA 13**

By the triangle inequality we can write

$$ \var(p_{T^{1:N}|V_{X}|Z}U_{1:N}V_{1:N}Z_i^{1:N}, \tilde{p}_{T^{1:N}|V_{X}|Z}U_{1:N}V_{1:N}Z_i^{1:N}) $$

$$ \leq \var(p_{T^{1:N}|V_{X}|Z}U_{1:N}V_{1:N}Z_i^{1:N}, p_{T^{1:N}|V_{X}|Z}U_{1:N}V_{1:N}Z_i^{1:N}) $$

$$ + \var(p_{T^{1:N}|V_{X}|Z}U_{1:N}V_{1:N}Z_i^{1:N}, \tilde{p}_{T^{1:N}|V_{X}|Z}U_{1:N}V_{1:N}Z_i^{1:N}) $$

$$ \leq \var(p_{T^{1:N}|V_{X}|Z}U_{1:N}V_{1:N}Z_i^{1:N}, p_{T^{1:N}|V_{X}|Z}U_{1:N}V_{1:N}Z_i^{1:N}) + 2\delta_N^{(P)} $$

$$ \leq \sqrt{2}\log 2(\sqrt{2}\log 2(\sqrt{2}\log 2(\sqrt{2}\log 2(N) + 2\delta_N^{(P)}))) $$

$$ = \sqrt{2}\log 2(\sqrt{2}\log 2(N) + 2\delta_N^{(P)}) $$

where (a) holds by the triangle inequality and Lemma 8, (b) holds by Pinsker’s inequality, (c) holds because using the fact that conditioning reduces entropy and $U - V - X$ we have

$$ I(T^{1:N}|X; Z_{1:N}U_{1:N}V_{1:N}) $$

$$ \leq |V_{X}|V_{Z} - \sum_{j \in V_{X}|V_{Z}} H(T^j|T^{1:j-1}Z_{1:N}U_{1:N}V_{1:N}) $$

$$ = |V_{X}|V_{Z} - \sum_{j \in V_{X}|V_{Z}} H(T^j|T^{1:j-1}Z_{1:N}V_{1:N}) $$

$$ \leq |V_{X}|V_{Z} + |V_{X}|V_{Z}|(\delta_N - 1) $$

$$ \leq N\delta_N.$$

November 4, 2014 DRAFT
Hence,
\[
\mathbb{V}(p_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}, \tilde{p}_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}) \\
\leq \mathbb{V}(p_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}, p_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}) \\
+ \mathbb{V}(p_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}, \tilde{p}_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}) \\
\leq \sqrt{2 \log 2 \sqrt{N \delta_N}} + 3 \delta^{(P)}_N,
\]
where (a) holds by the triangle inequality, (b) holds by Lemma 8, and (27).

Then, for $N$ large enough by [29],
\[
I(\Psi_{i}^{X} | V_{i}; Z_{i-1}^{1:N} S_{i} \Phi_{i}^{U} | \Psi_{i}^{U}) \\
= I(\tilde{T}_{i}^{1:N} | X_{i} | Z_{i}; Z_{i}^{1:N} \tilde{B}_{i}^{1:N} | H_{i}^{U} | V_{i}^{1:N} \Phi_{i}^{U} | \Psi_{i}^{U}) \\
\leq I(\tilde{T}_{i}^{1:N} | X_{i} | Z_{i}; Z_{i}^{1:N} \tilde{B}_{i}^{1:N} \tilde{U}_{i}^{1:N}) \\
\overset{(a)}{=} I(\tilde{T}_{i}^{1:N} | X_{i} | Z_{i}; Z_{i}^{1:N} \tilde{V}_{i}^{1:N} \tilde{U}_{i}^{1:N}) \\
\leq \mathbb{V}(\tilde{p}_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}, \tilde{p}_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}) \\
\times \log_2 \mathbb{V}(\tilde{p}_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}, \tilde{p}_{T_i}^{1:N} | X_{i} | Z_{i}; U_{i}^{1:N} V_{i}^{1:N} Z_{i}^{1:N}) \\
\overset{(b)}{=} \sqrt{2 \log 2 \sqrt{N \delta_N}} (1 + 6\sqrt{2} + 3\sqrt{3})(N - \log_2(\sqrt{2 \log 2 \sqrt{N \delta_N}} (1 + 6\sqrt{2} + 3\sqrt{3}))},
\]
where (a) holds by invertibility of $G_{n_i}$, (b) holds by (28) and because $x \mapsto x \log x$ is decreasing for $x > 0$ small enough.
Let \( i \in [1, k-1] \). We have

\[
\bar{L}_{i+1} - \bar{L}_i
= (S_{1:k}; \Psi_i^U \Phi_i^U Z_{1:i+1}^{1:N}) - (S_{1:k}; \Psi_i^U \Phi_i^U Z_{1:i}^{1:N})
= (S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N}|\Psi_i^U \Phi_i^U Z_{1:i}^{1:N})
= (S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N}|\Psi_i^U \Phi_i^U Z_{1:i}^{1:N})
+ I(S_{1:i+1}; Z_{1:i+1}^{1:N}|\Psi_i^U \Phi_i^U Z_{1:i}^{1:N} S_{i+1}^{1:N})
\]

(a)

\[
\leq (S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N}|\Psi_i^U ) + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N} S_{i+1}^{1:N})
\]

(b)

\[
\leq (S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N}|\Psi_i^U ) + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N} S_{i+1}^{1:N})
\]

(c)

\[
\leq \delta_{N}^{(a)} + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N}|\Psi_i^U ) + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N} S_{i+1}^{1:N})
\]

(d)

\[
\leq \delta_{N}^{(a)} + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N}|\Psi_i^U ) + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N} S_{i+1}^{1:N})
\]

(e)

\[
\leq \delta_{N}^{(a)} + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N}|\Psi_i^U ) + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N} S_{i+1}^{1:N})
\]

(f)

\[
\leq \delta_{N}^{(a)} + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N}|\Psi_i^U ) + I(S_{1:i+1}; \Phi_i^U Z_{1:i+1}^{1:N} S_{i+1}^{1:N})
\]

(g)

\[
\leq 3\delta_{N}^{(a)}
\]

where (a) holds by the chain rule and positivity of mutual information, (b) holds by independence of \( S_{i+2:k} \) with all the random variables of the previous blocks, (c) holds by Lemma 12, in (d) we introduce the random variable \( \Psi_i^{U|V} \) and \( \Psi_i^{X|V} \) to be able to break the dependencies between the random variables of block \( (i+1) \) and the random variables of the previous blocks, (e) holds because \( S_{1:i+2} \Phi_i^U Z_{1:i+1}^{1:N} \rightarrow \Psi_i^{U|V} \Psi_i^{X|V} \Psi_i^U \rightarrow \Phi_{i+1}^U Z_{i+1}^{1:N} S_{i+1}^{1:N} \), (f) holds because \( \Psi_i^{U|V} \Psi_i^{X|V} \Psi_i^U \) is independent of \( S_{i+1} \), (g)
holds by Lemmas 12, 13 and because $\Psi_i^{X|V}$ is constant equal to $\Psi_1^{X|V}$.

REFERENCES

[1] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, “Applications of LDPC codes to the wiretap channels,” IEEE Trans. Inf. Theory, vol. 53, no. 8, pp. 2933–2945, August 2007.
[2] A. Subramanian, A. Thangaraj, M. Bloch, and S. McLaughlin, “Strong secrecy on the binary erasure wiretap channel using large-girth LDPC codes,” IEEE Transactions on Information Forensics and Security, vol. 6, no. 3, pp. 585–594, September 2011.
[3] V. Rathi, R. Urbanke, M. Andersson, and M. Skoglund, “Rate-equivocation optimal spatially coupled LDPC codes for the bec wiretap channel,” in Proc. of IEEE Int. Symp. Info. Theory, Saint-Petersburg, Russia, August 2011, pp. 2393–2397.
[4] H. Mahdavifar and A. Vardy, “Achieving the Secrecy Capacity of Wiretap Channels using Polar Codes,” IEEE Trans. Inf. Theory, vol. 57, no. 10, pp. 6428–6443, 2011.
[5] E. Şaşoğlu and A. Vardy, “A New Polar Coding Scheme for Strong Security on Wiretap Channels,” in Proc. of IEEE Int. Symp. Info. Theory, 2013, pp. 1117–1121.
[6] J. M. Renes, R. Renner, and D. Sutter, “Efficient one-way secret-key agreement and private channel coding via polarization,” in Advances in Cryptology-ASIACRYPT 2013. Springer, 2013, pp. 194–213.
[7] M. Andersson, R. Schaefer, T. Oechtering, and M. Skoglund, “Polar coding for bidirectional broadcast channels with common and confidential messages,” IEEE Journal on Selected Areas in Communications, vol. 31, no. 9, pp. 1901–1908, 2013.
[8] M. Hayashi, “Exponential decreasing rate of leaked information in universal random privacy amplification,” IEEE Trans. Info. Theory, vol. 57, no. 6, pp. 3989–4001, 2011.
[9] M. Bellare and S. Tessaro, “Polynomial-time, semantically-secure encryption achieving the secrecy capacity,” arXiv preprint arXiv:1201.3160, 2012.
[10] A. D. Wyner, “The wire-tap channel,” The Bell System Technical Journal, The, vol. 54, no. 8, pp. 1355–1387, 1975.
[11] M. Mondelli, S. H. Hassani, I. Sason, and R. Urbanke, “Achieving the superposition and binning regions for broadcast channels using polar codes,” arXiv preprint arXiv:1401.6060, 2014.
[12] M. Mondelli, S. H. Hassani, and R. Urbanke, “How to achieve the capacity of asymmetric channels,” arXiv preprint arXiv:1406.7373, 2014.
[13] I. Csiszár and J. Korner, “Broadcast channels with confidential messages,” IEEE Trans. Inf. Theory, vol. 24, no. 3, pp. 339–348, 1978.
[14] S. Watanabe and Y. Oohama, “Broadcast channels with confidential messages by randomness constrained stochastic encoder,” in Proc. of IEEE Int. Symp. Info. Theory, 2012, pp. 61–65.
[15] M. Bloch and J. Kliewer, “On Secure Communication with Constrained Randomization,” in IEEE Int. Symp. Info. Theory. IEEE, 2012, pp. 1172–1176.
[16] T. Gulcu and A. Barg, “Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component,” arXiv preprint arXiv:1410.3422, 2014.
[17] Y. Wei and S. Ulukus, “Polar coding for the general wiretap channel,” arXiv preprint arXiv:1410.3812, 2014.
[18] E. Arikan, “Source Polarization,” in IEEE Int. Symp. Info. Theory, 2010, pp. 899–903.
[19] I. Csiszár, “Almost independence and secrecy capacity,” Problems of Information Transmission, vol. 32, no. 1, pp. 40–47, January-March 1996.
[20] J. Renes and R. Renner, “Noisy channel coding via privacy amplification and information reconciliation,” *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7377–7385, 2011.

[21] M. H. Yassaee, M. R. Aref, and A. Gohari, “Achievability proof via output statistics of random binning,” in *Proc. of IEEE Int. Symp. Info. Theory*, Boston, MA, July 2012, pp. 1044–1048.

[22] R. A. Chou, M. R. Bloch, and E. Abbe, “Polar coding for secret-key generation,” *arXiv preprint arXiv:1305.4746v2*, 2013.

[23] E. Şaşoğlu, “Polar codes for discrete alphabets,” in *IEEE Int. Symp. Info. Theory*, 2012, pp. 2137–2141.

[24] N. Goela, E. Abbe, and M. Gastpar, “Polar codes for broadcast channels,” *arXiv preprint arXiv:1301.6150*, 2013.

[25] S. Korada and R. Urbanke, “Polar Codes are Optimal for Lossy Source Coding,” *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751–1768, 2010.

[26] J. Honda and H. Yamamoto, “Polar coding without alphabet extension for asymmetric models,” *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, 2013.

[27] D. Aldous, “Random walks on finite groups and rapidly mixing markov chains,” in *Séminaire de Probabilités XVII 1981/82*. Springer, 1983, pp. 243–297.

[28] P. Cuff, “Communication in Networks for Coordinating Behavior,” Ph.D. dissertation, Stanford Univ., CA., 2009.

[29] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge Univ Pr, 1981.