Polar Coding for the Cognitive Interference Channel With Confidential Messages

Mengfan Zheng, Wen Chen, and Cong Ling

***Abstract*— In this paper, we propose a low-complexity, secrecy capacity achieving polar coding scheme for the cognitive interference channel with confidential messages (CICC) under the strong secrecy criterion. Existing polar coding schemes for interference channels rely on the use of polar codes for the multiple access channel, the code construction problem of which can be complicated. We show that the whole secrecy capacity region of the CICC can be achieved by simple point-to-point polar codes due to the cognitivity, and our proposed scheme requires the minimum rate of randomness at the encoder.**

***Index Terms*— Polar codes, cognitive interference channel, physical layer security, superposition coding.**

I. Introduction

COGNITIVE radio [1] has received increasing attention due to its capability of exploiting the under-utilized spectrum resource, as the scale of wireless networks has been growing drastically nowadays. The cognitive interference channel (CIC) is a typical model for the study of cognitive radios. In this model, a primary user (can be thought of as a licensed user of a frequency band) and a cognitive user (can be thought of as an unlicensed user wishing to share the same frequency band) who has non-causal knowledge of the primary user's message transmit their own messages to their own destinations at the same time. The problem of base station cooperation, in which base stations can share information via backhaul links of unlimited capacity, is a potential application scenario of the CIC. Another practical scenario of the CIC is in the problem of message retransmission, where the cognitive user can hear and decode the first transmission, while the primary user fails to do that. The communication problem in the CIC has been studied in [2]–[7]. The security issue of the CIC was first considered in [8], which gave the capacity-equivocation region of the CIC with confidential messages (CICC) under the weak secrecy criterion. A more general expression for the achievable rate region of the CICC with additional randomness constraints was derived in [9] under the strong secrecy criterion, which coincides with the result of [8].

In this paper, we aim to design a polar coding scheme to achieve the whole achievable rate region of [9]. Polar codes [10], originally targeted for achieving the symmetric capacity of point-to-point channels, have recently been shown to work for multi-user channels as well. It is shown that polar codes achieve the capacity regions or the known achievable rate regions of multiple access channels (MAC) [11]–[14], broadcast channels [15], [16], and interference channels (IC) [17], [18]. In the area of physical layer security, polar codes have been shown to achieve the secrecy capacity of wiretap channels [19]–[24], and the secrecy capacity regions or the known secrecy rate regions of MAC wiretap channels [23], [25], broadcast channels with confidential messages [22]–[24], IC with confidential messages [23], two-way wiretap channels [26], and bidirectional broadcast channels with common and confidential messages [27]. A capacity achieving secrecy polar coding scheme usually requires some eavesdropper channel information at the transmitter, which can be a drawback in practice. However, this is not a problem in the CICC since the assumption that the cognitive transmitter knows the channel information of both receivers is quite reasonable. Thus, the CICC can be a scenario where secrecy polar coding can be practically used.

The CICC differs from the IC with confidential messages considered by [23] in that only the cognitive user has confidential messages, and the cognitive user has non-causal knowledge about the primary user's messages. The secrecy capacity region of the CICC cannot be achieved by the scheme of [23] since the coding strategy lying behind is totally different. The scheme of [23] is more of an extension of wiretap polar codes while our scheme involves a more complicated design of auxiliary random variables and chaining scheme. We show that the cognitivity not only enlarges the achievable rate region of the IC, but also can help simplify the code design. As shown in [17], [18], and [23], existing polar code designs for ICs that can achieve optimal rate regions require the use of the permutation based MAC polarization [11], [14], as it is currently the only method that can achieve the whole achievable rate region of the MAC directly without time sharing or rate splitting. Although for some special types of permutations, MAC polar codes can be constructed using existing efficient methods [18], the practicality of this method remains open in general since

Manuscript received September 12, 2017; revised January 30, 2018; accepted February 16, 2018. Date of publication April 9, 2018; date of current version July 9, 2018. The work of W. Chen was supported in part by the National Natural Science Foundation of China under Grant 61671294, in part by the STCSM Project under Grant 16JC1402900 and Grant 17510740700, in part by the Natural Science Foundation of Guangxi Province under Grant 2015GXNSFDA139037, in part by the National Science and Technology Major Project under Grant 2017ZX03001002-005 and Grant 2018ZX03001009-002. *(Corresponding author: Wen Chen.)*

M. Zheng and W. Chen are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: zhengmengfan@sjtu.edu.cn; wenchen@sjtu.edu.cn).

C. Ling is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: c.ling@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2018.2825139

the induced random variable by the permutation can be very complicated. Existing efficient constructing methods for point-to-point polar codes (such as [28], [29]) may not be readily applied. In this paper, we show that the whole achievable rate region of the CICC can be achieved by point-to-point polar codes in conjunction with properly designed chaining schemes.

We summarize the contributions of this paper as follows.

- We propose a low-complexity polar coding scheme for the general CICC that achieves the whole secrecy capacity region under the strong secrecy criterion, without any assumption on channel symmetry or degradation.
- We avoid using MAC polarization, which is a common ingredient of polar code designs for ICs but may increase system complexity, and develop a secrecy capacity achieving scheme which only requires point-to-point polar codes. A novel, cross-transmitter chaining scheme is proposed to fulfill this task.
- Secrecy coding schemes require a large amount of randomness in order to protect the confidential message or to perform channel prefixing. As the transmission rate of today's communication systems increases drastically, the required generating rate of randomness in a secrecy coding scheme increases correspondingly, and thus cannot be considered as an unlimited resource. In this paper, we prove that our proposed scheme achieves the minimum generating rate of randomness.
- We show that our proposed scheme is a general solution for several other multi-user polar coding problems, including the CIC without secrecy requirement.

The rest of this paper is organized as follows. In Section II, we introduce the CICC model and the achievable rate region. In Section III, we review some background knowledge on polar codes. Details of our proposed scheme are presented in Section IV. We analyze the performance of our proposed scheme in Section V. In Section VI we discuss some extensions of our proposed scheme.

Notations: In this paper, $[N]$ denotes the index set of $\{1,2,\ldots,N\}$. For any $\mathcal{A}\subset[N]$, $X^{\mathcal{A}}$ denotes the subvector $\{X^i : i\in\mathcal{A}\}$ of $X^{1:N}\triangleq\{X^1,X^2,\ldots,X^N\}$. The generator matrix of polar codes is defined as $\mathbf{G}_N=\mathbf{B}_N\mathbf{F}^{\otimes n}$ [10], where $N=2^n$ with n being an arbitrary integer, $\mathbf{B}_N$ is the bit-reversal matrix, and $\mathbf{F}=\begin{bmatrix}1&0\\1&1\end{bmatrix}$. $H_q(X)$ stands for the entropy of X with q-based logarithm. $\delta_N=2^{-N^\beta}$ with some $\beta\in\ (0,1/2)$.

II. Problem Statement

A. Channel Model

Definition 1: A 2-user CIC consists of two input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, two output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$, and a probability transition function $P_{Y_1Y_2|X_1X_2}(y_1,y_2|x_1,x_2)$, where $x_1\in\mathcal{X}_1$ and $x_2\in\mathcal{X}_2$ are channel inputs of transmitter 1 and 2, respectively, and $y_1\in\mathcal{Y}_1$ and $y_2\in\mathcal{Y}_2$ are channel outputs of receiver 1 and 2, respectively. The conditional joint probability distribution of the 2-user CIC over N channel uses can be factored as

$$P_{Y_1^{1:N}Y_2^{1:N}|X_1^{1:N}X_2^{1:N}}(y_1^{1:N},y_2^{1:N}|x_1^{1:N},x_2^{1:N})=\prod_{j=1}^{N}P_{Y_1Y_2|X_1X_2}(y_1^j,y_2^j|x_1^j,x_2^j). \quad (1)$$

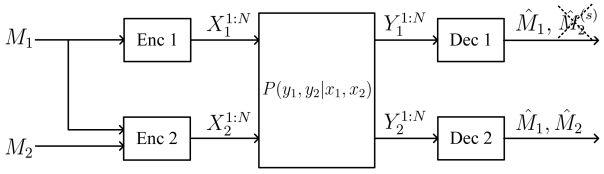

Fig. 1. The cognitive interference channel with confidential messages.

In this channel, transmitter i ($i=1,2$) sends message M_i to receiver i. Receiver 1 is required to decode M_1 only while receiver 2 is required to decode both M_1 and M_2[1]. Since transmitter 2 has non-causal knowledge about transmitter 1's message, M_1 can be jointly transmitted by the two transmitters. If transmitter 2 wishes to keep part of its message (denoted as $M_2^{(s)}$) secret from receiver 1, then this model is called the CICC, as shown in Fig. 1.

Definition 2: A $(2^{NR_1},2^{NR_2},N)$ code for the 2-user CICC consists of two message sets $\mathcal{M}_1=\{1,2,\ldots,[2^{NR_1}]\}$ and $\mathcal{M}_2=\{1,2,\ldots,[2^{NR_2}]\}$, two encoding functions

$$x_1^N(m_1):\mathcal{M}_1\mapsto\mathcal{X}_1^N \text{ and } x_2^N(m_1,m_2):\mathcal{M}_1\times\mathcal{M}_2\mapsto\mathcal{X}_2^N, \quad (2)$$

and two decoding functions

$$\hat{m}_1(\mathbf{y}_1^N):\mathcal{Y}_1^N\mapsto\mathcal{M}_1 \text{ and } \hat{m}_2(\mathbf{y}_2^N):\mathcal{Y}_2^N\mapsto\mathcal{M}_1\times\mathcal{M}_2. \quad (3)$$

For a given $(2^{NR_1},2^{NR_2},N)$ code for the 2-user CICC, reliability is measured by the average probability of error $P_e(N)$, defined as

$$P_e(N)=\frac{1}{2^{N(R_1+R_2)}}\sum_{(M_1,M_2)\in\mathcal{M}_1\times\mathcal{M}_2}\Pr\Big\{\big(\hat{m}_1(Y_1^{1:N}),\hat{m}_2(Y_2^{1:N})\big)\neq(M_1,M_1,M_2)|(M_1,M_2)\text{ sent}\Big\}, \quad (4)$$

where (M_1,M_2) is assumed to be uniformly distributed over $\mathcal{M}_1\times\mathcal{M}_2$. Secrecy is measured by the information leakage (strong secrecy)

$$L(N)=I(Y_1^{1:N};M_2^{(s)}), \quad (5)$$

or the information leakage rate (weak secrecy)

$$L_R(N)=\frac{1}{N}L(N). \quad (6)$$

[1]This is a case where the capacity and capacity-equivocation regions of a CIC is known [8]. In the general case, the capacity region of a CIC is still unknown [7].

B. Achievable Rate Region

Let R_{2s} be the transmission rate of confidential message $M_2^{(s)}$. A rate triple (R_1, R_2, R_{2s}) is said to be achievable for the 2-user CICC if there exists a coding scheme such that

$$\lim_{N\to\infty} P_e(N) = 0; \tag{7}$$

and

$$\lim_{N\to\infty} L(N) = 0 \text{ (strong secrecy), or} \tag{8}$$

$$\lim_{N\to\infty} L_R(N) = 0 \text{ (weak secrecy).} \tag{9}$$

The capacity-equivocation region of the CICC (under the constraint that the cognitive receiver must decode both transmitters' messages) was derived in [8] and is shown below.

Theorem 1 ([8]): The capacity-equivocation region of the CICC under the weak secrecy criterion is

$$\mathcal{R} = \bigcup_{P\in\mathcal{P}} \left\{ \begin{pmatrix} R_1 \\ R_2 \\ R_{2s} \end{pmatrix} \middle| \begin{array}{l} 0 \le R_1 \le \min\{I(U,X_1;Y_1), I(U,X_1;Y_2)\} \\ 0 \le R_2 \le I(U,V;Y_2|X_1) \\ R_1+R_2 \le \min\{I(U,X_1;Y_1), I(U,X_1;Y_2)\} \\ \qquad + I(V;Y_2|U,X_1) \\ 0 \le R_{2s} \le I(V;Y_2|U,X_1) - I(V;Y_1|U,X_1) \end{array} \right\},$$

where $\mathcal{P} = \{P_{UVX_1X_2}$ factorizing as: $P_{U,V,X_1}P_{X_2|V}\}$, and the cardinality bounds for auxiliary random variables U and V are

$$|\mathcal{U}| \le |\mathcal{X}_1|\cdot|\mathcal{X}_2| + 3,$$
$$|\mathcal{V}| \le |\mathcal{X}_1|^2\cdot|\mathcal{X}_2|^2 + 4|\mathcal{X}_1|\cdot|\mathcal{X}_2| + 3.$$

As we know, secrecy coding schemes require a large amount of randomness in the encoder. Reference [9] considered the generating rate of randomness needed by the stochastic encoder as a constraint and developed a more general achievable rate region under the strong secrecy criterion as shown in Theorem 2. In [9], besides the common message and the confidential message, transmitter 2's message was further divided into a private message, which should be decoded by receiver 2 but not necessarily be secret from receiver 1.

Theorem 2 ([9]): Let $\mathcal{R}^*$ be a closed convex set consisting of rate quadruples $(R_r, R_1, R_{2p}, R_{2s})$ for which there exist auxiliary random variables (U,V) such that

$$(U,X_1) \leftrightarrow V \leftrightarrow X_2, \tag{10}$$

$$(U,V) \leftrightarrow (X_1,X_2) \leftrightarrow (Y_1,Y_2), \tag{11}$$

and

$$R_1 \le \min\{I(U,X_1;Y_1), I(U,X_1;Y_2)\}, \tag{12}$$

$$R_{2p}+R_{2s} \le I(U,V;Y_2|X_1), \tag{13}$$

$$R_1+R_{2p}+R_{2s} \le I(V;Y_2|U,X_1) + \min\{I(U,X_1;Y_1), I(U,X_1;Y_2)\}, \tag{14}$$

$$R_{2s} \le I(V;Y_2|U,X_1) - I(V;Y_1|U,X_1), \tag{15}$$

$$R_{2p}+R_r \ge I(X_2;Y_1|U,X_1), \tag{16}$$

$$R_r \ge I(X_2;Y_1|U,V,X_1), \tag{17}$$

where R_r is the rate of randomness, and R_{2p} is transmitter 2's private message rate. Then $\mathcal{R}^*$ is an achievable rate region for the CICC. The cardinality bounds for auxiliary random variables U and V are the same as in Theorem 1.

III. Polar Coding Preliminaries

Polar codes were originally invented to achieve the symmetric capacity of discrete memoryless channels (DMC) [10]. To deal with non-uniform input distribution, one may apply Gallager's alphabet extension method [30, p. 208] as in [17], the chaining construction [31], or a more direct method which invokes results on polar coding for lossless compression [32]. Reference [33] proposed a refined method which overcomes the major drawback of the scheme in [32] that the encoder and decoder need to share a large amount of random mappings. Now we briefly review this method.

Let $W(Y|X)$ be a DMC with a q_X-ary input alphabet $\mathcal{X}$, where q_X is a prime number[2], and an arbitrary countable output alphabet $\mathcal{Y}$. Let $U^{1:N} = X^{1:N}\mathbf{G}_N$ and define $\mathcal{H}_X^{(N)}$, $\mathcal{L}_X^{(N)}$, $\mathcal{H}_{X|Y}^{(N)}$ and $\mathcal{L}_{X|Y}^{(N)}$ as follows:

$$\mathcal{H}_X^{(N)} = \{j\in[N] : H_{q_X}(U^j|U^{1:j-1}) \ge 1-\delta_N\}, \tag{18}$$

$$\mathcal{L}_X^{(N)} = \{j\in[N] : H_{q_X}(U^j|U^{1:j-1}) \le \delta_N\}, \tag{19}$$

$$\mathcal{H}_{X|Y}^{(N)} = \{j\in[N] : H_{q_X}(U^j|Y^{1:N},U^{1:j-1}) \ge 1-\delta_N\}, \tag{20}$$

$$\mathcal{L}_{X|Y}^{(N)} = \{j\in[N] : H_{q_X}(U^j|Y^{1:N},U^{1:j-1}) \le \delta_N\}, \tag{21}$$

which satisfy [22], [34]

$$\lim_{N\to\infty}\frac{1}{N}|\mathcal{H}_X^{(N)}| = H_{q_X}(X), \tag{22}$$

$$\lim_{N\to\infty}\frac{1}{N}|\mathcal{L}_X^{(N)}| = 1-H_{q_X}(X), \tag{23}$$

$$\lim_{N\to\infty}\frac{1}{N}|\mathcal{H}_{X|Y}^{(N)}| = H_{q_X}(X|Y), \tag{24}$$

$$\lim_{N\to\infty}\frac{1}{N}|\mathcal{L}_{X|Y}^{(N)}| = 1-H_{q_X}(X|Y). \tag{25}$$

Define the *information set* (or *reliable set*), *frozen set* and *almost deterministic set* respectively as follows:

$$\mathcal{I} \triangleq \mathcal{H}_X^{(N)} \cap \mathcal{L}_{X|Y}^{(N)}, \tag{26}$$

$$\mathcal{F} \triangleq \mathcal{H}_X^{(N)} \cap (\mathcal{L}_{X|Y}^{(N)})^C, \tag{27}$$

$$\mathcal{D} \triangleq (\mathcal{H}_X^{(N)})^C. \tag{28}$$

$\mathcal{D}$ is called almost deterministic because part of its indices, $(\mathcal{H}_X^{(N)})^C \cap (\mathcal{L}_X^{(N)})^C$, are not fully polarized. The encoding procedure goes as follows:

- $\{u^j\}_{j\in\mathcal{I}}$ carry information,
- $\{u^j\}_{j\in\mathcal{F}}$ are filled with uniformly distributed frozen symbols (shared between the transmitter and the receiver),
- $\{u^j\}_{j\in\mathcal{D}}$ are randomly generated according to conditional probability $P_{U^j|U^{1:j-1}}(u|u^{1:j-1})$.

[2]For the prime number case, polarization is similar to the binary case. For composite q_X, one needs to use some special types of operations to guarantee polarization [34]–[37]. We only consider the prime number case in this paper for simplicity.

In order for the receiver to decode successfully, [33] proposed to send part of the almost deterministic symbols, $\{u^j\}_{j\in(\mathcal{H}_X^{(N)})^C\cap(\mathcal{L}_{X|Y}^{(N)})^C}$, to the receiver with some reliable error-correcting code separately, the rate of which is shown to vanish as N goes to infinity.

Having received $y^{1:N}$ and recovered $\{u^j\}_{j\in(\mathcal{H}_X^{(N)})^C\cap(\mathcal{L}_{X|Y}^{(N)})^C}$, the receiver decodes $u^{1:N}$ with a successive cancellation decoder (SCD):

$$\bar{u}^j = \begin{cases} u^j, & \text{if } j \in (\mathcal{L}_{X|Y}^{(N)})^C, \\ \arg\max_{u\in\{0,1\}}, & \\ \quad P_{U^j|Y^{1:N}U^{1:j-1}}(u|y^{1:N}, u^{1:j-1}) & \text{if } j \in \mathcal{L}_{X|Y}^{(N)}. \end{cases}$$

The transmission rate of this scheme, $R = |\mathcal{I}|/N$, is shown to achieve channel capacity [32]

$$\lim_{N\to\infty} R = I(X;Y). \tag{29}$$

IV. Proposed Polar Coding Scheme

In this paper, we only discuss the case when random variables X_1, X_2, U and V all have prime alphabets. Suppose $q_{X_1} = |\mathcal{X}_1|$ and $q_{X_2} = |\mathcal{X}_2|$ are two prime numbers, $q_U = |\mathcal{U}|$ is the smallest prime number larger than $|\mathcal{X}_1| \cdot |\mathcal{X}_2| + 3$, and $q_V = |\mathcal{V}|$ is the smallest prime number larger than $|\mathcal{X}_1|^2\cdot|\mathcal{X}_2|^2+4|\mathcal{X}_1|\cdot|\mathcal{X}_2|+3$. Consider a random variable tuple $(U, V, X_1, X_2, Y_1, Y_2)$ with joint distribution $P_{UVX_1X_2Y_1Y_2}$ that satisfy (10) and (11). The goal of our proposed scheme is to achieve every equation in (12)–(17).

Our encoding scheme is illustrated in Fig. 2. Transmitter 1's message M_1 is split into two parts, $M_1^{(1)}$ and $M_1^{(2)}$, carried by transmitter 1's and transmitter 2's signals respectively. Transmitter 2's message M_2 is split into three parts, a common message $M_2^{(c)}$ intended for both receivers, a private message $M_2^{(p)}$ intended only for receiver 2, and a confidential message $M_2^{(s)}$ intended only for receiver 2 and must be secured from receiver 1. Details of transmitter 2's encoding procedure are as follows. $M_1^{(2)}$ and $M_2^{(c)}$ are encoded into $U^{1:N}$ first, $M_2^{(p)}$ and $M_2^{(s)}$ are then superimposed on $(U^{1:N}, X_1^{1:N})$ and encoded into $V^{1:N}$ (known as superposition coding). Finally, randomness M_R is added to $V^{1:N}$ to generate transmitter 2's final codeword $X_2^{1:N}$ (known as channel prefixing). Note that $X_1^{1:N}$ can be seen as the known interference to transmitter 2. Thus, this superposition coding scheme also involves the idea of dirty paper coding. In the rest of this section, the rates of $M_1^{(1)}$, $M_1^{(2)}$, $M_2^{(c)}$, $M_2^{(p)}$ and $M_2^{(s)}$ will be denoted by $R_1^{(1)}$, $R_1^{(2)}$, $R_2^{(c)}$, $R_2^{(p)}$ and $R_2^{(s)}$, respectively. Notice that in Theorem 2, R_{2p} is the sum of $R_2^{(c)}$ and $R_2^{(p)}$ defined here.

The chaining method [38] is a commonly adopted way to solve the problem unaligned polar indices in multi-user channels, which is also used in our scheme. Furthermore, we show that with properly designed chaining schemes, the secrecy capacity region of the CICC can be achieved with point-to-point polar codes. In our scheme, m ($m \geq 1$) encoding blocks are chained into a *frame*, and two receivers decode a frame in reverse orders. Our scheme is designed in such a way that receiver 1 (the primary receiver) decodes in the natural order (i.e., from block 1 to block m) while receiver 2 (the secondary receiver) decodes in the reverse order (i.e., from block m to block 1).

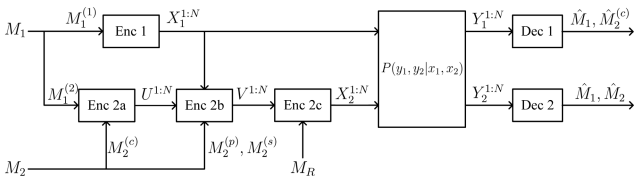

Fig. 2. Our encoding scheme for the CICC.

A. Common Message Encoding

Since common messages $M_1^{(1)}$, $M_1^{(2)}$ and $M_2^{(c)}$ are encoded into two sequences of random variables, $X_1^{1:N}$ and $U^{1:N}$, it is natural to consider designing a compound MAC polar code for the two synthetic MACs, $P(Y_1|X_1, U)$ and $P(Y_2|X_1, U)$. This approach requires the use of MAC polarization or rate splitting which will increase the complexity. Note that there is no major difference between $M_1^{(2)}$ and $M_2^{(c)}$ in regard to the encoding. They only affect the rate allocation between M_1 and M_2. Due to this flexibility and the cognitivity of the channel, cross-transmitter chaining schemes can be designed to do rate allocation between the two users instead of using MAC polar codes or rate splitting. For simplicity, define $R^{(c)} = R_1^{(2)} + R_2^{(c)}$. From (12) we have

$$R_1^{(1)} + R^{(c)} \leq \min\{I(U, X_1; Y_1), I(U, X_1; Y_2)\}. \tag{30}$$

We first show how to achieve (30) in this subsection and $R_2^{(p)} + R_2^{(s)} = I(U, V; Y_2|X_1)$ in the next subsection, and then prove in Section V-D that other rate pairs in the achievable rate region in Theorem 2 can be achieved by adjusting the ratio between $M_1^{(2)}$ and $M_2^{(c)}$.

Let $U_1^{1:N} = X_1^{1:N}\mathbf{G}_N$ and $U'^{1:N} = U^{1:N}\mathbf{G}_N$. Define the following polarized sets:

$$\begin{aligned}
\mathcal{H}_{X_1}^{(N)} &\triangleq \big\{j \in [N] : H_{q_{X_1}}(U_1^j|U_1^{1:j-1}) \geq 1 - \delta_N\big\}, \\
\mathcal{L}_{X_1|Y_1}^{(N)} &\triangleq \big\{j \in [N] : H_{q_{X_1}}(U_1^j|Y_1^{1:N}, U_1^{1:j-1}) \leq \delta_N\big\}, \\
\mathcal{L}_{X_1|Y_2}^{(N)} &\triangleq \big\{j \in [N] : H_{q_{X_1}}(U_1^j|Y_2^{1:N}, U_1^{1:j-1}) \leq \delta_N\big\}, \\
\mathcal{H}_{U|X_1}^{(N)} &\triangleq \big\{j \in [N] : H_{q_U}(U'^j|X_1^{1:N}, U'^{1:j-1}) \geq 1 - \delta_N\big\}, \\
\mathcal{L}_{U|Y_1X_1}^{(N)} &\triangleq \big\{j \in [N] : H_{q_U}(U'^j|Y_1^{1:N}, X_1^{1:N}, U'^{1:j-1}) \leq \delta_N\big\}, \\
\mathcal{L}_{U|Y_2X_1}^{(N)} &\triangleq \big\{j \in [N] : H_{q_U}(U'^j|Y_2^{1:N}, X_1^{1:N}, U'^{1:j-1}) \leq \delta_N\big\}.
\end{aligned} \tag{31}$$

Then define the following sets of indices for $U_1^{1:N}$:

$$\begin{aligned}
\mathcal{I}_{1c}^{(1)} &= \mathcal{H}_{X_1}^{(N)} \cap \mathcal{L}_{X_1|Y_1}^{(N)}, \\
\mathcal{I}_{1c}^{(2)} &= \mathcal{H}_{X_1}^{(N)} \cap \mathcal{L}_{X_1|Y_2}^{(N)}, \\
\mathcal{F}_{1c} &= \mathcal{H}_{X_1}^{(N)} \cap \big(\mathcal{L}_{X_1|Y_1}^{(N)}\big)^C \cap \big(\mathcal{L}_{X_1|Y_2}^{(N)}\big)^C, \\
\mathcal{D}_{1c} &= \big(\mathcal{H}_{X_1}^{(N)}\big)^C,
\end{aligned} \tag{32}$$

where $\mathcal{I}_{1c}^{(1)}$ and $\mathcal{I}_{1c}^{(2)}$ are the reliable sets for receiver 1 and 2 respectively, $\mathcal{F}_{1c}$ is the intersection of two receivers' frozen

sets, and $\mathcal{D}_{1c}$ is the almost deterministic set. Similarly define

$$\begin{aligned}\mathcal{I}_{2c}^{(1)} &= \mathcal{H}_{U|X_1}^{(N)} \cap \mathcal{L}_{U|Y_1X_1}^{(N)},\\ \mathcal{I}_{2c}^{(2)} &= \mathcal{H}_{U|X_1}^{(N)} \cap \mathcal{L}_{U|Y_2X_1}^{(N)},\\ \mathcal{F}_{2c} &= \mathcal{H}_{U|X_1}^{(N)} \cap \left(\mathcal{L}_{U|Y_1X_1}^{(N)}\right)^C \cap \left(\mathcal{L}_{U|Y_2X_1}^{(N)}\right)^C,\\ \mathcal{D}_{2c} &= \left(\mathcal{H}_{U|X_1}^{(N)}\right)^C. \end{aligned} \tag{33}$$

for $U'^{1:N}$. From (29) we have

$$\begin{aligned}&\lim_{N\to\infty}\frac{1}{N}|\mathcal{I}_{1c}^{(1)}| = I(X_1;Y_1), \quad \lim_{N\to\infty}\frac{1}{N}|\mathcal{I}_{2c}^{(1)}| = I(U;Y_1|X_1),\\ &\lim_{N\to\infty}\frac{1}{N}|\mathcal{I}_{1c}^{(2)}| = I(X_1;Y_2), \quad \lim_{N\to\infty}\frac{1}{N}|\mathcal{I}_{2c}^{(2)}| = I(U;Y_2|X_1).\end{aligned} \tag{34}$$

If we design two separate chaining schemes for $U^{1:N}$ and $U'^{1:N}$ respectively, it is easy to verify that the achievable common message rate is

$$\begin{aligned}R_1^{(1)} + R^{(c)} \le \min\{I(X_1;Y_1), I(X_1;Y_2)\}\\ + \min\{I(U;Y_1|X_1), I(U;Y_2|X_1)\}.\end{aligned} \tag{35}$$

Such a scheme achieves (30) only in the following two cases:

- (Case 1) $I(X_1;Y_1) \le I(X_1;Y_2)$ and $I(U;Y_1|X_1) \le I(U;Y_2|X_1)$,
- (Case 2) $I(X_1;Y_1) \ge I(X_1;Y_2)$ and $I(U;Y_1|X_1) \ge I(U;Y_2|X_1)$.

In the following two other cases of

- (Case 3) $I(X_1;Y_1) < I(X_1;Y_2)$ and $I(U;Y_1|X_1) > I(U;Y_2|X_1)$,
- (Case 4) $I(X_1;Y_1) > I(X_1;Y_2)$ and $I(U;Y_1|X_1) < I(U;Y_2|X_1)$,

the achievable rate in (35) is strictly smaller than that in (30). In these cases, the chaining scheme should be jointly designed for $U^{1:N}$ and $U'^{1:N}$, which we refer to as cross-transmitter chaining.

1) Case 1 and Case 2: Since Case 2 is similar to Case 1 by swapping the roles of two transmitters, we only describe the chaining scheme in Case 1 for brevity. From (34) we know that given sufficiently large N, we always have $|\mathcal{I}_{1c}^{(1)}| \le |\mathcal{I}_{1c}^{(2)}|$ and $|\mathcal{I}_{2c}^{(1)}| \le |\mathcal{I}_{2c}^{(2)}|$. Define

$$\begin{aligned}\mathcal{I}_{1c}^{(0)} &= \mathcal{I}_{1c}^{(1)} \cap \mathcal{I}_{1c}^{(2)}, \quad \mathcal{I}_{1c}^{(1a)} = \mathcal{I}_{1c}^{(1)} \setminus \mathcal{I}_{1c}^{(2)},\\ \mathcal{I}_{2c}^{(0)} &= \mathcal{I}_{2c}^{(1)} \cap \mathcal{I}_{2c}^{(2)}, \quad \mathcal{I}_{2c}^{(1a)} = \mathcal{I}_{2c}^{(1)} \setminus \mathcal{I}_{2c}^{(2)}.\end{aligned} \tag{36}$$

Choose an arbitrary subset $\mathcal{I}_{1c}^{(2a)}$ of $\mathcal{I}_{1c}^{(2)} \setminus \mathcal{I}_{1c}^{(1)}$ such that $|\mathcal{I}_{1c}^{(2a)}| = |\mathcal{I}_{1c}^{(1a)}|$, and an arbitrary subset $\mathcal{I}_{2c}^{(2a)}$ of $\mathcal{I}_{2c}^{(2)} \setminus \mathcal{I}_{2c}^{(1)}$ such that $|\mathcal{I}_{2c}^{(2a)}| = |\mathcal{I}_{2c}^{(1a)}|$. The chaining scheme goes as follows.

(I) In the 1st block, transmitter 1 encodes its common message as:

- $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(0)}\cup\mathcal{I}_{1c}^{(1a)}}$ store message symbols from $M_1^{(1)}$,
- $\{u_1^j\}_{j\in(\mathcal{I}_{1c}^{(0)}\cup\mathcal{I}_{1c}^{(1a)}\cup\mathcal{D}_{1c})^C}$ carry frozen symbols uniformly distributed over $\mathcal{X}_1$,
- $\{u_1^j\}_{j\in\mathcal{D}_{1c}}$ are randomly generated according to conditional probability $P_{U_1^j|U_1^{1:j-1}}(u|u_1^{1:j-1})$,

and transmitter 2 encodes its common message as:

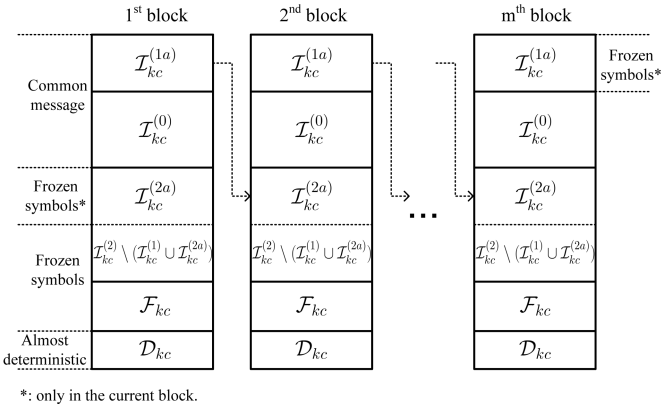

Fig. 3. The chaining scheme of transmitter k ($k = 1, 2$) for common messages in Case 1.

- $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(0)}\cup\mathcal{I}_{2c}^{(1a)}}$ store message symbols from $M_1^{(2)}$ and $M_2^{(c)}$,
- $\{u'^j\}_{j\in(\mathcal{I}_{2c}^{(0)}\cup\mathcal{I}_{2c}^{(1a)}\cup\mathcal{D}_{2c})^C}$ carry frozen symbols uniformly distributed over $\mathcal{U}$,
- $\{u'^j\}_{j\in\mathcal{D}_{2c}}$ are randomly generated according to conditional probability $P_{U'^j|U'^{1:j-1}}(u|u'^{1:j-1})$.

(II) In the ith ($1 < i < m$) block, transmitter 1 assigns $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(2a)}}$ with the same value of $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(1a)}}$ in block $i-1$, and transmitter 2 assigns $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(2a)}}$ with the same value of $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(1a)}}$ in block $i-1$. The rest of $u_1^{1:N}$ and $u'^{1:N}$ are determined in the same way as in (I).

(III) In the mth block, transmitter 1 assigns $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(1a)}}$ with frozen symbols uniformly distributed over $\mathcal{X}_1$, and transmitter 2 assigns $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(1a)}}$ with frozen symbols uniformly distributed over $\mathcal{U}$. The rest of $u_1^{1:N}$ and $u'^{1:N}$ are determined in the same way as in (II).

The chaining scheme in Case 1 is shown in Fig. 3. After each transmission block, transmitter 1 additionally sends a vanishing fraction of the almost deterministic symbols, $\{u_1^j\}_{j\in\,(\mathcal{H}_{X_1}^{(N)})^C\cap(\mathcal{L}_{X_1|Y_1}^{(N)})^C}$ and $\{u_1^j\}_{j\in(\mathcal{H}_{X_1}^{(N)})^C\cap(\mathcal{L}_{X_1|Y_2}^{(N)})^C}$, to receiver 1 and 2 respectively with some reliable error-correcting code. Similarly, transmitter 2 sends $\{u'^j\}_{j\in(\mathcal{H}_{U|X_1}^{(N)})^C\cap(\mathcal{L}_{U|Y_1X_1}^{(N)})^C}$ and $\{u'^j\}_{j\in(\mathcal{H}_{U|X_1}^{(N)})^C\cap(\mathcal{L}_{U|Y_2X_1}^{(N)})^C}$ to the two receivers respectively after each block. From Section III we know that the rate for transmitting these symbols vanishes as N increases. Thus, the cost for these extra transmissions can be made negligible.

Also note that frozen symbols at $\mathcal{F}_{kc} \cup \left(\mathcal{I}_{kc}^{(2)} \setminus (\mathcal{I}_{kc}^{(1)} \cup \mathcal{I}_{kc}^{(2a)})\right)$ ($k = 1, 2$) can be reused since they only need to be independently and uniformly distributed. Thus, the rate of frozen symbols which must be shared between transmitters and receivers can be made negligible as well by reusing them over sufficient number of blocks. In the secrecy analysis we will prove that the reuse of frozen symbols does not harm secrecy.

2) Case 3 and Case 4: Since Case 4 is similar to Case 3 by swapping the roles of two transmitters, we only describe the

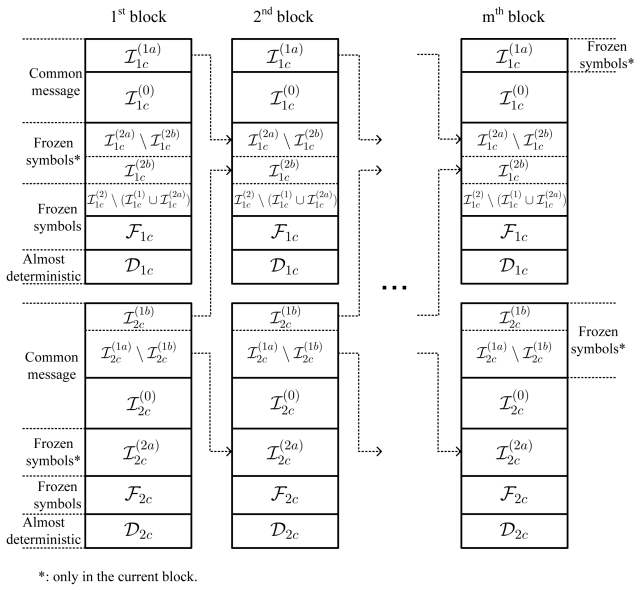

Fig. 4. The chaining scheme for common messages in Case 3-1.

chaining scheme in Case 3 for brevity. In this case, given sufficiently large N, we always have $|\mathcal{I}_{1c}^{(1)}| \leq |\mathcal{I}_{1c}^{(2)}|$ and $|\mathcal{I}_{2c}^{(1)}| \geq |\mathcal{I}_{2c}^{(2)}|$.

If $\min\{I(U,X_1;Y_1), I(U,X_1;Y_2)\} = I(U,X_1;Y_1)$, which we refer to as Case 3-1, we have $|\mathcal{I}_{1c}^{(2)}| - |\mathcal{I}_{1c}^{(1)}| \geq |\mathcal{I}_{2c}^{(1)}| - |\mathcal{I}_{2c}^{(2)}|$ for sufficiently large N. In this case, define $\mathcal{I}_{1c}^{(0)}$, $\mathcal{I}_{2c}^{(0)}$, $\mathcal{I}_{1c}^{(1a)}$ and $\mathcal{I}_{2c}^{(1a)}$ in the same way as in (36), and define

$$\mathcal{I}_{2c}^{(2a)} = \mathcal{I}_{2c}^{(2)} \setminus \mathcal{I}_{2c}^{(1)}. \tag{37}$$

Choose an arbitrary subset $\mathcal{I}_{2c}^{(1b)}$ of $\mathcal{I}_{2c}^{(1a)}$ with $|\mathcal{I}_{2c}^{(1b)}| = |\mathcal{I}_{2c}^{(1)}| - |\mathcal{I}_{2c}^{(2)}|$, and an arbitrary subset $\mathcal{I}_{1c}^{(2a)}$ of $\mathcal{I}_{1c}^{(2)} \setminus \mathcal{I}_{1c}^{(1)}$ with $|\mathcal{I}_{1c}^{(2a)}| = |\mathcal{I}_{1c}^{(1a)}| + |\mathcal{I}_{2c}^{(1b)}|$. Let $\mathcal{I}_{1c}^{(2b)}$ be a subset of $\mathcal{I}_{1c}^{(2a)}$ with the same size as $\mathcal{I}_{2c}^{(1b)}$. The chaining scheme in Case 3-1 goes as follows and is illustrated in Fig. 4.

(I) In the 1st block, the encoding procedure is similar to Case 1, except that $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(1b)}}$ of transmitter 2 are filled with message symbols from $M_1^{(1)}$ only, as they will be chained with transmitter 1's next encoding block.

(II) In the ith ($1 < i < m$) block, transmitter 1 assigns $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(2a)}\setminus\mathcal{I}_{1c}^{(2b)}}$ with the same value of $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(1a)}}$ in block $i-1$, and $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(2b)}}$ with the same value of $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(1b)}}$ in block $i-1$, while transmitter 2 assigns $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(2a)}}$ with the same value of $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(1a)}\setminus\mathcal{I}_{2c}^{(1b)}}$ in block $i-1$. The rest of $u_1^{1:N}$ and $u'^{1:N}$ are determined in the same way as in (I).

(III) In the mth block, transmitter 1 assigns $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(1a)}}$ with frozen symbols uniformly distributed over $\mathcal{X}_1$, and transmitter 2 assigns $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(1a)}}$ with frozen symbols uniformly distributed over $\mathcal{U}$. The rest of $u_1^{1:N}$ and $u'^{1:N}$ are determined in the same way as in (II).

Otherwise if $\min\{I(U,X_1;Y_1), I(U,X_1;Y_2)\} = I(U,X_1;Y_2)$, which we refer to as Case 3-2, we have $|\mathcal{I}_{1c}^{(2)}| - |\mathcal{I}_{1c}^{(1)}| \leq |\mathcal{I}_{2c}^{(1)}| - |\mathcal{I}_{2c}^{(2)}|$ given sufficiently large N. The chaining scheme in this case is similar to that in Fig. 4 with the two transmitters exchanging their roles.

Similar to Case 1, two transmitters send part of their almost deterministic symbols to the two receivers with some reliable error-correcting code after each block. Also, transmitter 1's frozen symbols at $\mathcal{F}_{1c} \cup \left(\mathcal{I}_{1c}^{(2)} \setminus (\mathcal{I}_{1c}^{(1)} \cup \mathcal{I}_{1c}^{(2a)})\right)$ and transmitter 2's frozen symbols at $\mathcal{F}_{2c}$ can be reused over different blocks.

B. Private and Confidential Message Encoding

Since private message $M_2^{(p)}$ and confidential message $M_2^{(s)}$ are superimposed on $(U^{1:N}, X_1^{1:N})$ by auxiliary random variable $V^{1:N}$, we treat $(U^{1:N}, X_1^{1:N})$ as side information when applying polarization on $V^{1:N}$. Let $V'^{1:N} = V^{1:N}\mathbf{G}_N$ and define

$$\begin{aligned}
\mathcal{H}_{V|X_1U}^{(N)} &\triangleq \big\{j \in [N] : H_{q_V}(V'^j | X_1^{1:N}, U^{1:N}, V'^{1:j-1}) \\
&\quad \geq 1-\delta_N\big\}, \\
\mathcal{H}_{V|Y_1X_1U}^{(N)} &\triangleq \big\{j \in [N] : H_{q_V}(V'^j | Y_1^{1:N}, X_1^{1:N}, U^{1:N}, V'^{1:j-1}) \\
&\quad \geq 1-\delta_N\big\}, \\
\mathcal{L}_{V|Y_2X_1U}^{(N)} &\triangleq \big\{j \in [N] : H_{q_V}(V'^j | Y_2^{1:N}, X_1^{1:N}, U^{1:N}, V'^{1:j-1}) \\
&\quad \leq \delta_N\big\}.
\end{aligned} \tag{38}$$

Partition the indices of $V'^{1:N}$ as follows:

$$\begin{aligned}
\mathcal{I}_{2s} &= \mathcal{H}_{V|X_1U}^{(N)} \cap \mathcal{L}_{V|Y_2X_1U}^{(N)} \cap \mathcal{H}_{V|Y_1X_1U}^{(N)}, \\
\mathcal{I}_{2p} &= \mathcal{H}_{V|X_1U}^{(N)} \cap \mathcal{L}_{V|Y_2X_1U}^{(N)} \cap \big(\mathcal{H}_{V|Y_1X_1U}^{(N)}\big)^C, \\
\mathcal{F}_2 &= \mathcal{H}_{V|X_1U}^{(N)} \cap \big(\mathcal{L}_{V|Y_2X_1U}^{(N)}\big)^C \cap \mathcal{H}_{V|Y_1X_1U}^{(N)}, \\
\mathcal{R}_2 &= \mathcal{H}_{V|X_1U}^{(N)} \cap \big(\mathcal{L}_{V|Y_2X_1U}^{(N)}\big)^C \cap \big(\mathcal{H}_{V|Y_1X_1U}^{(N)}\big)^C, \\
\mathcal{D}_2 &= \big(\mathcal{H}_{V|X_1U}^{(N)}\big)^C,
\end{aligned} \tag{39}$$

where $\mathcal{I}_{2s}$ is the reliable and secure set, $\mathcal{I}_{2p}$ is the reliable but insecure set, $\mathcal{R}_2$ is the unreliable and insecure set, $\mathcal{F}_2$ is the frozen set, and $\mathcal{D}_{2c}$ is the almost deterministic set.

The aim of using the chaining method here is to deal with the unreliable and insecure set $\mathcal{R}_2$. Consider the positive secrecy rate case (i.e., the right-hand-side of (15) is positive). In this case, $|\mathcal{I}_{2s}| > |\mathcal{R}_2|$ always holds for sufficiently large N. Choose a subset $\mathcal{I}_{2s}^{(2)}$ of $\mathcal{I}_{2s}$ such that $|\mathcal{I}_{2s}^{(2)}| = |\mathcal{R}_2|$. Denote $\mathcal{I}_{2s}^{(1)} = \mathcal{I}_{2s} \setminus \mathcal{I}_{2s}^{(2)}$. The chaining scheme for transmitter 2's private and confidential messages is also designed in such a way that receiver 2 decodes from block m to block 1, same as its decoding order for common messages. Details of the scheme are as follows and shown in Fig. 5.

(I) In the 1st block,

- $\{v'^j\}_{j\in\mathcal{I}_{2s}^{(1)}}$ store confidential message symbols from $M_2^{(s)}$,
- $\{v'^j\}_{j\in\mathcal{I}_{2p}\cup\mathcal{I}_{2s}^{(2)}\cup\mathcal{R}_2}$ carry private message symbols from $M_2^{(p)}$,
- $\{v'^j\}_{j\in\mathcal{F}_2}$ are filled with frozen symbols uniformly distributed over $\mathcal{V}$,
- $\{v'^j\}_{j\in\mathcal{D}_2}$ are randomly generated according to conditional probability $P_{V'^j|X_1^{1:N}U^{1:N}V'^{1:j-1}}$,

(II) In the ith ($1 < i < m$) block, $\{v'^j\}_{i\in\mathcal{I}_{2s}^{(2)}}$ are assigned with the same value as $\{v'^j\}_{i\in\mathcal{R}_2}$ in the $(i-1)$th block, and the rest of $v'^{1:N}$ are determined in the same way as in (I).

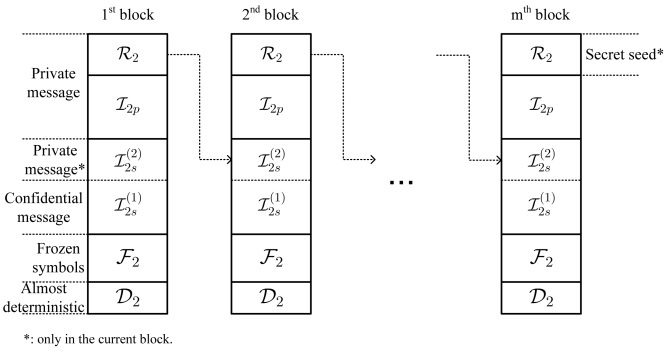

Fig. 5. The chaining scheme for transmitter 2's private and confidential messages.

(III) In the mth block, $\{v'^j\}_{j\in\mathcal{R}_2}$ carry some uniformly distributed random symbols that are shared only between transmitter 2 and receiver 2 (known as secret seed), and the rest of $v'^{1:N}$ are determined in the same way as in (II).

The secret seed rate can be made arbitrarily small by increasing the number of chained blocks in a frame. After each transmission block, transmitter 2 additionally sends a vanishing fraction of the almost deterministic symbols, $\{v'^j\}_{j\in\ (\mathcal{H}^{(N)}_{V|X_1U})^C\cap\ (\mathcal{L}^{(N)}_{V|Y_2X_1U})^C}$, to receiver 2 secretly with some reliable error-correcting code. Note that unlike in the common message encoding, the additional transmission here must be kept secret from receiver 1. Nevertheless, the rate of this transmission can be made arbitrarily small by increasing N. Similar to the common message encoding, frozen symbols at $\mathcal{F}_2$ can also be reused over different blocks.

C. Channel Prefixing

To generate the final codeword $X_2^{1:N}$ for transmitter 2, one can transmit $(X_1^{1:N}, U^{1:N}, V^{1:N})$ through a virtual channel with transition probability $P_{X_2|X_1UV}$. Also, one can consider X_2 and (X_1, U, V) as correlated sources and apply polar source coding to obtain the final codeword. To design a scheme that requires the minimum generating rate of randomness, we take the latter approach in this paper. Let $U_2^{1:N} = X_2^{1:N}\mathbf{G}_N$ and define

$$\begin{aligned}
&\mathcal{H}^{(N)}_{X_2|X_1UV}\\
&\triangleq \big\{j\in[N]: H_{q_{X_2}}(U_2^j|X_1^{1:N},U^{1:N},V^{1:N},U_2^{1:j-1})\\
&\qquad\geq 1-\delta_N\big\},\\
&\quad \mathcal{H}^{(N)}_{X_2|Y_1X_1UV}\\
&\triangleq \big\{j\in[N]: H_{q_{X_2}}(U_2^j|Y_1^{1:N},X_1^{1:N},U^{1:N},V^{1:N},U_2^{1:j-1})\\
&\qquad\geq 1-\delta_N\big\},\\
&\quad \mathcal{L}^{(N)}_{X_2|Y_1X_1UV}\\
&\triangleq \big\{j\in[N]: H_{q_{X_2}}(U_2^j|Y_1^{1:N},X_1^{1:N},U^{1:N},V^{1:N},U_2^{1:j-1})\\
&\qquad\leq \delta_N\big\}. \qquad (40)
\end{aligned}$$

Let w_r be a random sequence uniformly distributed over $\mathcal{X}_2$ and of length $|\mathcal{H}^{(N)}_{X_2|Y_1X_1UV}|$. Once $(X_1^{1:N}, U^{1:N}, V^{1:N})$ is determined, $X_2^{1:N}$ can be obtained as follows:

- $\{u_2^j\}_{j\in\mathcal{H}^{(N)}_{X_2|Y_1X_1UV}} = w_r$,
- $\{u_2^j\}_{j\in\mathcal{H}^{(N)}_{X_2|X_1UV}\setminus\mathcal{H}^{(N)}_{X_2|Y_1X_1UV}}$ are filled with random symbols uniformly distributed over $\mathcal{X}_2$,
- $\{u_2^j\}_{j\in(\mathcal{H}^{(N)}_{X_2|X_1UV})^C}$ are randomly generated according to conditional probability $P_{U_2^j|X_1^{1:N}U^{1:N}V^{1:N}U_2^{1:j-1}}$,
- Compute $x_2^{1:N} = u_2^{1:N}\mathbf{G}_N$.

An intuitive explanation for why random symbols in $\mathcal{H}^{(N)}_{X_2|Y_1X_1UV}$ can be reused but not those in $\mathcal{H}^{(N)}_{X_2|X_1UV}\setminus\mathcal{H}^{(N)}_{X_2|Y_1X_1UV}$ is that $\{u_2^j\}_{j\in\mathcal{H}^{(N)}_{X_2|Y_1X_1UV}}$ are very unreliable for receiver 1, thus reusing them does not harm security. We will show in the next section that with such a channel prefixing approach, our proposed scheme can achieve strong secrecy.

D. Decoding

1) Common Message Decoding: We first consider receiver 1, who decodes from block 1 to block m. Although we have considered different cases in Section IV-A, the decoding procedure can be summarized in a unified form as follows:

(I) In the 1st block, receiver 1 first decodes $\{u_1^j\}_{j\in\mathcal{I}_{1c}^{(0)}\cup\mathcal{I}_{1c}^{(1a)}}$ with a SCD and obtains an estimate of $\bar{u}_1^{1:N}$. Then it decodes $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(0)}\cup\mathcal{I}_{2c}^{(1a)}}$ with a SCD, in which $\bar{u}_1^{1:N}$ is treated as side information, and obtains an estimate of $\bar{u}'^{1:N}$.

(II) In the ith $(1 < i < m)$ block, $\{\bar{u}_1^j\}_{j\in\mathcal{I}_{1c}^{(2a)}}$ and $\{\bar{u}'^j\}_{j\in\mathcal{I}_{2c}^{(2a)}}$ are deduced from $\bar{u}_1^{1:N}$ and $\bar{u}'^{1:N}$ in block $i-1$ according to different cases (see Fig. 3 and Fig. 4), and the rest are decoded in the same way as in (I).

(III) In the mth block, $\{\bar{u}_1^j\}_{j\in\mathcal{I}_{1c}^{(1a)}}$ and $\{\bar{u}_1^j\}_{j\in\mathcal{I}_{2c}^{(1a)}}$ are decoded as frozen symbols, and the rest are decoded in the same way as in (II).

Receiver 2 decodes the common messages similarly, except that it decodes from block m to block 1.

2) Private and Confidential Messages Decoding: Receiver 2 decodes the private and confidential messages from block m to block 1 as follows.

(I) In the mth block, receiver 2 decodes $\{v'^j\}_{j\in\mathcal{I}_{2p}\cup\mathcal{I}_{2s}}$ with $\hat{u}_1^{1:N}$ and $\hat{u}'^{1:N}$ in the same block being treated as side information, where $\hat{u}_1^{1:N}$ and $\hat{u}'^{1:N}$ are its decoding results of the common messages, and obtains an estimate of $\hat{v}'^{1:N}$.

(II) In the ith $(1 \leq i < m)$ block, $\{v'^j\}_{i\in\mathcal{R}_2}$ are deduced from $\{\hat{v}'^j\}_{i\in\mathcal{I}_{2s}^{(2)}}$ in block $i+1$, and the rest are decoded in the same way as in (I).

V. Performance Analysis

A. Total Variation Distance

Let $\tilde{U}^{1:N}$, $\tilde{V}^{1:N}$, $\tilde{X}_1^{1:N}$, $\tilde{X}_2^{1:N}$, $\tilde{Y}_1^{1:N}$ and $\tilde{Y}_2^{1:N}$ be the vectors generated by our encoding scheme. The following lemma shows that the joint distribution of random variables induced by our encoding scheme is asymptotically indistinguishable from the target joint distribution of $P_{U^{1:N}V^{1:N}X_1^{1:N}X_2^{1:N}Y_1^{1:N}Y_2^{1:N}}$.

Lemma 1:

$$\begin{aligned}
\| P_{U^{1:N}V^{1:N}X_1^{1:N}X_2^{1:N}Y_1^{1:N}Y_2^{1:N}}\\
-P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}\tilde{X}_2^{1:N}\tilde{Y}_1^{1:N}\tilde{Y}_2^{1:N}}\|\leq\delta_N^c. \qquad (41)
\end{aligned}$$

where $\delta_N^c \triangleq \sqrt{2\log 2}\sqrt{N\delta_N}(2+2\sqrt{3})$.

Proof: See Appendix A. □

B. Error Performance

Lemma 2: The error probability of receiver 1 (resp. 2) in decoding all common messages in a whole frame can be upper bounded by

$$P_{e1}\ (\textit{resp. } P_{e2}^{(c)}) \le O(3^m\sqrt{N\delta_N}), \tag{42}$$

while the error probability of receiver 2 in decoding all private and confidential messages in a frame can be upper bounded by

$$P_{e2}^{(p,s)} \le O(m3^m\sqrt{N\delta_N}). \tag{43}$$

Proof: See Appendix B. □

C. Secrecy

We first introduce some notations used in this subsection. In the ith $(1 \le i \le m)$ block, the outputs of Enc 1, 2a and 2b (see Fig. 2) are denoted by $\mathbf{X}_{1,i}$, $\mathbf{U}_i$ and $\mathbf{V}_i$, respectively. Transmitter 2's confidential message at $\mathcal{I}_{2s}^{(1)}$ is denoted by M_i, and private message at $\mathcal{I}_{2s}^{(2)}$ (which is used for chaining) by E_i. Receiver 1's channel output is denoted by $\mathbf{Y}_{1,i}$. The additionally transmitted almost deterministic symbols in $U_1^{1:N}$ and $U'^{1:N}$ are denoted by $D_{1c,i}$ and $D_{2c,i}$, respectively. The reused frozen symbols in $U_1^{1:N}$, $U'^{1:N}$ and $V'^{1:N}$ are denoted by F_{1c}, F_{2c} and F_{2p}, respectively. The non-reused frozen symbols (see Fig. 3 and 4) in $U_1^{1:N}$ in the 1st and mth blocks are denoted by F_{11} and F_{1m} respectively, and those in $U'^{1:N}$ by F_{21} and F_{2m} respectively. The reused randomness at $\mathcal{H}_{X_2|Y_1X_1UV}^{(N)}$ in the channel prefixing scheme is denoted by W. For brevity, denote $F \triangleq \{F_{1c}, F_{2c}, F_{11}, F_{1m}, F_{21}, F_{2m}, F_{2p}\}$, $D_i \triangleq \{D_{1c,i}, D_{2c,i}\}$, $M^{i:m} \triangleq \{M_i, \ldots, M_m\}$, etc.

Lemma 3: For any $i \in [m]$, we have

$$I(M_i, E_i; \mathbf{Y}_{1,i}, D_i, F) \le O(N^2\sqrt{N\delta_N}). \tag{44}$$

Proof: See Appendix C. □

Lemma 4: For any $i \in [1, m-1]$,

$$\begin{aligned} &I(W; \mathbf{Y}_1^{i:m}, D^{i:m}, F | M^{i:m}, E_i) \\ &-I(E_{i+1}, W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, F | M^{i+1:m}) \le O(N^2\sqrt{N\delta_N}). \end{aligned} \tag{45}$$

Proof: See Appendix D. □

Lemma 5: For any $i \in [1, m-1]$, let

$$L_i = I(M^{i:m}, E_i, W; \mathbf{Y}_1^{i:m}, D^{i:m}, F). \tag{46}$$

Then we have

$$L_i - L_{i+1} \le O(N^2\sqrt{N\delta_N}). \tag{47}$$

Proof: See Appendix E. □

Suppose receiver 1 has perfect knowledge of the frozen symbols in each block. Then the information leakage is

$$\begin{aligned} L(N) &= I(M^{1:m}; \mathbf{Y}_1^{1:m}, D^{1:m}, F) \\ &\le I(M^{1:m}, \mathrm{E}_m, W; \mathbf{Y}_1^{1:m}, D^{1:m}, F). \end{aligned}$$

From the proof of Lemma 5 and the fact that receiver 1 has no knowledge about the secret seed we have $L_m \le O(N^{5/2}2^{-N^\beta/2})$. Thus, by induction hypothesis we have

$$L(N) \le \sum_{i=1}^{m-1}\big(L_i - L_{i+1}\big) + L_m \le O(mN^2\sqrt{N\delta_N}). \tag{48}$$

D. Achievable Rate Region

1) Randomness Rate: Since w_r is reused in a frame, the generating rate of randomness required by our channel prefixing scheme is

$$R_r = \frac{1}{mN}\big(|\mathcal{H}_{X_2|Y_1X_1UV}^{(N)}| + m|\mathcal{H}_{X_2|X_1UV}^{(N)} \setminus \mathcal{H}_{X_2|Y_1X_1UV}^{(N)}|\big). \tag{49}$$

From [39, Lemma 1] we have

$$\lim_{N\to\infty}\frac{1}{N}|\mathcal{H}_{X_2|Y_1X_1UV}^{(N)})^C \setminus \mathcal{L}_{X_2|Y_1X_1UV}^{(N)}| = 0. \tag{50}$$

Then we have

$$\begin{aligned} \lim_{N\to\infty,m\to\infty} R_r &= \lim_{N\to\infty}\frac{1}{N}|\mathcal{H}_{X_2|X_1UV}^{(N)} \cap (\mathcal{H}_{X_2|Y_1X_1UV}^{(N)})^C| \\ &= \lim_{N\to\infty}\frac{1}{N}|\mathcal{H}_{X_2|X_1UV}^{(N)} \cap \mathcal{L}_{X_2|Y_1X_1UV}^{(N)}| \\ &= I(X_2; Y_1|U, V, X_1). \end{aligned} \tag{51}$$

Thus, (17) is achievable with our proposed scheme.

2) Private and Confidential Message Rates: From Fig. 5 we can see that the private and confidential message rates in our proposed scheme are

$$R_2^{(p)} = \frac{1}{N}\big(|\mathcal{I}_{2p}| + |\mathcal{R}_2|\big), \quad R_2^{(s)} = \frac{1}{N}|\mathcal{I}_{2s}^{(1)}|, \tag{52}$$

respectively. By a similar analysis to the general wiretap polar code [24], we have

$$\lim_{N\to\infty} R_2^{(p)} = I(V; Y_1|U, X_1), \tag{53}$$

$$\lim_{N\to\infty} R_2^{(s)} = I(V; Y_2|U, X_1) - I(V; Y_1|U, X_1). \tag{54}$$

Thus, (15) is achieved. Since the difference between transmitter 2's private message and the randomness required by the encoder is just whether it carries information [9], from (51), (53) and the Markov chain of (11) we can see that (16) is achieved.

3) Common Message Rate: We first consider Case 1. In this case, the common message rates of the two transmitters in our proposed scheme are

$$\begin{aligned} R_1^{(1)} &= \frac{m|\mathcal{I}_{1c}^{(0)}| + (m-1)|\mathcal{I}_{1c}^{(1a)}|}{mN} = \frac{|\mathcal{I}_{1c}^{(1)}|}{N} - \frac{|\mathcal{I}_{1c}^{(1a)}|}{mN}, \\ R^{(c)} &= \frac{m|\mathcal{I}_{2c}^{(0)}| + (m-1)|\mathcal{I}_{2c}^{(1a)}|}{mN} = \frac{|\mathcal{I}_{2c}^{(1)}|}{N} - \frac{|\mathcal{I}_{2c}^{(1a)}|}{mN}, \end{aligned} \tag{55}$$

respectively. From (34) we have

$$\begin{aligned} \lim_{N\to\infty,m\to\infty} R_1^{(1)} &= I(X_1; Y_1), \\ \lim_{N\to\infty,m\to\infty} R^{(c)} &= I(U; Y_1|X_1). \end{aligned} \tag{56}$$

Since $\min\{I(U,X_1;Y_1), I(U,X_1;Y_2)\} = I(U,X_1;Y_1)$ in this case, if we allocate all of $R^{(c)}$ to transmitter 1's message, then (12) is achieved. No matter how we allocate two transmitters' common messages, the sum rate of all messages always achieves (14).

To prove the achievability of (13) in Case 1 requires some changes in the coding scheme. If we wish to maximize the sum rate of private and confidential messages, transmitter 2 will not help transmit M_1 at all. Therefore, whether receiver 1 can decode U does not matter. Then transmitter 2 can use all of $\mathcal{I}_{2c}^{(2)}$ to transmit its own message at any rate below $I(U;Y_2|X_1)$ (now this message becomes private message). Then from (53) and (54) we can see that (13) is achieved.

Another thing worth noting is that, in the above case, in order for both receivers to decode transmitter 1's message, $R_1 \leq \min\{I(X_1;Y_1), I(X_1;Y_2)\}$ must hold. Then the sum rate of all messages satisfies

$$R_1 + R_{2p} + R_{2s} \leq \min\{I(X_1;Y_1), I(X_1;Y_2)\} + I(U,V;Y_2|X_1),$$

which may seem to violate (14). However, since U in fact carries private message in this case, it is equivalent to remove auxiliary random variable U and simply design a code on V. We can also see this problem from the mutual information aspect. Due to the Markov chains of (10) and (11), we have

$$\begin{aligned} I(U,V;Y_2|X_1) &= I(V;Y_2|X_1) + I(U;Y_2|V,X_1) \\ &= I(V;Y_2|X_1). \end{aligned}$$

With auxiliary random variable U being removed, we can readily see that (14) still holds.

Next we consider Case 3-1. In this case, $R_1^{(1)}$ and $R^{(c)}$ are the same as in Case 1, thus (12) and (14) are achievable. As we have explained in Section IV-A.2, $\{u'^j\}_{j\in\mathcal{I}_{2c}^{(1b)}}$ must be assigned to transmitter 1's common message. Thus, in this case,

$$\begin{aligned} R_2^{(c)} &\leq I(U;Y_1|X_1) - \big(I(U;Y_1|X_1) - I(U;Y_2|X_1)\big) \\ &= I(U;Y_2|X_1). \end{aligned} \tag{57}$$

Then from (53), (54) and (57) we can see that (13) is achieved (R_{2p} in Theorem 2 is the sum of $R_2^{(c)}$ and $R_2^{(p)}$ here).

Since Case 2 (resp. 4) is similar to Case 1 (resp. 3), and Case 3-2 is similar to Case 3-1, we can now conclude that our proposed scheme achieves the whole region in Theorem 2 under the strong secrecy criterion with randomness constraint.

VI. Discussion

Although our proposed polar coding scheme is designed under secrecy constraints, it can be readily modified for the case without secrecy and achieve the capacity region of the CIC given by [8, Th. 4], since the capacity region is just a special case of the capacity-equivocation region when secrecy constraints are removed.

We note some relations between our work and [22], which considers polar coding for the broadcast channel with confidential messages. From Theorem 2 and [22, Th. 1] we can see that the rate region in [22, Th. 1] is a special case of that in Theorem 2 when transmitter 1 is removed. Also, as shown in [8], the region defined in Theorem 1 reduces to the capacity region of the MAC with degraded message sets if we set $Y_1 = Y_2$. Thus, our proposed scheme can be seen as a general solution for the aforementioned multi-user polar coding problems.

Appendix A
Proof of Lemma 1

Similar to the proof of [22, Lemma 5], we have

$$\begin{aligned} \mathbb{D}(P_{X_1^{1:N}} || P_{\tilde{X}_1^{1:N}}) &\leq N\delta_N, \\ \mathbb{D}(P_{U^{1:N}X_1^{1:N}} || P_{\tilde{U}^{1:N}\tilde{X}_1^{1:N}}) &\leq 2N\delta_N, \\ \mathbb{D}(P_{U^{1:N}V^{1:N}X_1^{1:N}} || P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}}) &\leq 3N\delta_N, \\ \mathbb{D}(P_{X_2^{1:N}V^{1:N}} || P_{\tilde{X}_2^{1:N}\tilde{V}^{1:N}}) &\leq 4N\delta_N. \end{aligned}$$

Then we have

$$\begin{aligned} &\| P_{U^{1:N}V^{1:N}X_1^{1:N}X_2^{1:N}Y_1^{1:N}Y_2^{1:N}} \\ &\quad - P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}\tilde{X}_2^{1:N}\tilde{Y}_1^{1:N}\tilde{Y}_2^{1:N}} \| \\ &= \| P_{X_2^{1:N}|V^{1:N}} P_{U^{1:N}V^{1:N}X_1^{1:N}} \\ &\quad - P_{\tilde{X}_2^{1:N}|\tilde{V}^{1:N}} P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}} \| && (58) \\ &\leq \| P_{X_2^{1:N}|V^{1:N}} P_{U^{1:N}V^{1:N}X_1^{1:N}} \\ &\qquad - P_{\tilde{X}_2^{1:N}|\tilde{V}^{1:N}} P_{U^{1:N}V^{1:N}X_1^{1:N}} \| \\ &\quad + \| P_{\tilde{X}_2^{1:N}|\tilde{V}^{1:N}} P_{U^{1:N}V^{1:N}X_1^{1:N}} \\ &\qquad - P_{\tilde{X}_2^{1:N}|\tilde{V}^{1:N}} P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}} \| && (59) \\ &= \| P_{X_2^{1:N}|V^{1:N}} P_{V^{1:N}} - P_{\tilde{X}_2^{1:N}|\tilde{V}^{1:N}} P_{V^{1:N}} \| \\ &\quad + \| P_{U^{1:N}V^{1:N}X_1^{1:N}} - P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}} \| && (60) \\ &\leq \| P_{X_2^{1:N}V^{1:N}} - P_{\tilde{X}_2^{1:N}\tilde{V}^{1:N}} \| \\ &\quad + \| P_{\tilde{X}_2^{1:N}\tilde{V}^{1:N}} - P_{\tilde{X}_2^{1:N}|\tilde{V}^{1:N}} P_{V^{1:N}} \| \\ &\quad + \| P_{U^{1:N}V^{1:N}X_1^{1:N}} - P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}} \| && (61) \\ &\leq \| P_{X_2^{1:N}V^{1:N}} - P_{\tilde{X}_2^{1:N}\tilde{V}^{1:N}} \| + \| P_{V^{1:N}} - P_{\tilde{V}^{1:N}} \| \\ &\quad + \| P_{U^{1:N}V^{1:N}X_1^{1:N}} - P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}} \| && (62) \\ &\leq \sqrt{2\log 2}\sqrt{N\delta_N}(2+2\sqrt{3}), && (63) \end{aligned}$$

where (58), (60) and (62) hold by [40, Lemma 17], and (59) and (61) hold by the triangle inequality.

Appendix B
Proof of Lemma 2

Denote receiver 1's error probability when decoding message $M_1^{(1)}$ in block i by $P_{e1,i}^{(1)}$, and that when decoding $(M_1^{(2)}, M_2^{(c)})$ by $P_{e1,i}^{(2)}$. For $i = [2,m]$, define the following error events

$$\begin{aligned} \mathcal{E}_{X_1Y_1,i} &\triangleq \{(X_1^{1:N}Y_1^{1:N}) \neq (\tilde{X}_1^{1:N}\tilde{Y}_1^{1:N})_i\}, \\ \mathcal{E}_{X_1,i-1}^{ch} &\triangleq \{(\bar{U}_1^{chaining})_{i-1} \neq (\tilde{U}_1^{chaining})_{i-1}\}, \\ \mathcal{E}_{U,i-1}^{ch} &\triangleq \{(\bar{U}'^{chaining})_{i-1} \neq (\tilde{U}'^{chaining})_{i-1}\}, \\ \mathcal{E}_{X_1,i} &\triangleq \{(\bar{X}_1^{1:N})_i \neq (\tilde{X}_1^{1:N})_i\}, \\ \mathcal{E}_i &\triangleq \mathcal{E}_{X_1Y_1,i} \cup \mathcal{E}_{X_1,i-1}^{ch} \cup \mathcal{E}_{U,i-1}^{ch}, \\ \mathcal{E}_i' &\triangleq \mathcal{E}_{X_1Y_1,i} \cup \mathcal{E}_{X_1,i-1}^{ch} \cup \mathcal{E}_{U,i-1}^{ch} \cup \mathcal{E}_{X_1,i}, \end{aligned}$$

where $(\cdot)_i$ denotes vectors in block i, $\bar{U}$ denotes the decoding result of U, and "chaining" in the superscript stands for the elements used for chaining. Using optimal coupling [41, Lemma 3.6] we have

$$P[\mathcal{E}_{X_1Y_1,i}] = \| P_{X_1^{1:N}Y_1^{1:N}} - P_{\tilde{X}_1^{1:N}\tilde{Y}_1^{1:N}} \| .$$

Then we have

$$\begin{aligned}
P_{e1,i}^{(1)} &\le P[(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i] \\
&= P[(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i|\mathcal{E}_i]P[\mathcal{E}_i] \\
&\quad + P[(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i|\mathcal{E}_i^C]P[\mathcal{E}_i^C] \\
&\le P[\mathcal{E}_i] + P[(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i|\mathcal{E}_i^C] \\
&\le P(\mathcal{E}_{X_1Y_1,i}) + P(\mathcal{E}_{X_1,i-1}^{ch}) + P(\mathcal{E}_{U,i-1}^{ch}) \\
&\quad + P[(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i|\mathcal{E}_i^C] \\
&\le \delta_N^c + N\delta_N + P[(\bar{X}_1^{1:N})_{i-1} \ne (\tilde{X}_1^{1:N})_{i-1}] \\
&\quad + P[(\bar{U}^{1:N})_{i-1} \ne (\tilde{U}^{1:N})_{i-1}], \qquad (64)
\end{aligned}$$

where (64) holds from (63) and the error probability of source polar coding [42]. Similarly we have

$$\begin{aligned}
P_{e1,i}^{(2)} &\le P[(\bar{U}^{1:N})_i \ne (\tilde{U}^{1:N})_i] \\
&\le P(\mathcal{E}_{X_1Y_1,i}) + P(\mathcal{E}_{X_1,i-1}^{ch}) + P(\mathcal{E}_{U,i-1}^{ch}) + P(\mathcal{E}_{X_1,i}) \\
&\quad + P[(\bar{U}^{1:N})_i \ne (\tilde{U}^{1:N})_i|\mathcal{E}_i'^C] \\
&\le \delta_N^c + N\delta_N + P[(\bar{X}_1^{1:N})_{i-1} \ne (\tilde{X}_1^{1:N})_{i-1}] \\
&\quad + P[(\bar{U}^{1:N})_{i-1} \ne (\tilde{U}^{1:N})_{i-1}] + P[(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i]. \qquad (65)
\end{aligned}$$

From (64) and (65) we have

$$\begin{aligned}
&P[(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i] + P[(\bar{U}^{1:N})_i \ne (\tilde{U}^{1:N})_i] \\
&\quad \le 3\Big(\delta_N^c + N\delta_N + P[(\bar{X}_1^{1:N})_{i-1} \ne (\tilde{X}_1^{1:N})_{i-1}] \\
&\qquad + P[(\bar{U}^{1:N})_{i-1} \ne (\tilde{U}^{1:N})_{i-1}]\Big),
\end{aligned}$$

By induction we have

$$\begin{aligned}
&P[(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i] + P[(\bar{U}^{1:N})_i \ne (\tilde{U}^{1:N})_i] \\
&\quad \le \sum_{k=1}^{i-1} 3^k(\delta_N^c + N\delta_N) + 3^{i-1}P[(\bar{X}_1^{1:N})_1 \ne (\tilde{X}_1^{1:N})_1] \\
&\qquad + 3^{i-1}P[(\bar{U}^{1:N})_1 \ne (\tilde{U}^{1:N})_1]. \qquad (66)
\end{aligned}$$

From the above analysis and the assumption that receivers have perfect knowledge of frozen symbols we have

$$\begin{aligned}
&P[(\bar{X}_1^{1:N})_1 \ne (\tilde{X}_1^{1:N})_1] + P[(\bar{U}^{1:N})_1 \ne (\tilde{U}^{1:N})_1] \\
&\qquad \le 3(\delta_N^c + N\delta_N).
\end{aligned}$$

Thus, the overall error probability of receiver 1 in a frame can be upper bounded by

$$\begin{aligned}
P_{e1} &\le \sum_{k=1}^{m} \left(P_{e1,k}^{(1)} + P_{e1,k}^{(2)}\right) \\
&\le \sum_{i'=1}^{m}\sum_{k=1}^{i'} 3^k(\delta_N^c + N\delta_N) \\
&= O(3^m\sqrt{N\delta_N}). \qquad (67)
\end{aligned}$$

For receiver 2, the error probability of decoding common messages in a frame can be similarly upper bounded by

$$P_{e2}^{(c)} \le O(3^m\sqrt{N\delta_N}). \qquad (68)$$

To estimate receiver 2's error probability in decoding its private and confidential messages in block i, $P_{e2,i}^{(p,s)}$, we define the following error events:

$$\begin{aligned}
\mathcal{E}_{UVX_1Y_2,i} &\triangleq \{(U^{1:N}X_1^{1:N}V^{1:N}Y_2^{1:N}) \\
&\qquad \ne (\tilde{U}^{1:N}\tilde{X}_1^{1:N}\tilde{V}^{1:N}\tilde{Y}_2^{1:N})_i\}, \\
\mathcal{E}_{U,i} &\triangleq \{(\bar{U}^{1:N})_i \ne (\tilde{U}^{1:N})_i\}, \\
\mathcal{E}_{X_1,i} &\triangleq \{(\bar{X}_1^{1:N})_i \ne (\tilde{X}_1^{1:N})_i\}, \\
\mathcal{E}_{V,i+1} &\triangleq \{(\bar{V}^{1:N})_{i+1} \ne (\tilde{V}^{1:N})_{i+1}\}, \\
\mathcal{E}_i &\triangleq \mathcal{E}_{UVX_1Y_2,i} \cup \mathcal{E}_{U,i} \cup \mathcal{E}_{X_1,i} \cup \mathcal{E}_{V,i+1}.
\end{aligned}$$

Using optimal coupling [41, Lemma 3.6] we have

$$\begin{aligned}
&P[\mathcal{E}_{UVX_1Y_2,i}] \\
&\qquad = \| P_{U^{1:N}V^{1:N}X_1^{1:N}Y_1^{1:N}} - P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}\tilde{Y}_1^{1:N}} \| .
\end{aligned}$$

Similar to the analysis for common message decoding, $P_{e2,i}^{(p,s)}$ can be upper bounded by

$$\begin{aligned}
P_{e2,i}^{(p,s)} &\le P[\mathcal{E}_{V,i}] \\
&\le \delta_N^c + N\delta_N + P[\mathcal{E}_{U,i}] + P[\mathcal{E}_{X_1,i}] + P[\mathcal{E}_{V,i+1}] \\
&\le \sum_{k=i}^{m}(3^{m-k}+1)(\delta_N^c + N\delta_N) + P[\mathcal{E}_{V,i+1}].
\end{aligned}$$

By induction and the counterpart of (66) for receiver 2 we have

$$P_{e2,i}^{(p,s)} \le \textstyle\sum_{i'=i}^{m}\sum_{k=i'}^{m}(3^{m-k}+1)(\delta_N^c + N\delta_N)). \qquad (69)$$

Then

$$P_{e2}^{(p,s)} \le \sum_{k=1}^{m} P_{e2,k}^{(p,s)} = O(m3^m\sqrt{N\delta_N}). \qquad (70)$$

APPENDIX C
PROOF OF LEMMA 3

Let $t = |\mathcal{I}_{2s}| + |\mathcal{I}_{2p}|$ and $w = |\mathcal{F}_2|$. Denote $\{a_1, a_2, \ldots, a_t\} = \mathcal{I}_{2s}$ with $a_1 < \ldots < a_t$, $\{b_1, b_2, \ldots, b_w\} = \mathcal{F}_2$ with $b_1 < \ldots < b_w$, and $\{c_1, c_2, \ldots, c_{t+w}\} = \{a_1, \ldots, a_t, b_1, \ldots, b_w\}$ with $c_1 < \ldots < c_{t+w}$. Let F_c be short for $\{F_{1c}, F_{2c}, F_{11}, F_{1m}, F_{21}, F_{2m}\}$. Then we have

$$\begin{aligned}
&I(M_i, E_i; \mathbf{Y}_{1,i}, D_i, F) \\
&= H_{q_V}(M_i, E_i) - H_{q_V}(M_i, E_i|\mathbf{Y}_{1,i}, D_i, F) \\
&= H_{q_V}(M_i, E_i) - H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, D_i, F_c) \\
&\quad + H_{q_V}(F_{2p}|\mathbf{Y}_{1,i}, D_i, F_c) \\
&\le t + w - H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, \mathbf{X}_{1,i}, \mathbf{U}_i) \qquad (71) \\
&= t + w - \sum_{j=1}^{t+w} H_{q_V}(\tilde{V}'^{c_j}|\tilde{Y}_1^{1:N}, \tilde{X}_1^{1:N}, \tilde{U}^{1:N}, \tilde{V}'^{\{c_1,\ldots,c_{j-1}\}}) \\
&\le \sum_{j=1}^{t+w}\left(1 - H_{q_V}(\tilde{V}'^{c_j}|\tilde{Y}_1^{1:N}, \tilde{X}_1^{1:N}, \tilde{U}^{1:N}, \tilde{V}'^{1:c_{j-1}})\right), \qquad (72)
\end{aligned}$$

where (71) holds because

$$H_{q_V}(M_i, E_i) \le t, \quad H_{q_V}(F_{2p}|\mathbf{Y}_{1,i}, D_i, F_c) \le w,$$

and

$$H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, D_i, F_c) \ge H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, \mathbf{X}_{1,i}, \mathbf{U}_i), \quad (73)$$

which is shown in more details as follows. For $i = 1$,

$$H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, D_i, F_c) = H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, D_i, F_{1c}, F_{2c}, F_{11}, F_{21})$$

because (F_{1m}, F_{2m}) is independent of the rest items in the left-hand-side of (73). Similarly, for $i = m$, we have

$$H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, D_i, F_c) = H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, D_i, F_{1c}, F_{2c}, F_{1m}, F_{2m}).$$

For $1 < i < m$, we have

$$H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, D_i, F_c) = H_{q_V}(M_i, E_i, F_{2p}|\mathbf{Y}_{1,i}, D_i, F_{1c}, F_{2c})$$

Thus, (73) always holds.

Note that the entropies above are calculated under the induced distribution by the encoding scheme. Under the target distribution $P_{U^{1:N}V^{1:N}X_1^{1:N}Y_1^{1:N}}$, from (39) we have

$$H_{q_V}(V'^{c_j}|Y_1^{1:N}, X_1^{1:N}, U^{1:N}, V'^{1:c_{j-1}}) \ge 1 - \delta_N. \quad (74)$$

From [43, Theorem 17.3.3] we have

$$\begin{aligned}
&|H_{q_V}(\tilde{V}'^{c_j}|\tilde{Y}_1^{1:N}, \tilde{X}_1^{1:N}, \tilde{U}^{1:N}, \tilde{V}'^{1:c_{j-1}}) \\
&\quad - H_{q_V}(V'^{c_j}|Y_1^{1:N}, X_1^{1:N}, U^{1:N}, V'^{1:c_{j-1}})| \\
&\le \| P_{U^{1:N}V^{1:N}X_1^{1:N}Y_1^{1:N}} - P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}\tilde{Y}_1^{1:N}} \| \\
&\quad \times \log \frac{|\mathcal{U}|^N|\mathcal{V}|^N|\mathcal{X}_1|^N|\mathcal{Y}_1|^N}{\| P_{U^{1:N}V^{1:N}X_1^{1:N}Y_1^{1:N}} - P_{\tilde{U}^{1:N}\tilde{V}^{1:N}\tilde{X}_1^{1:N}\tilde{Y}_1^{1:N}} \|} \\
&= O(N\sqrt{N\delta_N}). \quad (75)
\end{aligned}$$

From (72), (74) and (75) we have

$$I(M_i, E_i; \mathbf{Y}_{1,i}, D_i, F) \le O(N^2\sqrt{N\delta_N}). \quad (76)$$

APPENDIX D
PROOF OF LEMMA 4

To prove Lemma 4, we first prove the following two lemmas.

Lemma 6: For any $i \in [1, m]$,

$$I(E_i; W|\mathbf{Y}_1^{i:m}, D^{i:m}, M^{i:m}) \le O(N^2\sqrt{N\delta_N}). \quad (77)$$

Proof: Since E_i is independent of $(\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m})$, we have

$$\begin{aligned}
&I(E_i; W|\mathbf{Y}_1^{i:m}, D^{i:m}, M^{i:m}) \\
&= I(E_i; W|\mathbf{Y}_{1,i}, D_i, M_i) \\
&= H_{q_{X_2}}(W|\mathbf{Y}_{1,i}, D_i, M_i) - H_{q_{X_2}}(W|\mathbf{Y}_{1,i}, D_i, M_i, E_i) \\
&\le H_{q_{X_2}}(W) - H_{q_{X_2}}(W|\mathbf{Y}_{1,i}, \mathbf{X}_{1,i}, \mathbf{U}_i, \mathbf{V}_i).
\end{aligned}$$

Then we can prove (77) similar to the proof of Lemma 3. □

Lemma 7: For any $i \in [1, m-1]$,

$$\begin{aligned}
&I(\mathbf{Y}_{1,i}, D_i, F, M_i, E_i; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}|W) \\
&\quad - I(E_{i+1}, W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, F, M^{i+1:m}) \\
&\quad - I(\mathbf{Y}_{1,i}, D_i, F, M_i, E_i; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad + I(W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\le O(N^2\sqrt{N\delta_N}).
\end{aligned}$$

Proof:

$$\begin{aligned}
&I(\mathbf{Y}_{1,i}, D_i, F, M_i, E_i; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}|W) \\
&\quad - I(E_{i+1}, W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, F, M^{i+1:m}) \\
&\le I(\mathbf{Y}_{1,i}, D_i, F, M_i, E_i, E_{i+1}; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}|W) \\
&\quad - I(E_{i+1}, W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, F, M^{i+1:m}) \\
&= I(\mathbf{Y}_{1,i}, D_i, M_i, E_i; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}|W, F, E_{i+1}) \\
&\quad + I(F, E_{i+1}; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}|W) \\
&\quad - I(E_{i+1}, W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, F, M^{i+1:m}) \\
&= I(F, E_{i+1}; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}, W) \\
&\quad - I(E_{i+1}, W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, F, M^{i+1:m}) \quad (78) \\
&= I(F, E_{i+1}; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad + I(F, E_{i+1}; W|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(E_{i+1}, W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(E_{i+1}, W; F|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&= I(F; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}|E_{i+1}) \\
&\quad - I(W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}|E_{i+1}) \\
&\quad + I(F, E_{i+1}; W|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(E_{i+1}, W; F|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&= I(F; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}, E_{i+1}) \\
&\quad - I(W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}, E_{i+1}) \\
&\quad + I(F, E_{i+1}; W|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(E_{i+1}, W; F|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \quad (79) \\
&= I(F; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad + I(F; E_{i+1}|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}, E_{i+1}) \\
&\quad + I(F, E_{i+1}; W|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(E_{i+1}, W; F|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&= I(F; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}, E_{i+1}) \\
&\quad + I(E_{i+1}; W|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(E_{i+1}; F|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \quad (80) \\
&\le I(F; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \\
&\quad - I(W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}, E_{i+1}), \\
&\quad + I(E_{i+1}; W|\mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m})
\end{aligned}$$

where (78) holds because block i and the next $m - i$ blocks are independent conditioned on (W, F, E_{i+1}) and W is independent of (F, E_{i+1}), (79) and (80) hold because E_{i+1} and (F, W) are independent.

Since

$$I(F; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}) \le I(\mathbf{Y}_{1,i}, D_i, F, M_i, E_i; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m})$$

and

$$I(W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}, E_{i+1}) \ge I(W; \mathbf{Y}_1^{i+1:m}, D^{i+1:m}, M^{i+1:m}),$$

by Lemma 6 we can readily prove Lemma 7. □

Now we can prove Lemma 4. Since W, M_i and E_i are independent of one another, we have

$$\begin{aligned}
&I(W;\mathbf{Y}_1^{i:m},D^{i:m},F|M^{i:m},E_i)\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F|M^{i+1:m})\\
&=I(W;\mathbf{Y}_1^{i:m},D^{i:m},F,M^{i:m},E_i)\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F,M^{i+1:m})\\
&=I(W;\mathbf{Y}_{1,i},D_i,F,M_i,E_i)\\
&\quad +I(W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},M^{i+1:m}|\mathbf{Y}_{1,i},D_i,F,M_i,E_i)\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F,M^{i+1:m})\\
&=I(W;\mathbf{Y}_{1,i},D_i,F,M_i,E_i)+I(W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},M^{i+1:m})\\
&\quad +I(\mathbf{Y}_{1,i},D_i,F,M_i,E_i;\mathbf{Y}_1^{i+1:m},D^{i+1:m},M^{i+1:m}|W)\\
&\quad -I(\mathbf{Y}_{1,i},D_i,F,M_i,E_i;\mathbf{Y}_1^{i+1:m},D^{i+1:m},M^{i+1:m})\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F,M^{i+1:m})\\
&\le I(W;\mathbf{Y}_{1,i},D_i,F,M_i,E_i)+O(N^{5/2}2^{-N^\beta/2}),
\end{aligned}\tag{81}$$

where (81) holds due to Lemma 7. Similarly to the proof of Lemma 3, we have

$$\begin{aligned}
I(W;\mathbf{Y}_{1,i},D_i,F,M_i,E_i)&\le I(W;\mathbf{Y}_{1,i},\mathbf{X}_{1,i},\mathbf{U}_i,\mathbf{V}_i)\\
&\le O(N^2\sqrt{N\delta_N}).
\end{aligned}$$

This proves Lemma 4.

Appendix E
Proof of Lemma 5

$$\begin{aligned}
&L_i-L_{i+1}\\
&=I(M_i,E_i,W;\mathbf{Y}_1^{i:m},D^{i:m},F|M^{i+1:m})\\
&\quad +I(M^{i+1:m};\mathbf{Y}_{1,i},D_i|\mathbf{Y}_1^{i+1:m},D^{i+1:m},F)\\
&\quad +I(M^{i+1:m};\mathbf{Y}_1^{i+1:m},D^{i+1:m},F)\\
&\quad -I(M^{i+1:m},E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F)\\
&=I(M_i,E_i,W;\mathbf{Y}_1^{i:m},D^{i:m},F|M^{i+1:m})\\
&\quad +I(M^{i+1:m};\mathbf{Y}_{1,i},D_i|\mathbf{Y}_1^{i+1:m},D^{i+1:m},F)\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F|M^{i+1:m})\\
&=I(M_i,E_i,W;\mathbf{Y}_1^{i:m},D^{i:m},F|M^{i+1:m})\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F|M^{i+1:m})\qquad(82)\\
&=I(M_i,E_i;\mathbf{Y}_1^{i:m},D^{i:m},F|M^{i+1:m})\\
&\quad +I(W;\mathbf{Y}_1^{i:m},D^{i:m},F|M^{i:m},E_i)\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F|M^{i+1:m})\\
&=I(M_i,E_i;\mathbf{Y}_{1,i},D_i,F|M^{i+1:m})\\
&\quad +I(W;\mathbf{Y}_1^{i:m},D^{i:m},F|M^{i:m},E_i)\\
&\quad +I(M_i,E_i;\mathbf{Y}_1^{i+1:m},D^{i+1:m}|M^{i+1:m},\mathbf{Y}_{1,i},D_i,F)\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F|M^{i+1:m})\\
&=I(M_i,E_i;\mathbf{Y}_{1,i},D_i,F)+I(W;\mathbf{Y}_1^{i:m},D^{i:m},F|M^{i:m},E_i)\\
&\quad -I(E_{i+1},W;\mathbf{Y}_1^{i+1:m},D^{i+1:m},F|M^{i+1:m}),\qquad(83)
\end{aligned}$$

where (82) holds due to the independence between $M^{i+1:m}$ and $(\mathbf{Y}_{1,i},D_i)$, and (83) holds because (M_i,E_i) and $(\mathbf{Y}_1^{i+1:m},D^{i+1:m})$ are independent, and $M^{i+1:m}$ is independent of both (M_i,E_i) and $(\mathbf{Y}_{1,i},D_i,F)$, thus $I(M_i,E_i;\mathbf{Y}_{1,i},D_i,F|M^{i+1:m}) = I(M_i,E_i;\mathbf{Y}_{1,i},D_i,F)$. Then by Lemma 3 and 4 we have

$$L_i-L_{i+1}\le O(N^2\sqrt{N\delta_N}).\tag{84}$$

References

[1] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," Ph.D. dissertation, Roy. Inst. Technol., Stockholm, Sweden, 2000.

[2] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[3] N. Devroye, P. Mitran, and V. Tarokh, "Limits on communications in a cognitive radio channel," *IEEE Commun. Mag.*, vol. 44, no. 6, pp. 44–49, Jun. 2006.

[4] W. Wu, S. Vishwanath, and A. Arapostathis, "Capacity of a class of cognitive radio channels: Interference channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4391–4399, Nov. 2007.

[5] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3945–3958, Sep. 2009.

[6] J. Jiang, Y. Xin, and H. K. Garg, "Interference channels with common information," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 171–187, Jan. 2008.

[7] S. Rini, D. Tuninetti, and N. Devroye, "New inner and outer bounds for the memoryless cognitive interference channel and some new capacity results," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4087–4109, Jul. 2011.

[8] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai (Shitz), and S. Verdú, "Capacity of cognitive interference channels with and without secrecy," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 604–619, Feb. 2009.

[9] S. Watanabe and Y. Oohama, "Cognitive interference channels with confidential messages under randomness constraint," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7698–7707, Dec. 2014.

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

[11] E. Arıkan, "Polar coding for the slepian-wolf problem based on monotone chain rules," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 566–570.

[12] E. Şaşoğlu, E. Telatar, and E. Yeh, "Polar codes for the two-user multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6583–6592, Oct. 2013.

[13] E. Abbe and I. Telatar, "Polar codes for the m-user multiple access channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5437–5448, Aug. 2012.

[14] H. Mahdavifar, M. El-Khamy, J. Lee, and I. Kang, "Achieving the uniform rate region of general multiple access channels by polar coding," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 467–478, Feb. 2016.

[15] N. Goela, E. Abbe, and M. Gastpar, "Polar codes for broadcast channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 758–782, Feb. 2015.

[16] M. Mondelli, S. H. Hassani, I. Sason, and R. L. Urbanke, "Achieving Marton's region for broadcast channels using polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 783–800, Feb. 2015.

[17] L. Wang, "Channel coding techniques for network communication," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. California, San Diego, San Diego, CA, USA, 2015.

[18] M. Zheng, C. Ling, W. Chen, and M. Tao, "A new polar coding scheme for the interference channel," *CoRR*, vol. abs/1608.08742, Aug. 2016. [Online]. Available: http://arxiv.org/abs/1608.08742

[19] 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, Oct. 2011.

[20] O. O. Koyluoglu and H. El Gamal, "Polar coding for secure transmission and key agreement," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 5, pp. 1472–1483, Oct. 2012.

[21] E. Şaşoğlu and A. Vardy, "A new polar coding scheme for strong security on wiretap channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 1117–1121.

[22] R. A. Chou and M. R. Bloch, " Polar coding for the broadcast channel with confidential messages: A random binning analogy," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2410–2429, May 2016.

[23] Y.-P. Wei and S. Ulukus, "Polar coding for the general wiretap channel with extensions to multiuser scenarios," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, pp. 278–291, Feb. 2016.
[24] T. C. Gulcu and A. Barg, "Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1311–1324, Feb. 2017.
[25] R. A. Chou and A. Yener, "Polar coding for the multiple access wiretap channel via rate-splitting and cooperative jamming," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 983–987.
[26] M. Zheng, M. Tao, W. Chen, and C. Ling, "Secure polar coding for the two-way wiretap channel," *CoRR*, vol. abs/1612.00130, Dec. 2016. [Online]. Available: http://arxiv.org/abs/1612.00130
[27] M. Andersson, R. F. Schaefer, T. J. Oechtering, and M. Skoglund, "Polar coding for bidirectional broadcast channels with common and confidential messages," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1901–1908, Sep. 2013.
[28] R. Mori and T. Tanaka, "Performance of polar codes with the construction using density evolution," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 519–521, Jul. 2009.
[29] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6562–6582, Oct. 2013.
[30] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.
[31] M. Mondelli, R. Urbanke, and S. H. Hassani, "How to achieve the capacity of asymmetric channels," in *Proc. 52nd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep./Oct. 2014, pp. 789–796.
[32] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric models," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, Dec. 2013.
[33] R. A. Chou and M. R. Bloch, "Using deterministic decisions for low-entropy bits in the encoding and decoding of polar codes," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep./Oct. 2015, pp. 1380–1385.
[34] E. Şaşoğlu, "Polar coding theorems for discrete systems," Ph.D. dissertation, École Polytechn. Féd. Lausanne, Lausanne, Switzerland, 2011.
[35] R. Nasser and E. Telatar, "Polar codes for arbitrary DMCs and arbitrary MACs," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 2917–2936, Jun. 2016.
[36] R. Nasser, "An ergodic theory of binary operations—Part I: Key properties," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 6931–6952, Dec. 2016.
[37] R. Nasser, "An ergodic theory of binary operations—Part II: Applications to polarization," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1063–1083, Feb. 2017.
[38] S. H. Hassani and R. Urbanke, "Universal polar codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun./Jul. 2014, pp. 1451–1455.
[39] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6213–6237, Nov. 2015.
[40] P. W. Cuff, "Communication in networks for coordinating behavior," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, 2009.
[41] D. Aldous, "Random walks on finite groups and rapidly mixing Markov chains," in *Séminaire de Probabilités XVII 1981/82*. Berlin, Germany: Springer, 1983, pp. 243–297.
[42] E. Arikan, "Source polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 899–903.
[43] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.

Mengfan Zheng received the B.S. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2012, where he is currently pursuing the Ph.D. degree in electronic engineering. His research interests include channel coding, information theory, and physical layer security.

Wen Chen received the B.S. and M.S. degrees from Wuhan University, China, in 1990 and 1993, respectively, and the Ph.D. degree from the University of Electro-Communications, Tokyo, Japan, in 1999. In 2001, he joined the University of Alberta, Edmonton, AB, Canada, as a Post-Doctoral Fellow and then a Research Associate. Since 2006, he has been a Full Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China, where he is also the Director of the Institute for Signal Processing and Systems. From 2014 to 2015, he was the Dean of the School of Electronics Engineering and Automations, Guilin University of Electronic Technology. He was a Research Fellow of the Japan Society for the Promotion of Sciences from 1999 to 2001. Since 2016, he has been the Chairman of the SJTU Intellectual Property Management Corporation.

Dr. Chen received the Ariyama Memorial Research Prize in 1997, the PIMS Post-Doctoral Fellowship in 2001. He received the honors of the New Century Excellent Scholar in China in 2006 and the Pujiang Excellent Scholar in Shanghai in 2007. He is the Vice General Secretary of the Shanghai Institute of Electronics in 2008. He received the Best Service Award of the China Institute of Electronics in 2013, the Best Paper Awards of the Chinese Information Theory Society in 2014, the Innovate 5G Competition Award in 2015.He is invited to deliver a keynote speech in the IEEE APCC 2016, and tutorials in the IEEE ICCC 2016 and the IEEE VTC 2017.

He has authored 89 papers in IEEE journals and over 110 papers in IEEE conferences. His research interests include multiple access, coded cooperation, and green heterogenous networks. He is the Chair of the IEEE Vehicular Technology Society Shanghai Chapter. He is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON COMMUNICATIONS, and the IEEE ACCESS.

Cong Ling received the B.S. and M.S. degrees in electrical engineering from the Nanjing Institute of Communications Engineering, Nanjing, China, in 1995 and 1997, respectively, and the Ph.D. degree in electrical engineering from Nanyang Technological University, Singapore, in 2005.

He had been on the Faculties of the Nanjing Institute of Communications Engineering and King's College London. He is currently a Senior Lecturer with the Electrical and Electronic Engineering Department, Imperial College London. His research interests include coding, signal processing, and security, especially lattices.

Dr. Ling is currently an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS. He has also served as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.