Polar Coding for the General Wiretap Channel With Extensions to Multiuser Scenarios

Yi-Peng Wei, *Student Member, IEEE*, and Sennur Ulukus, *Fellow, IEEE*

***Abstract*—Information-theoretic work for wiretap channels is mostly based on random coding schemes. Designing practical coding schemes to achieve information-theoretic secrecy is an important problem. By applying two recently developed techniques for polar codes, namely, universal polar coding and polar coding for asymmetric channels, we propose a polar coding scheme to achieve the secrecy capacity of the general wiretap channel. We then apply this coding scheme to achieve the best-known inner bounds for the multiple access wiretap channel (MAC-WTC), and the broadcast and interference channels with confidential messages (BC-CM and IC-CM).**

***Index Terms*—Wiretap channel, broadcast channel with confidential messages, interference channel with confidential messages, multiple access wiretap channel, universal polar coding, chaining construction.**

I. Introduction

THE WIRETAP channel was first introduced by Wyner [1], in which a legitimate transmitter (Alice) wishes to send messages to a legitimate receiver (Bob) secretly in the presence of an eavesdropper (Eve). Wyner [1] characterized the capacity equivocation region for the degraded wiretap channel, in which the received signal at Eve is a degraded version of the received signal at Bob. Later, Csiszár and Körner [2] characterized the capacity equivocation region for general, not necessarily degraded, wiretap channels. These works are based on information-theoretic random coding schemes.

Polar coding, invented by Arıkan [3], is the first code that provably achieves the capacity of the binary-input discrete symmetric output channels (B-DMC). The idea of polar coding has been extended to lossless source coding [4], lossy source coding [5], and to multi-user scenarios, such as, multiple access channel [6]–[8], broadcast channel [9], [10], interference channel [11], and Slepian-Wolf coding problem [12].

On a B-DMC, polarization results in two kinds of sub-channels [3][1]. The first kind is good sub-channels. The capacity for these sub-channels approaches 1 bit per channel use. The second kind is bad sub-channels. The channel output for these sub-channels is independent of the channel input; therefore the capacity for these sub-channels approaches 0. In particular, if a B-DMC A is degraded with respect to a B-DMC B, then the good sub-channels of A must be a subset of the good sub-channels of B [13]. We call this the *subset property*.

Polar coding schemes for *degraded* wiretap channels with *symmetric* main and eavesdropper channels are developed using the subset property in [14]–[17]. For degraded wiretap channels, the good sub-channels of Eve is a subset of the good sub-channels of Bob. The polar coding scheme is designed to transmit the confusion messages (random bits) on the sub-channels simultaneously good for Bob and Eve, and to transmit the secret messages on the sub-channels only good for Bob. However, for non-degraded wiretap channels, the subset property no longer holds [18]–[22], i.e., the good sub-channels of Eve is not necessary a subset of the good sub-channels of Bob. Moreover, the secrecy capacity achieving input distribution is not necessarily a uniform distribution. Therefore, the polar coding schemes in [14]–[17] cannot directly extend to the non-degraded wiretap channel.

By applying two recently developed techniques for polar codes, we can achieve the secrecy capacity of the general wiretap channel. The first technique is *universal polar codes* [21], [22]. Universal polar coding allows us to align the good sub-channels of Bob and Eve together. Therefore, we can artificially construct the subset property for the non-degraded wiretap channel. Then, Alice transmits the random bits on the sub-channels simultaneously good for Bob and Eve, and the secret message on the sub-channels only good for Bob. The second technique is *polar coding for asymmetric models* [23], which allows us to deal with the non-uniform input distribution. Different from B-DMC, polarization for asymmetric channels results in three different kinds of sub-channels.

Another polar coding scheme for the general wiretap channel is provided in [24], which uses a concatenated code consisting of two polar codes. The inner layer ensures that the transmitted message can be reliably decoded by Bob, and the outer layer guarantees that the message is kept secret from Eve. Our work jointly handles these two goals in one shot. Hence, the decoding error probability of our scheme is approximately $O(2^{-n^{1/2}})$, whereas it is $O(\sqrt{n}2^{-n^{1/4}})$ in [24]. Although the scheme in [24] does not require to share randomness, for practical code construction, there is still no efficient way to characterize the outer index set [24, Sec. III. C.], while our coding scheme can be efficiently constructed by [19].

Next, we extend our coding scheme to several multiuser scenarios: multiple access wiretap channel (MAC-WTC) [25],

Manuscript received March 25, 2015; revised September 5, 2015; accepted October 23, 2015. Date of publication December 3, 2015; date of current version January 14, 2016. This work was supported by the NSF under Grant CNS 13-14733, Grant CCF 14-22111, and Grant CCF 14-22129. This paper was presented in part at the IEEE ITW, Jerusalem, Israel, April 2015.

The authors are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: ypwei@umd.edu; ulukus@umd.edu).

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.2015.2504275

[1] Here, and later in the discussion of asymmetric channels, we neglect partially polarized sub-channels, which are of order $o(n)$.

[26], broadcast channel with confidential messages (BC-CM) [27], and interference channel with confidential messages (IC-CM) [27]. We are motivated by wireless communications scenarios for these extensions. Wireless communications environment is naturally a multi-user environment, where multiple users access the channel simultaneously in transmitting data, and the signal is received simultaneously by multiple receivers. In addition, the wireless environment is particularly susceptible to eavesdropping attacks [28]–[30] due to its inherent openness. The three models considered represent the most basic network structures with multiple transmitters and receivers. In the MAC-WTC, two transmitters wish to send independent messages to the legitimate receiver in the presence of an eavesdropper. In the BC-CM[2], the transmitter wishes to send independent messages to two receivers, while keeping the messages secret from the unintended receiver. In the IC-CM, two transmitters wish to send independent messages to their respective receivers, and keep the messages confidential from the other receiver.

In each of these models, multiple messages need to be protected from eavesdroppers. To the best of our knowledge, there are no practical coding schemes for these multiuser scenarios. We develop polar coding schemes to achieve the best-known secrecy rates achievable by random coding schemes in each one of these channel models. For the MAC-WTC, we achieve the entire dominant face of the best-known achievable region by combining the coding scheme for the general wiretap channel we introduce here with the *monotone chain rule* [12]. For the BC-CM, we introduce a *double chaining* construction to achieve the best-known inner bound. Finally, we extend the coding scheme for the general wiretap channel to the setting of IC-CM.

We acknowledge independent and concurrent papers which present similar results on polar coding for general wiretap channels at the same conference; see [31]–[33]. Reference [31] generalizes the polar coding scheme for strong secrecy in [34], while in our work, we artificially construct the subset property to extend the polar coding scheme in [14]–[17]. Interestingly, these two points of view lead to the same chaining construction method [33]. Moreover, references [31], [33] provide a strong secrecy proof, while in our work, we provide a weak secrecy proof. The remaining parts of these three works are different. References [31], [33] mainly deal with the broadcast channel with a confidential component [2]. However, we not only achieve the secrecy capacity of [2] but also propose coding schemes to achieve the best-known inner bounds of the multiuser models of MAC-WTC, BC-CM and IC-CM, which require different constructions.

II. System Model

A. Wiretap Channel Model

A wiretap channel consists of a legitimate transmitter who wishes to send messages to a legitimate receiver secretly in the presence of an eavesdropper. Let X denote the single-letter input to the main and eavesdropper channels. Let Y and Z denote the corresponding single-letter outputs of the main and the eavesdropper channels, respectively. W represents the message to be sent to Bob and kept secret from Eve with $W \in \mathcal{W} = \{1, \cdots, 2^{nR}\}$. Let $P_e = \Pr(\hat{\mathrm{W}} \neq \mathrm{W})$ denote the probability of error for Bob's decoding.

The equivocation rate is given by

$$\frac{1}{n}H(W|Z^n), \tag{1}$$

which reflects the uncertainty of the message given the eavesdropper's channel observation. A rate pair (R, R_e) is achievable if for any $\epsilon > 0$, as $n \to \infty$,

$$\Pr(\hat{\mathrm{W}} \neq \mathrm{W}) \leq \epsilon, \qquad \frac{1}{n}H(W|Z^n) \geq R_e - \epsilon. \tag{2}$$

Perfect (weak) secrecy is achieved if $R = R_e$ [2]. Therefore, perfect secrecy is achieved if $\frac{1}{n}I(W; Z^n) \to 0$, and the *secrecy capacity* C_s is the highest achievable perfect secrecy rate R, which is also the highest possible equivocation rate [2]. Csiszár and Körner characterized the secrecy capacity for the general wiretap channel as [2]

$$C_s = \max_{V \to X \to Y,Z} I(V; Y) - I(V; Z). \tag{3}$$

B. Multiple Access Wiretap Channel

A MAC-WTC consists of two transmitters, one receiver and an eavesdropper. For $k \in \{1, 2\}$, the two transmitters, with channel inputs X_k, wish to send independent messages $W_k \in \mathcal{W}_k = \{1, \cdots, 2^{nR_k}\}$ to the legitimate receiver, with channel output Y, in the presence of an eavesdropper, with channel output Z. A rate pair (R_1, R_2) is achievable if for any $\epsilon > 0$, as $n \to \infty$,

$$\begin{aligned} &\Pr(\hat{\mathrm{W}}_\mathrm{k} \neq \mathrm{W}_\mathrm{k}) \leq \epsilon, \\ &\frac{1}{n}H(W_1, W_2|Z^n) \geq R_1 + R_2 - \epsilon. \end{aligned} \tag{4}$$

The secrecy capacity region of the MAC-WTC is still an open problem. The best-known achievable rate region is [25], [26] (see also [29], [35], [36]):

$$\begin{aligned} R_1 &\leq [I(V_1; Y|V_2, T) - I(V_1; Z|T)]^+, \\ R_2 &\leq [I(V_2; Y|V_1, T) - I(V_2; Z|T)]^+, \\ R_1 + R_2 &\leq [I(V_1, V_2; Y|T) - I(V_1, V_2; Z|T)]^+, \end{aligned} \tag{5}$$

for any distribution of the form

$$P(t)P(v_1|t)P(v_2|t)P(x_1|v_1)P(x_2|v_2)P(y, z|x_1, x_2). \tag{6}$$

C. Broadcast Channel With Confidential Messages

A BC-CM consists of a transmitter and two receivers. For $k \in \{1, 2\}$, the transmitter wishes to send independent messages, $W_k \in \mathcal{W}_k = \{1, \cdots, 2^{nR_k}\}$, to their respective receiver k, while keeping the messages secret from the unintended receiver. Let X, Y_1, Y_2 denote the single-letter input and outputs of the

[2] Although the naming of BC-CM is similar to [2], these two channel models are different. In particular, [2] is a "single-user" wiretap channel, in the sense that there is only one message to be secured; it is a generalization of [1] to non-degraded channels, together with the introduction of a common message to be sent (insecurely) to both Bob and Eve. BC-CM [27], on the other hand, has two messages each to be secured from the unintended receiver.

broadcast channel. A rate pair (R_1, R_2) is achievable if for any $\epsilon > 0$, as $n \to \infty$,

$$\begin{aligned} &\Pr(\hat{\mathrm{W}}_\mathrm{k} \neq \mathrm{W}_\mathrm{k}) \leq \epsilon, \\ &\frac{1}{n} H(W_1|Y_2^n) \geq R_1 - \epsilon, \\ &\frac{1}{n} H(W_2|Y_1^n) \geq R_2 - \epsilon. \end{aligned} \tag{7}$$

The secrecy capacity region of the BC-CM is still an open problem. The best-known achievable rate region [27] is:

$$\begin{aligned} R_1 &\leq I(V_1; Y_1|T) - I(V_1; V_2|T) - I(V_1; Y_2|V_2, T), \\ R_2 &\leq I(V_2; Y_2|T) - I(V_1; V_2|T) - I(V_2; Y_1|V_1, T), \end{aligned} \tag{8}$$

over all distributions of the form

$$P(t)P(v_1, v_2|t)P(x|v_1, v_2)P(y_1, y_2|x). \tag{9}$$

D. Interference Channel With Confidential Messages

An IC-CM consists of two transmitters and two receivers. The two transmitters wish to send independent messages to their respective receivers, and keep the messages confidential from the other receiver. For $k \in \{1, 2\}$, let X_k, Y_k denote the single-letter input and output of the interference channel with messages $W_k \in \mathcal{W}_k = \{1, \cdots, 2^{nR_k}\}$. A rate pair (R_1, R_2) is achievable if for any $\epsilon > 0$, as $n \to \infty$,

$$\begin{aligned} &\Pr(\hat{\mathrm{W}}_\mathrm{k} \neq \mathrm{W}_\mathrm{k}) \leq \epsilon, \\ &\frac{1}{n} H(W_1|Y_2^n) \geq R_1 - \epsilon, \\ &\frac{1}{n} H(W_2|Y_1^n) \geq R_2 - \epsilon. \end{aligned} \tag{10}$$

The secrecy capacity region of the IC-CM is still an open problem. The best-known achievable rate region [27] is:

$$\begin{aligned} R_1 &\leq I(V_1; Y_1|T) - I(V_1; Y_2|V_2, T), \\ R_2 &\leq I(V_2; Y_2|T) - I(V_2; Y_1|V_1, T), \end{aligned} \tag{11}$$

over all distribution of the form

$$P(t)P(v_1|t)P(v_2|t)P(x_1|v_1)P(x_2|v_2)P(y_1, y_2|x_1, x_2). \tag{12}$$

III. Existing Random Coding Schemes for Secure Communication

In this section, we summarize the well-known random coding techniques for secure communication. We first show how to achieve the secrecy rate, $I(X; Y) - I(X; Z)$, through the *stochastic encoding* technique introduced in [1] for the degraded wiretap channel. We then show how to apply *channel prefixing* introduced in [2] for the general wiretap channel to achieve the secrecy capacity in (3). We next summarize some relevant extensions to multiuser scenarios.

To achieve the secrecy rate $I(X; Y) - I(X; Z)$, we fix the input distribution $P(x)$ and generate a random codebook by using independent and identically distributed realizations according to $P(x)$. The random codebook consists of $2^{n(R_s+\tilde{R}_s)}$ n-length codewords. We take $R_s = I(X; Y) - I(X; Z)$ and $\tilde{R}_s = I(X; Z)$. Let $W_s \in \{1, 2, \ldots, 2^{nR_s}\}$ denote the secret message, and let $\tilde{W}_s \in \{1, 2, \ldots, 2^{n\tilde{R}_s}\}$ denote the confusion message. $\tilde{W}_s$ carries no information and only serves to protect W_s. In the encoding procedure, after we choose the secure message W_s, we randomly pick the confusion message $\tilde{W}_s$ to determine the codeword for transmission. Therefore, W_s and $\tilde{W}_s$ together determine the transmitted codeword $x^n(W_s, \tilde{W}_s)$. This *stochastic encoding* procedure enables secure communication.

Since the code rate is $R_s + \tilde{R}_s = I(X; Y)$, Bob decodes both W_s and $\tilde{W}_s$ reliably. In order to prove secrecy against Eve, we evaluate the equivocation rate $\frac{1}{n} H(W_s|Z^n)$ at Eve, i.e., the entropy of the secure message given Eve's observation (similar steps in (45)-(48)) [1], [2]:

$$\begin{aligned} \frac{1}{n} H(W_s|Z^n) \geq &\frac{1}{n} H(W_s) + \frac{1}{n} H(\tilde{W}_s) \\ &- \frac{1}{n} I(X^n; Z^n) - \frac{1}{n} H(\tilde{W}_s|W_s, Z^n), \end{aligned} \tag{13}$$

where $\frac{1}{n} H(\tilde{W}_s) \approx I(X; Z) \approx \frac{1}{n} I(X^n; Z^n)$, and $\frac{1}{n} H(\tilde{W}_s|W_s, Z^n) \approx 0$ through Fano's inequality. Therefore, $\frac{1}{n} H(W_s|Z^n) \geq \frac{1}{n} H(W_s) - \epsilon$, and the (weak) secrecy constraint is satisfied.

To achieve the secrecy capacity in (3) for the general wiretap channel, we create an artificial channel $P_{X|V}$, which is called *channel prefixing* in [2]. Although from data processing inequality, $I(V; Y) \leq I(X; Y)$ and $I(V; Z) \leq I(X; Z)$, the difference $I(V; Y) - I(V; Z)$ may be larger than $I(X; Y) - I(X; Z)$, and channel prefixing, in general, is useful. For degraded channels, optimum V equals X, and the secrecy capacity is $C_s = \max_X I(X; Y) - I(X; Z)$ [1], [2].

For the achievable rate regions for multiuser scenarios in (5), (8) and (11), T serves as the time-sharing random variable, and V_1 and V_2 denote the channel prefixing auxiliary random variables. For MAC-WTC in (5), both users apply stochastic encoding, with sacrificed confusion message rates of $\tilde{R}_k \leq I(V_k; Z|T)$ for $k \in \{1, 2\}$, with $\tilde{R}_1 + \tilde{R}_2 = I(V_1, V_2; Z|T)$. For BC-CM in (8), each user $k \in \{1, 2\}$ sacrifices the rate $I(V_k; Y_j|V_j, T), k \neq j$ for stochastic encoding, and each user uses the rate $I(V_1, V_2; Z|T)$ for binning. For IC-CM in (11), each user sacrifices the rate of $I(V_k; Y_j|V_j, T), k \neq j$ for stochastic encoding.

IV. Existing Polar Coding Techniques

A. Polar Codes for Asymmetric Channels

Let P_{XY} be the joint distribution of a pair of random variables (X, Y), where X is a binary random variable and Y is any finite-alphabet random variable. Let us define the Bhattacharyya parameter as follows:

$$Z(X|Y) = 2 \sum_y P_Y(y) \sqrt{P_{X|Y}(0|y) P_{X|Y}(1|y)}. \tag{14}$$

Let $U^n = X^n G_n$, where X^n denotes n independent copies of the random variable X with $X \sim P_X$, and $G_n = G^{\otimes k}$ where

$G = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $\otimes$ denotes the Kronecker product of matrices for $n = 2^k$. Reference [4] shows that as $n \to \infty$, U_i is almost independent of U^{i-1} and uniformly distributed, or otherwise U_i is almost determined by U^{i-1}. Therefore, $[n]$, the index set $\{1, 2, \ldots, n\}$, is almost polarized into two sets $\mathcal{H}_X$ and $\mathcal{L}_X$ [10]:

$$\mathcal{H}_X = \{i \in [n] : Z(U_i|U^{i-1}) \geq 1 - \delta_n\}, \tag{15}$$

$$\mathcal{L}_X = \{i \in [n] : Z(U_i|U^{i-1}) \leq \delta_n\}, \tag{16}$$

where $\delta_n = 2^{-n^\beta}$ and $\beta \in (0, 1/2)$. Moreover,

$$\lim_{n\to\infty} \frac{1}{n}|\mathcal{H}_X| = H(X), \tag{17}$$

$$\lim_{n\to\infty} \frac{1}{n}|\mathcal{L}_X| = 1 - H(X). \tag{18}$$

Let P be a discrete memoryless channel with a binary input X and finite alphabet output Y. Here, P does not have to be a symmetric channel. Fix a distribution P_X for X. Reference [23] generalizes the above argument to achieve a rate close to $I(X; Y)$. Consider two subsets of $[n]$, $\mathcal{H}_{X|Y}$ and $\mathcal{L}_{X|Y}$,

$$\mathcal{H}_{X|Y} = \{i \in [n] : Z(U_i|U^{i-1}, Y^n) \geq 1 - \delta_n\}, \tag{19}$$

$$\mathcal{L}_{X|Y} = \{i \in [n] : Z(U_i|U^{i-1}, Y^n) \leq \delta_n\}. \tag{20}$$

Similar to (17) and (18), we have

$$\lim_{n\to\infty} \frac{1}{n}|\mathcal{H}_{X|Y}| = H(X|Y), \tag{21}$$

$$\lim_{n\to\infty} \frac{1}{n}|\mathcal{L}_{X|Y}| = 1 - H(X|Y). \tag{22}$$

With (15) and (20), we define the following three sets

$$\mathcal{I} = \mathcal{H}_X \cap \mathcal{L}_{X|Y}, \tag{23}$$

$$\mathcal{F}_r = \mathcal{H}_X \cap \mathcal{L}^{\mathrm{c}}_{X|Y}, \tag{24}$$

$$\mathcal{F}_d = \mathcal{H}^{\mathrm{c}}_X. \tag{25}$$

In the following, we call the set $\mathcal{I}$ the *information set*, and sets $\mathcal{F}_r$ and $\mathcal{F}_d$ the *frozen set*. Although we call them the *frozen set*, $\mathcal{F}_r$ and $\mathcal{F}_d$ have different operational meanings which will be illustrated below. Note that for the symmetric channel capacity achieving code design, $\mathcal{F}_d$ is an empty set [3].

To achieve rate $I(X; Y)$ for channel P, let us consider the following coding scheme. First, the encoder transmits the information bits in the index set $\mathcal{I}$. For $i \in \mathcal{I}$ in (23), since $i \in \mathcal{H}_X$, U_i is almost independent of U^{i-1} and uniformly distributed. Therefore, the encoder can freely assign values to $U_\mathcal{I}$, where $U_\mathcal{I}$ denotes a sub-vector $\{U_i\}_{i\in\mathcal{I}}$. Moreover, since $i \in \mathcal{L}_{X|Y}$, U_i is almost determined by U^{i-1} and Y^n, which means that given the channel output Y^n, U_i can be decoded in a successive manner.

Second, for $i \in \mathcal{F}_r$ in (24), U_i is almost independent of U^{i-1} and uniformly distributed, and given the channel output Y^n, U_i cannot be reliably decoded. The encoder transmits $U_{\mathcal{F}_r}$ with a uniformly random sequence and the randomness is shared between the transmitter and receiver.

Last, for $i \in \mathcal{F}_d$ in (25), U_i is almost determined by U^{i-1}. The values of $U_{\mathcal{F}_d}$ are computed in successive order through the following mapping:

$$u_i = \arg\max_{u\in\{0,1\}} P_{U_i|U^{i-1}}(u|u^{i-1}). \tag{26}$$

By (17) and (21), it is easy to verify that

$$\lim_{n\to\infty} \frac{1}{n}|\mathcal{I}| = I(X; Y). \tag{27}$$

Moreover, by applying successive cancellation decoder, the block error probability P_e can be upper bounded by [37]

$$P_e \leq \sum_{i\in\mathcal{I}} Z(U_i|U^{i-1}, Y^n) = O(2^{-n^\beta}) \tag{28}$$

for any $\beta \in (0, 1/2)$, with complexity $O(n \log n)$. Therefore, the rate $I(X; Y)$ is achieved.

B. Universal Polar Coding

Consider two B-DMCs $P : X \to Y$ and $Q : X \to Z$, and assume that these two channels have identical capacities, i.e., $C(P) = C(Q)$. Let $U^n = X^n G_n$, and denote $\mathcal{P}$ and $\mathcal{Q}$ as the information set defined in (23), i.e.,

$$\mathcal{P} = \{i \in [n] : Z(U_i|U^{i-1}, Y^n) \leq \delta_n\}, \tag{29}$$

$$\mathcal{Q} = \{i \in [n] : Z(U_i|U^{i-1}, Z^n) \leq \delta_n\}, \tag{30}$$

where $\delta_n = 2^{-n^\beta}$ and $\beta \in (0, 1/2)$. Since we assume $C(P) = C(Q)$, we also have $|\mathcal{P}| = |\mathcal{Q}|$.

In general, the differences $\mathcal{P} \setminus \mathcal{Q}$ and $\mathcal{Q} \setminus \mathcal{P}$ are not empty sets [18]–[20]; therefore, it is not straightforward to apply standard polar coding to achieve the capacity of the compound channel consisting of P and Q. Reference [21] proposes a method, called *chaining construction*, to solve this problem; see also [34].

Definition 1: (Chaining construction [21]) Let $m \geq 2$. The m-chain of $\mathcal{P}$ and $\mathcal{Q}$ is a code of length mn that consists of m polar blocks of length n. In each of the m blocks, the set $\mathcal{P} \cap \mathcal{Q}$ is set to be an information set. In the ith block, $1 \leq i < m$, the set $\mathcal{P} \setminus \mathcal{Q}$ is also set to be an information set. Moreover, the set $\mathcal{P} \setminus \mathcal{Q}$ in the ith block is chained to the set $\mathcal{Q} \setminus \mathcal{P}$ in the $(i + 1)$th block in the sense that the information is repeated in these two sets. All other indices are frozen. Therefore, in each block, the set $(\mathcal{P} \cup \mathcal{Q})^c$ is frozen, and the set $\mathcal{Q} \setminus \mathcal{P}$ in the 1st block and the set $\mathcal{P} \setminus \mathcal{Q}$ in the mth block are frozen, too. The rate of the chaining construction is

$$\frac{|\mathcal{P} \cap \mathcal{Q}| + \frac{m-1}{m}|\mathcal{P} \setminus \mathcal{Q}|}{n}. \tag{31}$$

Next, we discuss the decoding procedure for the compound channel consisting of P and Q. If channel P is used, then we decode from the first block. On the other hand, if channel Q is used, then we decode from the last block.

First, suppose that channel P is used and a code of length mn has been received. For this case, we decode from the first block. In the 1st block, all the information bits are put in the set $\mathcal{P}$; thus, the decoder can decode correctly. For the 2nd block, through chaining construction, the set $\mathcal{P} \setminus \mathcal{Q}$ in the 1st block is chained to the set $\mathcal{Q} \setminus \mathcal{P}$ in the 2nd block, and the set $(\mathcal{P} \cup \mathcal{Q})^c$ is frozen. Equivalently, the decoder only needs to decode the bits in the set $\mathcal{P}$, which can be correctly decoded. The same procedure holds until the $(m-1)$th block. For the mth block, the information bits are only put in the set $\mathcal{P} \cap \mathcal{Q}$, and the remaining part has been determined. Hence, information bits can be reliably decoded.

Second, consider the case that channel Q is used. In this case, we decode from the last block. In the mth block, since the information bits are put in the set $\mathcal{Q}$, reliable decoding is guaranteed. For the $(m-1)$th block, due to the chaining process, the set $\mathcal{Q} \setminus \mathcal{P}$ in the mth block is chained to the set $\mathcal{P} \setminus \mathcal{Q}$ in the $(m-1)$th block, and note that the set $(\mathcal{P} \cup \mathcal{Q})^c$ is frozen. The decoder only needs to decode the information bits in the set $\mathcal{Q}$, thus correct decoding is ensured. This procedure is applied until the 2nd block. For the 1st block, information bits which have not been determined fall in the set $\mathcal{P} \cap \mathcal{Q}$, thus the decoder can decode them correctly.

In summary, for a fixed m, if we let $n \to \infty$, we can achieve the rate in (31) with arbitrary small error probability, which also means that the rate $C(P) - \frac{1}{m}\frac{|\mathcal{P}\setminus\mathcal{Q}|}{n}$ can be achieved. Additionally, if we let $m \to \infty$, then the rate $C(P)$, which is the capacity of the compound channel consisting of channels P and Q, can be achieved.

C. Polar Coding for MAC Based on Monotone Chain Rules

Consider a two-user MAC $(\mathcal{X}_1 \times \mathcal{X}_2, P(y|x_1, x_2), \mathcal{Y})$ with binary input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$. The capacity region of this channel is the union of convex hull of all rate pairs satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2), \\ R_2 &\leq I(X_2; Y|X_1), \\ R_1 + R_2 &\leq I(X_1, X_2; Y), \end{aligned} \tag{32}$$

over the distributions of the form $P(x_1)P(x_2)$. The rate pairs satisfying $R_1 + R_2 = I(X_1, X_2; Y)$ are said to be on the *dominant face* of the rate region.

Reference [12] gives a polar coding scheme that achieves the entire dominant face based on the monotone chain rules. Consider $U_1^n = X_1^n G_n$ and $U_2^n = X_2^n G_n$. We call J^{2n} as a monotone permutation of $U_1^n U_2^n$ if the elements of both U_1^n and U_2^n appear in increasing order in J^{2n}. When we expand the mutual information term $I(U_1^n, U_2^n; Y^n)$ according to the monotone permutation, we say that it follows the monotone chain rule

$$I(U_1^n, U_2^n; Y^n) = \sum_{i=1}^{2n} I(J_i; Y^n | J^{i-1}). \tag{33}$$

Moreover, define the rates as follows (similar to [11], [12]):

$$\begin{aligned} R_x &= \frac{1}{n} \sum_{\{i \in [2n]: J_i = U_{1,k}, k \in [n]\}} I(J_i; Y^n | J^{i-1}), \\ R_y &= \frac{1}{n} \sum_{\{i \in [2n]: J_i = U_{2,k}, k \in [n]\}} I(J_i; Y^n | J^{i-1}). \end{aligned} \tag{34}$$

Reference [12] shows that the rate pair (R_x, R_y) in (34) can be set arbitrarily close to the rate pairs on the dominant face of (32) by the permutations of the form $J^{2n} = (U_1^i, U_2^n, U_1^{i+1:n})$, where $U_1^{i+1:n}$ denotes $U_{1,i+1}, \ldots, U_{1,n}$.

V. Polar Coding for the General Wiretap Channel

Assume now that we know the optimal distributions [38] to achieve the secrecy capacity C_s in (3), i.e., we know the optimal V and X. For illustration, we consider the case of a binary input channel, i.e., $|\mathcal{X}| = 2$. The cardinality bound for channel prefixing, V, is $|\mathcal{V}| \leq 2$. Although we focus on developing a coding scheme for binary inputs below, there is no difficulty to extend the work to q-ary inputs [39]–[42].

A. The Scheme

Let $U^n = V^n G_n$. Consider the following sets:

$$\begin{aligned} \mathcal{H}_V &= \{i \in [n] : Z(U_i | U^{i-1}) \geq 1 - \delta_n\}, \\ \mathcal{L}_{V|Y} &= \{i \in [n] : Z(U_i | U^{i-1}, Y^n) \leq \delta_n\}, \\ \mathcal{L}_{V|Z} &= \{i \in [n] : Z(U_i | U^{i-1}, Z^n) \leq \delta_n\}, \end{aligned} \tag{35}$$

where $\delta_n = 2^{-n^\beta}$ and $\beta \in (0, 1/2)$.

The set $[n]$ can be partitioned into the following four sets:

$$\begin{aligned} G_{Y \wedge Z} &= \mathcal{H}_V \cap \mathcal{L}_{V|Y} \cap \mathcal{L}_{V|Z}, \\ G_{Y \setminus Z} &= \mathcal{H}_V \cap \mathcal{L}_{V|Y} \cap \mathcal{L}_{V|Z}^c, \\ G_{Z \setminus Y} &= \mathcal{H}_V \cap \mathcal{L}_{V|Y}^c \cap \mathcal{L}_{V|Z}, \\ B_{Y \wedge Z} &= \mathcal{H}_V^c \cup (\mathcal{L}_{V|Y}^c \cap \mathcal{L}_{V|Z}^c). \end{aligned} \tag{36}$$

From a successive decoding point of view, the sub-channels corresponding to the set $G_{Y \wedge Z}$ are simultaneously good for Bob and Eve. The sub-channels in the set $G_{Y \setminus Z}$ are good for Bob but bad for Eve. On the other hand, the sub-channels in the set $G_{Z \setminus Y}$ are good for Eve but bad for Bob. Last, the sub-channels in the set $B_{Y \wedge Z}$ are bad for both Bob and Eve.

Similar to (23)–(25), we have:

$$\begin{aligned} \mathcal{I}_Y &= \mathcal{H}_V \cap \mathcal{L}_{V|Y}, \\ \mathcal{I}_Z &= \mathcal{H}_V \cap \mathcal{L}_{V|Z}, \\ \mathcal{F}_r^Y &= \mathcal{H}_V \cap \mathcal{L}_{V|Y}^c, \\ \mathcal{F}_r^Z &= \mathcal{H}_V \cap \mathcal{L}_{V|Z}^c, \\ \mathcal{F}_d &= \mathcal{H}_V^c. \end{aligned} \tag{37}$$

By (27), we have

$$\begin{aligned} \lim_{n \to \infty} \frac{1}{n} |\mathcal{I}_Y| &= I(V; Y), \\ \lim_{n \to \infty} \frac{1}{n} |\mathcal{I}_Z| &= I(V; Z). \end{aligned} \tag{38}$$

For the *symmetric* and *degraded* wiretap channel [14]–[17], $G_{Z \setminus Y}$ is an empty set, since the degraded property of the channel causes $\mathcal{I}_Z \subset \mathcal{I}_Y$ [13]. However, for the general wiretap

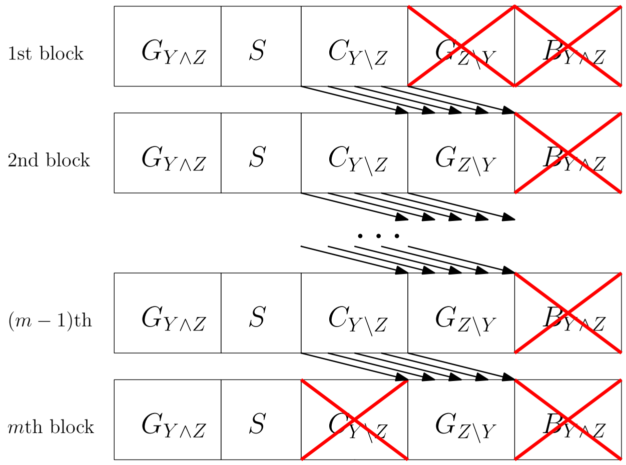


Fig. 1. Chaining construction for the general wiretap channel.

channel, $G_{Z\setminus Y}$ is no longer an empty set, and $|G_{Z\setminus Y}|$ cannot be negligible [18]–[20].

Here, we consider the positive secrecy capacity case, thus, we have $|G_{Y\setminus Z}| > |G_{Z\setminus Y}|$. Choose a set, $C_{Y\setminus Z}$, such that $C_{Y\setminus Z} \subset G_{Y\setminus Z}$ and $|C_{Y\setminus Z}| = |G_{Z\setminus Y}|$. Define the set S as:

$$S = G_{Y\setminus Z} \setminus C_{Y\setminus Z}. \tag{39}$$

From (38), we have

$$\lim_{n\to\infty} \frac{1}{n}|S| = I(V;Y) - I(V;Z). \tag{40}$$

We construct the code as follows. Consider an *m-chain* polar code in Definition (1). For $1 \le i < m$, the set $C_{Y\setminus Z}$ in the ith block is chained to $G_{Z\setminus Y}$ in the $(i+1)$th block as in Fig. 1. For each of the m blocks, the set $B_{Y\wedge Z}$ is set to be frozen. Moreover, the set $G_{Z\setminus Y}$ in the 1st block is set to be frozen in the sense that $G_{Z\setminus Y} \subseteq \mathcal{F}_r^Y$, and the set $C_{Y\setminus Z}$ in the mth block is also set to be frozen in the sense that $C_{Y\setminus Z} \subseteq \mathcal{F}_r^Z$. In Fig. 1, we use a red cross to denote a frozen set.

We put the secret information bits in the set S in each block. Therefore, the set S is used for secret message transmission. For blocks $1 \le i < m$, we put uniformly distributed random bits to $C_{Y\setminus Z}$ to serve as the confusion messages. Through the chaining construction, the confusion messages are also chained to the set $G_{Z\setminus Y}$ in block $1 < i \le m$. Moreover, the set $G_{Y\wedge Z}$ in each block are also filled with random bits to serve as confusion message. For the frozen sets, if the index belongs to $\mathcal{F}_r^Y$ or $\mathcal{F}_r^Z$, then we put uniformly distributed random bits and share the randomness with the decoder (Bob and Eve). Last, if the index belongs to $\mathcal{F}_d$, then we determine the value according to the mapping defined in (26). We summarize the encoding procedure as follows.

Encoding procedure:

For each block, put the secret information bits in U_S, and determine the bits in $U_{\mathcal{F}_d}$ by (26).

For the 1st block,

1) Put uniformly distributed random bits to $U_{G_{Y\wedge Z}\cup C_{Y\setminus Z}}$.
2) Put uniformly distributed random bits to $U_{\mathcal{F}_r^Y}$, and share the randomness with the decoder.

For the jth block, $2 \le j < m$,

1) Put uniformly distributed random bits to $U_{G_{Y\wedge Z}\cup C_{Y\setminus Z}}$.
2) Chaining construction: repeat the bits in $C_{Y\setminus Z}$ of the $(j-1)$th block to the bits in $U_{G_{Z\setminus Y}}$.
3) Put uniformly distributed random bits to $U_{\mathcal{F}_r^Y \cap \mathcal{F}_r^Z}$, and share the randomness with the decoder.

For the mth block,

1) Put uniformly distributed random bits to $U_{G_{Y\wedge Z}}$.
2) Chaining construction: repeat the bits in $C_{Y\setminus Z}$ of the $(m-1)$th block to the bits in $U_{G_{Z\setminus Y}}$.
3) Put uniformly distributed random bits to $U_{\mathcal{F}_r^Z}$, and share the randomness with the decoder.

Note that in the chaining construction we require the bits in $U_{G_{Z\setminus Y}}$ equal the bits in $U_{C_{Y\setminus Z}}$. Since we fill uniformly distributed random bits to $U_{C_{Y\setminus Z}}$, we simultaneously fill random bits to $U_{G_{Z\setminus Y}}$. Due to the fact that $G_{Z\setminus Y} \cap \mathcal{F}_d = \emptyset$, we can freely choose the bits in this set.

Decoding procedure:

Bob decodes from the 1st block. If $i \in \mathcal{F}_d$, then $\hat{u}_i = \arg\max_{u\in\{0,1\}} P_{U_i|U^{i-1}}(u|\hat{u}^{i-1})$. For the 1st block,

$$\hat{u}_i = \begin{cases} u_i, \text{if } i \in \mathcal{F}_r^Y, \\ \arg\max_{u\in\{0,1\}} P_{U_i|U^{i-1},Y^n}(u|\hat{u}^{i-1}, y^n), \\ \text{if } i \in G_{Y\wedge Z} \cup C_{Y\setminus Z} \cup S. \end{cases} \tag{41}$$

For the jth block, $2 \le j < m$,

$$\hat{u}_i = \begin{cases} u_i, \text{if } i \in \mathcal{F}_r^Y \cap \mathcal{F}_r^Z, \\ \arg\max_{u\in\{0,1\}} P_{U_i|U^{i-1},Y^n}(u|\hat{u}^{i-1}, y^n), \\ \text{if } i \in G_{Y\wedge Z} \cup C_{Y\setminus Z} \cup S, \\ \hat{u}_{i'} \text{ in the } (j-1)\text{th block, where } i' \in C_{Y\setminus Z}, \\ \text{if } i \in G_{Z\setminus Y}. \end{cases} \tag{42}$$

For the mth block,

$$\hat{u}_i = \begin{cases} u_i, \text{if } i \in \mathcal{F}_r^Z, \\ \arg\max_{u\in\{0,1\}} P_{U_i|U^{i-1},Y^n}(u|\hat{u}^{i-1}, y^n), \\ \text{if } i \in G_{Y\wedge Z} \cup S, \\ \hat{u}_{i'} \text{ in the } (m-1)\text{th block, where } i' \in C_{Y\setminus Z}, \\ \text{if } i \in G_{Z\setminus Y}. \end{cases} \tag{43}$$

Theorem 1: For any $\beta \in (0, 1/2)$, there exists an *m-chain* polar coding scheme developed in Section V-A, such that as $n \to \infty$, the *m-chain* polar coding scheme achieves the secrecy capacity for the general wiretap channel in (3), and the block error probability decays as $O(2^{-n^{\beta}})$.

The proof of Theorem 1 has two parts: proof of reliability at Bob is given in Section V-B and the equivocation calculation (proof of secrecy at Eve) is given in Section V-C.

B. Reliability

From (40), we know as $n \to \infty$, our coding scheme can achieve the secrecy rate in (3). Moreover, when Bob applies the decoding procedure described in Section V-A, according to (28), the block error probability of the whole *m-chain* block can be upper bounded by

$$P_e^B \leq (m-1) \sum_{i \in C_{Y \setminus Z}} Z(U_i|U^{i-1}, Y^n) + m \sum_{i \in G_{Y \wedge Z} \cup S} Z(U_i|U^{i-1}, Y^n) = O(2^{-n^\beta}) \tag{44}$$

for any $\beta \in (0, 1/2)$ with complexity $O(n \log n)$. Thus, the secrecy rate in (3) is achieved reliably.

C. Equivocation Calculation

We first introduce necessary notation for the calculation of the equivocation rate. In the encoding process, we consider m blocks each with block length n. Let Z^{mn} denote what Eve receives. For each block, we perform $U^n = V^n G_n$, therefore, for the total of m blocks, we have V^{mn} and U^{mn}.

Let W_s denote the secret message, and $\tilde{W}_s$ denote the confusion message. Let the subscript i of a set denote the set in the ith block. For example, S_i denotes the set S in the ith block, and $G_{Y \wedge Zj}$ denotes the set $G_{Y \wedge Z}$ in the jth block. Since secret message is put in S_i, $1 \leq i \leq m$, we have $W_s = \cup_{1 \leq i \leq m} U_{S_i}$. Also, the confusion message is put in $G_{Y \wedge Zi}$, $1 \leq i \leq m$ and $C_{Y \setminus Zj}$, $1 \leq j < m$. Therefore, we have $\tilde{W}_s = \cup_{1 \leq i \leq m, 1 \leq j < m} U_{G_{Y \wedge Zi}} U_{C_{Y \setminus Zj}}$.

We can calculate the equivocation rate as follows:

$$\begin{aligned} & H(W_s|Z^{mn}) \\ &= H(W_s, \tilde{W}_s|Z^{mn}) - H(\tilde{W}_s|W_s, Z^{mn}) && (45) \\ &= H(W_s, \tilde{W}_s) - I(W_s, \tilde{W}_s; Z^{mn}) - H(\tilde{W}_s|W_s, Z^{mn}) && (46) \\ &\geq H(W_s, \tilde{W}_s) - I(V^{mn}; Z^{mn}) - H(\tilde{W}_s|W_s, Z^{mn}) && (47) \\ &= H(W_s) + H(\tilde{W}_s) - I(V^{mn}; Z^{mn}) - H(\tilde{W}_s|W_s, Z^{mn}) && (48) \end{aligned}$$

which is equivalent to

$$\frac{1}{mn} I(W_s; Z^{mn}) \leq \frac{1}{mn} I(V^{mn}; Z^{mn}) + \frac{1}{mn} H(\tilde{W}_s|W_s, Z^{mn}) - \frac{1}{mn} H(\tilde{W}_s). \tag{49}$$

Note that in (45), to keep the notation concise we do not list the randomness shared with the decoder (see the encoding procedure in Section V-A) in the expression of the conditional entropy. Here, (45) is due to the chain rule of conditional entropy, (46) is due to the definition of mutual information, (47) comes from the data processing inequality, (48) is due to the independence of the secret message and the confusion message. In (49), we bound each term on the right hand side as follows:

For the first term, we have $I(V^{mn}; Z^{mn}) \leq \sum_{i=1}^{mn} I(V_i; Z_i) \leq mnI(V; Z)$. Therefore, $\frac{1}{mn} I(V^{mn}; Z^{mn}) \leq I(V; Z)$.

To bound the second term, suppose Eve obtains W_s and Z^{mn}, and wants to decode $\tilde{W}_s$. By symmetry of chaining construction, Eve can apply similar decoding rule as described in Section V-A. However, this time Eve decodes from the mth block, then the block error probability of the whole *m-chain* block can be upper bounded by

$$P_e^E \leq (m-1) \sum_{i \in G_{Z \setminus Y}} Z(U_i|U^{i-1}, Y^n) + m \sum_{i \in G_{Y \wedge Z}} Z(U_i|U^{i-1}, Y^n) = O(2^{-n^\beta}) \tag{50}$$

for $\beta \in (0, 1/2)$. Hence, by applying Fano's inequality, we have

$$\begin{aligned} H(\tilde{W}_s|W_s, Z^{mn}) &\leq H(P_e^E) + P_e^E \log |\tilde{W}_s| \\ &< H(P_e^E) + P_e^E [mnI(V; Z)]. \end{aligned} \tag{51}$$

Therefore, as $n \to \infty$, $\frac{1}{mn} H(\tilde{W}_s|W_s, Z^{mn}) \to 0$.

For the last term, as $n \to \infty$, by (31) and (38), we have $(m-1)nI(V; Z) < H(\tilde{W}_s) < mnI(V; Z)$. Hence, as $m \to \infty$, $\frac{1}{mn} H(\tilde{W}_s) \to I(V; Z)$.

From the above, we know as $n \to \infty$ and $m \to \infty$, $\frac{1}{mn} I(W_s; Z^{mn}) \to 0$. Thus, the weak secrecy constraint is achieved.

VI. Polar Coding for the Multiple Access Wiretap Channel

In this section, instead of achieving the corner point of (5) through standard polar coding techniques [6], we show how to achieve the rate pairs on the dominant face of (5), since reference [43] shows the former scheme is strictly suboptimal[3]. Here, we consider the positive rate case in (5), i.e., $R_1 > 0$, $R_2 > 0$ and $R_1 + R_2 > 0$. We first consider a constant T in (5). Following the method given in [11, Sec. III. B.], we can generalize the result to a T with arbitrary distribution. For $k \in \{1, 2\}$, let $\mathcal{V}_k$ be the corresponding alphabet of the channel prefixing V_k. As in Section V, we assume the cardinality for the channel prefixing V_k is $|\mathcal{V}_k| = 2$ for illustration.

A. The Scheme

For a fixed input distribution in (6), consider two different MACs, the first MAC, P, consisting of two users and Bob and the second MAC, Q, consisting of the two users and Eve. In Fig. 2, we use a solid line to show the achievable region for the first MAC, P, and a dotted line to represent the second MAC, Q. Consider two rate pairs on the dominant faces of the channels P and Q, which we use green and red points to denote in Fig. 2.

Reference [12] shows that there exist monotone permutations J^{2n} and K^{2n} for channels P and Q to achieve the green and red points in Fig. 2. Since the green rate pair is greater than the red rate pair in the sense of both rate of user 1 and rate of user 2, we can also achieve the red rate pair for channel P by the same monotone chain J^{2n}. In the following, we present a polar coding scheme such that we set the rate of the confusion

[3] Coding schemes for MAC-WTC are related to coding schemes for compound MAC, since in a MAC-WTC there are two MACs, one to Bob and one to Eve. Reference [43] considers ICs. In the best-known coding scheme for ICs, i.e., the Han-Kobayashi (HK) coding scheme, each transmitter divides its message into two, a private part and a common part. The common parts need to be decoded by both receivers. Therefore, if private messages are ignored, IC with HK coding scheme becomes a compound MAC. [43] shows that rate-splitting may not achieve the optimal compound rates in such channels in general.

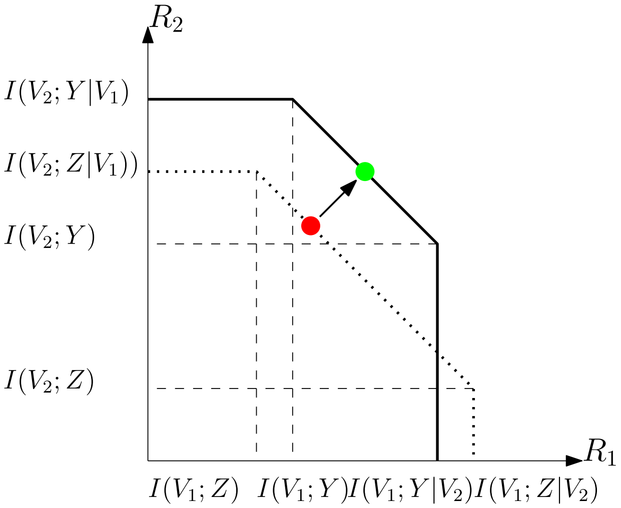


Fig. 2. General MAC regions.

message as the red rate pair and the rate of the secret message as the difference of the green and red rate pairs.

For $k \in \{1,2\}$, let $U_k^n = V_k^n G_n$. Once we determine the distribution in (6), similar to (35), we can define $\mathcal{H}_{V_k}$. According to different monotone permutations, J^{2n}, we have different index sets for $\mathcal{L}_{V_k|Y,J}$. We define them as follows:

$$\mathcal{L}_{V_k|Y,J} = \{i \in [n] : Z(U_{k,i}|Y^n, J^{j-1}) \le \delta_n,\ J_j = U_{k,i}\}, \tag{52}$$

where $\delta_n = 2^{-n^\beta}$ and $\beta \in (0, 1/2)$. Similarly, we can also define $\mathcal{L}_{V_k|Z,K}$ for another monotone permutation, K^{2n}.

The set $[n]$ for the user k can be partitioned into the following sets:

$$\begin{aligned} G^{(k)}_{Y\wedge Z} &= \mathcal{H}_{V_k} \cap \mathcal{L}_{V_k|Y,J} \cap \mathcal{L}_{V_k|Z,K}, \\ G^{(k)}_{Y\setminus Z} &= \mathcal{H}_{V_k} \cap \mathcal{L}_{V_k|Y,J} \cap \mathcal{L}^c_{V_k|Z,K}, \\ G^{(k)}_{Z\setminus Y} &= \mathcal{H}_{V_k} \cap \mathcal{L}^c_{V_k|Y,J} \cap \mathcal{L}_{V_k|Z,K}, \\ B^{(k)}_{Y\wedge Z} &= \mathcal{L}_{V_k} \cup (\mathcal{L}^c_{V_k|Y,J} \cap \mathcal{L}^c_{V_k|Z,K}). \end{aligned} \tag{53}$$

Since we consider the positive rate case in (5), we have $|G^{(k)}_{Y\setminus Z}| > |G^{(k)}_{Z\setminus Y}|$. Pick $C^{(k)}_{Y\setminus Z} \subset G^{(k)}_{Y\setminus Z}$, such that $|C^{(k)}_{Y\setminus Z}| = |G^{(k)}_{Z\setminus Y}|$. Define the set $S^{(k)}$ as follows:

$$S^{(k)} = G^{(k)}_{Y\setminus Z} \setminus C^{(k)}_{Y\setminus Z}. \tag{54}$$

According to the result in [12], we have

$$\lim_{n\to\infty} \frac{1}{n}(|S^{(1)}| + |S^{(2)}|) = I(V_1, V_2; Y) - I(V_1, V_2; Z). \tag{55}$$

The encoding procedure for the two users are similar. We show the encoding procedure in Fig. 3 for user 1. For each user, we put the secret bits in the set $S^{(k)}$ and put random bits as the confusion message in the sets $G^{(k)}_{Y\wedge Z}$ and $C^{(k)}_{Y\setminus Z}$. Moreover, we chain the bits in the set $C^{(k)}_{Y\setminus Z}$ in the ith block to the set $G^{(k)}_{Z\setminus Y}$ in the $(i+1)$th block. To guarantee correct decoding, we freeze the sets $B^{(k)}_{Y\wedge Z}$ in each block, $G^{(k)}_{Z\setminus Y}$ in the 1st block, and $C^{(k)}_{Y\setminus Z}$ in the mth block. We use red crosses in Fig. 3 to denote the frozen sets.

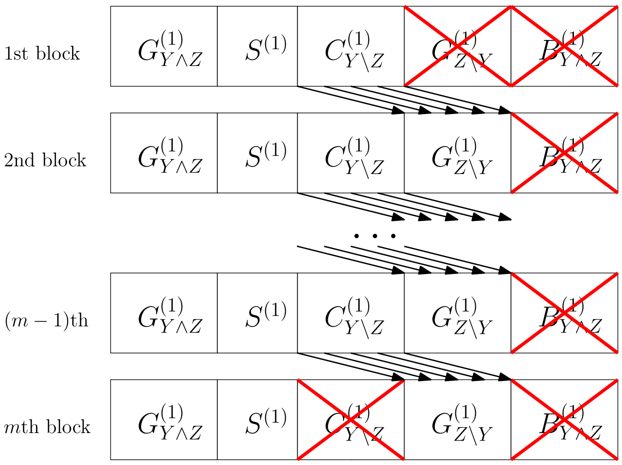


Fig. 3. Chaining construction for the MAC-WTC for user 1.

The decoding procedure is from the 1st block to the mth block according to the monotone permutation J^{2n} for Bob. For the 1st block, since the bits Bob needs to decode are all in the sets $G^{(k)}_{Y\wedge Z}$ or $G^{(k)}_{Y\setminus Z}$, they all can be decoded reliably. For the 2nd block, due to the chaining construction in the encoding procedure, the remaining bits Bob needs to decode are also in the sets $G^{(k)}_{Y\wedge Z}$ or $G^{(k)}_{Y\setminus Z}$. Therefore, the correct decoding can also be guaranteed. The same procedure holds to the mth block. Since the confusion message and the secret message can be decoded reliably, we can guarantee that the rate in (55) can be achieved.

Theorem 2: For any $\beta \in (0, 1/2)$, there exists an *m-chain* polar coding scheme developed in Section VI-A, such that as $n \to \infty$, the *m-chain* polar coding scheme achieves the secrecy rate pairs on the dominant face of (5) for the MAC-WTC, and the block error probability decays as $O(2^{-n^\beta})$.

The proof of reliability at Bob is similar to the proof in Section V-B. The equivocation rate calculation (proof of secrecy at Eve) is given in Section VI-B.

B. Equivocation Calculation

Following the notation given in Section V-A, we show the equivocation rate calculation. For $k \in \{1,2\}$, let $W_s^{(k)}$ and $\tilde{W}_s^{(k)}$ denote the secret message and the confusion message sent by user k. Since we put the secret message in the set $S^{(k)}$ in each block, we have $W_s^{(k)} = \cup_{1\le i\le m} U_{k,S_i^{(k)}}$. For the confusion message, $\tilde{W}_s^{(k)}$, we have $\tilde{W}_s^{(k)} = \cup_{1\le i\le m, 1\le j\le (m-1)} U_{k,G^{(k)}_{Y\wedge Zi}} U_{k,C^{(k)}_{Y\setminus Zj}}$. For simplicity of notation, we let $W_s = W_s^{(1)} \cup W_s^{(2)}$ and $\tilde{W}_s = \tilde{W}_s^{(1)} \cup \tilde{W}_s^{(2)}$.

Similar to (45)–(48), we can calculate the equivocation rate as follows:

$$\begin{aligned} H(W_s|Z^{mn}) \ge & H(W_s) + H(\tilde{W}_s) \\ & - I(V_1^{mn}, V_2^{mn}; Z^{mn}) - H(\tilde{W}_s|W_s, Z^{mn}), \end{aligned} \tag{56}$$

which is equivalent to

$$\frac{1}{mn}I(W_s;Z^{mn}) \leq \frac{1}{mn}I(V_1^{mn},V_2^{mn};Z^{mn}) + \frac{1}{mn}H(\tilde{W}_s|W_s,Z^{mn}) - \frac{1}{mn}H(\tilde{W}_s). \tag{57}$$

To bound each term in (57), we only consider the second term since the first and third terms are similar to bounding in (49). These two terms can be upper bounded by ϵ, and $\epsilon \to 0$ as $n \to \infty$ and $m \to \infty$. For the second term, suppose Eve obtains W_s and Z^{mn}, and wants to decode $\tilde{W}_s$. This time Eve decodes from the mth block to the 1st block, and note that Eve decodes according to the monotone permutation K^{2n}. For the mth block, the bits that Eve needs to decode are in the set $G^{(k)}_{Y\wedge Z}$ and $G^{(k)}_{Z\backslash Y}$. Therefore, Eve can do the correct decoding. For the $(m-1)$th block, due to the chaining construction, the remaining bits that Eve needs to decode are also in the set $G^{(k)}_{Y\wedge Z}$ and $G^{(k)}_{Z\backslash Y}$. The same procedure holds to the 1st block. Since Eve can do the correct decoding, we can bound this term through Fano's inequality. Therefore, we can guarantee the conditions in (4).

VII. Polar Coding for the Broadcast Channel With Confidential Messages

Before we show how to achieve the corner points of the rate region given in (8) by double chaining method, we briefly review the result in [10], which shows how to apply polar coding to achieve the rate pair $(R_1,R_2)=(I(V_1;Y_1),I(V_2;Y_2)-I(V_2;V_1))$ of the binning region. We first consider a constant T in (8). This result can be generalized to T with arbitrary distribution [11, Sec. III. B.]. Again, we consider binary code design for illustration.

A. Polar Coding for the Binning Region

Applying polar coding to achieve $R_1 = I(V_1;Y_1)$ is described in Section IV-A. Now, we discuss how to achieve $R_2 = I(V_2;Y_2)-I(V_2;V_1)$ following [10]. Let $U_2^n = V_2^n G_n$. Similar to (35), we can define $\mathcal{H}_{V_2}$ and $\mathcal{L}_{V_2|Y_2}$. Since V_1 and V_2 are dependent, by thinking of V_1 as the side information of V_2, we can further define the set $\mathcal{L}_{V_2|V_1}$. Similar to (36), the set $[n]$ can be partitioned into the following sets:

$$\begin{aligned}
G_{Y_2\wedge V_1} &= \mathcal{H}_{V_2}\cap\mathcal{L}_{V_2|Y_2}\cap\mathcal{L}_{V_2|V_1},\\
G_{Y_2\backslash V_1} &= \mathcal{H}_{V_2}\cap\mathcal{L}_{V_2|Y_2}\cap\mathcal{L}^c_{V_2|V_1},\\
G_{V_1\backslash Y_2} &= \mathcal{H}_{V_2}\cap\mathcal{L}^c_{V_2|Y_2}\cap\mathcal{L}_{V_2|V_1},\\
B_{Y_2\wedge V_1} &= \mathcal{H}^c_{V_2}\cup(\mathcal{L}^c_{V_2|Y_2}\cap\mathcal{L}^c_{V_2|V_1}).
\end{aligned} \tag{58}$$

Roughly speaking, once the values for V_1 is known, the bits corresponding to the sets $G_{Y_2\wedge V_1}$ and $G_{V_1\backslash Y_2}$ can be determined. Since the second receiver observes Y_2, it can decode the set $G_{Y_2\wedge V_1}$ and $G_{Y_2\backslash V_1}$. To guarantee that the second receiver obtains the information bits in the set $G_{V_1\backslash Y_2}$, pick $C_{Y_2\backslash V_1}\subset G_{Y_2\backslash V_1}$ such that $|C_{Y_2\backslash V_1}| = |G_{V_1\backslash Y_2}|$ to serve the chaining purpose of repeating the information in the set $G_{V_1\backslash Y_2}$. Last, we put the information bits for the second user in the set $S = G_{Y_2\backslash V_1}\setminus C_{Y_2\backslash V_1}$. It can be verified that the rate of the second user is:

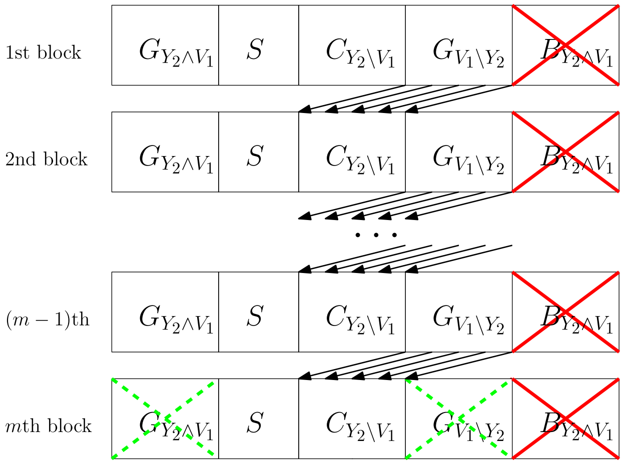


Fig. 4. Chaining construction for the second user to achieve the binning region in a broadcast channel.

$$\lim_{n\to\infty}\frac{1}{n}|S| = I(V_2;Y_2)-I(V_2;V_1). \tag{59}$$

Consider the encoding procedure in Fig. 4. The information for the first receiver, V_1, is determined first. Since V_1 has been determined, the set $G_{Y_2\wedge V_1}$ and $G_{V_1\backslash Y_2}$ can also be determined from the 1st block to the mth block. It is important to note that V_1 in the mth block is frozen and shared with the two receivers; therefore, the sets $G_{Y_2\wedge V_1}$ and $G_{V_1\backslash Y_2}$ can be decoded with the information of V_1 for the mth block, which we use dashed green crosses to denote in Fig. 4. Same as before, the red crosses denote the frozen sets in Fig. 4. By the chaining construction, for $1\leq i<m$, we repeat the determined value in the set $G_{V_1\backslash Y_2}$ in the ith block to the set $C_{Y_2\backslash V_1}$ in the $(i+1)$th block. Last, we put the information bits for the second receiver in the set S in each block.

Decoding procedure for the second receiver starts from the mth block. For the mth block, the second user only needs to decode the information in the set S and $C_{Y_2\backslash V_1}$. To decode the $(m-1)$th block, since the bits in the set $G_{V_1\backslash Y_2}$ can be obtained from the mth block due to the chaining construction of the encoding process, the second user only needs to decode the bits in the set $G_{Y_2\wedge V_1}$ and $G_{Y_2\backslash V_1}$. The same procedure holds till the 1st block, and the information in the set S can be decoded reliably.

B. The Scheme

Here, we introduce a *double chaining* method to achieve the *double binning* rate pair $(R_1,R_2)=(I(V_1;Y_1)-I(V_1;V_2)-I(V_1;Y_2|V_2), I(V_2;Y_2)-I(V_2;V_1)-I(V_2;Y_1|V_1))$, which is the corner point of (8) when T is a constant. Let $U_2^n = V_2^n G_n$. Once we determine the distribution in (9), we can define $\mathcal{H}_{V_2}$, $\mathcal{L}_{V_2|Y_2}$ and $\mathcal{L}_{V_2|V_1}$. We can further define $\mathcal{L}_{V_2|Y_1,V_1}$ as in Section VII-A. The set $[n]$ can be partitioned into the following sets:

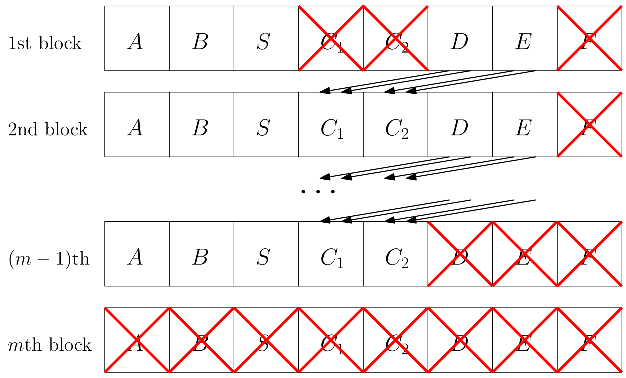


Fig. 5. Chaining construction for the BC-CM for user 1.

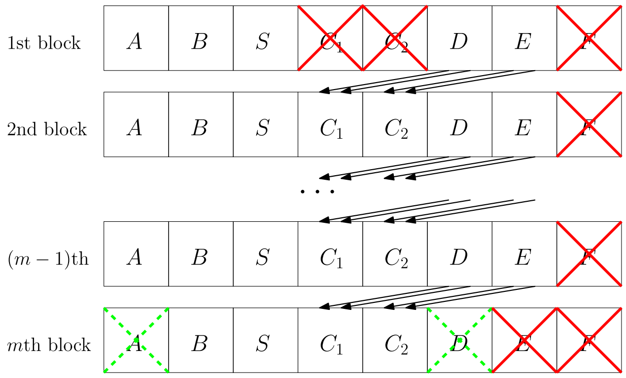


Fig. 6. Chaining construction for the BC-CM for user 2.

$$\begin{aligned}
A &= \mathcal{H}_{V_2} \cap \mathcal{L}_{V_2|Y_2} \cap \mathcal{L}_{V_2|V_1} \cap \mathcal{L}_{V_2|V_1,Y_1},\\
B &= \mathcal{H}_{V_2} \cap \mathcal{L}_{V_2|Y_2} \cap \mathcal{L}^c_{V_2|V_1} \cap \mathcal{L}_{V_2|V_1,Y_1},\\
C &= \mathcal{H}_{V_2} \cap \mathcal{L}_{V_2|Y_2} \cap \mathcal{L}^c_{V_2|V_1} \cap \mathcal{L}^c_{V_2|V_1,Y_1},\\
D &= \mathcal{H}_{V_2} \cap \mathcal{L}^c_{V_2|Y_2} \cap \mathcal{L}_{V_2|V_1} \cap \mathcal{L}_{V_2|V_1,Y_1},\\
E &= \mathcal{H}_{V_2} \cap \mathcal{L}^c_{V_2|Y_2} \cap \mathcal{L}^c_{V_2|V_1} \cap \mathcal{L}_{V_2|V_1,Y_1},\\
F &= \mathcal{H}^c_{V_2} \cup (\mathcal{L}^c_{V_2|Y_2} \cap \mathcal{L}^c_{V_2|V_1} \cap \mathcal{L}^c_{V_2|V_1,Y_1}). \qquad (60)
\end{aligned}$$

Similarly, let $U_1^n = V_1^n G_n$. We can partition the set $[n]$ for user 1 as (60) by changing the subscript 2 to 1 and 1 to 2.

Similar to (37) and (38), we have

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n}|A \cup B \cup C| = I(V_2; Y_2),\\
&\lim_{n\to\infty} \frac{1}{n}|A \cup D| = I(V_2; V_1),\\
&\lim_{n\to\infty} \frac{1}{n}|B \cup E| = I(V_2; Y_1|V_1). \qquad (61)
\end{aligned}$$

Here, we consider the case $R_1 > 0$ and $R_2 > 0$. Therefore, we can pick $C_1 \subset C$ with $|C_1| = |D|$, $C_2 \subset C$ with $|C_2| = |E|$, and $C_1 \cap C_2 = \emptyset$. Define the set S as follows:

$$S = C \setminus (C_1 \cup C_2). \qquad (62)$$

By (61), we also have

$$\lim_{n\to\infty} \frac{1}{n}|S| = I(V_2; Y_2) - I(V_2; V_1) - I(V_2; Y_1|V_1). \qquad (63)$$

Now, we consider the encoding procedure. Assume we determine the information for the first receiver, V_1, at first. As described in Section VII-A, to guarantee the correct decoding of the second user, V_1 in the mth block is frozen and shared with the two receivers. As shown in Fig. 5, the red crosses denote the frozen sets. We put the secret message in the set S from the 1st block to the $(m-1)$th block. Later, we will show that the rate

$$R_1 = \left(\frac{m-1}{m}\right)[I(V_1; Y_1) - I(V_1; V_2) - I(V_1; Y_2|V_2)] \qquad (64)$$

can be achieved. To guarantee the secrecy, we put the random bits in the set A, B, D and E in the 1st block. To ensure the reliability for the user 1, we chain the message in the sets D and E to the sets C_1 and C_2 in the 2nd block. The same procedure holds till the $(m-2)$th block. For the $(m-1)$th block, we still chain the sets D and E from the $(m-2)$th block to the sets C_1 and C_2; however, we freeze the set D and E in the $(m-1)$th block to guarantee correct decoding for user 1.

For the second user, we put the secret message to the set S from the 1st block to the mth block, and will show that the rate

$$R_2 = I(V_2; Y_2) - I(V_2; V_1) - I(V_2; Y_1|V_1) \qquad (65)$$

can be achieved. To guarantee the secrecy, we put the random bits to the sets B and E as the confusion message from the 1st block to the $(m-1)$th block. Since V_1 has been determined, the sets A and D can also be determined with the knowledge of V_1. For the first chaining construction, for $1 \le i < m$, we repeat the determined value in the set D in the ith block to the set C_1 in the $(i+1)$th block. For the second chaining construction, for $1 \le i < m$, we repeat the determined value in the set E in the ith block to the set C_2 in the $(i+1)$th block. As described in Section VII-A, V_1 in the mth block is frozen and shared with the two receivers; thus, the sets A and D can be decoded with the information of V_1 for the mth block, which we use dashed green crosses to denote in Fig. 6. Same as before, the red crosses denote the frozen sets in Fig. 6. For the 1st block, we freeze the sets C_1 and C_2, and for the mth block, we freeze the set E, to guarantee the reliability.

The decoding procedure for the two users are similar. They both decode from the mth block to the 1st block. Let us use user 2 for illustration. For the mth block, since user 2 knows V_1, it can decode the sets A, B, C and D. Through the chaining construction, the decoder only needs to decode the sets A, B and C in the $(m-1)$th block. The same procedure holds till the 2nd block. For the 1st block, due to the chaining construction and the frozen sets, the decoder only needs to decode the sets A, B and S, which can be done reliably.

Theorem 3: For any $\beta \in (0, 1/2)$, there exists an *m-chain* polar coding scheme developed in Section VII-B, such that as $n \to \infty$, the *m-chain* polar coding scheme achieves the secrecy rate region in (8) for the BC-CM, and the block error probability decays as $O(2^{-n^\beta})$.

The reliability and secrecy proofs for Theorem 3 are given in Section VII-C and VII-D.

C. Reliability

The block error probability of the first and second user can be upper bounded by

$$\begin{aligned} P_{e,1} \leq &(m-2) \sum_{i \in A \cup B \cup C} Z(U_{1,i}|U_1^{i-1}, Y_1^n) \\ &+ \sum_{i \in A \cup B \cup S} Z(U_{1,i}|U_1^{i-1}, Y_1^n) = O(2^{-n^{\beta}}), \\ P_{e,2} \leq &(m-2) \sum_{i \in A \cup B \cup C} Z(U_{2,i}|U_2^{i-1}, Y_2^n) \\ &+ \sum_{i \in A \cup B \cup S} Z(U_{1,i}|U_2^{i-1}, Y_2^n) \\ &+ \sum_{i \in B \cup C} Z(U_{1,i}|U_2^{i-1}, Y_2^n) = O(2^{-n^{\beta}}) \end{aligned} \tag{66}$$

for any $\beta \in (0, 1/2)$ with complexity $O(n \log n)$. Therefore, the rate pair in (64) and (65) can be achieved reliably. Thus, as $m \to \infty$, we can achieve the rate pair in (8).

D. Equivocation Calculation

Following the notation given in Section V-C, we show the equivocation calculation for receiver 2, and this result can be extended to receiver 1 by symmetry. Since we put the secret message in the set S in each block, we have $W_{s,1} = \cup_{1 \leq i < m} U_{1,S_i}$. For the confusion message, $\tilde{W}_{s,1}$, we have $\tilde{W}_{s,1} = \cup_{1 \leq i < m, 1 \leq j < (m-1)} U_{1,(A \cup B)_i} U_{1,(D \cup E)_j}$.

We can calculate the equivocation rate as follows:

$$\begin{aligned} &H(W_{s,1}|Y_2^{mn}) \\ &\geq H(W_{s,1}|Y_2^{mn}, V_2^{mn}, T^{mn}) && (67) \\ &= H(W_{s,1}, Y_2^{mn}|V_2^{mn}, T^{mn}) - H(Y_2^{mn}|V_2^{mn}, T^{mn}) && (68) \\ &= H(W_{s,1}, V_1^{mn}, Y_2^{mn}|V_2^{mn}, T^{mn}) \\ &\quad - H(V_1^{mn}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \\ &\quad - H(Y_2^{mn}|V_2^{mn}, T^{mn}) && (69) \\ &= H(W_{s,1}, V_1^{mn}|V_2^{mn}, T^{mn}) \\ &\quad + H(Y_2^{mn}|V_1^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \\ &\quad - H(V_1^{mn}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \\ &\quad - H(Y_2^{mn}|V_2^{mn}, T^{mn}) && (70) \\ &= H(W_{s,1}, V_1^{mn}|V_2^{mn}, T^{mn}) + H(Y_2^{mn}|V_1^{mn}, V_2^{mn}, T^{mn}) \\ &\quad - H(V_1^{mn}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \\ &\quad - H(Y_2^{mn}|V_2^{mn}, T^{mn}) && (71) \\ &= H(W_{s,1}, V_1^{mn}|V_2^{mn}, T^{mn}) \\ &\quad - H(V_1^{mn}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \\ &\quad - I(V_1^{mn}; Y_2^{mn}|V_2^{mn}, T^{mn}) && (72) \\ &\geq \underbrace{H(V_1^{mn}|V_2^{mn}, T^{mn})}_{\text{first term}} - \underbrace{H(V_1^{mn}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1})}_{\text{second term}} \\ &\quad - \underbrace{I(V_1^{mn}; Y_2^{mn}|V_2^{mn}, T^{mn})}_{\text{third term}} && (73) \end{aligned}$$

where (67) is due to conditioning reduces entropy, and (68), (69) and (70) are due to the chain rule of entropy. Due to the Markov chain $W_{s,1} \to (V_1^{mn}, V_2^{mn}, T^{mn}) \to Y_2^{mn}$, we have $I(W_{s,1}; Y_2^{mn}|V_1^{mn}, V_2^{mn}, T^{mn}) = 0$. Hence, (71) holds. (72) is due to the definition of conditional mutual information, and (73) is due to the chain rule of entropy.

Consider the first term in (73)

$$\begin{aligned} &H(V_1^{mn}|V_2^{mn}, T^{mn}) \\ &\quad = H(V_1^{mn}|T^{mn}) - I(V_1^{mn}; V_2^{mn}|T^{mn}). \end{aligned} \tag{74}$$

Therefore, we can lower bound the sum of the first and the third term in (73) as

$$\begin{aligned} &(m-2)nI(V_1; Y_1|T) - mnI(V_1; V_2|T) \\ &\quad - mnI(V_1; Y_2|V_2, T). \end{aligned} \tag{75}$$

For the second term, $H(V_1^{mn}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) = H(\tilde{W}_{s,1}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1})$. Suppose receiver 2 knows Y_2^{mn}, V_2^{mn} and $W_{s,1}$, and tries to decode $\tilde{W}_{s,1}$. From Fig. 5, it can decode from the 1st block to the $(m-1)$th block, and the block error probability can be upper bounded by $O(2^{-n^{\beta}})$ for $\beta \in (0, 1/2)$. By applying Fano's inequality, we have $H(\tilde{W}_{s,1}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \leq mn\epsilon$. After we bound the three terms as above, we have

$$\begin{aligned} &\frac{1}{mn} H(W_{s,1}|Y_2^{mn}) \\ &\quad \geq \left(1 - \frac{2}{m}\right) I(V_1; Y_1|T) - I(V_1; V_2|T) \\ &\qquad - I(V_1; Y_2|V_2, T) - \epsilon. \end{aligned} \tag{76}$$

Therefore, as $n \to \infty$ and $m \to \infty$, the secrecy constraints in (7) hold.

VIII. Polar Coding for the Interference Channel With Confidential Messages

In the following, we show how to achieve the corner points of the rate region given in (11). By simple modification, this method can achieve the entire rate region. Note that given T, V_1 and V_2 are independent as in (12). Therefore, achieving (11) is also equivalent to achieving the rate pair $(R_1, R_2) = (I(V_1; Y_1) - I(V_1; Y_2, V_2), I(V_2; Y_2) - I(V_2; Y_1, V_1))$. We consider a constant T in (11), and binary code design for illustration.

A. The Scheme

Here, we discuss the code design for user 1 only, as the code design method for the two users is similar. Let $U_1^n = V_1^n G_n$. Once we determine the distribution in (12), similar to (35), we can define $\mathcal{H}_{V_1}$ and $\mathcal{L}_{V_1|Y_1}$. We can further define

$$\mathcal{L}_{V_1|Y_2,V_2} = \{i \in [n] : Z(U_{1,i}|U_1^{i-1}, Y_2^n, V_2^n) \leq \delta_n\}, \tag{77}$$

where $\delta_n = 2^{-n^{\beta}}$ and $\beta \in (0, 1/2)$.

By thinking of Y_1 as Y and $[Y_2, V_2]$ as Z in (36), we can partition the set $[n]$ into the following:

$$\begin{aligned} G_{Y_1\wedge[Y_2,V_2]} &= \mathcal{H}_{V_1} \cap \mathcal{L}_{V_1|Y_1} \cap \mathcal{L}_{V_1|[Y_2,V_2]}, \\ G_{Y_1\setminus[Y_2,V_2]} &= \mathcal{H}_{V_1} \cap \mathcal{L}_{V_1|Y_1} \cap \mathcal{L}^c_{V_1|[Y_2,V_2]}, \\ G_{[Y_2,V_2]\setminus Y_1} &= \mathcal{H}_{V_1} \cap \mathcal{L}^c_{V_1|Y_1} \cap \mathcal{L}_{V_1|[Y_2,V_2]}, \\ B_{Y_1\wedge[Y_2,V_2]} &= \mathcal{H}^c_{V_1} \cup (\mathcal{L}^c_{V_1|Y_1} \cap \mathcal{L}^c_{V_1|[Y_2,V_2]}). \end{aligned} \tag{78}$$

Similar to (37), we also have

$$\begin{aligned} \mathcal{I}_{Y_1} &= \mathcal{H}_{V_1} \cap \mathcal{L}_{V_1|Y_1}, \\ \mathcal{I}_{[Y_2,V_2]} &= \mathcal{H}_{V_1} \cap \mathcal{L}_{V_1|[Y_2,V_2]}, \\ \mathcal{F}_r^{Y_1} &= \mathcal{H}_{V_1} \cap \mathcal{L}^c_{V_1|Y_1}, \\ \mathcal{F}_r^{[Y_2,V_2]} &= \mathcal{H}_{V_1} \cap \mathcal{L}^c_{V_1|[Y_2,V_2]}, \\ \mathcal{F}_d &= \mathcal{H}^c_{V_1}. \end{aligned} \tag{79}$$

Same as (38), we have

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n}|\mathcal{I}_{Y_1}| &= I(V_1; Y_1), \\ \lim_{n\to\infty} \frac{1}{n}|\mathcal{I}_{[Y_2,V_2]}| &= I(V_1; Y_2, V_2). \end{aligned} \tag{80}$$

Here, we consider the case $R_1 > 0$; therefore, we have $|G_{Y_1\setminus[Y_2,V_2]}| > |G_{[Y_2,V_2]\setminus Y_1}|$. Pick a set, $C_{Y_1\setminus[Y_2,V_2]}$, such that $C_{Y_1\setminus[Y_2,V_2]} \subset G_{Y_1\setminus[Y_2,V_2]}$ and $|C_{Y_1\setminus[Y_2,V_2]}| = |G_{[Y_2,V_2]\setminus Y_1}|$. Last, we define the set S similar to (39) as

$$S = G_{Y_1\setminus[Y_2,V_2]} \setminus C_{Y_1\setminus[Y_2,V_2]}. \tag{81}$$

From (80), we have

$$\lim_{n\to\infty} \frac{1}{n}|S| = I(V_1; Y_1) - I(V_1; Y_2, V_2). \tag{82}$$

The polar coding scheme construction for IC-CM is almost the same as the code design for the wiretap channel in Section V-A. By replacing Y by Y_1 and Z by $[Y_2, V_2]$ in Section V-A, we can construct the codebook for user 1 shown in Fig. 7, where the red crosses indicate that the sub-channels are frozen. Same as before, we put the secret message in the set S, and put the random bits in the sets $G_{Y_1\wedge[Y_2,V_2]}$ and $C_{Y_1\setminus[Y_2,V_2]}$ as the confusion message. By replacing U by U_1, $U_{\mathcal{F}_r^Y}$ by $U_{1,\mathcal{F}_r^{Y_1}}$, and $U_{\mathcal{F}_r^Z}$ by $U_{1,\mathcal{F}_r^{[Y_2,V_2]}}$ as defined in (79), we can follow the same encoding and decoding procedures given in Section V-A. The secrecy rate $R_1 = I(V_1; Y_1) - I(V_1; Y_2, V_2)$ can be achieved reliably since the secret message in the set S can be correctly decoded as described in Section V-B, where the set S ensures the rate given in (82).

Theorem 4: For any $\beta \in (0, 1/2)$, there exists an *m-chain* polar coding scheme developed in Section VIII-A, such that as $n \to \infty$, the *m-chain* polar coding scheme achieves the secrecy rate region in (11) for the IC-CM, and the block error probability decays as $O(2^{-n^\beta})$.

The proof reliability at the receivers is similar to the proof in Section V-B. The equivocation rate calculation (proof of secrecy) is given in Section VIII-B.

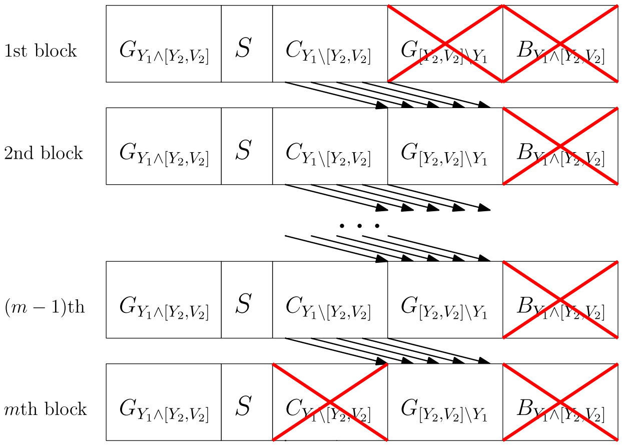


Fig. 7. Chaining construction for the IC-CM for user 1.

B. Equivocation Calculation

Following the notation given in Section V-C, we show the equivocation calculation for receiver 2, and this result can be extended to receiver 1 by symmetry. Since we put the secret message in the set S in each block, we have $W_{s,1} = \cup_{1\le i\le m} U_{1,S_i}$. For the confusion message, $\tilde{W}_{s,1}$, we have $\tilde{W}_{s,1} = \cup_{1\le i\le m, 1\le j<m} U_{1,G_{Y_1\wedge[Y_2,V_2]i}} U_{1,C_{Y_1\setminus[Y_2,V_2]j}}$.

We can calculate the equivocation rate as follows (see (67)–(73)):

$$\begin{aligned} H(W_{s,1}|Y_2^{mn}) \geq & H(V_1^{mn}|V_2^{mn}, T^{mn}) \\ & - H(V_1^{mn}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \\ & - I(V_1^{mn}; Y_2^{mn}|V_2^{mn}, T^{mn}). \end{aligned} \tag{83}$$

Now, we discuss each term in (83). Since given $T^{mn} = t^{mn}$, V_1^{mn} and V_2^{mn} are independent, we have $H(V_1^{mn}|V_2^{mn}, T^{mn}) = H(V_1^{mn}|T^{mn})$, and $I(V_1^{mn}; Y_2^{mn}|V_2^{mn}, T^{mn}) = I(V_1^{mn}; Y_2^{mn}, V_2^{mn}|T^{mn})$. Then, we can lower bound the sum of the first and third term as

$$(m-1)nI(V_1; Y_1|T) - mnI(V_1; Y_2, V_2|T). \tag{84}$$

For the second term, $H(V_1^{mn}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) = H(\tilde{W}_{s,1}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1})$. Suppose receiver 2 knows Y_2^{mn}, V_2^{mn} and $W_{s,1}$, and tries to decode $\tilde{W}_{s,1}$. From Fig. 7, it can decode from the mth block to the 1st block, and the block error probability can be upper bounded by

$$\begin{aligned} P_e \leq & (m-1) \sum_{i\in G_{[Y_2,V_2]\setminus Y_1}} Z(U_{1,i}|U_1^{i-1}, Y_2^n) \\ & + m \sum_{i\in G_{Y_1\wedge[Y_2,V_2]}} Z(U_{1,i}|U_1^{i-1}, Y_2^n) = O(2^{-n^\beta}) \end{aligned} \tag{85}$$

for $\beta \in (0, 1/2)$. Hence, by applying Fano's inequality, we have

$$\begin{aligned} & H(\tilde{W}_{s,1}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \\ & \quad \leq H(P_e) + P_e \log|\tilde{W}_s| \\ & \quad < H(P_e) + P_e[mnI(V_1; Y_2, V_2|T)]. \end{aligned} \tag{86}$$

Therefore, as $n \to \infty$, $H(\tilde{W}_{s,1}|Y_2^{mn}, V_2^{mn}, T^{mn}, W_{s,1}) \to 0$.

Finally, considering (84) and (86), we know that as $n \to \infty$ and $m \to \infty$, the secrecy constraints in (10) hold.

IX. Conclusion

We propose practical coding schemes based on polar coding for the general wiretap channel, multiple access wiretap channel (MAC-WTC), broadcast channel with confidential messages (BC-CM), and interference channel with confidential messages (IC-CM). By applying the chaining construction and polar coding for asymmetric channels, we propose a polar coding scheme to achieve the secrecy capacity of the general wiretap channel. Compared to the previous work, our construction has better decoding error probability and it can be constructed more efficiently. For the MAC-WTC, we combine our coding scheme for the general wiretap channel with the technique of monotone chain rule. For the BC-CM, we introduce double chaining construction to guarantee the secrecy and achieve the binning rate. For the IC-CM, we view the output of the channel as the actual output and the intended message carrying signal, and apply our coding scheme for the general wiretap channel.

References

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[3] 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.

[4] E. Arıkan, "Source polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010.

[5] S. Korada and R. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751–1768, Apr. 2010.

[6] E. Şaşoğlu, İ. 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.

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

[8] S. Önay, "Successive cancellation decoding of polar codes for the two-user binary-input MAC," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013.

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

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

[11] L. Wang and E. Şaşoğlu, "Polar coding for interference networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 29/Jul. 4, 2014, pp. 311–315.

[12] E. Arıkan, "Polar coding for the Slepian-Wolf problem based on monotone chain rules," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012.

[13] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, Dept. Comput. Commun. Sci., EPFL, Lausanne, Switzerland, May 2009.

[14] 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.

[15] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested polar codes for wiretap and relay channels," *IEEE Commun. Lett.*, vol. 14, no. 8, pp. 752–754, Aug. 2010.

[16] O. O. Koyluoglu and H. E. Gamal, "Polar coding for secure transmission and key agreement," *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 5, pp. 1472–1483, Sep. 2012.

[17] E. Hof and S. Shamai, "Secrecy-achieving polar-coding," in *Proc. IEEE Inf. Theory Workshop*, Aug. 2010.

[18] S. H. Hassani, S. Korada, and R. Urbanke, "The compound capacity of polar codes," in *Proc. 47th Annu. Allerton Conf. Commun. Control Comput.*, Sep. 2009.

[19] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6562–6582, Oct. 2013.

[20] D. Sutter and J. M. Renes, "Universal polar codes for more capable and less noisy channels and sources," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014.

[21] S. H. Hassani and R. Urbanke, "Universal polar codes," in *Proc. IEEE Int. Symp. Inf. Theory Proc.*, Jun. 2014.

[22] E. Şaşoğlu and L. Wang, "Universal polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014.

[23] 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.

[24] 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*. New York, NY, USA: Springer, 2013, pp. 194–213.

[25] E. Tekin and A. Yener, "The Gaussian multiple access wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5747–5755, Dec. 2008.

[26] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[27] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[28] Y. Liang *et al.*, "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, 2009.

[29] E. Ekrem and S. Ulukus, "Cooperative secrecy in wireless communications," in *Securing Wireless Communications at the Physical Layer*, W. Trappe and R. Liu, Eds. New York, NY, USA: Springer, 2009, pp. 143–172.

[30] R. Bassily *et al.* "Cooperative security at the physical layer: A summary of recent advances," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 16–28, Sep. 2013.

[31] T. C. Gulcu and A. Barg, "Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component," in *Proc. IEEE Inf. Theory Workshop*, Apr. 26/May 1, 2015, pp. 1–5.

[32] Y.-P. Wei and S. Ulukus, "Polar coding for the general wiretap channel," in *Proc. IEEE Inf. Theory Workshop*, Apr. 26/May 1, 2015, pp. 1–5.

[33] R. A. Chou and M. R. Bloch, "Polar coding for the broadcast channel with confidential messages and constrained randomization," http://arxiv.org/abs/1411.0281, Nov. 2014.

[34] 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.

[35] E. Ekrem and S. Ulukus, "On the secrecy of multiple access wiretap channel," in *Proc. 46th Annu. Allerton Conf. Commun. Control Comput.*, Sep. 2008.

[36] R. Bassily and S. Ulukus, "Ergodic secret alignment," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1594–1611, Mar. 2012.

[37] M. Ye and A. Barg, "Polar codes for distributed hierarchical source coding," *Adv. Math. Commun.*, vol. 9, no. 1, pp. 87–103, Feb. 2015.

[38] O. Ozel and S. Ulukus, "Wiretap channels: Implications of the more capable condition and cyclic shift symmetry," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2153–2164, Apr. 2013.

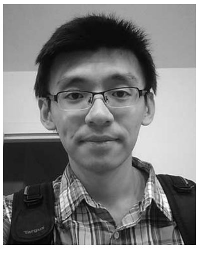
[39] E. Şaşoğlu and I. E. Telatar, "Polarization for arbitrary discrete memoryless channels," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2009.

[40] R. Mori and T. Tanaka, "Channel polarization on *q*-ary discrete memoryless channels by arbitrary kernel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010.

[41] E. Şaşoğlu, "Polar codes for discrete alphabets," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012.

[42] W. Park and A. Barg, "Polar codes for *q*-ary channels $q = 2^r$," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 955–969, Feb. 2013.

[43] L. Wang, E. Şaşoğlu, and Y.-H. Kim, "Sliding-window superposition coding for interference networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014.



Yi-Peng Wei (S'15) received the B.Sc. degree in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, and the M.Sc. degree in communication engineering from the National Taiwan University, Taipei, Taiwan, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree in electrical and computer engineering at the University of Maryland, College Park, MD, USA. His research interests include information theoretic physical layer security.

Sennur Ulukus (S'90–M'98–SM'15–F'16) received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, and the Ph.D. degree in electrical and computer Engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, New Brunswick, NJ, USA. She is a Professor of Electrical and Computer Engineering with the University of Maryland at College Park, College Park (UMD), MD, USA, where she also holds a joint appointment with the Institute for Systems Research (ISR). Prior to joining UMD, she was a Senior Technical Staff Member at AT&T Labs-Research. Her research interests include wireless communications, information theory, signal processing, networking, information theoretic physical layer security, and energy harvesting communications. She served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY (2007–2010) and the IEEE TRANSACTIONS ON COMMUNICATIONS (2003–2007). She served as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the special issue on wireless communications powered by energy harvesting and wireless energy transfer (2015), *Journal of Communications and Networks* for the special issue on energy harvesting in wireless networks (2012), the IEEE TRANSACTIONS ON INFORMATION THEORY for the special issue on interference networks (2011), the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the special issue on multiuser detection for advanced communication systems and networks (2008). She was the recipient of the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, the 2005 NSF CAREER Award, the 2010–2011 ISR Outstanding Systems Engineering Faculty Award, and the 2012 George Corcoran Education Award.