Finite-Length Analysis of Polar Secrecy Codes for Wiretap Channels

Hessam Mahdavifar and Fariba Abbasi

***Abstract*— In a classical wiretap channel setting, Alice communicates with Bob through a main communication channel, while her transmission also reaches an eavesdropper Eve through a wiretap channel. In this paper, we consider a general class of polar secrecy codes for wiretap channels and study their finite-length performance. In particular, bounds on the normalized mutual information security (MIS) leakage, a fundamental measure of secrecy in information-theoretic security frameworks, are presented for polar secrecy codes. The bounds are utilized to characterize the finite-length scaling behavior of polar secrecy codes, where scaling here refers to the non-asymptotic behavior of both the gap to the secrecy capacity as well as the MIS leakage. Furthermore, the bounds are shown to facilitate characterizing numerical bounds on the secrecy guarantees of polar secrecy codes in finite block lengths of practical relevance, where directly calculating the MIS leakage is in general infeasible.**

I. Introduction

The wiretap channel model was first introduced by Wyner in 1975 [1]. Wyner's seminal work paved the way for an entire area of research encompassing hundreds of papers by now, often referred to as physical layer security. In Wyner's setting, demonstrated in Figure 1, a transmitter (Alice) communicates with a legitimate receiver (Bob) through a main communication channel W^*, while her transmission also reaches an eavesdropper Eve through a wiretap channel W. The goal is to design a coding scheme that makes it possible for Alice to communicate both reliably and securely. Reliability is measured in terms of Bob's probability of error in estimating the message $\boldsymbol{U}$, denoted by $\hat{\boldsymbol{U}}$, same as in a classical channel coding problem. The security condition is expressed in terms of the mutual information between the message $\boldsymbol{U}$ and Eve's observation $\boldsymbol{Z}$, often referred to as the mutual information security (MIS) leakage. More specifically, the *weak* secrecy condition, introduced by Wyner [1], requires that $\lim_{k\to\infty} I(\boldsymbol{U};\boldsymbol{Z})/k = 0$, where k is the length of the message $\boldsymbol{U}$. Wyner characterized the fundamental limit on the communication rate in this regime, referred to as the *secrecy capacity*. Later, Maurer extended the security condition to strong secrecy, that is, to require that $\lim_{k\to\infty} I(\boldsymbol{U};\boldsymbol{Z}) = 0$ [2]. Remarkably, Maurer showed that the secrecy capacity remains the same under strong secrecy.

H. Mahdavifar is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 (email: h.mahdavifar@northeastern.edu.) F. Abbasi is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48104 (email: fabbasia@umich.edu.)

This work was supported in part by the Center for Ubiquitous Connectivity (CUbiC), sponsored by Semiconductor Research Corporation (SRC) and Defense Advanced Research Projects Agency (DARPA) under the JUMP 2.0 program.

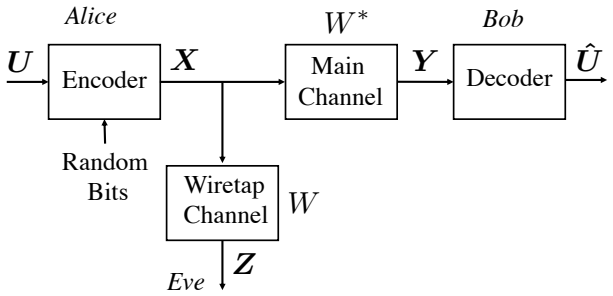

Fig. 1. Wyner's wiretap channel setting.

Polar secrecy codes, built upon Arıkan's polar codes [3], were first introduced in [4] to achieve the secrecy capacity of wiretap channels when W and W^* are both binary-input memoryless symmetric (BMS) channels and W is *degraded with respect to* W^*. The latter means that W can be written as a cascade of W^* with another channel. Under this assumption the secrecy capacity is $C_s = C(W^*) - C(W)$, where $C(.)$ is the channel capacity [1]. Note that the results in [4] are established under the weak secrecy condition. A scheme together with empirical observations are provided in [4] for strong secrecy, however, it remains an open problem whether *plain* polar codes are strong enough to achieve the secrecy capacity with strong secrecy.

There have been various extensions of the scheme in [4], e.g., to establish different variations of secrecy [5]–[7]. It has also found applications in quantum communication [8], and quantum key distribution [9], [10]. Also, an extension of this scheme together with a *chaining* constructions is used in [11] to establish secrecy capacity-achieving property, under the same assumptions as in [4], with strong secrecy. Due to the nature of the chaining construction in [11], the results are only asymptotic and the techniques are not *strong* enough to characterize the *scaling* behavior of the chained polar secrecy codes, i.e., how fast the gap-to-secrecy capacity and the MIS leakage of the scheme approach zero.

In this paper, we go back to the basic scheme of [4], which only involves *plain* polar coding, i.e., no concatenation or chaining constructions are applied on top of the scheme. We provide new bounds on the MIS leakage of polar secrecy codes. We show how these bounds, together with leveraging existing results on finite-length scaling of polar codes by Hassani *et al.* [12], lead to a characterization of finite-length scaling of polar secrecy codes. In other words, we characterize how the gap to secrecy capacity of polar secrecy codes as well as their

MIS leakage scale with the block length as the rate approaches the secrecy capacity. The bounds presented in this paper not only enables us to characterize the finite-length scaling of polar secrecy codes, but also to numerically bound the actual MIS leakage in practical settings.

The rest of this paper is organized as follows. In Section II some preliminaries are presented for polar codes and secrecy codes. In Section III, we present new bounds on the MIS leakage of polar secrecy codes. Finite-length scaling behavior of polar secrecy codes are characterized in Section IV, and some numerical results are presented in Section V.

II. Preliminaries

A. Notation Convention

First, let us briefly introduce the notations. For $n \in \mathbb{N}$, let $[n] := \{1, 2, \dots, n\}$. The parameter n is reserved for the code block length throughout the paper. Vectors are denoted by bold lowercase letters. For a vector $\boldsymbol{v} = (v_1, v_2, \dots, v_n)$, and $\mathcal{A} \in [n]$, let $\boldsymbol{v}_\mathcal{A}$ denote the sub-vector of $\boldsymbol{v}$ with entries v_i's for $i \in \mathcal{A}$. Following the convention, uppercase letters represent random variables and lowercase letters represent the instances of random variables, e.g., $\boldsymbol{U} = (U_1, U_2, \dots, U_n)$ represents a vector of random variables U_i's and $\boldsymbol{u}$ represents the instances u_i's of random variables U_i's.

Let G denote the 2×2 matrix $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ throughout this paper. Let $n = 2^m$, for some positive integer m. Then $G^{\otimes m}$ denotes the m-th Kronecker power of G, which is an $n \times n$ matrix.

B. Polar Codes

Polar codes belong to a general family of codes whose generator matrices, for the block length $n = 2^m$, are sub-matrices of $G^{\otimes m}$. Any code $\mathcal{C}$ belonging to this family can be essentially specified by a set of *information* bits $\mathcal{A} \subseteq [n]$. In other words, $\mathcal{A}$ denotes the set of indices of the rows of $G^{\otimes m}$ to be included in the generator matrix of $\mathcal{C}$. Alternatively, $\mathcal{C}$ can be specified by $\mathcal{B} = \mathcal{A}^c$, where, in the context of polar coding, $\mathcal{B}$ is referred to as the set of frozen indices. To unify the terminology, we simply refer to any code in this family as a polar code. More specifically, we refer to $\mathcal{C}$ as an (n, k) polar code associated with $\mathcal{A}$, where $k = |\mathcal{A}|$.

The encoder for an (n, k) binary polar code $\mathcal{C}$ associated with $\mathcal{A}$ is as follows. The input to the encoder is $\boldsymbol{u} \in \{0,1\}^k$. The encoder first forms the vector $\boldsymbol{v} \in \{0,1\}^n$, by setting $\boldsymbol{v}_\mathcal{A} = \boldsymbol{u}$ and $\boldsymbol{v}_\mathcal{B} = \boldsymbol{0}$. Then, the encoder outputs $\boldsymbol{x} = \boldsymbol{v}G^{\otimes m}$.

Let us recall the terminology of *good* bit-channels from the polar coding literature. In general, the polarization transform, introduced by Arıkan [3], with n inputs and n outputs over a BMS channel W is split into n synthesized bit-channels, denoted by $W_1, W_2, \dots, W_n$. Roghly speaking, the i-th bit-channel W_i is the channel that the i-th input bit observes in the polarization transform. The *quality* of these bit-channels, i.e., a metric to measure how *good* they are, is defined through certain underlying parameters of the bit-channels, e.g., the Bhattacharyya parameter. More specifically, the set of *good* bit-channels in the polarization transform, with an underlying threshold ϵ, is denoted by $\mathcal{G}_n(W, \epsilon)$ and is defined as

$$\mathcal{G}_n(W, \epsilon) \stackrel{\text{def}}{=} \{i : Z(W_i) < \epsilon\}, \tag{1}$$

where $Z(.)$ is the Bhattacharyya parameter of the channel.

C. Polar Secrecy Codes

The general form of the scheme in [4], which we refer to as a polar secrecy code (or the plain scheme, as referred to earlier), is specified by splitting $[n]$, where n is the block length, into three subsets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{R}$ of indices for information bits, frozen bits, and random bits, respectively. Let $|\mathcal{A}| = k$, and $|\mathcal{R}| = r$. Consequently, $|\mathcal{B}| = n - k - r$. In particular, a polar secrecy code is defined through its encoder, formally defined as follows:

Definition 1. *The encoder for the polar secrecy code of length n associated with $(\mathcal{A}, \mathcal{R})$ is a function $\mathcal{E} : \{0,1\}^k \times \{0,1\}^r \to \{0,1\}^n$. It accepts as input a message $\boldsymbol{u} \in \{0,1\}^k$ and a vector $\boldsymbol{e} \in \{0,1\}^r$. It is assumed that the entries of $\boldsymbol{e}$ are selected independently and uniformly at random by the encoder. The encoder first forms the vector $\boldsymbol{v} \in \{0,1\}^n$, by setting $\boldsymbol{v}_\mathcal{A} = \boldsymbol{u}$, $\boldsymbol{v}_\mathcal{R} = \boldsymbol{e}$, and $\boldsymbol{v}_\mathcal{B} = \boldsymbol{0}$. The encoder then outputs $\mathcal{E}(\boldsymbol{u}, \boldsymbol{e}) := \boldsymbol{v}G^{\otimes m}$. The secrecy rate of this scheme is $R_s := \frac{k}{n}$.*

An upper bound on the MIS leakage of polar secrecy codes is presented in [4, Lemma 15] which we recall here under a slight variation:

Proposition 1. *The following upper bound holds on the MIS leakage $I(\boldsymbol{U}; \boldsymbol{Z})$ of a polar secrecy code associated with $(\mathcal{A}, \mathcal{R})$, as defined in Definition 1:*

$$I(\boldsymbol{U}; \boldsymbol{Z}) \leqslant \min(\sum_{i \in \mathcal{R}^c} C_i, k), \tag{2}$$

where C_i is the capacity of the i-th bit-channel W_i in the polarization transform of W.

III. Bounds on the MIS Leakage of Polar Secrecy Codes

In this section, we first present a new lower bound on the MIS leakage $I(\boldsymbol{U}; \boldsymbol{Z})$ for polar secrecy codes in wiretap settings.

Note that while the upper bound in Proposition 1 holds regardless of the distribution on $\boldsymbol{U}$, our new lower bound holds when the input message $\boldsymbol{U}$ has a uniform distribution. Note that without any assumption on the distribution of $\boldsymbol{U}$, $I(\boldsymbol{U}; \boldsymbol{Z})$ can be as small as zero. However, in practice, the message bits often come from a uniform distribution which is a common assumption in designing digital communication systems.

Before proceeding to demonstrate and prove our lower bound on the MIS leakage, we present the following lemma which will be used in the proof of the lower bound, and could be of independent interest. The proof of lemma is omitted due to space constraints.

Lemma 2. *Consider the polarization transform of length n over the BMS channel W with input $\boldsymbol{V}$, and output $\boldsymbol{Z}$. Then for a subset of indices $\mathcal{D} \subseteq \{1, 2, \dots, n\}$, we have*

$$I(\boldsymbol{V}_\mathcal{D}; \boldsymbol{Z} | \boldsymbol{V}_{\mathcal{D}^c}) \geqslant \sum_{i \in \mathcal{D}} C_i, \tag{3}$$

where C_i is the capacity of the i-th bit-channel W_i.

Proposition 3. *The following lower bound holds on the MIS leakage $I(\boldsymbol{U};\boldsymbol{Z})$ of a polar secrecy code associated with $(\mathcal{A},\mathcal{R})$, as defined in Definition 1:*

$$I(\boldsymbol{U};\boldsymbol{Z}) \geqslant \max(\sum_{i\in\mathcal{A}} C_i + \sum_{j\in\mathcal{R}} C_j - r, 0), \tag{4}$$

where $r=|\mathcal{R}|$, and C_i is the capacity of the i-th bit-channel W_i.

Proof: First, we have

$$\begin{aligned} I(\boldsymbol{U};\boldsymbol{Z}) &= I(\boldsymbol{V}_{\mathcal{A}};\boldsymbol{Z}|\boldsymbol{V}_{\mathcal{B}}=\mathbf{0}) \\ &= I(\boldsymbol{V}_{\mathcal{A}};\boldsymbol{Z}|\boldsymbol{V}_{\mathcal{B}}=\boldsymbol{v}_{\mathcal{B}}), \end{aligned} \tag{5}$$

for any $\boldsymbol{v}_{\mathcal{B}} \in \{0,1\}^{n-k-r}$, where the equality holds by the symmetry of the channel. Hence, one can assume that $\boldsymbol{v}_{\mathcal{B}}$ is an instance of $\boldsymbol{V}_{\mathcal{B}}$ of i.i.d. uniform binary random variables with the observation given to Eve. In other words, we have

$$I(\boldsymbol{U};\boldsymbol{Z}) = I(\boldsymbol{V}_{\mathcal{A}};\boldsymbol{Z}|\boldsymbol{V}_{\mathcal{B}}). \tag{6}$$

The rest of the proof is by a series of equalities and inequalities as follows.

$$\begin{aligned} I(\boldsymbol{U};\boldsymbol{Z}) &\overset{(a)}{=} I(\boldsymbol{V}_{\mathcal{A}};\boldsymbol{Z}|\boldsymbol{V}_{\mathcal{B}}) \\ &\overset{(b)}{=} I(\boldsymbol{V}_{\mathcal{A}\cup\mathcal{R}};\boldsymbol{Z}|\mathbf{V}_{\mathbf{B}}) - I(\boldsymbol{V}_{\mathcal{R}};\boldsymbol{Z}|\mathbf{V}_{\mathcal{A}\cup\mathcal{B}}) \\ &\overset{(c)}{\geqslant} \sum_{i\in\mathcal{A}} C_i + \sum_{j\in\mathcal{R}} C_j - I(\boldsymbol{V}_{\mathcal{R}};\boldsymbol{Z}|\mathbf{V}_{\mathcal{A}\cup\mathcal{B}}) \\ &\overset{(d)}{\geqslant} \sum_{i\in\mathcal{A}} C_i + \sum_{j\in\mathcal{R}} C_j - r, \end{aligned} \tag{7}$$

where (a) holds as in (6), (b) holds by the chain rule of mutual information, (c) holds by Lemma 2, and (d) holds since $\boldsymbol{V}_{\mathcal{R}}$ is a binary vector of length r. Furthermore, we know that the mutual information $I(\boldsymbol{U};\boldsymbol{Z})$ is a non-negative value. This concludes the proof of the proposition. ∎

In the next corollary, we rearrange the terms in the lower and upper bounds in order to unify them and present them in terms of the capacity of the wiretap channel C. Note that the sum of capacities of all bit-channels in the polarization transform of length n equals nC, i.e.,

$$\sum_{i\in\mathcal{A}} C_i + \sum_{j\in\mathcal{B}} C_j + \sum_{l\in\mathcal{R}} C_l = \sum_{i=1}^{n} C_i = nC. \tag{8}$$

This leads to the following corollary which combines the results from Proposition 1 and Proposition 3.

Corollary 4. *For the MIS leakage $I(\boldsymbol{U};\boldsymbol{Z})$ we have*

$$n(C-\frac{r}{n}) - \sum_{i\in\mathcal{B}} C_i \leqslant I(\boldsymbol{U};\boldsymbol{Z}) \leqslant n(C-\frac{r}{n}) + \sum_{l\in\mathcal{R}}(1-C_l). \tag{9}$$

Note that the term $C-\frac{r}{n}$ is the gap to capacity of a polar code of length n associated with $\mathcal{R}$. There is already an extensive literature on studying this quantity, also referred to as finite-length scaling of polar codes, started by Hassani *et al.* in [12]. This is the main reason for representing the bounds as in Corollary 4 using which together with the already known results on the finite-length scaling of polar codes we will characterize the scaling behavior of the polar secrecy codes.

IV. Finite-Length Scaling of Polar Secrecy Codes

First, let us recall the known results on finite-length scaling of polar codes. In general, a widely accepted conjecture is that for a sequence of polar codes guaranteeing a probability of error upper bounded by a fixed value P_e and of increasing length n, the gap to capacity $C-\frac{r}{n}$, where $\frac{r}{n}$ is the rate and C is the capacity, scales as $n^{-1/\mu}$, for some $\mu > 2$ referred to as the scaling exponent, i.e.,

$$C - \frac{r}{n} = \alpha n^{-1/\mu} + o(n^{-1/\mu}), \tag{10}$$

for some constant $\alpha \in \mathbb{R}^+$, where μ depends on the underlying channel. This problem was first studied by Hassani *et al.* [12]. It is shown in [12] that there exists $\overline{\mu}$ and $\underline{\mu}$ such that

$$\underline{\alpha} n^{-1/\underline{\mu}} \leqslant C - \frac{r}{n} \leqslant \overline{\alpha} n^{-1/\overline{\mu}}, \tag{11}$$

where $\underline{\alpha}, \overline{\alpha} \in \mathbb{R}^+$, for large enough, yet finite, values of n. Furthermore, numerical values for $\underline{\mu}$ and $\overline{\mu}$ are characterized in [12].

A. Bounds on the Scaling Exponent

In this section, we consider polar secrecy codes associated with $(\mathcal{A},\mathcal{R})$ set as follows:

$$\mathcal{R} = \mathcal{G}_n(W,\epsilon), \text{ and } \mathcal{A} = \mathcal{G}_n(W^*,\epsilon) \setminus \mathcal{R}, \tag{12}$$

where $\epsilon = P_e/n$ and P_e is a fixed value representing a bound on the probability of error at Bob's side. Note that by the degraded assumption of the wiretap setting, we have $\mathcal{G}_n(W,\epsilon) \subseteq \mathcal{G}_n(W^*,\epsilon)$ [13, Lemma 4.7]. Recall that the number of information bits in this scheme is $k=|\mathcal{A}|$, and the secrecy rate is $R_s = k/n$.

Let $\underline{\mu}$ and $\overline{\mu}$ denote the lower bound and upper bound on the scaling exponent of polar codes over the wiretap channel W. Also, let $\underline{\alpha}$ and $\overline{\alpha}$ denote the corresponding constants. Also, let $\underline{\mu}^*$ and $\overline{\mu}^*$, $\underline{\alpha}^*$ and $\overline{\alpha}^*$ be defined similarly for the main channel W^*.

Now, utilizing (11) one can express the lower and upper bound on the MIS leakage of the above polar secrecy code, as presented in Corollary 4, as follows:

$$I(\boldsymbol{U};\boldsymbol{Z}) \geqslant \underline{\alpha} n^{1-1/\underline{\mu}} + o(n^{1-1/\underline{\mu}}) - \sum_{i\in\mathcal{B}} C_i, \tag{13}$$

and

$$I(\boldsymbol{U};\boldsymbol{Z}) \leqslant \overline{\alpha} n^{1-1/\overline{\mu}} + o(n^{1-1/\overline{\mu}}) + \sum_{l\in\mathcal{R}}(1-C_l). \tag{14}$$

Note that we prefer to use the small-o notation to avoid repeatedly saying the inequalities hold for large enough n.

In the next theorem, we characterize bounds on the finite-length scaling behavior of polar secrecy codes, where scaling here refers to the non-asymptotic behavior of both the gap to the secrecy capacity as well as the MIS leakage. More

specifically, it is shown that the gap to the secrecy capacity is upper bounded by $O(n^{-1/\overline{\mu}^*})$ and the normalized MIS leakage is upper bounded by $O(n^{-1/\overline{\mu}})$.

Theorem 5. *The polar secrecy code associated with $(\mathcal{A}, \mathcal{R})$, as specified in (12), has the following properties:*

(i) The gap to secrecy capacity:

$$C_s - R_s \leqslant \overline{\alpha}^* n^{-1/\overline{\mu}^*} + o(n^{-1/\overline{\mu}^*}). \tag{15}$$

(ii) Security condition:

$$\frac{I(\boldsymbol{U};\boldsymbol{Z})}{k} \leqslant \overline{\alpha} R_s^{-1} n^{-1/\overline{\mu}} + o(n^{-1/\overline{\mu}}). \tag{16}$$

(iii) Relibility condition: Bob's probability of error under successive cancellation decoding is upper bounded by P_e.

Proof: Note that $C_s = C(W^*) - C(W)$ for degraded wiretap channels [1], and by (12), we have

$$R_s = \frac{|\mathcal{A}|}{n} = \frac{|\mathcal{G}_n(W^*, \epsilon)|}{n} - \frac{|\mathcal{G}_n(W, \epsilon)|}{n}$$

By invoking this together with the result on the scaling exponent of polar codes, expressed in (11), for W^* we get:

$$C_s - R_s = C(W^*) - \frac{|\mathcal{G}_n(W^*, \epsilon)|}{n} - \big(C(W) - \frac{|\mathcal{G}_n(W, \epsilon)|}{n}\big) \tag{17}$$

$$\leqslant \overline{\alpha}^* n^{-1/\overline{\mu}^*} + o(n^{-1/\overline{\mu}^*}), \tag{18}$$

which completes the proof for (i).

To show (ii), we first upper bound the term $\sum_{l\in\mathcal{R}}(1 - C_l)$ that appears in the upper bound on the MIS leakage in Corollary 4. Note that for any BMS channel with capacity C_l and Bhattacharyya parameter Z_l we have [3]:

$$C_l + Z_l \geqslant 1. \tag{19}$$

Note also that the finite-length scaling results of Hassani et al. is derived by upper bounding the probability of error as the sum of Bhattacharyya parameters of the good bit-channel. More specifically, the r good bit-channels for transmitting the information bits satisfy the following for the target upper bound on the probability of error P_e:

$$\sum_{l\in\mathcal{R}} Z_l \leqslant P_e. \tag{20}$$

Hence, combining (14) together with (19) and (20) we arrive at the following upper bound on the MIS leakage:

$$\begin{aligned} I(\boldsymbol{U};\boldsymbol{Z}) &\leqslant \overline{\alpha} n^{1-1/\overline{\mu}} + o(n^{1-1/\overline{\mu}}) + \sum_{l\in\mathcal{R}} Z_l \\ &\leqslant \overline{\alpha} n^{1-1/\overline{\mu}} + o(n^{1-1/\overline{\mu}}) + P_e \\ &= \overline{\alpha} n^{1-1/\overline{\mu}} + o(n^{1-1/\overline{\mu}}). \end{aligned} \tag{21}$$

This together by noting that $R_s = k/n$ completes the proof of (ii). And, finally, (iii) is by the original result on the probability of error of polar codes under the successive cancellation (SC) decoding [3]. In fact, Bob treats the entire $\mathcal{A}\cup\mathcal{R}$ as the set of information bits and invokes SC decoder to decode them. The decoded bits on $\mathcal{R}$ are simply discarded and the decoded bits on $\mathcal{A}$ are output as the decoded message. ■

Remark 1. The discussions on this section are based upon polar coding with Arıkan's 2×2 kernel. There is an extensive research on enhancing the performance of polar codes, including improving their scaling exponent, by utilizing larger $l \times l$ kernels. In particular, it is shown in [14] that by increasing l one can approach the optimal scaling exponent of 2. Such results can be incorporated here in a straightforward fashion. More specifically, by invoking the results in [14], and by building upon arbitrarily large kernels, one can obtain polar secrecy codes with large kernels whose gap to the secrecy capacity is upper bounded by $O(n^{-1/2+\epsilon})$ and whose normalized MIS leakage is upper bounded by $O(n^{-1/2+\epsilon})$, for any fixed $\epsilon > 0$.

Remark 2. Non-asymptotic fundamental limits for wiretap channels are characterized in [15]. However, [15] considers a non-asymptotic regime where the secrecy leakage, though in terms of the total variation distance (TVD), is treated as a small constant, same as the reliability measure P_e. However, the polar secrecy codes in this section operate in a regime where the leakage also approaches zero polynomially in n. We are not aware if this regime is addressed in the literature. However, one can naturally expect that there is a sequence of random secrecy codes, together with a converse result stating they are optimal, with the gap to secrecy capacity as well as the normalized MIS leakage of $\Theta(n^{-1/2})$.

B. Operating above the Secrecy Capacity

Next, we explore an interesting situation where it is possible to operate slightly above secrecy capacity, with a gap polynomial in n and, of course, vanishing as n grows large while the weak secrecy condition is still satisfied. Note that, in general, one can operate at rates above secrecy capacity at the expense of having a non-zero normalized MIS leakage. In fact, the entire *rate-equivocation region* for degraded wiretap channels is characterized by Wyner [1]. However, to the best of authors' knowledge, how this region exactly scales when operating slightly above (i.e., with a gap vanishing in n) the secrecy capacity is not well understood.

Next, we discuss how the polar secrecy code associated with $(\mathcal{A}, \mathcal{R})$, specified in (12), can be modified to operate above the secrecy capacity while the weak secrecy condition is still satisfied. Let $\delta > \overline{\mu}^*$ be fixed. For instance any $\delta > 4.714$ would always work [16]. Then we pick an arbitrary subset $\mathcal{R}' \subset \mathcal{R}$ with $|\mathcal{R}'| = n^{1-\delta}$ and exclude it from $\mathcal{R}$. More specifically, we set the new $\mathcal{R}$ and $\mathcal{A}$ as follows:

$$\mathcal{R} = \mathcal{G}_n(W, \epsilon) \setminus \mathcal{R}', \text{ and } \mathcal{A} = \mathcal{G}_n(W^*, \epsilon) \setminus \mathcal{R}, \tag{22}$$

Theorem 6. *The polar secrecy code associated with $(\mathcal{A}, \mathcal{R})$, as specified in (22), has the following properties:*

(i) The gap to secrecy capacity:

$$R_s - C_s = n^{-1/\delta} + o(n^{-1/\delta}). \tag{23}$$

(ii) Security condition:

$$\frac{I(\boldsymbol{U};\boldsymbol{Z})}{k} \leqslant R_s^{-1} n^{-1/\delta} + o(n^{-1/\delta}). \tag{24}$$

(iii) Relibility condition: Bob's probability of error under successive cancellation decoding is upper bounded by P_e.

Proof: To show (i), we have

$$\begin{aligned} R_s - C_s &= n^{-1/\delta} + C(W) - \frac{|\mathcal{G}_n(W,\epsilon)|}{n} \\ &\quad - \big(C(W^*) - \frac{|\mathcal{G}_n(W^*,\epsilon)|}{n}\big) \\ &= n^{-1/\delta} + O(n^{-1/\overline{\mu}^*}) + O(n^{-1/\overline{\mu}}) \\ &= n^{-1/\delta} + o(n^{-1/\delta}). \end{aligned}$$

which completes the proof for (i).

To show (ii), by Corollary 4 we have

$$I(\boldsymbol{U};\boldsymbol{Z}) \leqslant n(C(W) - \frac{|\mathcal{R}|}{n}) + \sum_{l\in\mathcal{R}}(1-C_l)$$

Note that the dominating term in $C(W) - \frac{|\mathcal{R}|}{n}$ is $n^{-1/\delta}$. Also, same is as in the proof of Theorem 5, we have $\sum_{l\in\mathcal{R}}(1-C_l) < P_e$. This completes the proof of (ii). Finally, the argument for part (iii) is exactly same as that of the part (iii) of Theorem 5. ■

V. Numerical Results

Most existing studies on information-theoretic security, such as wiretap channels and secret key generation, primarily focus on asymptotic results. This involves establishing fundamental limits on secrecy rates or devising schemes to approach these limits in the asymptotic sense. However, to validate the practicality of these schemes, it is also critical to present numerical results regarding security guarantees of these schemes. This approach mirrors how performance curves, showing decoder error rates, are frequently showcased and compared to ensure the reliability guarantees of channel coding techniques. There are only a few prior work that have shown finite-length secrecy performance of codes, in terms of MIS leakage or other security metrics. Some numerical results for the polar secrecy codes are presented in [4] for large block lengths. More recently, [17] considers the special case where main channel is noiseless and the wiretap channel is BEC and present numerical bounds on the TVD secrecy measure of certain polar and Reed-Muller secrecy codes. In another related work, numerical results on the secrecy performance of randomized convolutional codes for the wiretap channel are presented in [18].

The upper and lower bounds presented in this paper, summarized in Corollary 4, not only enables us to characterize the finite-length scaling of polar secrecy codes, as discussed in Section IV, but also to numerically bound the actual MIS leakage in practical settings. Note that directly computing the MIS leakage $I(\boldsymbol{U};\boldsymbol{Z})$ is in general infeasible even for short block lengths n. However, the bit-channels' capacities C_i's can be efficiently computed for BECs and can be also well-approximated for general BMS channels [19] leading to efficient numerical methods to approximate both the lower and the upper bounds on the MIS leakage.

Furthermore, the actual MIS leakage $I(\boldsymbol{U};\boldsymbol{Z})$ can be approximated using a Monte Carlo method when wiretap channel is a BEC. First, let $\tilde{G}_{(k+r)\times n}$ denote the generator matrix for a polar code associated with $\mathcal{A}\cup\mathcal{R}$, where $\mathcal{A}$ and $\mathcal{R}$ are the subsets of indices for the information bits and random bits in the considered polar secrecy code. Let $\mathcal{A}', \mathcal{R}' \subseteq [k+r]$ denote the indices of information bits and random bits, after discarding the frozen bits. For $i \in \mathcal{A}'$, let $\boldsymbol{u}_i$ denote the indicator vector corresponding to index i, i.e., its i-th entry is 1 and the rest of entries are zeros. Let also $\mathcal{U}_{\mathcal{A}'}$ denote the subspace of $\mathbb{F}_2^{k+r}$ spanned by $\boldsymbol{u}_i$'s, for $i \in \mathcal{A}'$.

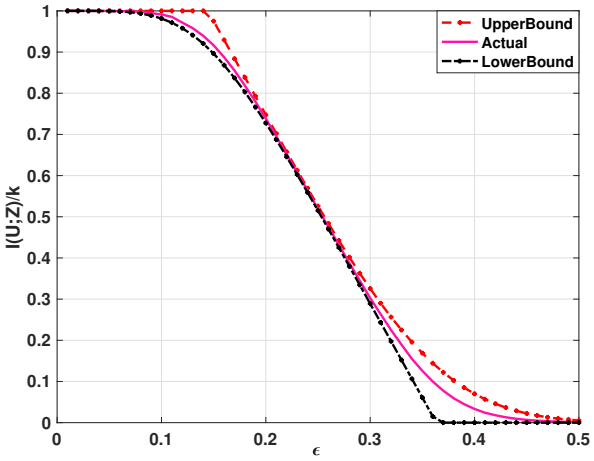

Fig. 2. Upper and lower Bounds versus the Monte-Carlo Approximation of the normalized MIS leakage over wiretap BEC(ϵ).

The Monte Carlo-based approximation of the MIS leakage for wiretap BECs is as follows. One can generate random instances of the erasure patterns for the wiretap channel repeatedly. For each such an instance, let $\tilde{\mathcal{G}}'$ denote the subspace spanned by the columns of $\tilde{G}$ indexed by non-erasures. Then it is straightforward to observe that $\dim(\tilde{\mathcal{G}}' \cap \mathcal{U}_{\mathcal{A}'})$ is the corresponding *instance* of the MIS leakage. One can average this quantity across a large number of random erasure patterns in order to obtain an approximation for the MIS leakage.

In Figure 2 we present numerical results for a polar secrecy code of length $n = 256$ with $k = 56$, and $r = 163$ over wiretap BECs for a range of erasure parameters for the wiretap channel W. Both the lower bound of Proposition 3 as well as the upper bound of Proposition 1 on the normalized MIS leakage $I(\boldsymbol{U};\boldsymbol{Z})/k$ are demonstrated together with the approximation of the actual values obtained via the method explained above. The results show that the bounds are actually tight especially for *medium* ranges of the MIS leakage.

Acknowledgment

We would like to thank Hamed Hassani for helpful discussions.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] U. Maurer and S. Wolf, "Information theoretic key agreement: From weak to strong secrecy for free," *Advances in Cryptology-EUROCRYPT*, p. 351–368, 2000.

[3] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[4] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6428–6443, 2011.

[5] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8077–8098, 2013.

[6] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *IEEE Transactions on Information Theory*, vol. 61, no. 11, pp. 6213–6237, 2015.

[7] O. Günlü, P. Trifonov, M. Kim, R. F. Schaefer, and V. Sidorenko, "Randomized nested polar subcode constructions for privacy, secrecy, and storage," in *2020 International Symposium on Information Theory and Its Applications (ISITA)*. IEEE, 2020, pp. 475–479.

[8] M. M. Wilde and S. Guha, "Polar codes for degradable quantum channels," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4718–4729, 2013.

[9] P. Jouguet and S. Kunz-Jacques, "High performance error correction for quantum key distribution using polar codes," *Quantum Information & Computation*, vol. 14, no. 3-4, pp. 329–338, 2014.

[10] J. Fang, Z. Yi, J. Li, Z. Liang, Y. Wu, W. Lei, Z. L. Jiang, and X. Wang, "Improved polar-code-based efficient post-processing algorithm for quantum key distribution," *Scientific Reports*, vol. 12, no. 1, pp. 1–11, 2022.

[11] E. Şaşoğlu and A. Vardy, "A new polar coding scheme for strong security on wiretap channels," in *2013 IEEE International Symposium on Information Theory*, 2013, pp. 1117–1121.

[12] S. H. Hassani, K. Alishahi, and R. L. Urbanke, "Finite-length scaling for polar codes," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 5875–5898, 2014.

[13] S. B. Korada, "Polar codes for channel and source coding," EPFL, Tech. Rep., 2009.

[14] A. Fazeli, H. Hassani, M. Mondelli, and A. Vardy, "Binary linear codes with optimal scaling: Polar codes with large kernels," *IEEE Transactions on Information Theory*, vol. 67, no. 9, pp. 5693–5710, 2020.

[15] W. Yang, R. F. Schaefer, and H. V. Poor, "Wiretap channels: Nonasymptotic fundamental limits," *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4069–4093, 2019.

[16] M. Mondelli, S. H. Hassani, and R. L. Urbanke, "Unified scaling of polar codes: Error exponent, scaling exponent, moderate deviations, and error floors," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 6698–6712, 2016.

[17] M. Shakiba-Herfeh, L. Luzzi, and A. Chorti, "Finite blocklength secrecy analysis of polar and Reed-Muller codes in BEC semi-deterministic wiretap channels," in *2021 IEEE Information Theory Workshop (ITW)*, 2021, pp. 1–6.

[18] A. Nooraiepour and T. M. Duman, "Randomized convolutional codes for the wiretap channel," *IEEE Transactions on Communications*, vol. 65, no. 8, pp. 3442–3452, 2017.

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