# The Un-Polarized Bit-Channels in the Wiretap Polar Coding Scheme

Hamed Mirghasemi and Jean-Claude Belfiore

***Abstract*—Polar coding theorems state that as the number of channel use, $n$, tends to infinity, the fraction of un-polarized bit-channels (the bit-channels whose $Z$ parameters are in the interval $(\delta(n), 1-\delta(n))$, tends to zero. Consider two BEC channels $W(z_1)$ and $W(z_2)$. Motivated by polar coding scheme proposed for the wiretap channel, we investigate the number of bit-channels which are simultaneously un-polarized for both of $W(z_1)$ and $W(z_2)$. We show that for finite values of $n$, there is a considerable regime of $(z_1, z_2)$ where the set of (joint)un-polarized bit-channels is empty. We also show that for $\gamma \leq 1/2$ and $\delta(n) = 2^{-n^\gamma}$, the number of un-polarized bit-channels is lower bounded by $2\gamma \log(n)$.**

## I. Introduction

The most basic channel model studied in information-theoretic security is the wiretap channel which was first introduced by Wyner [7]. In this situation, Alice wants to communicate a message $\mathbf{U} \in \{0,1\}^k$ to Bob through the main channel $W^B$ while this communication is being eavesdropped by Eve through the eavesdropper channel $W^E$. The encoding of $\mathbf{U}$ by Alice should be done such that:

- Eve asymptotically obtains no information about $\mathbf{U}$, i.e:
$$\lim_{k\to\infty} (I(\mathbf{U};\mathbf{Z})) \to 0\,,$$
where $\mathbf{Z}$ is Eve's received signal, and $I(\mathbf{U};\mathbf{Z})$ stands for the mutual information between message and Eve output, and;
- Bob is able to decode $\mathbf{U}$ reliably, i.e.:
$$\lim_{k\to\infty} \Pr\{\tilde{\mathbf{U}}(\mathbf{Y}) \neq \mathbf{U}\} \to 0\,,$$
where $\tilde{\mathbf{U}}(\mathbf{Y})$ is the decoded message, and $\mathbf{Y}$ is Bob's received signal.

After the introduction of Polar codes by Arikan [2], some polar coding schemes are proposed for the wiretap channel. In [1, 4, 6], it is shown that there exist secrecy-capacity achieving polar coding schemes which satisfy both the reliable and the weak security conditions [1]. The idea of using polar code for wiretap channel comes from the fact that due to polarization phenomena, almost all bit-channels are either perfect or completely noisy. Therefore, we can simply put information bits in the bit-channels which are simultaneously perfect for Bob and completely noisy for Eve. However, there are some bit-channels which are neither perfect nor completely noisy. We call this bit-channels "un-polarized bit-channels". Since the notion of weak security is not practical, in [5], a polar coding scheme was proposed to ensure the strong security. Although this proposed polar code scheme satisfies the strong security condition, the reliable condition is not satisfied. The reason that this coding scheme can not satisfy both the reliable and security conditions is due to the possible existence of un-polarized bit-channels. In this paper, we study the number of un-polarized bit-channels when the main and the eavesdropper channels are BECs.

This work was supported in part by FP7 project PHYLAWS (EU FP7-ICT 317562). H. Mirghasemi and J-C. Belfiore are with the Communications and Electronics Department, Telecom ParisTech, 75634 Paris Cedex 13. Emails: {mirghasemi,belfiore@telecom-paristech.fr}.

[1] Weak security condition: $\lim_{k\to\infty} \frac{I(\mathbf{U};\mathbf{Z})}{k} \to 0$.

### A. Strong-Security Coding Scheme

Let $W : \mathcal{X} = \{0,1\} \to \mathcal{Y}$ be an arbitrary binary-input DMC channel with transition probabilities $\{W(y|x) : x \in \{0,1\}, y \in \mathcal{Y}\}$. The Bhattacharyya parameter of $W$, denoted by $Z(W)$ is given by

$$Z(W) = \sum_{y\in\mathcal{Y}} \sqrt{W(y|0)W(y|1)}\,.$$

Let $I(W)$ be the symmetric capacity of $W$ defined as the mutual information between the input and the output of $W$ when the input is chosen from the uniform distribution over $\{0,1\}$. The channel obtained by $n = 2^m$ independent channel use is given by $W^n(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n W(y_i|x_i)$. Let

$$G = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

and let $G^{\otimes m}$ denote the $m$-th Kronecker power of $G$. In polar coding, the vector $\mathbf{v} = (v_1, \cdots, v_n)$ is encoded as $\mathbf{x} = \mathbf{v}P_nG^m$, where $P_n$ is the $n \times n$ bit-reversal permutation matrix. For any $1 \leq i \leq n$, the $i$-th bit-channel is defined as

$$W_i(\mathbf{y}, v_1^{i-1}|v_i) \stackrel{\text{def}}{=} \frac{1}{2^{n-1}} \sum_{\mathbf{v}\in\{0,1\}^{n-1}} \tilde{W}(\mathbf{y}|(v_1^{i-1}, v_i, v_{i+1}^n))\,,$$

where $\tilde{W}(\mathbf{y}|\mathbf{v}) \stackrel{\text{def}}{=} W^n(\mathbf{y}|\mathbf{v}P_nG^{\otimes m})$. The set of $n$ bit-channels can be partitioned into $Z$-good and $Z$-bad channels as follows:

$$\mathcal{ZG}_n(W,\delta_1) \stackrel{\text{def}}{=} \{i \in [n] : Z(W_i) < \delta_1(n)\}$$
$$\mathcal{ZB}_n(W,\delta_1) \stackrel{\text{def}}{=} \{i \in [n] : Z(W_i) \geq 1 - \delta_1(n)\}\,.$$

Also, for a given sequence of $\delta_2(n)$, we can define the set of $I$-bad channels as the set of bit-channels whose symmetric

capacities are below than $\delta_2(n)$, i.e.

$$\mathcal{IB}_n(W,\delta_2) \stackrel{\text{def}}{=} \{i\in[n] : I(W_i)\le \delta_2(n)\}\,.$$

In [5], three sets have been defined as follows:

- $\mathcal{R} \stackrel{\text{def}}{=} \mathcal{IG}_n(W^E,\delta_2)$
- $A \stackrel{\text{def}}{=} \mathcal{ZG}_n(W^B,\delta_1)\cup\mathcal{BI}_n(W^E,\delta_2)$
- $B \stackrel{\text{def}}{=} \mathcal{ZB}_n(W^B,\delta_1)\cup\mathcal{BI}_n(W^E,\delta_2)\,.$

The following theorem asserts that if we transmit information bits, uniformly random bits and frozen bits over $\mathcal{A}$, $\mathcal{R}$, and $\mathcal{B}$ respectively, we can achieve the strong security condition [[5, Theorems. 17, 21]].

**Theorem 1.** *Let $\mathbf{U}\in\{0,1\}^k$ be the message intended to be transmitted securely to Bob. We construct the vector $V$ with $\mathbf{V}_{\mathcal{A}}=U$, $\mathbf{V}_{\mathcal{B}}=\mathbf{0}$ and $\mathbf{V}_{\mathcal{R}}$ uniform over $\{0,1\}^{|\mathcal{R}|}$. Then, with the polar encoding/decoding scheme*

- *the amount of leaked information can be bounded by*

$$I(\mathbf{U};\mathbf{Z})\le n\delta_2(n)\,;$$

- *Bob's decoding error probability can be bounded by*

$$Pr\{\tilde{\mathbf{U}}\neq\mathbf{U}\}\le n\delta_1(n)+\sum_{i\in\mathcal{X}_n(W^B,W^E)} Z_i(W^B)\,,$$

*where the set $\mathcal{X}_n(W^B,W^E,\delta_1,\delta_2)$ is defined as*

$$\mathcal{X}_n(W^B,W^E,\delta_1,\delta_2) \stackrel{def}{=} \{i\in[n] : Z(W_i^B)\ge\delta_1(n),\ I(W_i^E)\ge\delta_2(n)\}$$

If we choose $\delta_1(n),\delta_2(n)=o(n)$ as $n\to\infty$, the polar code can achieve the strong security condition, however the sum $\sum_{i\in\mathcal{X}}(Z(W_i^B))$ can be larger than zero. The aim of this paper is to analyze this set and the conditions where the size of this set is zero. From [2, Propositions 1, 11], we know that for any B-DMC channel, we have $1-I(W)\le Z(W)\le\sqrt{1-I^2(W)}$ and therefore, $|\mathcal{X}|$ can be bounded as follows

$$\begin{aligned}|\mathcal{D}_n(W^B,W^E,\delta_1,\delta_2)|&\le|\mathcal{X}|\\ &\le|\mathcal{D}_n(W^B,W^E,\delta_1,1-\sqrt{1-\delta_2^2})|\,,\end{aligned}\tag{1}$$

where the set $\mathcal{D}_n(W^B,W^E,\delta_1,\delta_2)$ is defined as

$$\begin{aligned}\mathcal{D}_n(W^B,W^E,\delta_1,\delta_2)\stackrel{\text{def}}{=}\{i\in[n] : Z(W_i^B)\ge\delta_1,\\ Z(W_i^E)\le 1-\delta_2\}\,.\end{aligned}\tag{2}$$

For the sake of simplicity, in this paper, we study $\mathcal{D}_n$ instead of $\mathcal{X}_n$. Any results about $|\mathcal{D}_n|$ can be related to $|\mathcal{X}_n|$ by Equation (1).

## II. $D_n(W^B,W^E,\delta_1,\delta_2)$

The following lemma states that $|\mathcal{D}_n(W^B,W^E,\delta_1,\delta_2)|$ for any two BMS channels $W^B$ and $W^E$ can be bounded by $|\mathcal{D}_n^{BEC}(z_1,z_2,\delta_1,\delta_2)|$ for suitable choice of $z_1$ and $z_2$.

**Lemma 1.**
- *For any pair of BSC channels with cross-over probabilities $p_1$ and $p_2$, we have*

$$\begin{aligned}&\mathcal{D}_n(BSC(p_1),BSC(p_2),\delta_1,\delta_2)\\ &\qquad=\mathcal{D}_n^{BEC}(z_1,z_2,\delta_1^2,2\delta_2-\delta_2^2)\,,\end{aligned}$$

  *where $z_i=4p_i(1-p_i)$ for $i=\{1,2\}$.*
- *For any pair of $BMS$ channels $W^B$ and $W^E$, we have*

$$\begin{aligned}&|\mathcal{D}_n(W^B,W^E,\delta_1,\delta_2)|\\ &\qquad\le|\mathcal{D}_n^{BEC}(z_1,z_2^2,\delta_1,2\delta_2-\delta_2^2)|\,,\end{aligned}$$

  *where $Z(W^B)=z_1$ and $Z(W^E)=z_2$.*

*Proof:* For any $i\in[n]$, we denote $\mathbf{w_i}=[w_{i,1},\cdots,w_{i,m}]$ as the binary expansion of $i-1$. We define two function $f_1(x)=x^2$ and $f_0(x)=2x-x^2$. For BEC$(z)$ channel, we have

$$Z_i^{BEC}(z)=f_{w_{i,m}}\circ f_{w_{i,m-1}}\circ\cdots f_{w_{i,1}}(z)\tag{3}$$

We also define two function $g_1(x)=x^2$ and $g_0(x)=x\sqrt{2-x^2}$. For BSC$(p)$ channel, we have

$$Z_i^{BSC}=g_{w_{i,m}}\circ g_{w_{i,m-1}}\circ\cdots g_{w_{i,1}}(2\sqrt{p(1-p)})\,.\tag{4}$$

Noting that $f_b(x)=g_b^2(x^2)$ for any $b=0,1$, we can easily see that

$$Z_i^{BSC}(p)=[Z_i^{BEC}(4p(1-p))]^{0.5}\,.$$

The second part follows from the fact that for any BMS channel $W$ with $Z(W)=z$, the $Z$ parameters of bit-channel $W_n$ can be lower and upper bounded by

$$Z_i^{BSC}(p)\le Z(W_i)\le Z_i^{BEC}(z)\,,$$

where $p=\frac{1-\sqrt{1-z^2}}{2}$. Therefore, if $i\in\mathcal{D}_n(W^B,W^E,\delta_1,\delta_2)$, we have $Z_i^{BEC}(z_1)\ge\delta_1$ and $Z_i^{BSC}(p_2)\le 1-\delta_2(n)$ where $p_2=\frac{1-\sqrt{1-z_2^2}}{2}$, which means that $Z^{BEC}(z_2^2)\le 1-2\delta_2+\delta_2^2$. ∎

### A. The Numerical Evaluation of $|\mathcal{D}^{BEC}|$

The numerical evaluation of $|\mathcal{D}_n^{\text{BEC}}(z_1,z_2,\delta)|$ is plotted in Figures 1 to 4. We note that this function[2] is symmetric which means than $|\mathcal{D}_n(z_1,z_2)|=|\mathcal{D}_n(1-z_2,1-z_1)|$. From these plots, we can see that if we choose large values of $\delta$, we can enlarge the region of $(z_1,z_2)$ where $|\mathcal{D}(z_1,z_2)|=0$. Hence, there exists a trade-off in $\delta$ selection: among all sequences of order $o(n)$, if we choose the smallest ones, the amount of leaked information decreases, however, $|\mathcal{D}|$, and therefore, the upper bound on Bob's decoding error probability increase.

For any given $z_1$ and $\delta(n)$, we define $z_2^*(z_1,\delta)$ as the minimum of $z_2$ such that $|D_n(z_1,z_2,\delta)|=0$. This quantity provides us the minimum required distance between the erasure probabilities of the main and eavesdropper channels such that there exists no un-polarized bit-channel, and hence, coding scheme proposed in [5] can satisfy the both constraints. The numerical values of the minimum $z_2^*(z_1,\delta)$ are plotted in Figures 5 and 6. Interestingly, we see that as $z_1$ increases, the growth rate of $z_2^*$ decreases.

[2]In the rest of paper, we omit the superscript BEC for the sake of simplicity.

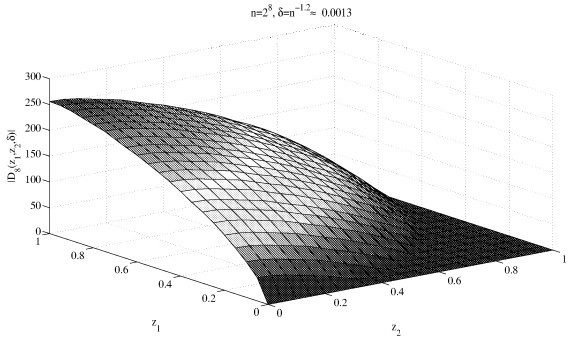


Fig. 1. $|D_n^{BEC}(z_1, z_2, \delta)|$ for $n = 2^8$ and $\delta = n^{-1.2}$.

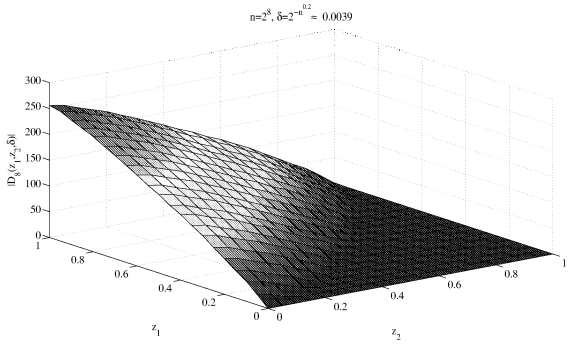


Fig. 2. $|D_n^{BEC}(z_1, z_2, \delta)|$ for $n = 2^8$ and $\delta = 2^{-n^{0.2}}$.

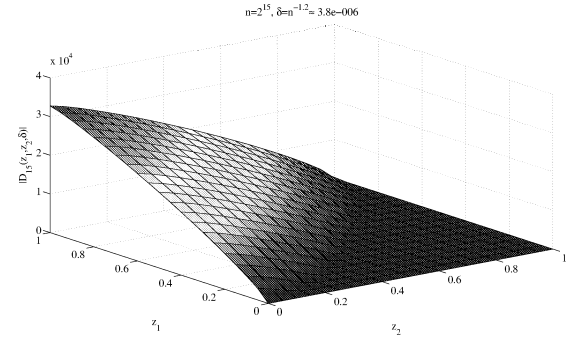


Fig. 3. $|D_n^{BEC}(z_1, z_2, \delta)|$ for $n = 2^{15}$ and $\delta = n^{-1.2}$.

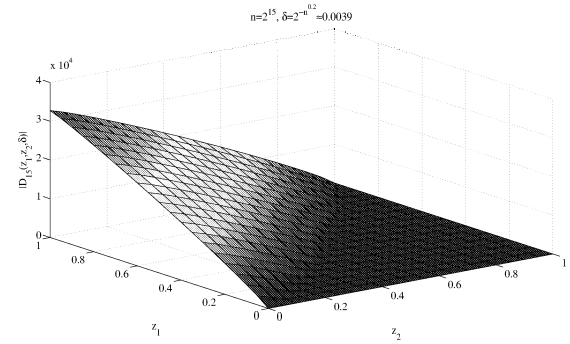


Fig. 4. $|D_n^{BEC}(z_1, z_2, \delta)|$ for $n = 2^8$ and $\delta = 2^{-n^{0.2}}$.

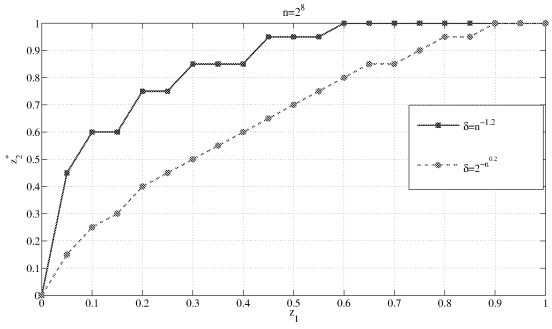


Fig. 5. $z_2^*(z_1, \delta)$ for $n = 2^8$.

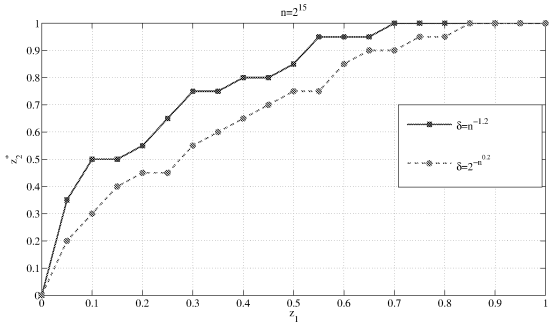


Fig. 6. $z_2^*(z_1, \delta)$ for $n = 2^{15}$.

## III. Safeguard

For any bit-channel $W_i$, the "un-polarized interval", denoted by $\mathcal{I}(W_i, \delta(n))$, is defined as the interval in $(0, 1)$ where $W_i$ is not polarized, i.e.,

$$\mathcal{I}(W_i, \delta(n)) \stackrel{\text{def}}{=} \{z \in [0, 1] : Z(W_i)(z) \in [\delta(n), 1 - \delta(n)]\}.$$

Therefore, for any given $z_1$ and $z_2$, if we can find one bit-channel $W_i$ such that $z_1, z_2 \in \mathcal{I}(W_i, \delta)$, then $|D_n(z_1, z_2, \delta)| \neq 0$. We define the maximum value of $|\mathcal{I}(W_i, \delta(n))|$ over all bit-channels as the "safeguard"

$$z_n^*(\delta(n)) \stackrel{\text{def}}{=} \max_{i \in [n]} |\mathcal{I}(W_i, \delta(n)|.$$

Hence for all $0 \leq z_1 < z_2 \leq 1$ such that $z_2 - z_1 \geq z_n^*$, we have $|\mathcal{D}_n(z_1, z_2, \delta(n)| = 0$. From Figure 7, we can see that for $n = 2^{15}$ and $\delta = 3.8 \times 10^{-6}$, for any pair of $(z_1, z_2)$ such that $z_2 - z_1 > 0.44$, we have $|\mathcal{D}_n(z_1, z_2) = 0|$.

### A. The Expected Value of Un-Polarized Interval Length

We define two functions $h_0(y) = 1 - \sqrt{1 - y}$ and $h_1(y) = \sqrt{y}$. For any $i \in [n]$, we define

$$Y_i(y) \stackrel{\text{def}}{=} h_{w_{i,m}} \circ h_{w_{i,m-1}} \cdots \circ h_{w_{1,i}}(y), \tag{5}$$

where $\mathbf{w}_i$ is the binary expansion of $i - 1$. We can see that $Y_i(Z_i(z)) = z$. Therefore

$$z^*(\delta(n)) = \max_{i \in [n]} (Y_i(1 - \delta) - Y_i(\delta)).$$

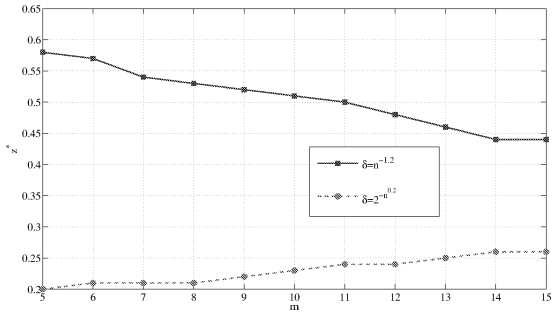


Fig. 7. $z^*(\delta)$ for $m = [5, \cdots, 15]$ and $n = 2^m$.

The following lemmas states the $\mathbb{E}[Y_n(b) - Y_n(a)] \xrightarrow[n\to\infty]{a.e} 0$ for any two constants $a, b \in (0, 1)$.

**Lemma 2.** *We have*

$$E[Y_n(y)] \xrightarrow[n\to\infty]{a.e} 0.5 + (y - 0.5)e^{-nln(2)-0.5}\,. \tag{6}$$

*Proof of Lemma 2:* First, we show that for any $n \geq 1$, we have $E[Y_{w_n}(0.5)] = 0.5$. From 5, if for a given $k$, we have $E[Y_k(0.5)] = 0.5$, then

$$\begin{aligned} E[Y_{k+1}(0.5)] &= 0.5 + \frac{1}{2}[E(\sqrt{Y_k(0.5)}) - E(\sqrt{1 - Y_k(0.5)})] \\ &= 0.5\,. \end{aligned}$$

Noting that $E[Y_1(0.5)] = 0.5$ completes the proof. For any $y > 0.5$, we have

$$Y_{w_n}(y) = Y_{w_n}(0.5) + (y - 0.5)Y^{'}_{w_n}(c)\,, \tag{7}$$

for some $c \in (0.5, y)$. From the chin rule we have

$$\ln(Y^{'}_{w_n}(c)) = \sum_{j=1}^{n} Y^{'}_{B_j}(Y_{w_{j-1}}(y))\,.$$

From [3], we know that the uniform measure on $[0, 1]$ is the unique and invariant measure for $Y_n$, and hence, from Ergodic theorem, we have

$$\frac{1}{n}\ln(Y^{'}_{w_n}(c)) \to E[\ln(Y^{'}(y))] = 0.5 - \ln(2)\,.$$

■

In Figure 8, the numerical values of $\mathbb{E}(Y_n(1-\delta) - Y_n(\delta)) = \mathbb{E}[|\mathcal{I}_n(\delta)|]$ is plotted. As $n \to \infty$, it seems that $\mathbb{E}[|\mathcal{I}_n(\delta)|]$ goes to zero. Also, these curves suggests that the expected value of un-polarized interval length is not a good lower bound on $z^*$ for large values of $n$.

### *B. Two-Runs Sequences*

In this subsection, we study the most "un-polarized" bit-channels, i.e. the bit-channels which have the maximum un-polarized length interval among all the bit-channels. The numerical values of $\{|\mathcal{I}(W_i, \delta(n))|\colon i \in [n]\}$ for $m = \{8, 15\}$ and for $\delta(n) = \{2^{-n^{0.2}}, n^{-1.2}\}$ are plotted in Figures 9 to 12. This numerical evaluation suggests that for finite values of $n \leq 20$, the binary sequence $\mathbf{w}_i$ whose un-polarized length interval is the largest has always two runs.

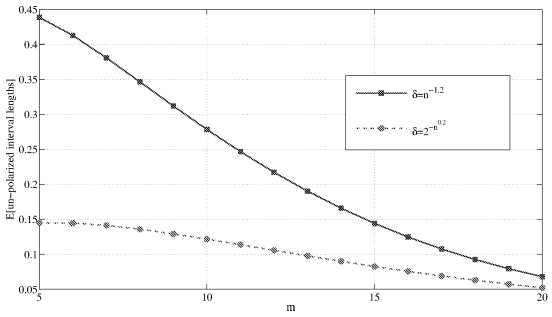


Fig. 8. $\mathbb{E}[|\mathcal{I}_n(\delta)|]$ for $m = [5, \cdots, 20]$ and $n = 2^m$.

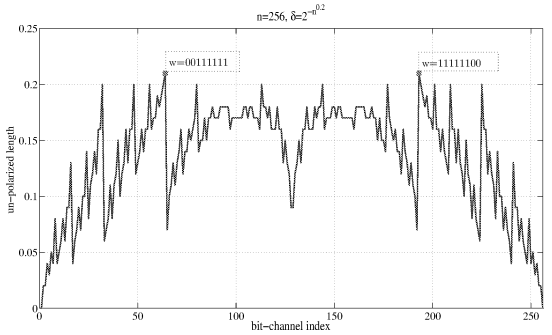


Fig. 9. Un-polarized lengths of bit-channels for $n = 256$ and $\delta = 2^{-n^{0.2}}$.

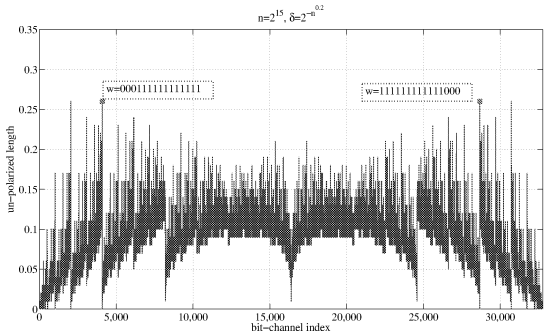


Fig. 10. Un-polarized lengths of bit-channels for $n = 2^{15}$ and $\delta = 2^{-n^{0.2}}$.

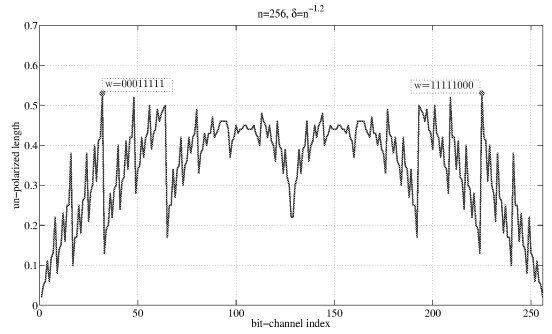


Fig. 11. Un-polarized lengths of bit-channels for $n = 256$ and $\delta = n^{-1.2}$.

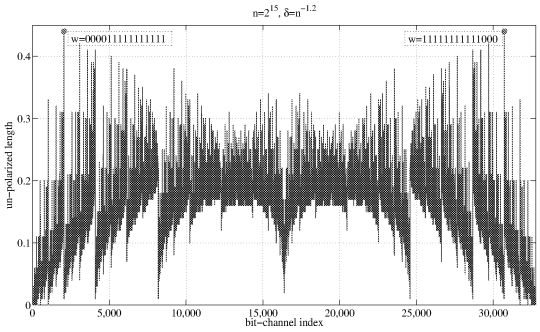


Fig. 12. Un-polarized lengths of bit-channels for $n = 2^{15}$ and $\delta = n^{-1.2}$.

**Theorem 2.**
- *For* $\delta(n) = 2^{-n^\gamma}$ *and* $\gamma \leq 0.5$, $\lim_{n\to\infty} z^* = 1$*, and hence, for all pairs of* $(z_1, z_2)$*, we have* $\lim_{n\to\infty} |\mathcal{D}_n(z_1, z_2, \delta)| > 2\gamma \log n$.
- *For any* $\delta(n)$ *such that* $\log(-\log\delta(n)) = o(m)$*, the "un-polarized" interval length of any* 2*-run sequence goes to 0.*

*Proof.* We consider a 2-run sequence $\mathbf{w}$ of length $m$, such that $w_1^r = 1$ and $w_{r+1}^m = 0$. Let $r_2 = m - r$. For any $y_0$, we denote $y_1(y_0) = (y_0)^{1/2^{r_1}}$, and $y_2(y_0) = 1-(1-y_1(y_0))^{1/2^{r_2}}$. First, we consider the case $y_0 = \delta(m) \stackrel{\text{def}}{=} 2^{-f(m)}$. We have $y_1 = (\delta(m))^{1/2^{r_1}} = 2^{-\frac{f(m)}{2^{r_1}}}$. Therefore, if $\frac{f(m)}{2^{r_1}} \to 0$, we have $y_1 \to 1$ and if $\frac{f(m)}{2^{r_1}} \to \infty$, then we have $y_1 \to 0$. Now we consider $y_2 = 1-(1-y_1)^{1/2^{r_2}}$. We have $\ln(1-y_2) = \frac{\ln(1-y_1)}{2^{r_2}}$. If $y_1 \to 1$, we have

$$\ln(1-y_2) \approx \frac{[\ln(f(m)/2^{r_1}) + \ln(\ln 2)]}{2^{r_2}}\,,$$

where the last approximation follows from $\ln(1-2^{-x}) \approx \ln(x\ln 2)$ as $x \to 0$. Also, if $y_1 \to 0$, then we have

$$\ln(1-y_2) \approx \frac{y_1}{2^{r_2}}\,,$$

where the last approximation follows from $\ln(1-x) \approx x$ as $x \to 0$. Hence, we can conclude that

$$y_2(\delta(m)) = \begin{cases} 1-[f(m)e^{-r_1\ln 2}\ln 2]^{1/2^{r_2}} & \text{if}\frac{f(m)}{2^{r_1}} \to 0 \\ 1-e^{-2^{-\frac{f(m)}{2^{r_1}}-r_2}} & \text{if}\frac{f(m)}{2^{r_1}} \to \infty \end{cases}$$

Therefore, to have $y_2(\delta(m)) \to 0$, we should have one of these two conditions:

$$\mathrm{C}_1 : \frac{f(m)}{2^{r_1}} \to 0\,, \frac{\ln f(m) - r_1\ln 2 + \ln(\ln 2)}{2^{r_2}} \to 0\,; \quad (8)$$

or

$$\mathrm{C}_2 : \frac{f(m)}{2^{r_1}} \to \infty\,, \frac{f(m)}{2^{r_1}} + r_2 \to \infty\,. \quad (9)$$

Now we consider the case $y_0 = 1-\delta(m)$. Then we have $y_1 = (1-\delta(m))^{1/2^{r_1}} \approx e^{-\delta(m)/2^{r_1}} \to 1$. Thus, we have $y_2 = 1-(1-y_1)^{1/2^{r_2}}$. Therefore $\ln(1-y_2) \approx \frac{\ln(\delta(m)) - r_1\ln 2}{2^{r_2}}$ and

$$y_2(1-\delta(m)) = 1-[e^{-r_1\ln 2}\delta(m)]^{1/2^{r_2}}$$

Therefore to have $y_2(1-\delta(m)) \to 1$, we should have

$$\mathrm{C}_3 : \frac{r_1 + f(m)\ln 2}{2^{r_2}} \to \infty\,. \quad (10)$$

Let denote $k_i(m) \stackrel{\text{def}}{=} \frac{m}{r_i}$ for $i = \{1, 2\}$. Now we consider the combination of $\mathrm{C}_1$ and $\mathrm{C}_3$. If $\log(-\log(\delta(m))) = o(m)$, to satisfy the condition $\mathrm{C}_3$, we should have $k_2(m) \to 1$ which contradicts the condition $\frac{\ln f(m) - r_1\ln 2}{2^{r_2}} \to 0$. Similarly, if we consider the conditions $\mathrm{C}_2$ and $\mathrm{C}_3$, we see that for $\log(-\log(\delta(m))) = o(m)$, we can not have simultaneously $y_2(\delta(m)) \to 0$ and $y_2(1-\delta(m)) \to 1$. Now we consider the case $\delta(m) = 2^{-2^{m^\gamma}}$. The combination of $\mathrm{C}_1$ and $\mathrm{C}_3$ yields that for any $k_1(m) \geq 1-\beta$, we have $y_2(\delta(m)) \to 0$ and $y_2(1-\delta(m)) \to 1$. Therefore, we have $2m\gamma$ two-run sequences whose un-polarized interval lengths go to 1. □

The above theorem states that for $\delta(n) = 2^{-n^\gamma}$, $z^*$ goes to 1 but, for any $\delta(n)$ such that $\log(-\log\delta) = o(m)$, we only proved that the "un-polarized" interval length of any 2-run sequences tends to 0. However, the numerical simulation (Figures 11 and 12) and the above theorem strengthen the possibility that as $n \to \infty$, the most "un-polarized" sequences are the 2-run sequences. If true, we can conclude that for $\delta(n)$ such that $\log(-\log\delta) = o(m)$, $|\mathcal{D}_n(z_1, z_2, \delta)| \to 0$ for any pair of $z_1 < z_2$.

## REFERENCES


[1] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund. Nested polar codes for wiretap and relay channels. *Comm. Letters.*, 14(8):752–754, August 2010.

[2] E. Arikan. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *Information Theory, IEEE Transactions on*, 55(7):3051 –3073, july 2009.

[3] S. H. Hassani, K Alishahi, and R. L. Urbanke. Finite-length scaling of polar codes. *CoRR*, abs/1304.4778, 2013.

[4] O. O. Koyluoglu and H El Gamal. Polar coding for secure transmission and key agreement. *IEEE Transactions on Information Forensics and Security*, 7(5):1472–1483, 2012.

[5] H Mahdavifar and A Vardy. Achieving the secrecy capacity of wiretap channels using polar codes. *IEEE Transactions on Information Theory*, 57(10):6428–6443, 2011.

[6] E. Sasoglu and A. Vardy. A new polar coding scheme for strong security on wiretap channels. In *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pages 1117–1121, July 2013.

[7] A. D. Wyner. The wire-tap channel. *Bell System Technical Journal*, 54(8):1355–1387, 1975.