Achieving the Uniform Rate Region of Multiple Access Channels Using Polar Codes

Hessam Mahdavifar, Mostafa El-Khamy, Jungwon Lee, Inyup Kang

Abstract—Designing polar codes for transmission over two-user multiple access channels is considered. In the proposed scheme, both users encode their messages using a polar encoder, while a joint successive cancellation decoder is deployed at the receiver. The decoding is done separately, while the codes are constructed jointly. This is done by treating the whole polar transformation on both users as a single polar transformation, wherein the multiple access channel (MAC) is regarded as one more level of polarization. We prove that our scheme achieves the whole uniform rate region by changing the decoding order in the joint successive cancellation decoder. Various simulation results over binary-additive Gaussian noise MAC are provided. At the end, a comparison is made with the existing results on polar codes for multiple access channels to emphasize the differences and the main advantages of our scheme.

Index Terms—polar code, multiple access channel, uniform rate region

I. INTRODUCTION

Polar codes were introduced by Arikan in the seminal work of [1]. They are the first family of codes for the class of binary-input symmetric discrete memoryless channels that are provable to be capacity-achieving with low encoding and decoding complexity. Construction of polar codes is based on a phenomenon called the channel polarization. Arikan proves that as the block length goes to infinity the channels seen by individual bits through a certain transformation called the polar transformation start polarizing which means that they approach either a noise-less channel or a pure-noise channel. In general, polar codes achieve the symmetric capacity of all binary-input memoryless channels, where the symmetric capacity of a binary-input channel is the mutual information between the input and output of the channel assuming that the input distribution is uniform.

Polar codes and polarization phenomenon have been successfully applied to various problems such as wiretap channels [2], data compression [3], [4] and multiple access channels [5], [6]. The notion of channel polarization has been extended to two-user multiple access channels (MAC) [5] and later to \( m \)-user MAC [6], wherein a technique is described to polarize a given binary-input MAC same as in Arikan’s groundbreaking work of [1]. The authors convert multiple uses of this MAC into single uses of extremal MACs. In the case of single user channel, there are only two extremal cases, one is for the noise-less channel and the other one is the pure-noise channel whereas in the two-user MAC, there are five possible cases as we will discuss in Section VII.

The capacity region of multiple access channels is fully characterized by Ahlswede [7] and Liao [8] for the case that the sources transmit independent messages. However, in this paper, we are only interested in the uniform rate region. For a multiple access channel, the uniform rate region is the achievable region corresponding to the case that the input distributions are uniform. The single-user counterpart of uniform rate region is indeed symmetric capacity. It is shown that at least one point on the dominant face of the uniform rate region can be achieved by the polar code constructed based on MAC-polarization [5], [6].

In this paper, we aim at achieving the whole uniform rate region by deploying polar-based schemes. Our approach for constructing MAC-polar codes is totally different from that of [5], [6]. Multiple access channels with two users are considered and independent polar encoders are deployed by both users. The construction is done jointly built upon recursively polarizing a certain MAC polarization building block. The polarization building block which consists of a certain number of independent uses of the underlying MAC is split into single-user bit-channels by extending the definition of bit-channels from [1] in this context. This is done in such a way that the MAC can be regarded as one more level of polarization. A proposed joint successive cancellation decoding is used at the decoder. The bits transmitted by both users are decoded successively in a certain order that is determined a priori. We show the direct relation between the decoding order and the polarization building block which guarantees the channel polarization. Furthermore, we show how to change the decoding order to approach all the points in the uniform rate region using the single-user channel polarization theorem.

The rest of this paper is organized as follows. In Section II we review some background on polar codes and multiple access channels. In Section III the straightforward way of achieving the whole uniform rate region using time sharing is discussed. Also, some simulation results are provided for binary-additive Gaussian noise MAC. In Section IV we explain how to use the previously developed compound polar codes [9] in order to take advantage of the longer transmission blocks in time sharing. The proposed MAC-polar codes with joint successive cancellation decoding are discussed in Section V for the special case when the length of the polarization building block is four. This is extended to the general case in Section VI and we prove that the whole uniform rate region is achieved by our scheme.

We compare the decoding complexity and capacity-achieving properties of the proposed scheme with the previous work of [5] in Section VII. Also, discussions on comparison with some other related works is made in Section VII. At the end, we conclude the paper by mentioning some directions for future work in Section VIII.
II. PRELIMINARIES

A. Polar codes

In this subsection, we provide a brief overview of the groundbreaking work of Arikan [1] and others [10]–[12] on polar codes and channel polarization.

Polar codes are constructed based upon a phenomenon called channel polarization discovered by Arikan [1]. The basic polarization matrix is given as

\[
G = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}
\]

(1)

The Kronecker powers of \(G\) are defined by induction. Let \(G^{\otimes 1} = G\) and for any \(n > 1\):

\[
G^{\otimes(n)} = \begin{bmatrix} G^{\otimes(n-1)} & 0 \\ G^{\otimes(n-1)} & G^{\otimes(n-1)} \end{bmatrix}
\]

(2)

It can be observed that \(G^{\otimes(n)}\) is a \(2^n \times 2^n\) matrix. Let \(N = 2^n\). Then \(G^{\otimes n}\) is the \(N \times N\) polarization matrix. Let \((U_1, U_2, \ldots, U_N)\), denoted by \(U_N\), be a block of \(N\) independent and uniform binary random variables. The polarization matrix \(G^{\otimes n}\) is applied to \(U_N\) to get \(X_N = U_N G^{\otimes n}\). Then \(X_i\)'s are transmitted through \(N\) independent copies of a binary-input discrete memoryless channel (B-DMC) \(W\). The output is denoted by \(Y_i^N\). This transformation with input \(U_N\) and output \(Y_i^N\) is called the polar transformation. In this transformation, \(N\) independent uses of \(W\) is transformed into \(N\) bit-channels, described next. Following the convention, random variables are denoted by capital letters and their instances are denoted by small letters. Let \(W_N : \mathcal{B}^N \to \mathcal{Y}^N\) denote the channel consisting of \(N\) independent copies of \(W\) i.e.

\[
W_N(y^N_i|x^N_i) \overset{\text{def}}{=} W(y_i|x_i)
\]

(3)

The combined channel \(\hat{W}\) is defined with transition probabilities given by

\[
\hat{W}(y^N_1|u^N_1) \overset{\text{def}}{=} W_N(y^N_1|u^N_1 G^{\otimes n})
\]

(4)

For \(i = 1, 2, \ldots, N\), the bit-channel \(W_N^{(i)}\) is defined as follows:

\[
W_N^{(i)}(y^N_i,u^{i-1}_i|u_i) \overset{\text{def}}{=} \frac{1}{2^{n-1}} \sum_{u^{i-1}_i \in \{0,1\}^{n-1}} \hat{W}(y^N_i|u^N_i)
\]

(5)

Intuitively, this is the channel that bit \(u_i\) observes through a successive cancellation decoder, deployed at the output. Under this decoding method, proposed by Arikan for polar codes [1], all the bits \(u^{i-1}_i\) are already decoded and are assumed to be available at the time that \(u_i\) is being decoded. The channel polarization theorem states that as \(N\) goes to infinity, the bit-channels start polarizing meaning that they either become a noise-less channel or a pure-noise channel.

In order to measure how good a binary-input channel \(W\) is, Arikan uses the Bhattacharyya parameter of \(W\), denoted by \(Z(W)\) [1], defined as

\[
Z(W) \overset{\text{def}}{=} \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}
\]

(6)

It is easy to show that the Bhattacharyya parameter \(Z(W)\) is always between 0 and 1. Channels with \(Z(W)\) close to zero are almost noiseless, while channels with \(Z(W)\) close to one are almost pure-noise channels. More precisely, it can be proved that the probability of error of a binary symmetric memoryless channel (BSM) is upperbounded by its Bhattacharyya parameter. Let \([N]\) denotes the set of positive integers less than or equal to \(N\). The set of good bit-channels \(G_N(W, \beta)\) is defined for any \(\beta < \frac{1}{2}\) [10], [11]:

\[
G_N(W, \beta) \overset{\text{def}}{=} \left\{ i \in [N] : Z(W_N^{(i)}) < 2^{-N\beta}/N \right\}
\]

(7)

Then the channel polarization theorem is proved by showing that the fraction of good bit-channels approaches the symmetric capacity \(I(W)\), as \(N\) goes to infinity, where \(I(W)\) is the mutual information between the input and output of \(W\) that given the input distribution is uniform [10]. This theorem readily leads to a construction of capacity-achieving polar codes. The idea is to transmit the information bits over the good bit-channels while freezing the input to the other bit-channels to a priori known values, say zeros. The decoder for this constructed code is the successive cancellation decoder of Arikan [1], where it is further proved that the frame error probability under successive cancellation decoding is upperbounded by the sum of Bhattacharyya parameters of the selected good bit-channels, which is \(2^{-N\beta}\) by the particular choice of good bit-channels in (5).

B. Multiple access channel

Let \(W : \mathcal{X}^2 \to \mathcal{Y}\) be a two-user MAC, where \(\mathcal{X}\) is the binary alphabet and \(\mathcal{Y}\) is the output alphabet. With slight abuse of notation, the channel is described by the transition probability \(W(y|u,v)\) for any \(u,v \in \mathcal{X}\) and \(y \in \mathcal{Y}\). Namely, \(W(y|u,v)\) denote the probability of receiving \(y\) given that \(u\) is transmitted by the first user and \(v\) is transmitted by the second user. The capacity region is given by

\[
C(W) \overset{\text{def}}{=} \text{Conv}(\cup_{U,V,R(U,V)})
\]

where \(\text{Conv}(S)\) denote the convex hull of the set \(S\), for any set of points \(S\), and \(R(U,V)\) is the set of all pairs \((R_1,R_2)\), where \(R_1\) and \(R_2\) are the rates of the two user, respectively, satisfying

\[
0 \leq R_1 \leq I(U;Y,V),
\]

\[
0 \leq R_2 \leq I(V;U,Y),
\]

\[
R_1 + R_2 \leq I(U,V;Y)
\]

and the union is over all random variables \(U,V \in \mathcal{X}\) and \(Y \in \mathcal{Y}\) jointly distributed as

\[
p_{U,V,Y}(u,v,y) = p_U(u)p_V(v)W(y|u,v)
\]

(8)

The uniform rate region of \(W\) is defined to be \(R(U,V)\), when \(U\) and \(V\) are independent and uniformly distributed over \(\{0,1\}\). This region is depicted in Figure 1. The segment specified by the equation \(R_1 + R_2 = I(U,V;Y)\) is referred to as the dominant face of the uniform rate region. Also, \(I(U,V;Y)\) is denoted by \(I(W)\).

Let \(A_W\) and \(B_W\) denote the two corner points indicated in Figure 1. More precisely, let the inputs \(U\) and \(V\) be two independent uniform binary random variables. Then

\[
A_W \overset{\text{def}}{=} (I(U;Y),I(V;Y|U))
\]

(9)
and similarly,
\[ B_W \overset{\text{def}}{=} (I(U; Y|V), I(V; Y)) \] (7)
The segment connecting \( A_W \) and \( B_W \) is actually the dominant face of the uniform rate region.

III. ACHIEVING THE UNIFORM RATE REGION USING TIME SHARING

In this section, we discuss the straightforward way of achieving the whole uniform rate region using time sharing. This is actually the onion peeling scheme of [13], [14]. From one point of view, a two-user MAC can be split into two single user channels as described next. This observation will be helpful later when we establish the main results of this paper based upon treating the MAC as one more level of polarization. For a given 2-user MAC \( W \), let \( \hat{W} : \mathcal{X} \to \mathcal{Y} \) denote a B-DMC whose transition probability for any \( y \in \mathcal{Y} \) and \( u \in \mathcal{X} \) is given by
\[ \hat{W}(y|u) = \frac{1}{2} \sum_{v \in \mathcal{X}} W(y|u, v) \] (8)
Intuitively, \( \hat{W} \) denotes the channel observed by the first user treating the bit transmitted by the second user as noise i.e. drawn from a uniform binary distribution. Let also \( \tilde{W} : \mathcal{X} \to \mathcal{Y} \times \mathcal{Y} \) denote another B-DMC whose transition probability for any \( y \in \mathcal{Y} \) and \( (u, v) \in \mathcal{X}^2 \) is given by
\[ \tilde{W}(y, u|v) = \frac{1}{2} W(y|u, v) \] (9)
Intuitively, \( \tilde{W} \) denotes the channel observed by the second user assuming that the bit transmitted by the first user is drawn from a uniform binary distribution and is given at the output of the channel. Notice that for independent and uniform binary inputs \( U \) and \( V \), \( I(W) = I(\hat{W}) + I(\tilde{W}) \). In fact, \( (I(\hat{W}), I(\tilde{W})) \) is \( A_W \). One can regard this as splitting the two-user MAC \( W \) into two single-user channels \( \hat{W} \) and \( \tilde{W} \). This is shown in Figure 2.

A. Achievability with point to point polar codes

Lemma 1: The points \( A_W \) and \( B_W \) are achievable using point-to-point polar codes.

Proof: Consider the first case of achieving \( A_W \). The other case is exactly similar to this case if we only change the role of the first user and the second user. Basically the decoder first decodes the message transmitted by the first user treating the message transmitted by the second user as noise. In fact, the first user observes the channel \( \hat{W} \) as defined in (8). There is a family of polar codes to achieve the symmetric capacity \( I(\hat{W}) \) as proved in [1]. If the message transmitted by the first user is decoded successfully, the second user observes the channel \( \tilde{W} \). Again, there is a family of polar codes to achieve the symmetric capacity \( I(\tilde{W}) \) as proved in [1]. The probability of error of the total decoding procedure is bounded by the sum of probability of errors for each of the users. As the probability of error goes to zero for each of the users, the total probability of error of the scheme also goes to zero. This completes the proof.

Corollary 2: The points on the segment connecting \( A_W \) and \( B_W \) in Figures 1 are also achievable using time-sharing.

The proof follows by the well-known method of time sharing between the points \( A_W \) and \( B_W \). See for example [13], [14].

B. Polar code construction and decoding

In this paper, the simulation system model assumes that the multiple access channel \( W \) is a binary-additive Gaussian noise channel where inputs \( u, v \in \{0, 1\} \) are modulated using BPSK (0 is mapped to \(-1\) and 1 is mapped to \(+1\)) into \( \pi \) and \( \bar{\pi} \), respectively. The output of the channel is denoted by \( y \), where \( y = \pi + \bar{\pi} + N \) and \( N \) is the Gaussian noise of unit variance \( N_0 \). For this channel, the capacity region is same as the uniform rate region and is given by the set of all possible pairs \((R_1, R_2)\) that satisfies
\[ R_1, R_2 \leq I(U; Y|V) = H(Y|V) - H(Y|U, V) = 0.7215 \]
\[ R_1 + R_2 \leq I(U; Y; V) = H(Y) - H(Y|U, V) = 1.11 \]

For the fixed channel described above, achievable rates at finite block lengths are considered, given the total probability of frame error error \( 10^{-2} \) on decoding both \( u \) and \( v \). Polar codes of length \( N = 2^n \), for \( n = 9, 10, 11, 12 \), are constructed for the channels \( \hat{W} \) and \( \tilde{W} \).

To initiate the successive cancellation decoder for \( \hat{W} \), the likelihood ratio of the channel \( \hat{W} \) given the channel output \( y \) and the noise variance \( \sigma^2 \) can be computed as follows:
\[ P(y|u = 0) = e^{-\frac{y^2}{\sigma^2}} e^{-\frac{(-u+y)^2}{\sigma^2}} \]
\[ P(y|u = 1) = e^{-\frac{y^2}{\sigma^2}} e^{-\frac{(-u-y)^2}{\sigma^2}} \] (10)

Given the transmitted message by the first user, the second user observes an AWGN channel i.e. \( \tilde{W} \) is an AWGN channel. The
log-likelihood ratio of the channel \( \hat{W} \) given the channel output \( y \), the noise variance \( \sigma^2 \) and \( u \) is given as follows:

\[
\log P(y, u | v = 0) - \log P(y, u | v = 1) = \frac{(y - 2u)^2 - (y - 2u + 2)^2}{\sigma^2}
\]

(11)

The codes with respect to the channels \( \hat{W} \) and \( \check{W} \) are constructed using a simulation-based method in which the probability of error of all the bit-channels are estimated as follows. For \( i = 1, 2, \ldots, N \), it is assumed that all the previous input bits indexed by \( 1, 2, \ldots, i - 1 \) are provided to the decoder by a genie at the time that the \( i \)-th bit is decoded. The decoder is run for sufficiently large number of uniformly distributed independent inputs (in this case \( 10^5 \)) to get an estimate of the probability of the event that the \( i \)-th bit is not decoded successfully given that the bits indexed by \( 1, 2, \ldots, i - 1 \) are successfully decoded. For each of the channels, the best bit-channels are picked such that the sum of probabilities of errors of the selected bit-channels does not exceed \( 5 \times 10^{-3} \). That way the frame error probability (FEP) of \( 5 \times 10^{-3} \) is guaranteed for each of the channels. Therefore, the total probability of error is bounded by \( 10^{-2} \).

The achievable rates on the MAC \( W \) with point to point polar codes for different block sizes are depicted in Figure 3.

![Fig. 3: Achievable rates for binary-additive AWGN channel](image)

IV. IMPROVING THE PERFORMANCE OF TIME SHARING BY INCREASING DECODING LATENCY

In the foregoing section, we explained how the entire uniform rate region can be achieved using time sharing between two polar codes designed for the certain single-user channels. The problem of time sharing is that it requires a large transmission block. For instance, let \( W \) be a given symmetric MAC and suppose that the points \( (R_1, R_2) \) and \( (R_2, R_1) \), corresponding to \( W \) and \( \check{W} \), are achievable with probability of error \( P_e \) using polar codes of length \( N \). In order to achieve the middle point \( \left( \frac{R_1 + R_2}{2}, \frac{R_2 + R_1}{2} \right) \), a transmission block of length \( 2N \) is required. In this section, we discuss how to take advantage of this larger transmission block to improve polarization by deploying compound polar codes [9], at the cost of increasing decoding latency. Compound polar code is a capacity-achieving scheme based on polar codes for reliable communication over multi-channels [9]. A multi-channel consists of several binary-input channels through which the encoded bits are transmitted in parallel. Instead of encoding separately across the individual constituent channels, which requires multiple encoders and decoders, we took advantage of the recursive structure of polar codes to construct a unified scheme with a single encoder and considerably improved performance. For more details, please refer to [9].

Suppose that in the first block of length \( N \) the first user observes \( \hat{W} \) and the second user observes \( \check{W} \), i.e. the decoder first decodes the message transmitted by the first user and then decodes the message transmitted by the second user. Consequently, in the second block of length \( N \), the users exchange roles. It is possible to use a compound polar code of length \( 2N \) for the second user that is constructed for the multi-channel consisting of \( W \) and \( \hat{W} \). This is done at the cost of increasing the decoding latency for the second user, as it has to wait for the whole block of length \( 2N \) to be decoded, whereas the first user's message is decoded as blocks of length \( N \).

The scheme is shown in Figure 4. The first user transmits a polar codeword constructed with respect to \( W \), denoted by \( C_1 \), in the first transmission block and transmits a polar codeword constructed with respect to \( \check{W} \), denoted by \( C_2 \), in the second transmission block. User 2 constructs a compound polar code of length \( 2N \) with respect to the \( 2 \)-multi-channel \((W, \hat{W})\). At the decoder, the codeword \( C_1 \) is decoded first. Therefore, the second user observes channel \( \hat{W} \) at the first transmission block and \( W \) at the second transmission block. In the next step, the received word of length \( 2N \), transmitted as \( C \) by the second user, is decoded. Therefore, in the second transmission block, the first user observes \( \check{W} \). In the last step, the second received word of length \( N \), transmitted as \( C_2 \) by the first user, is decoded.

![Fig. 4: The proposed scheme for using compound polar codes in time sharing](image)
its transmission rate.

| $N$ | length $2N$ | length $N$ | compound rate |
|-----|-------------|------------|---------------|
| 512 | 0.378       | 0.357      | 0.374         |
| 1024| 0.4         | 0.378      | 0.396         |
| 2048| 0.418       | 0.4        | 0.415         |

V. A PROPOSED MAC-POLAR CODE

Any two-user multiple access channel has a uniform rate region as depicted in Figure 5. The point $A_1W$, and likewise $B_1W$, is achieved by constructing and decoding polar codes across $W$ and $\bar{W}$. This can be regarded as separate construction and decoding. The goal of this section is to propose methods to achieve other points on the segment connecting the two corner points at the same block length, thereby avoiding time sharing which requires larger transmission blocks. The idea is to do joint decoding instead of simple separate decoding for the users. For instance, in its simplest case, suppose that each user is transmitting codewords of length $N$. Then instead of decoding all $N$ bits of the first user, and then decoding all $N$ bits of the second user, we can decode the first $\frac{N}{2}$ bits of the first user, then the first $\frac{N}{2}$ bits of the second user, then the second $\frac{N}{2}$ bits of the first user followed by the second $\frac{N}{2}$ bits of the second user. In order to show the polarization, one may intuitively think of the MAC as one level of polarization on the input bits, as depicted in Figure 5. The idea of joint decoding as well as the channel polarization will be elaborated throughout the rest of this paper.

A. Proposed construction with $4 \times 2$ polarization building blocks

The construction is based on the MAC polarization building block depicted in Figure 5. There are two independent copies of a given MAC $W$. Bits $U_1$ and $U_2$ are assigned to the first user and bits $V_1$ and $V_2$ are assigned to the second user. Then the recursion steps of polar transformation are applied to this block as Arikan’s polar code, for each user separately. More precisely, let $U_1^N = (U_1, U_2, \ldots, U_N)$ and $V_1^N = (V_1, V_2, \ldots, V_N)$ be two vectors of independent and uniformly distributed bits that are independently generated by the first user and the second user. Then let $X_1^N = U_1^N G_1^N$ and $X_2^N = V_1^N G_2^N$. For $i = 1, 2, \ldots, N$, the pair $(X_i, X'_i)$ is transmitted through $i$-th independent copy of $W$ and the output is denoted by $Y_i$.

For notational convenience, let $D_2^{2N} = (U_1^{N/2}, V_1^{N/2}, U_{N/2+1}^N, V_{N/2+1}^N)$. This will be the order of successive decoding as will be described in the next subsection. In fact, $D_2^{2N}$ is just a vector of length $2N$ of independent uniformly distributed binary bits. Let $\tilde{W}_{2N}$ denote the channel from $d_1^{2N}$ to $y_1^N$. For this particular MAC polarization building block, depicted in Figure 5 and its corresponding decoding order, the individual bit-channels are defined as follows. For $i = 1, 2, \ldots, N$,

$$W_{2N}^{(i)}(y_1^N, d_1^{i-1} | d_i) = \frac{1}{2^{2N-1}} \sum_{d_1^{i-1} \in (0,1)^{2N-i}} \tilde{W}_{2N}(y_1^N | d_1^{2N})$$

(12)

Intuitively, $W_{2N}^{(i)}$ is the channel that bit $d_i$ observes under successive cancellation decoding, while $d_1^{i-1}$ are assumed to be given at the output of the decoder and $d_1^{2N}$ are treated as noise.

The next lemma is similar to the Proposition 3 of [1] that still holds for this particular polar transformation defined on the MAC $W$ and its corresponding bit-channels. Recall the channel combining operations from [1]. For any two single-user binary-input DMCs $W_1 : \mathcal{X} \rightarrow \mathcal{Y}_1$ and $W_2 : \mathcal{X} \rightarrow \mathcal{Y}_2$ (indeed $\mathcal{Y} = \{0,1\}$), $W_1 \oplus W_2 : \mathcal{X} \rightarrow \mathcal{Y}_1 \times \mathcal{Y}_2$ denote another B-DMC whose transition probability for any $(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ and $x \in \mathcal{X}$ is given by

$$W_1 \oplus W_2(y_1, y_2 | u) = \frac{1}{2} \sum_{x \in \mathcal{X}} W_1(y_1 | u \oplus x) W_2(y_2 | u)$$

(13)

Also, $W_1 \odot W_2 : \mathcal{X} \rightarrow \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{X}$ denote another B-DMC whose transition probability for any $(y_1, y_2, x) \in \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{X}$ and $x, u \in \mathcal{X}$ is given by

$$W_1 \odot W_2(y_1, y_2, x | u) = \frac{1}{2} W_1(y_1 | u \oplus x) W_2(y_2 | u)$$

(14)

Lemma 3: For any multiple access channel $W$ and any $n \geq 2$, $N = 2^n$, $i \leq N$,

$$W_{2N}^{(2i-1)} = W_N^{(i)} \oplus W_N^{(i)}$$

and

$$W_{2N}^{(2i)} = W_N^{(i)} \odot W_N^{(i)}$$

Proof: Let $d_1^{2N}$ and $d_1^{2N}$ denote the even-indexed and odd-indexed subvectors, respectively. Then

$$W_{2N}^{(2i)} = \frac{1}{2^{2N-1}} \sum_{d_{2i+1}^{2N}} \tilde{W}_N(y_1^N d_1^{2N} d_{2i+1}^{2N}) \tilde{W}_N(y_1^N d_1^{2N})$$

$$= \frac{1}{2^{2N-1}} \sum_{d_{2i+1}^{2N}} \tilde{W}_N(y_1^N d_{2i+1}^{2N})$$

$$= W_N^{(i)} \odot W_N^{(i)}$$

where we used the definitions of bit-channels in [12], the channel operation $\odot$ in [14] and the recursive structure of polar
transformation. The other equation can be derived using the similar arguments.

**Corollary 4:** For a given $i \in \{1, 2, 3, 4\}$, let $\mathcal{W}$ denote the single user channel $W_i^{(j)}$. Let also $n \geq 1$ and $N = 2^n$. Then for any $\frac{(i-1)N}{2} + 1 \leq j \leq \frac{iN}{2}$,

$$W_{2N}^{(j)} = \mathcal{W}^{(j - \frac{(i-1)N}{2})}$$

where the bit-channels with respect to $\mathcal{W}$ are defined as in [4].

**Proof:** The proof is by induction on $n$. The base of induction is trivial for $n = 1$. The induction steps are by the fact that channel combining recursion steps in Arikan’s polar transformation, as proved in Proposition 3 of [1], match with Lemma [3].

In the MAC polarization building block depicted in Figure 5, assume that $U_1, U_2, V_1, V_2$ are independent uniform binary random variables. Then let $I_1 = (U_1; Y_1^2), I_2 = (V_1; Y_1^2, U_1), I_3 = (U_2; Y_2, V_1^2, U_1), \text{and } I_4 = (V_2; Y_2, U_2, V_1^2, U_1)$.

**Lemma 5:** For any two-user binary-input MAC $W$

$$I_1 + I_2 + I_3 + I_4 = 2I(W)$$

**Proof:** The proof is simply by chain rule on the mutual information, given independent and uniform channel inputs $U_1^2$ and $V_1^2$:

$$2I(W) = I(U_1 \oplus U_2, V_1 \oplus V_2, U_2, V_2; Y_1^2)$$

$$= I(U_1, V_1, U_2, V_2; Y_1^2) = I_1 + I_2 + I_3 + I_4$$

For $N = 2^n$, which will be the length of code for each of the users, and $\beta < \frac{1}{4}$, we define the set of good bit-channels as follows:

$G^{(1)}_N(W, \beta) \overset{\text{def}}{=} \left\{ i \in [2N] : \left\lfloor \frac{2i}{N} \right\rfloor \in \{1, 3\} \text{ and } Z(W_{2N}^{(i)}) < 2^{-N^3}/N \right\}$

$G^{(2)}_N(W, \beta) \overset{\text{def}}{=} \left\{ i \in [2N] : \left\lfloor \frac{2i}{N} \right\rfloor \in \{2, 4\} \text{ and } Z(W_{2N}^{(i)}) < 2^{-N^3}/N \right\}$

**Theorem 6:** For any two-user binary-input discrete MAC $W$ and any constant $\beta < \frac{1}{2}$ we have

$$\lim_{N \to \infty} \frac{|G^{(1)}_N(W, \beta)|}{N} = \frac{I_1 + I_3}{2}$$

$$\lim_{N \to \infty} \frac{|G^{(2)}_N(W, \beta)|}{N} = \frac{I_2 + I_4}{2}$$

**Proof:** By Corollary 4,

$$G^{(1)}_N(W, \beta) = \mathcal{G}_N^2(W_4^{(1)}, \beta) \cup \mathcal{G}_N^2(W_4^{(3)}, \beta)$$

and similarly,

$$G^{(2)}_N(W, \beta) = \mathcal{G}_N^2(W_4^{(2)}, \beta) \cup \mathcal{G}_N^2(W_4^{(4)}, \beta)$$

There exists $\beta' < \frac{1}{4}$ such that $N^3 < (N/2)^{\beta'}$ for large enough $N$. Therefore, channel polarization theorem for the single-user case [10] imply that

$$\lim_{N \to \infty} \frac{|G^{(1)}_N(W, \beta)|}{N} = \frac{1}{2} \lim_{N \to \infty} \frac{|G_N^2(W_4^{(1)}, \beta')|}{N/2}$$

$$+ \frac{1}{2} \lim_{N \to \infty} \frac{|G_N^2(W_4^{(3)}, \beta')|}{N/2} = \frac{I_1 + I_3}{2}$$

and similarly,

$$\lim_{N \to \infty} \frac{|G^{(2)}_N(W, \beta)|}{N} = \frac{1}{2} \lim_{N \to \infty} \frac{|G_N^2(W_4^{(2)}, \beta')|}{N}$$

$$+ \frac{1}{2} \lim_{N \to \infty} \frac{|G_N^2(W_4^{(4)}, \beta')|}{N} = \frac{I_2 + I_4}{2}$$

But both inequalities should be equality as the fraction of all good bit-channels can not exceed $I(W) = (I_1 + I_2 + I_3 + I_4)/2$. This completes the proof of theorem.

**Corollary 7:** The polar scheme constructed with respect to the MAC polarization building block shown in Figure 5 achieves the point ($(I_1 + I_3)/2, (I_2 + I_4)/2$) on the dominant face.

**Proof:** The polar codes for the first and second user are constructed by putting information bits on the positions in $G^{(1)}_N(W, \beta)$ and $G^{(2)}_N(W, \beta)$, respectively, and fixing the rest to a random choice which is revealed to the receiver. Then by Theorem 6 the rates of codes approach ($(I_1 + I_3)/2, (I_2 + I_4)/2$) which is on the dominant face by Lemma [5].

Similarly, ($(I_2 + I_4)/2, (I_1 + I_3)/2$) can be achieved if we replace the two users with each other. More points on the dominant face can be achieved if we modify the decoding order. Consider the MAC polarization building block shown in Figure 5 with decoding order $(u_1, v_1, v_2, u_2)$. As a result, after $n - 1$ more levels of polarization, the decoding order is as follows. The first $N/2$ bits of the first user is decoded, where $N = 2^n$, then all the bits of the second user are decoded, and at the end, the remaining $N/2$ bits of the first user are decoded. Using the same arguments discussed in this subsection we can show that the new code achieves the point ($(I_1 + I_3)/2, (I_2 + I_3)/2$) on the dominant face of the uniform rate region. ($(I_1 + I_4)/2, (I_2 + I_3)/2$) is also achieved by symmetry, if the first user and the second user are replaced with each other.

### B. Finite length improvement and simulation results

In this subsection, we provide simulation results of the proposed construction based on $4 \times 2$ MAC polarization building block. As we will see, 6 different points on the dominant face of the capacity region can be achieved asymptotically with no need of time sharing. Also, the finite length performance is improved by allowing the rates transmitted by each users vary while fixing the target FER.

The same model as in Section [III] is picked for simulation, a binary-additive two-user Gaussian channel $W$ with BPSK modulation and noise of unit variance. Using the $4 \times 2$ polarization building blocks, one can achieve four different points on the dominant face of the capacity region, two symmetric points for each of the building blocks. There are also two corner points achievable by the separated scheme i.e. corresponding to the
decoding order \((U_1, U_2, V_1, V_2)\) and \((V_1, V_2, U_1, U_2)\). In fact the total number of achievable points using building block of length up to four, i.e. either two or four, is \(6 = \binom{4}{2}\), subject to the constraint that \(u_1\) (or \(v_1\)) is decoded before \(u_2\) (or \(v_2\)). The parameters \(I_1, I_2, I_3\) and \(I_4\), defined in the previous subsection, are calculated as follows:

\[
I_1 = I(U_1; Y_1^2) = H(Y_1^2) - H(Y_1^2 | U_1) = 0.1550
\]
\[
I_2 = I(U_1; Y_1^2 | U_1) = H(Y_1^2 | U_1) - H(Y_1^2 | U_1, V_1) = 0.4646
\]
\[
I_3 = I(U_2; Y_2^2 | U_1, V_1) = H(Y_2^2 | U_1, V_1) - H(Y_2^2 | U_1^2, V_1) = 0.6889
\]
\[
I_4 = I(V_2; Y_2^2 | U_1^2, V_1) = H(Y_2^2 | U_1^2, V_1) - H(Y_2^2 | U_1^2, V_1^2) = 0.9115
\]

The channel is fixed as described above and we look at achievable rates at finite block lengths given the total probability of frame error \(10^{-2}\) on decoding both \(U\) and \(V\). Polar codes of length \(N = 2^n\) for \(n = 9, 10, 11, 12\), are constructed for the channels \(W\) and \(\hat{W}\). For each block length, let \(R_1\) and \(R_2\) be the rates assigned to the first and second user, respectively. For each of the six possible decoding orders, the probability of error for all the bit-channels are estimated by running the successive cancellation decoder over a sample set of size \(10^6\). We let the rates vary as the probability of error for each of the users varies between 0 and \(10^{-2}\). The probability of error for each user is found by summing over the estimated probability of error of the best selected bit-channels. Let \(P_1\) and \(P_2\) be the probability of error for the first and the second user, respectively. The target probability of error is \(10^{-2}\). Therefore, the only constraint is \(P_1 + P_2 \leq 10^{-2}\). As a result, each decoding order results in a different rate region. The achievable region is simply the union of all the 6 regions. The achievable regions for \(N = 2^n\), \(n = 9, 10, 11, 12\), are shown in Figure 6. In an asymptotic sense, all the bit-channels polarize and they are either good or bad. Therefore, the capacity region becomes a collection of step functions.

**Remark.** The proposed technique for achieving more points on the dominant face of the uniform rate region also enables more efficient time sharing with larger achievable region. For instance, in the separated scheme, by increasing the transmission block from \(N\) to \(2N\), only one more point on the dominant face can be achieved asymptotically. On the other hand, given the 6 achievable points on the dominant face by our scheme, \(\binom{6}{2} = 15\) more points can be achieved by increasing the transmission block to \(2N\), while the constituent codes have length \(N\). □

**VI. ACHIEVING THE UNIFORM RATE REGION WITH JOINT DECODING**

In this section, we extend the construction proposed in Section V to construct schemes with larger polarization building blocks. Intuitively, by extending the size of the polarization building block one should be able to get more points on the dominant face of the uniform rate region with smaller resolution. Our aim is to get arbitrarily close to all the points on the dominant face of the capacity region as the block length goes to infinity.

**A. Proposed construction with general polarization building block**

Let \(W\) be a given two-user binary-input discrete multiple access channel. Let also \(l \geq 1\) be a positive integer and \(L = 2^l\). Assume that \(U_l^L = (U_i, U_2, \ldots, U_L)\) and \(V_l^L = (V_1, V_2, \ldots, V_L)\) are two vectors of independent and uniformly distributed bits that are independently generated by the first user and the second user, respectively. Let \(X_l^L = U_l^L \cdot G_l^L\) and \(Y_l^L = V_l^L \cdot G_l^L\). For \(i = 1, 2, \ldots, L\), the pair \((X_i, X_i')\) is transmitted through \(i\)-th independent copy of \(W\) and the output is denoted by \(Y_i\). There are several ways to form an ordered sequence from the vector \((U_l^L, V_l^L)\). In general, there are \(2L!\) permutations that can be applied to the sequence \((U_l^L, V_l^L)\). However, we limit our consideration to a small subset of all possible permutations that is sufficient for us to establish a bound on the resolution of achievable points on the dominant face.

For \(i = 1, 2, \ldots, L + 1\), let

\[
P^{(i)} \equiv (V_i^L, U_i^L, V_i^L)
\]

Then the mutual information \(I(U_l^L, V_l^L; Y_l^L)\) can be expanded with respect to the ordered sequence \(P^{(i)}\) using the chain rule as follows:

\[
I(U_l^L, V_l^L; Y_l^L) = \sum_{j=1}^{i-1} I(V_j; Y_j^L, V_j^{i-1}) + \sum_{j=i}^{L} I(U_j; Y_j^L, V_j^{i-1}, U_j^{i-1}) + \sum_{j=i}^{L} I(V_j; Y_j^L, V_j^{i-1}, U_j^{i-1}, U_L^L)
\]

Then for \(j = 1, 2, \ldots, 2L\), let \(I_{j,i}\) denote the \(j\)-th term in the above equation i.e.

\[
I_{j,i} \equiv \begin{cases} I(V_j; Y_j^L, V_j^{i-1}) & \text{if } j \leq i - 1 \\ I(U_{j-i+1}; Y_j^L, V_j^{i-1}, U_j^{i-1}) & \text{if } i \leq j \leq i + L - 1 \\ I(V_j; Y_j^L, U_j^{i-1}, U_L^L) & \text{if } L + i \leq j \leq 2L \\
\end{cases}
\]

Fig. 6: Achievable regions for various block lengths
Also, let

\[ R_{i_1}^{(1)} \overset{\text{def}}{=} \frac{\sum_{j=2}^{L+i-1} I_{j,i}}{L} \]  

(17)

and

\[ R_{i_1}^{(2)} \overset{\text{def}}{=} \frac{\sum_{j=1}^{i-1} I_{j,i} + \sum_{j=L+1}^{2L} I_{j,i}}{L} \]  

(18)

As we will see, \((R_{i_1}^{(1)}, R_{i_1}^{(2)})\) is the achievable pair of rates using the \(2L \times L\) polarization building block with decoding order \(P^{(i)}\).

**Lemma 8:** For \(i = 1, 2, \ldots, L\),

\[(R_{i+1}^{(1)} - R_{i+1}^{(2)}, R_{i+1}^{(2)} - R_{i+1}^{(1)}) \in \left[-\frac{1}{L}, 0\right] \times \left[0, \frac{1}{L}\right].\]

**Proof:** By definition of \(R_{i_1}^{(2)}\) and \(I_{j,i}\) in (18) and (16),

\[ L(R_{i_1}^{(2)} - R_{i+1}^{(2)}) = \sum_{j=1}^{i-1} I(V_j; Y_1^L, Y_1^{i-1}) \]

\[ + \sum_{j=1}^{L} I(V_j; Y_1^L, V_1^{i-1}, U_1^L) \]

\[ - \sum_{j=1}^{i} I(V_j; Y_1^L, V_1^{i-1}) \]

\[ - \sum_{j=L+1}^{2L} I(V_j; Y_1^L, V_1^{i-1}, U_1^L) \]

\[ = I(V_i; Y_1^L, V_1^{i-1}, U_1^L) - I(V_i; Y_1^L, V_1^{i-1}) \]  

(19)

Also,

\[ 0 \leq I(V_i; Y_1^L, V_1^{i-1}) \leq I(V_i; Y_1^L, V_1^{i-1}, U_1^L) \leq 1 \]  

(20)

(19) and (20) together imply that

\[ 0 \leq R_{i_1}^{(2)} - R_{i_1+1}^{(2)} \leq \frac{1}{L}. \]

Observe that \(R_{i_1}^{(1)} + R_{i_1}^{(2)} = R_{i+1}^{(1)} + R_{i+1}^{(2)} = \frac{1}{L} I(U_1^L; Y_1^L; Y_1^L)\)

which completes the proof of lemma. \(\blacksquare\)

**Theorem 9:** For \(i \in [L + 1]\), \((R_{i_1}^{(1)}, R_{i_1}^{(2)})\) is located on the dominant face of the uniform rate region. Furthermore, \((R_{i_1}^{(1)}, R_{i_1}^{(2)})\) is equal to \(A_W\) for \(i = 1\), and is equal to \(B_W\) for \(i = L + 1\).

**Proof:** Observe that

\[ R_{i_1}^{(1)} + R_{i_1}^{(2)} = \frac{I(U_1^L; Y_1^L; Y_1^L)}{L} = I(W) \]  

(21)

Hence, \((R_{i_1}^{(1)}, R_{i_1}^{(2)})\) is located on the line connecting \(A_W\) and \(B_W\). Therefore, it is sufficient to prove the second part only for one coordinate. Let \(X_1^L = U_1^L, G^L\) and \(X_1^L = V_1^L, G^L\).

\[ LR_{i_1}^{(1)} = \sum_{j=1}^{L} I(U_j; Y_1^L, U_1^{L-1}) = I(U_1^L; Y_1^L) = I(X_1^L; Y_1^N) \]

\[ = LI(W) \]

Therefore, \(R_{i_1}^{(1)}\) is equal to the first coordinate of \(A_W\). Similarly, \(R_{i_1}^{(2)}\) is equal to the first coordinate of \(B_W\) if we replace \(U\)’s with \(V\)’s. \(\blacksquare\)

For the first part of lemma, notice that by Lemma 8,

\[ R_{i_1}^{(1)} \leq R_{i_1}^{(2)} \leq \ldots \leq R_{i_1+1}^{(1)} \]

This together with (21) and second part of lemma, show that all the points \((R_{i_1}^{(1)}, R_{i_1}^{(2)})\) are located on the segment connecting \(A_W\) and \(B_W\).

**Corollary 10:** For any point \(Q = (Q_1, Q_2)\) on the dominant face of the uniform rate region of \(W\), there exists \(i \in [L + 1]\) such that

\[ |Q_1 - R_{i_1}^{(1)}| \leq \frac{1}{L} \quad \text{and} \quad |Q_2 - R_{i_1}^{(2)}| \leq \frac{1}{L} \]

**Proof:** Since \(Q\) is located on the dominant face, by Theorem 9

\[ R_{i_1}^{(1)} \leq Q_1 \leq R_{i_1+1}^{(1)} \]

Let \(i\) be the largest element in \([L + 1]\) such that

\[ R_{i_1}^{(1)} \leq Q_1 \]

Therefore, \(R_{i_1+1} > Q_1\) and by Lemma 8

\[ 0 \leq Q_1 - R_{i_1}^{(1)} \leq R_{i_1+1}^{(1)} - R_{i_1}^{(1)} \leq \frac{1}{L} \]

Since \((R_{i_1}^{(1)}, R_{i_1}^{(2)})\) is also located on the dominant face by Theorem 9, the second part of lemma follows. \(\blacksquare\)

Next, we show that all the pairs of rates \((R_{i_1}^{(1)}, R_{i_1}^{(2)})\), defined in (17) and (18) can be achieved using the \(2L \times L\) polarization building block with decoding order \(P^{(i)}\). For any \(n \geq l\), let \(N = 2^n\) and suppose that \(n - l\) recursion steps of polar transformation are deployed on top of the polarization building block, same as Arikan’s polar code. To make this statement more precise, consider the polar transformation discussed Section V-A and recall the channel \(\tilde{W}_{2N}\) with input \((u_1^N, v_1^N)\) and output \(y_1^N\).

Let \(d_1^N\) denote the ordered vector \((v_1^N, u_1^N, v_1^N, \ldots, u_1^N, v_1^N)\). The bit-channels are defined with respect to this ordered sequence, which will also enforce the decoding order in the joint successive decoding. For \(j = 1, 2, \ldots, 2N\), the \(j\)-th bit-channel is defined as

\[ W_{2N}(y_1^N, d_1^{j-1} | d_j) \overset{\text{def}}{=} \frac{1}{2^{2N-I}} \sum_{d_1^{j-1} \in \{0, 1\}^{2N-j}} \tilde{W}_{2N}(y_1^N | d_1^{j-1}) \]  

(22)

The decoding order is in fact given as follows: \(v_1^{N-(n-l)}\) are decoded first. Then \(v_1^N\) are decoded and at the end the remaining \(v_j^N\)’s, that is \(v_1^{N-(n-l)+1}\), are decoded. As we will see, this is the result of \(n - l\) levels of polarization of the \(2L \times L\) polarization building block with decoding order enforced by \(P^{(i)}\) defined in (15). The decoding order is assumed to be fixed for the rest of this section. Actually the results do not depend on the choice of \(i \in [L + 1]\).

The following lemma is similar to the Proposition 3 of [1] and Lemma 3 that still holds for this particular polar transformation defined here and its corresponding bit-channels.

**Lemma 11:** Given \(N\) independent copies of \(W\) and the bit-channels defined in (22), for any \(j\) with \(1 \leq j \leq N\),

\[ W_{2N}^{(j-1)} = W_{N}^{(j)} \oplus W_{N}^{(j)} \]
and

\[ W_{2N}^{(2j)} = W_{N}^{(j)} \oplus W_{N}^{(j)} \]

The proof is similar to the proof of Lemma 3. The following corollary is also a generalization of Corollary 4.

**Corollary 12:** For a given \( k \in [2L] \), let \( W \) denote the single user channel \( W_{2L}^{(k)} \). Let also \( n \geq l \) and \( N = 2^n \). Then for any \( \frac{(k-1)N}{L} + 1 \leq j \leq \frac{kN}{L} \),

\[ W_{2N}^{(j)} = W_{N}^{(j-\frac{(k-1)N}{L})} \]

where the bit-channels with respect to \( W \) are defined as in (4).

The set of indices \([2N]\) is split into two subsets \( S_{N}^{(1)} \) and \( S_{N}^{(2)} \) as follows:

\[ S_{N}^{(1)} = \left\{ j \in [2N] : i - 1 < \frac{Lj}{N} \leq L + i - 1 \right\} \]

\[ S_{N}^{(2)} = \left\{ j \in [2N] : \frac{Lj}{N} \leq i - 1 \text{ or } L + i - 1 < \frac{Lj}{N} \right\} \]

The set of good bit-channels for both users are defined as follows. For any \( \beta < \frac{1}{2} \) and \( N = 2^n \)

\[ G_{N}^{(1)}(W, \beta) = \left\{ j \in S_{N}^{(1)} : Z(W_{2L}^{(j)}) < 2^{-N^\beta} \right\} \]

\[ G_{N}^{(2)}(W, \beta) = \left\{ j \in S_{N}^{(2)} : Z(W_{2L}^{(j)}) < 2^{-N^\beta} \right\} \]

Next theorem generalizes the results of Theorem 3 for the construction based on \( 2L \times L \) MAC polarization building block with arbitrary \( L \).

**Theorem 13:** For any two-user binary-input discrete MAC \( W \) and any constant \( \beta < \frac{1}{2} \) we have

\[ \lim_{N \to \infty} \frac{|G_{N}^{(1)}(W, \beta)|}{N} = R_{i}^{(1)} \]

\[ \lim_{N \to \infty} \frac{|G_{N}^{(2)}(W, \beta)|}{N} = R_{i}^{(2)} \]

**Proof:** By Corollary 12

\[ G_{N}^{(1)}(W, \beta) = \bigcup_{j \in S_{N}^{(1)}} G_{2L}^{(j)}(W_{2L}^{(j)}, \beta) \]

and similarly,

\[ G_{N}^{(2)}(W, \beta) = \bigcup_{j \in S_{N}^{(2)}} G_{2L}^{(j)}(W_{2L}^{(j)}, \beta) \]

There exists \( \beta' < \frac{1}{2} \) such that \( \gamma_{\beta'} < \left( \frac{N}{2} \right)^{\beta} \) for large enough \( N \). Therefore, channel polarization theorem for the single-user case [10] imply that

\[ \lim_{N \to \infty} \frac{|G_{N}^{(1)}(W, \beta)|}{N} \geq \frac{1}{L} \sum_{j \in S_{N}^{(1)}} \lim_{N \to \infty} \frac{|G_{2L}^{(j)}(W_{2L}^{(j)}, \beta)|}{N/L} \]

\[ = \frac{1}{L} \sum_{j=i}^{L+i-1} I_{j,i} = R_{i}^{(1)} \]

and similarly,

\[ \lim_{N \to \infty} \frac{|G_{N}^{(2)}(W, \beta)|}{N} \geq \frac{1}{L} \sum_{j \in S_{N}^{(2)}} \lim_{N \to \infty} \frac{|G_{2L}^{(j)}(W_{2L}^{(j)}, \beta)|}{N/L} \]

\[ = \frac{1}{L} \sum_{j=i}^{L+i-1} I_{j,i} = R_{i}^{(2)} \]

Since the fraction of all good bit-channels can not exceed \( I(W) = R_{i}^{(1)} + R_{i}^{(2)} \), both the above inequalities should be equality. This completes the proof of theorem.

**B. Encoding, decoding and asymptotic performance**

The encoding of the proposed scheme in the previous subsection is similar to that of original polar codes for each of the users. In fact, regardless of the choice of polarization building block the encoder for any block length \( N \) is fixed, where each user multiplies a vector of length \( N \) by the polarization matrix \( G^{\otimes n} \). The only thing that depends on the polarization building block is the indices of information bits. For a fixed length of the polarization building block \( L = 2^l \) and decoding order \( P^{(i)} \) defined in (15), the set of good bit-channels for both users \( G_{N}^{(1)}(W, \beta) \) and \( G_{N}^{(2)}(W, \beta) \) are defined in (25). In the polar encoder for the first user, \( u_j \) is an information bit for any \( j \in G_{N}^{(1)}(W, \beta) \). Otherwise, \( u_j \) is frozen to a fixed value. Therefore, \( R_{i}^{(1)} \) is indeed the rate of polar code for the first user. Similarly, in the polar encoder for the second user, \( v_j \) is an information bit for any \( j \in G_{N}^{(2)}(W, \beta) \). Otherwise, \( v_j \) is frozen to a fixed value. Therefore, \( R_{i}^{(2)} \) is indeed the rate of polar code for the second user. The frozen bits are chosen uniformly at random, revealed to the receiver a priori and fixed during the course of communication. If the underlying channels are symmetric, then the frozen bits can be set to zeros.

For joint decoding, the successive cancellation decoding defined in (1) is performed, with some modification. The order of decoding is determined by \( P^{(i)} \), defined in (15). Bits \( v_{1}^{(i-1)} \) are decoded first, then \( v_{1}^{N} \) followed by \( v_{1}^{N-1} \) are decoded. The low-complex implementation of successive cancellation decoder invented by Arikan in (1) can be extended to our scheme for MAC. The recursive steps of computing likelihood ratios can be done in a similar way using a 2\( N \times (n - l + 1) \) trellis. The likelihood ratios of the bits in the decoding building block of length \( 2L \) can be computed using a naive way with complexity \( O(2L2^{2L}) \). Therefore, the total complexity of decoding is given by \( O(2N(n - l + 1 + 2L)) \). \( L \) is regarded as a fixed parameter in the scheme. Therefore, the decoding complexity is asymptotically \( O(N \log N) \), similar to separate polar decoding.

For the asymptotic performance of polar codes, the following theorem follows similar to (1) and (10).

**Theorem 14:** For any \( \beta < \frac{1}{2} \), any two-user MAC \( W \), any \( \epsilon > 0 \) and any point \( Q \) on the dominant face of the uniform rate region, there exists a family of polar codes that approaches a point on the dominant face within distance \( \epsilon \) from \( Q \). Furthermore, the probability of frame error under successive cancellation decoding is less than \( 2^{-N^\beta} \), where \( N \) is the block length of the code for each of the users.
Proof: The choice of \( L \) for the polarization building block depends on \( \epsilon \). Fix \( L = 2^l \) such that \( \epsilon^{2^l} < \epsilon \). Then there exists \( i \) such that \( (R^{(1)}_i, R^{(2)}_i) \) satisfies the condition in Corollary 10. Consequently, the distance between \( Q \) and \( (R^{(1)}_i, R^{(2)}_i) \) is less than \( \epsilon \). For any block length \( N = 2^n \gg L \), the polar codes for first user and second user are constructed with respect to the set of good bit-channels \( G^{(1)}_N(W, \beta) \) and \( G^{(2)}_N(W, \beta) \), defined in (22), respectively. Then by Theorem 13, the rates approach \( (R^{(1)}_i, R^{(2)}_i) \) as \( N \) goes to infinity. The probability of frame error is bounded by the sum of the Bhattacharyya parameters of the selected bit-channels. Therefore, it is less than \( 2^{-N^\alpha} \) for both users by definition of \( G^{(1)}_N(W, \beta) \) and \( G^{(2)}_N(W, \beta) \).

VII. COMPARISON WITH PRIOR AND RELATED WORKS

In this section, we first review the scheme given by Saçedoğlu, Telatar and Yeh in [5] which is based on MAC polarization in order to make a comparison with our scheme. The approaches are essentially different. The authors of [5] propose a framework for MAC polarization by extending the notion of channel splitting from the single-user case to the two-user case and identifying all the possible extremes as the block length goes to infinity. On the other hand, we propose a new criterion for the decoding order using which, along with the single-user channel polarization, we manage to approach all the points in the uniform rate region. We compare both schemes in terms of the decoding complexity and achievable regions. Lastly, other related works are mentioned [15], [16] which are independent and contemporaneous to our work. While some of the main results in [15] and in this paper are similar, there are also important differences that we will emphasize.

A. Overview of Saçedoğlu-Telatar-Yeh scheme

In this subsection, we review the scheme for two-user multiple access channel polarization proposed in [5]. Let the inputs \( U \) and \( V \) to a given two-user MAC \( W \) are independent and uniform. Then define

\[
I^{(1)}(W) \overset{\text{def}}{=} I(U; Y, V)
\]

\[
I^{(2)}(W) \overset{\text{def}}{=} I(V; Y, U)
\]

and let \( K(W) = (I^{(1)}(W), I^{(2)}(W), I(W)) \in \mathbb{R}^3 \). The notion of channel combining operations are extended from the single-user case to the case of two-user MAC. This is depicted in Figure 7. Two independent uses of \( W \) is turned into two MACs \( W^- \) and \( W^+ \). Consecutively, by doing one more level of polarization, there will be four bit-channels \( W^-, W^-, W^+ \) and \( W^+ \) etc.

Let

\[
M = \{(0, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 1), (1, 1, 2)\} \subset \mathbb{R}^3
\]

Then it is proved in [5] that as \( n \) goes to infinity, for any string \( S \in \{+, -\}^n \) chosen uniformly at random, \( K(W^S) \) approaches one of the points in the set \( M \) with probability 1. It means that there are five extreme cases for a two-user MAC while there are only two extremes, either noise-less or pure noise channel, in the single user case. The point \((0, 0, 0)\) corresponds to the case that the output does not provide any information about neither of its inputs. The points \((0, 1, 1)\) and \((1, 0, 1)\) correspond to the cases that the output provides complete information about one of the inputs and nothing about the other. \((1, 1, 2)\) corresponds to the case that the output gives complete information about both of the inputs. The unexpected point \((1, 1, 1)\) is a pure contention channel. In that case, the uniform rate region is the triangle with vertices \((0, 0), (0, 1), (1, 0)\). For this channel, if any of the users communicate at zero rate, the other user observes a noise-less channel.

The idea for the code construction is similar to the single-user case i.e. either freeze the input bits or assign an information bit to them depending on the corresponding bit-channel. This can be done in a straightforward way for the four cases except the pure contention case \((1, 1, 1)\). For the pure contention channel, one of the inputs is picked arbitrarily to carry the information bit while the other one is frozen. In [5], the first user is always picked to transmit on the pure contention channel. The frozen bits are chosen uniformly at random and then fixed. They are revealed to the receiver a priori. It is proved in [5] that the constructed polar code achieves a point on the dominant face of the uniform rate region. The achievable region by this MAC polarization approach is not discussed in [5]. It is just pointed out that in general, the achievable region is a subset of the uniform rate region. Later in this section, we will show an example in which only a single point on the dominant face can be achieved using this approach.

B. Decoding complexity

Both schemes use a successive cancellation decoding, originally proposed by Arikan in [1]. As discussed in the foregoing section, we have to combine the low-complex Arikan’s decoder with a basic decoder for the building block. As a result, we have an extra term \( O(N^2 L^3) \) term in the complexity which is dominated by \( O(N \log N) \) as \( N \) goes to infinity, while \( L \) is assumed to be constant. On the other hand, in the scheme proposed in [5], the likelihood of a vector of length two needs to be tracked along the decoding trellis. Therefore, instead of a simple likelihood ratio, a vector of length 4 for the probability of all 4 possible cases has to be computed recursively. This increases the decoding complexity by a factor of 2. In fact, the decoding complexity is still \( O(N \log N) \), but if we look at the actual number of operations needed to complete the decoding, the decoding complexity of the scheme in [5] is two times more than the decoding complexity of our scheme, asymptotically.
C. Achievable region

In Theorem 14 we proved that all the points on the dominant face of the uniform rate region can be achieved with arbitrary small resolution. As pointed out before, the scheme proposed in [5] does not necessarily achieve the whole uniform rate region. It is only guaranteed that one point on the dominant face is achievable. In this subsection, we provide an example for a two-user MAC such that the scheme proposed in [5], only achieves one point on the dominant face.

Let \( W' : \{0, 1\} \rightarrow \mathcal{Y} \) be an arbitrary single user binary-input DMC. Then the two-user multiple access channel \( W : \{0, 1\} \rightarrow \mathcal{Y} \) is constructed as follows. Let \( \mathcal{Y} = \mathcal{Y} \times \mathcal{Y} \).

For any two binary inputs \( u \) and \( v \) corresponding to the first and second user, \( u \oplus v \) and \( v \) are transmitted through two independent copies of \( W' \) with outputs \( y_1 \) and \( y_2 \). Then the output of \( W \) is defined to be \( y = (y_1, y_2) \in \mathcal{Y} \). The diagram of \( W \) is depicted in Figure 8.

![Fig. 8: A constructed example for two-user MAC](image)

Let \( I = I(W') \), \( I^- = I(W' \oplus W') \) and \( I^+ = I(W' \ominus W') \).

Let also \( U \) and \( V \) be two independent and uniformly distributed binary random variables. Then

\[
I(U; Y|V) = I(U; Y; V) = I(U \oplus V; Y; V) = I(U \oplus V; Y_1) = I(W') = I
\]

On the other hand

\[
I(V; Y|U) = I(V; Y_1, Y_2, U) = I(W' \ominus W') = I^+
\]

And

\[
I(U, V; Y) = I(U \oplus V, V; Y_1, Y_2) = 2I(W') = 2I
\]

Notice that \( I^+ + I^- = 2I \). Therefore, the uniform rate region of \( W \) is as shown in Figure 9.

It is easy to observe that a polar transformation of length \( N \) on two-user MAC \( W \), in the sense defined in [5], is equivalent to a single-user polar transformation of length \( 2N \) over \( W' \). Let \( U_1^N \) and \( V_1^N \) be the input bits by the first and the second user, respectively. Let \( W_{N}^{(i)} \) be the \( i \)-th two-user bit-channel observed by \( (U_i, V_i) \). Then

\[
I^{(2)}(W_{N}^{(i)}) = I(U_i; Y_1^i, U_1^i, V_1^{i-1}) = I(W_{2N}^{(2i)})
\]

Also,

\[
I^{(1)}(W_{N}^{(i)}) = I(U_i; Y_1^i, U_1^{i-1}, V_1^i) \geq I(U_i; Y_1^i, U_1^{i-1}, V_1^{i-1}) = I(W_{2N}^{(2i-1)})
\]

![Fig. 9: The uniform rate region for the constructed example](image)

Notice that as \( N \) goes to infinity, \( W_{2N}^{(2i-1)} \) and \( W_{2N}^{(2i)} \) are both good or both bad, for most of the \( i \)'s. In fact, the fraction of \( i \)'s for which only one of them is good approaches zero. Therefore, the triple \( (I^{(1)}(W_{N}^{(i)}), I^{(2)}(W_{N}^{(i)}), I(W_{N}^{(i)})) \) approaches either \( (0, 0, 0) \) or \( (1, 1, 2) \) among the five possible cases discussed in Section 11. Thus, the only achievable point on the dominant face by this scheme is the point \( (I, I) \), and the achievable region is the square with vertices \( (0, 0), (0, I), (I, 0), (I, I) \).

On the other hand, as proved in the previous subsection, the whole region shown in Figure 9 can be achieved by our proposed scheme with joint successive decoder. For instance, in the simplest case which also matches with the separated scheme, the point \( (I^-, I^+) \) is achieved if \( u_N^1 \) is decoded first and then \( v_1^1 \) is decoded.

D. Other related works

Related works on polar codes for the dual problem of Slepian-Wolf [16] and two-user multiple access channels [15] are appeared which are independent and contemporaneous to our work. While some of the underlying ideas are similar, there are some essential differences as we explain.

Arikman gives a polar-based scheme that achieves the full admissible rate region in the Slepian-Wolf problem with time-sharing [16]. The method is based on monotone chain rule expansions for source polarization. Although the duality between the Slepian-Wolf problem and the two-user MAC is mentioned in [16], no explicit formulation is given. The technique of monotone chain rule expansion is actually applied to the problem of two-user MAC in [15]. This expansion is similar to our expansion of mutual information between the inputs and output given in [16]. However, the focus of [15] is on the decoding side and improving it using the list decoding algorithm of polar codes proposed in [17], while we provide comprehensive study.
on polarization and achievability of the whole uniform rate region based on proposed MAC-polarization building blocks. We also provide bounds on the error-decay rate and discuss the joint construction as well as the joint successive cancellation decoding. Several simulation results are provided along with the full characterization of achievable regions in finite length binary additive Gaussian channel. Furthermore, in the case of time-sharing, the application of the compound polar code of [9] is discussed to improve the performance by taking advantage of increased latency.

VIII. DISCUSSIONS AND CONCLUSION

In this paper, we considered the problem of designing polar codes for transmission over two-user multiple access channels. We started by the straightforward method of time-sharing to achieve the uniform rate region. We explained how to apply compound polar codes to this case in order to improve the finite-length performance. Then we proposed a new scheme with a joint successive cancellation decoder on a designed MAC polarization building block. Furthermore, we proved that all the points in the uniform rate region can be achieved by changing the decoding order of the MAC polarization building block.

Since the individual codes for each of the users can be regarded as single-user polar codes, all the existing methods for improving the performance of single-user polar codes can be applied on top of our proposed MAC-polar codes. For instance, the individual polar codes can be made systematic as suggested by Arikan [18] in order to improve the bit error rate.

A direction for future work is to extend the proposed construction to $m$-user multiple access channel. This is not very straightforward as the dominant face of the uniform rate region is a polygon in an $m-1$-dimensional plane in the $m$-dimensional space. Then the permutation on the decoding order must be changed in such a way that all the points on the dominant face can be approached with small enough resolution.

REFERENCES

[1] E. Arikan, “Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,” IEEE Trans. Inform. Theory, vol. 55, no. 7, pp. 3051–3073, July 2009.

[2] H. Mahdavifar and A. Vardy, “Achieving the secrecy capacity of wiretap channels using polar codes,” IEEE Trans. Inform. Theory, vol. 57, no. 10, pp. 6428–6443, October 2011.

[3] E. Arikan “Source polarization,” in Proc. IEEE Intern. Symp. on Information Theory, pp. 899–903, Austin, TX., June 2010.

[4] E. Abbe, “Extracting randomness and dependencies via a matrix polarization,” [arXiv:1102.1247], 2011.

[5] E. Sa¸so˘glu, E. Telatar, and Edmund Yeh, “Polar codes for the two-user binary-input multiple-access channels,” Proc. 2010 Information Theory Workshop (ITW), pp. 1–5, Dublin, Ireland, 2010.

[6] E. Abbe, E. Telatar, “Polar codes for the $m$-user multiple access channel,” IEEE Trans. Inform. Theory, vol. 58, no. 8, pp. 5437–5448, August 2012.

[7] R. Ahlswede, “Multi-way communication channels,” Proc. IEEE Intern. Symp. Information Theory, pp. 23–52, Budapest, Hungary, 1973.

[8] H. Liao, “Multiple Access Channels,” Ph.D. dissertation, Dept. Elec. Eng., Univ. Hawaii, Honolulu, 1972.

[9] H. Mahdavifar, M. El-Khamy, J. Lee and I. Kang, “Compound polar codes,” Proc. of Information Theory and Applications Workshop (ITA), pp. 1–6, San Diego, California, February 2013.

[10] E. Arikan and E. Telatar, “On the rate of channel polarization,” preprint of July 24, 2008, [arXiv.org/abs/0807.3806].

[11] S.B. Korada, Polar Codes for Channel and Source Coding, Ph.D. dissertation, EPFL, Lausanne, Switzerland, May 2009.