Polar Coding for Multi-level 3-Receiver Broadcast Channels

Karthik Nagarjuna Tunuguntla
CMRR, University of California, San Diego
Email: tkarthik@eng.ucsd.edu

Paul H. Siegel
CMRR, University of California, San Diego
Email: psiegel@ucsd.edu

Abstract—We consider achieving the rates in the capacity region of a multi-level 3-receiver broadcast channel, in which the second receiver is degraded with respect to the first receiver, with degraded message sets. We propose a two-level chaining strategy based on polar codes that achieves the capacity region of the considered setting without time-sharing. We also look at a slight variation of this problem, where the first receiver only requires to decode its own private message and the other two receivers require to decode another private message common to them. We observe that the capacity region does not enlarge and so the proposed polar coding strategy achieves the capacity region for this problem as well.

I. INTRODUCTION

A. Background

Arikan [1] constructed capacity-achieving polar codes for binary input symmetric channels. Since then, many coding strategies have been introduced for multi-user settings using the polarization method. Goela, Abbe and Gastpar [4] introduced polar codes for m-user deterministic broadcast channels. They also introduced polar coding for 2-user noisy broadcast channels. They implemented superposition and Marton schemes which involve some assumptions of degradation on the channel parameters to align the polar indices. Mondelli, Hassani, Sason, and Urbanke [6] proposed schemes to remove such constraints using a polar-based channelling construction. Chou and Bloch [2] proposed a polar coding scheme for a broadcast channel with confidential messages. Alos and Fonollosa [10] proposed a polar coding scheme for a broadcast channel with two legitimate receivers, that receive a confidential and private message, and one eavesdropper.

In this paper, we consider the problem of achieving the rates in the capacity region for a discrete memoryless (DM) 3-receiver broadcast channel with degraded message sets [7] [3]. The second receiver is degraded with respect to the first receiver. The problem is to transmit a public message intended for all three receivers and a private message intended for the first receiver. We discuss the motivation of the problem by providing an application in the following subsection.

B. Motivation with a client-server model

We consider a client server model, in which the server sends data to its three clients that are computer, phone-1 and phone-2. Computer supports both video and audio applications whereas the two phones only support audio application. Suppose that server has one audio and one video file to send its clients. All three clients are interested in receiving the audio file and computer is the only client interested in receiving both the audio and video files. This is shown in Fig.1. In some scenarios, computer may just want to receive the video file. The internet link and bluetooth link add noise to the signal received at the clients. These scenarios can be modelled as coding problems on a multi-level 3-receiver broadcast channel.

C. Coding problem of DM multi-level broadcast channel with degraded message sets

The 3-receiver multi-level broadcast channel, shown in Fig. 2, consists of a finite input alphabet \( \mathcal{X} \) and arbitrary output alphabets \( \mathcal{Y}_j \) for each output at the receiver-\( j \) for \( j \in \{1, 2, 3\} \). The conditional distribution of outputs at receiver-1 and receiver-3 given the input, i.e. \( p_{Y_1, Y_3|X}(y_1, y_3|x) \), along with the conditional distribution of output at receiver-2 given the output at receiver-1, i.e. \( p_{Y_2|Y_1}(y_2|y_1) \), are given for this broadcast channel setting, where \( X \) is the input, \( Y_j \) is output at receiver-\( j \) for \( j = 1, 2, 3 \), \( x \in \mathcal{X} \) and \( y_j \in \mathcal{Y}_j \) for each \( j \in \{1, 2, 3\} \). These two conditional distributions define the channel model as we have the degradation assumption on receiver-2.

Now we define the coding problem where the goal is to transmit a public message for all the receivers and a private message intended only for receiver-1.

A \( (2^{N_{R_0}}, 2^{N_{R_1}} \times N) \) code consists of

- a message set for public message: \( \{1, 2, \ldots, 2^{N_{R_0}}\} \)
- a message set for private message of receiver-1: \( \{1, 2, \ldots, 2^{N_{R_1}}\} \)
We use a chaining construction at two levels, one of which is the block length, \( R_0 \) is the rate of the public message and \( R_1 \) is the rate of the private message. Let \( M_0 \) be the public message which is chosen uniformly from the set \( \{1, 2, \ldots, 2^{N R_0}\} \) and \( M_1 \) be the private message of receiver-1 which is chosen uniformly from the set \( \{1, 2, \ldots, 2^{N R_1}\} \). Let \( Y_1^{1:N} \) be the output vector at receiver-\( j \) where \( j \in \{1, 2, 3\} \).

Let \( P_{e}(N) = P( (h_1(Y_1^{1:N}) \neq (M_0, M_1)) \cup (h_2(Y_2^{1:N}) \neq M_0) \cup (h_3(Y_3^{1:N}) \neq M_0) ) \) be the probability of error. If there is a sequence of \( (2^{N R_0}, 2^{N R_1}, N) \) codes, for which \( P_{e}(N) \) goes to zero, then the rate \( (R_0, R_1) \) is achieved. The closure of all such achievable rate pairs is the capacity region.

**D. Contribution**

In this paper, we use a polar coding strategy to achieve the rates in the capacity region for the multi-level 3-receiver broadcast with degraded message sets without time-sharing. This represents the first time in the literature that polar coding for 3-receiver broadcast channels without eavesdropper is considered.

Three-layered polarization results are established using auxiliary random variables that characterize the capacity region. We do a suitable rate splitting of the private message of receiver-1 for the implementation of our polar coding strategy. We use a chaining construction at two levels, one of which is within first and second layers whereas the second level of chaining is done within the second layer. The two-level chaining construction that we provide essentially translates into polar coding strategy the ideas of three layered superposition coding and more importantly, indirect-coding [7] with rate splitting of the private message.

The two-level chaining construction is new in the context of reliable decoding at three receivers. In particular, the first level of chaining is done to recover the public message by all the receivers. The second level of chaining helps to recover the split of private message reliably at receiver-1 while translating indirect coding of public message for receiver-3. In contrast, note that Marton’s coding [6] uses a two-level chaining construction, where the first level of chaining is to align good bit-channels of the two receivers and the second level of chaining is to maintain the joint distribution of auxiliary random variables involved.

We also consider a slight variation to the problem of degraded message sets. Suppose that receiver-1 requires to decode only \( M_1 \). Then \( M_1 \) becomes private message to receiver-1 and \( M_0 \) is common private message to receiver-2 and receiver-3. We show that the capacity region does not enlarge by relaxing the constraint at receiver-1. So the same polar coding strategy achieves the capacity region of the modified problem. This is an interesting observation, as we know that for any 2-receiver broadcast channel, superposition coding is not optimal in general, unless it is a problem with degraded message sets.

**II. Preliminaries**

We denote the set \( \{1, 2, \ldots, n\} \) as \( [n] \) where \( n \in \mathbb{Z}^+ \). Let \( G_N \) be the conventional polar transform [1], represented by a binary matrix of dimension \( N \times N \) where \( N = 2^n, n \in \mathbb{Z}^+ \).

Let \( X \) be a binary random variable. Let the random variable pair \( (X, Y) \) be distributed as \( P_{X,Y}(x, y) \), then the Bhattacharya parameter is defined as

\[
Z(X|Y) = 2 \sum_y P_Y(y) \sqrt{P_{X|Y}(1|y)P_{X|Y}(0|y)}.
\]

The **capacity region** for this multi-level 3-receiver broadcast problem [3] [7] is as follows:

\[
R_0 < \min(I(W;Y_2), I(V;Y_3)) \tag{1}
\]

\[
R_1 < I(X;Y_1|W) \tag{2}
\]

\[
R_0 + R_1 < I(V;Y_2) + I(X;Y_1|V) \tag{3}
\]

for some joint distribution \( p(w, v)p(x|v) \) with \( |W| \leq |X| + 4 \) and \( |V| \leq (|X| + 1)(|X| + 4) \). Here \( W \) and \( V \) are random variables over the alphabets \( W \) and \( V \), respectively, and \( Y_j \) is the output at receiver-\( j \) when \( X \) is input for \( j = 1, 2, 3 \).

Let \( (W_i, V_i, X_i)_{i=1}^N \) be the binary triplet random variable sequence that is i.i.d. (identical and independently distributed) according to the joint distribution \( p(w, v)p(x|v) \). Let \( (W, V, X) \) also be the binary random triplet distributed according to \( p(w, v)p(x|v) \). Let \( Y_j^{1:N} \) be the received vector at receiver-\( j \) when the random variable sequence \( X_j^{1:N} \) is transmitted over the 3-receiver discrete memoryless broadcast channel and let \( Y_j \) be the output at receiver-\( j \) when \( X \) is input for \( j = 1, 2, 3 \).

Now we establish three-level polarization results that are going to be used in the code construction.

Let \( \beta < 0.5 \). Let \( (U_w)^{1:N} = W^{1:N}G_N \), and define the following bit-channel subsets as follows, where \( j = 1, 2, 3 \).

\[
H_W = \{ i \in [N] : Z((U_w)_i)|(|(U_w)_i|^{1:(i-1)}) \geq 1 - \delta_n \},
\]

\[
L_W = \{ i \in [N] : Z((U_w)_i)|(|(U_w)_i|^{1:(i-1)}) \leq \delta_n \},
\]

\[
L_W|Y_j = \{ i \in [N] : Z((U_w)_i)|(|(U_w)_i|^{1:(i-1)}Y_j^{1:N}) \leq \delta_n \},
\]

where \( \delta_n = 2^{-N^\beta} \). Note that \( L_W|Y_j \subseteq L_W|Y_1 \) from Lemma 7 in [4] due to the degradation assumption on receiver-2. Then,

\[
\lim_{N \to \infty} \frac{|H_W|}{N} = H(W), \quad \lim_{N \to \infty} \frac{|L_W|}{N} = 1 - H(W),
\]

\[
\lim_{N \to \infty} \frac{|H_W|}{N} = H(W|Y_j), \quad \lim_{N \to \infty} \frac{|L_W|}{N} = 1 - H(W|Y_j).
\]
Let \((U_v)^{1:N} = V^{1:N} G_N\). Similarly, we now define bit-channel subsets \( \mathcal{H}_{V|W} \) and \( \mathcal{L}_{V|W} \) based on the Bhattacharyya parameter \( Z((U_v)_i | (U_v)_j)^{(1-i)W^{1:N}} \) as we did above. We also define \( \mathcal{H}_{V|W_1} \) and \( \mathcal{L}_{V|W_1} \) based on the value of Bhattacharyya parameter \( Z((U_v)_i | (U_v)_j)^{(1-i)W_1^{1:N} Y_j^{1:N}} \) for \( j = 1, 3 \). Then,

\[
\lim_{N \to \infty} \frac{|\mathcal{H}_{V|W}|}{N} = H(W|V), \quad \lim_{N \to \infty} \frac{|\mathcal{L}_{V|W}|}{N} = 1 - H(W|V),
\]
\[
\lim_{N \to \infty} \frac{|\mathcal{H}_{V|W_1}|}{N} = H(W|V Y_1), \quad \lim_{N \to \infty} \frac{|\mathcal{L}_{V|W_1}|}{N} = 1 - H(W|V Y_1).
\]

Let \((U_v)^{1:N} = X^{1:N} G_N\). We define the bit-channel subsets \( \mathcal{H}_{X|V} \), \( \mathcal{L}_{X|V} \) and also \( \mathcal{H}_{X|V_1} \), \( \mathcal{L}_{X|V_1} \) based on the values of Bhattacharyya parameters \( Z((U_v)_i | (U_v)_j)^{(1-i)V^{1:N}} \) and \( Z((U_v)_i | (U_v)_j)^{(1-i)V_1^{1:N} Y_j^{1:N}} \), respectively, as we did above. Then,

\[
\lim_{N \to \infty} \frac{|\mathcal{H}_{X|V}|}{N} = H(X|V), \quad \lim_{N \to \infty} \frac{|\mathcal{L}_{X|V}|}{N} = 1 - H(X|V),
\]
\[
\lim_{N \to \infty} \frac{1}{N} |\mathcal{H}_{X|V_1}| = H(X|V Y_1), \quad \lim_{N \to \infty} \frac{1}{N} |\mathcal{L}_{X|V_1}| = 1 - H(X|V Y_1).
\]

Define \( I^{w}_j = \mathcal{L}_{V|Y_1} \cap \mathcal{H}_W \), \( I^{v}_j = \mathcal{L}_{V|Y_1} \cap \mathcal{H}_W \) and \( I^{f}_j = \mathcal{L}_{X|V_1} \cap \mathcal{H}_X \). Note that \( \lim_{N \to \infty} \frac{|I^{w}_j|}{N} = I(W;Y_j) \), \( \lim_{N \to \infty} \frac{|I^{v}_j|}{N} = I(V;Y_j) \) and \( \lim_{N \to \infty} \frac{|I^{f}_j|}{N} = I(X;Y_j|V) \). We refer to \( I^{w}_j \), \( I^{v}_j \) and \( I^{f}_j \) as information bit-channels of receiver-\( j \) in \((U_w)^{1:N}\), \((U_v)^{1:N}\) and \((U_x)^{1:N}\) respectively for \( j = 1, 2, 3 \).

Define \( R^{w} = (\mathcal{H}_W \cup \mathcal{L}_W)^c \), \( R^{v} = (\mathcal{H}_{V|W} \cup \mathcal{L}_{V|W})^c \) and \( R^{f} = (\mathcal{H}_{X|V} \cup \mathcal{L}_{X|V})^c \). We refer to \( R^{w} \), \( R^{v} \) and \( R^{f} \) as not-completely polarized bit-channels in \((U_w)^{1:N}\), \((U_v)^{1:N}\) and \((U_x)^{1:N}\) respectively.

We denote the subvector of \( U^{1:N} \) corresponding to the bit-channel set \( A \subseteq [N] \) by \( U^A \).

### III. Polar Coding for the DM Multi-Level 3-receiver Broadcast Channel

In this section, we are going to discuss the polar coding scheme that achieves the rate pairs that satisfy equations (1), (2) and (3) for a joint distributions on random variables over the alphabets of the required size mentioned in the definition of the capacity region. We first consider the case when \(|X| = |Y| = |V| = 2\) to describe the polar coding scheme and then discuss the extension to higher alphabets. We now discuss the rate-splitting of the private message of receiver.

#### A. Rate splitting of the private message in typical set coding

The random coding method using typical sets [7] uses rate splitting of the private message, \( R_1 \) into \( R_{11} \) and \( R_{12} \). If \( R_0, R_{11}, R_{12} \) satisfy the following:

\[
R_0 < I(W;Y_2),
\]
\[
R_{12} < I(X;Y_1|V),
\]
\[
R_{11} + R_{12} < I(X;Y_1|W)
\]
\[
R_0 + R_{11} + R_{12} < I(X;Y_1),
\]
\[
R_0 + R_{11} < I(V;Y_3),
\]

then reliable recovery of the intended messages at each of the receivers is ensured [7]. After eliminating variables \( R_{11} \) and \( R_{12} \) by Fourier-Motzkin [3] procedure by substituting \( R_1 = R_{11} + R_{12} \), we get the region defined by equations (1), (2) and (3) that defines the capacity region. The intuition behind the rate splitting is that if we want to achieve a private message rate satisfying \( R_1 > I(X;Y_1|V) \) and \( R_1 < I(X;Y_1|W) \), then we rate split \( R_1 \) into \( R_{11} \) and \( R_{12} \) such that \( R_{11} < I(X;Y_1|V) \) and \( R_{12} < I(X;Y_1|W) \). As we recover public message indirectly [7], the sum of public message rate \( R_0 \) and \( R_{11} \) should be less than \( I(V;Y_3) \). So, if we make \( R_{11} \) small while rate splitting, then it can be noticed that the public message rate can be improved, provided the reliability constraint at receiver-2, \( R_0 < I(W;Y_2) \) [7], is loose.

#### B. Rate splitting of the private message for polar coding

Notice that a point in the region satisfied by equations (4), (5), (6), (7) and (8) does not always satisfy the constraint \( R_{11} < I(Y_1|W) \). We impose the new additional constraint \( R_{11} < I(Y_1|W) \) for the rate split in the implementation of our polar coding strategy through the following lemma.

**Lemma 1.** For any rate pair \((R_0, R_1)\) that satisfies equations (1), (2) and (3) and for a particular joint distribution \( p(w,v|x) \) on \((W,V,X)\), there exist rates \( R_{11} \) and \( R_{12} \) such that \( R_1 = R_{11} + R_{12} \) (rate split of \( R_1 \) and the following three identities hold.

\[
R_{11} < I(V;Y_1|W)
\]
\[
R_{12} < I(X;Y_1|V)
\]
\[
R_0 + R_{11} < I(V;Y_3)
\]

**Proof:** Refer to full version [9] for the proof.

In our polar coding strategy, the private message bits for receiver-1 are given in bits \( I^{w}_1 \) and \( I^{f}_1 \) in the chaining construction. The rate split in Lemma 1 allows us to associate the private message bits encoded in \( I^{w}_1 \) and \( I^{f}_1 \) to split rates of the private message \( R_{11} \) and \( R_{12} \), respectively. We also involve the bits corresponding to \( R_{11} \), which are private message bits encoded in \( I^{w}_1 \), in the chaining procedure to translate the indirect coding method at receiver-3 into polar coding. Now we provide our code construction in the following subsection.

#### C. Code construction

We consider \( k \) polar blocks of size \( N \) large enough so that the polarization happens. We propose a chaining construction with these \( k \) polar blocks for the rate pair \((R_0, R_1)\) by using
the rate split given by the Lemma 1. While encoding each polar block, we first construct \((U_w)_{1:N}^{1}\) and compute \(W_{1:N} = (U_w)_{1:N}^{1:N} G_N\). We next construct \(V_{1:N} = (U_w)_{1:N}^{1:N} G_N\) given \(W_{1:N}\) and apply polar transform to obtain \(V_{1:N}\). Lastly, we construct \((U_v)_{1:N}\) given \(V_{1:N}\) and apply polar transform to obtain \(X_{1:N}\) (codeword).

We first give the construction for the case where \(R_0 \geq I(W; Y_3)\). This is the case where we translate the indirect coding into polar coding strategy. We assume \(NR_1 > |I_3^\cap I_3^1|\) to demonstrate the code construction. The construction we give under this assumption gives the general idea of the chaining construction which can easily be extended to the case where this assumption does not hold.

Note that public message bits have to be recovered at all the receivers. If we give \(NR_0\) public message bits in \(I_w^0\), receiver-2 and receiver-1 (due to degradation condition) can recover these bits. But receiver-3 may not be able to decode in that case. On the other hand we can recover these bits at receiver-3, if we place these bits into \(I_w^0\) and remaining \(NR_0 - |I_w^0|\) bits in \(I_w^1\), as \(NR_0 > |I_w^1|\). In this case, receiver-1 and receiver-2 may not be able to decode. We do a chaining to resolve the alignment of the bit-channel set in \(I_w^1\) with bit-channels sets in \(I_w^0\) and \(I_w^1\) to allocate the public message bits for reliable recovery at all the receivers.

Since we are assigning a portion of public message bits in \((U_v)_{1:N}\) vectors for receiver-3, we need to recover \((U_v)_{1:N}\) vectors at receiver-3. But we also use \((U_v)_{1:N}\) vectors for encoding private message bits corresponding to the rate \(R_1\). If we give these private message bits in \(I_v^0\), receiver-3 cannot recover these bits, which blocks receiver-3 from recovering \((U_v)_{1:N}\) vectors for decoding the portion of intended public message bits. Here is where we need to do a second level of chaining for aligning bit-channel set in \(I_v^1\) with bit-channels set in \(I_v^0\) where we provide private message bits corresponding to \(R_1\). This summarizes the main idea behind the construction that translates indirect coding at receiver-3.

Fig. 3 shows the two-level chaining done in the first two layers, \((U_w)_{1:N}\) and \((U_v)_{1:N}\) vectors, when \(k = 3\), allocating public and private message bits. The links between vectors in Fig. 3 indicate the copying of bits between bit-channel sets of successive blocks. Now we provide detailed steps of the encoding and decoding methods in the two-level chaining construction for this case, \(R_0 \geq I(W; Y_3)\).

Encoding:

- Encoding \((k - 1)NR_1 + |I_3 \cap I_3^1|\) bits of the public message, first level of chaining:
  - We first place \(|I_3 \cap I_3^1|\) bits in \((U_w)_{1:N}^{I_3 \cap I_3^1}\) for all the blocks \(t = 1 : k\). Note that \(NR_0\) is the sum of \(|I_3^0 \cap I_3^2| + |I_3^1 \cap I_3^2| + (NR_0 - |I_3^0|)|\).
  - We place \(|I_3 \cap I_3^1|\) bits in \((U_w)_{1:N}^{I_3 \cap I_3^1}\) and \(NR_0 - |I_3^1|\) in \((U_v)_{1:N}^{I_3 \cap I_3^1}\) for the blocks \(t = 1 : k\) where \(I_3\) is partitioned as \(I_3^0 \cup I_3^1\) and \(|I_3^1| = NR_0 - |I_3^0|\). Note that \(NR_0 + NR_1 < |I_3^0| + |I_3^1|\) due to Lemma 1. As \(NR_1 > |I_3^1 \cap I_3^0|\), it can be deduced that \(I_3^1 \subset I_3^0 \cap I_3^1\).

- We copy bits in \((U_w)_{1:N}^{I_3 \cap I_3^1}\) and \((U_v)_{1:N}^{I_3 \cap I_3^1}\) of block \(t\) to \((U_w)_{1:N}^{I_3 \cap I_3^1}\) of block \(t + 1\) for \(t = 1 : k - 1\) where \(|I_3 \cap I_3^1|\) is partitioned as \(I_3^2 \cap I_3^1\). Note that \(NR_1 > |I_3^1 \cap I_3^0|\) due to Lemma 1.

- We copy the bits in \((U_v)_{1:N}^{I_3 \cap I_3^1}\) of block \(t\) to \((U_v)_{1:N}^{I_3 \cap I_3^1}\) for all these blocks \(t = 1 : k\).

- Encoding \((k - 1)NR_1 + |I_3 \cap I_3^1|\) bits of the private message for receiver-1, second level of chaining:
  - We first place \(|I_3 \cap I_3^1|\) private message bits in \((U_v)_{1:N}^{I_3 \cap I_3^1}\) for all the blocks \(t = 1 : k\).
  - We place \(NR_1 - |I_3 \cap I_3^1|\) bits in \((U_v)_{1:N}^{I_3 \cap I_3^1}\) for the blocks \(t = 1 : k - 1\) where \(|I_3 \cap I_3^1|\) is partitioned as \((I_3^2 \cap I_3^1)\) and \(|I_3^2| = NR_1 - |I_3^1|\).

- Encoding \(kNR_2\) bits of the private message for receiver-1:
  - We place \(NR_2\) bits in \((U_v)_{1:N}^{I_3 \cap I_3^1}\) for all these blocks \(t = 1 : k\).

- We place randomly chosen frozen bits with i.i.d. uniform distribution in \((U_w)_{1:N}^{I_3 \cap I_3^1}\) for block \(t\) and in \((U_v)_{1:N}^{I_3 \cap I_3^1}\) for block \(t = k\). We also place randomly chosen bits with i.i.d. uniform distribution in the remaining positions of \((U_w)_{1:N}^{I_3 \cap I_3^1}\) and \((U_v)_{1:N}^{I_3 \cap I_3^1}\), which are not filled by private or public message bits, in all the blocks. We share these remaining bits that are in \((U_v)_{1:N}^{I_3 \cap I_3^1}\) for each block with the receiver-j for \(j = 1, 2, 3\), in all the blocks.

- We have constructed \((U_w)_{1:N}^{I_3 \cap I_3^1}\) and \((U_v)_{1:N}^{I_3 \cap I_3^1}\) for all the blocks. Now we encode other positions in \((U_w)_{1:N}^{I_3 \cap I_3^1}\) using an appropriate argmax rule for all blocks [5, 8]. We encode positions in \((U_w)_{1:N}^{I_3 \cap I_3^1}\) using common randomness or boolean functions as we do for single asymmetric channel case [5, 8]. Refer to full version [9] for details.

- Now we transmit \(X_{1:N} = (U_x)_{1:N}^{I_3 \cap I_3^1}\) for all \(k\) blocks. The rate pair of this scheme approaches \((R_0, R_1)\) as \(k\) ap-
proaches $\infty$.

**Decoding at receiver-3 with $Y^{1:N}_i$ of each block:**

- We decode the blocks $t = 1 : k$, in forward direction.
- For blocks $t = 1 : k - 1$, $NR_0$ public message bits $(U_w)^{I_{NR_0}}$, $(U_v)^{I_{NR_0}}$ and $NR_{12}$ bits of $(U_w)^{I_{NR_0}}$ of the private message of receiver-1 will be recovered at each block $t$ during the successive cancellation decoding. Note that some of these recovered bits give bits in $(U_w)^{I_{NR_0}}$ and $(U_v)^{I_{NR_0}}$ of its successive block $(t + 1)$ required in successive cancellation decoding.
- For block $k$, the public message bits $(U_w)^{I_{NR_0} \cap I_N^1}$ will be recovered during the successive cancellation decoding.

**Decoding at receiver-1 and receiver-2:**

- We decode the blocks $t = k : 1$, using $Y^{1:N}_i$, in reverse direction at receiver-$j$, $j = 1, 2$. We skip the steps in parentheses-() for decoding at receiver-2.
- For blocks $t = k : 1$, $NR_0$ public message bits $(U_w)^{I_{NR_0} \cap I_N^1}$ and $(U_v)^{I_{NR_0}}$, (and $NR_{12}$ bits $(U_v)^{I_{NR_0} \cap (I_{NR_0} \cup I_{NR_0}^2)}$, $NR_{12}$ bits in $(U_v)^{I_{NR_0}}$ of private message) are recovered during the successive cancellation decoding. Note that some of these recovered bits give bits in $(U_w)^{I_{NR_0} \cap I_N^1}$(as well as $(U_v)^{I_{NR_0}}$ and $(U_v)^{I_{NR_0} \cup I_{NR_0}^2}$) of the successive block $t - 1$ required in successive cancellation decoding.
- For block 1, the public message bits $(U_w)^{I_{NR_0} \cap I_N^1}$ (the bits $(U_v)^{I_{NR_0}}$ and bits in $(U_v)^{I_{NR_0} \cup I_{NR_0}^2}$ of private message of receiver-1) will be recovered.

During the successive cancellation decoding, we recover needed bits in $(U_w)^{I_{NR_0} \cap I_N^1}$, $(U_v)^{I_{NR_0}}$, $(U_v)^{I_{NR_0} \cup I_{NR_0}^2}$ by an appropriate arg-max rule, at each receiver. For the remaining bits in $(U_w)^{I_{NR_0}}$, $(U_v)^{I_{NR_0}}$ and $(U_v)^{I_{NR_0} \cup I_{NR_0}^2}$ we use shared boolean functions/common randomness. Refer to [9] for details.

Suppose if $NR_{12} \leq |I_N^1 \cap I_N^2|$, we just need to perform chaining at first level. For the other case, $R_0 < I(W; Y_0)$, we just require to do chaining at first level so that all receivers recover the public message. We omit the details due to space constraint. Please refer to full version [9] for details. Theorem 1 provides the analysis of the probability of decoding error in the chaining construction.

**Theorem 1.** Let $P_e(\mathcal{C})$ be the probability of error for a given code in the random ensemble, denoted by random variable $\mathcal{C}$. The average probability of error for the random code construction $E_C[P_e(\mathcal{C})]$ is $O(k2^{-N\beta'})$ for $\beta' < \beta < 0.5$.

**Proof:** Refer to full version [9] for the proof. □

Both encoding and decoding complexities will become $O(N \log N)$ per block [5].

We can adapt our multi-level polar code construction when $|X|$, $|V|$, $|X'|$ are of arbitrary size [11] [12]. Let $|X| = \Pi_{i=1}^{P_j}$ $|V| = \Pi_{j=1}^{P_j}$, $|W| = \Pi_{i=1}^{P_j}$, $\{r_j\}$, $\{q_j\}$ and $\{p_j\}$ are prime factors of $W$, $V$ and $X'$, respectively. Then random variables $W, V$ and $X$ can be represented by random vectors $(W_1, \ldots, W_k)$, $(V_1, \ldots, V_l)$ and $(X_1, \ldots, X_m)$ where $W_j$, $V_j$ and $X_j$ are supported over the set $\{0, 1, \ldots, r_j - 1\}$, $\{0, 1, \ldots, q_j - 1\}$ and $\{0, 1, \ldots, p_j - 1\}$, respectively. By chain rule of entropy, we get $H(W; V, X) = \sum_{j=1}^{N} H(W_j | W_1^{j-1}) + \sum_{j=1}^{N} H(V_j | W_1^{j-1}) + \sum_{j=1}^{N} H(X_j | W_1^{j-1}).$ We can use the polarization for prime alphabets for each term in the above identity and construct an appropriate polar coding scheme.

**D. Extension:** receiver-1 requires only $M_1$

For a $(2^{N R_0}, 2^{N R_1}, N)$ code of a setting with degraded messages sets, the converse proof of the capacity region just uses the fact that $H(M_1|Y_1^{1:N}), H(M_0|Y_2^{2:N})$ and $H(M_0|Y_3^{3:N})$ are $o(N)$ [7]. We do not have to use the stronger fact that $H(M_1, M_0|Y_1^{1:N})$ is $o(N)$ to complete the converse proof. This means that the same proof becomes the converse proof of the capacity region for the problem when receiver-1 requires to recover only $M_1$. Hence, the capacity region does not enlarge and remains the same. So the same polar coding method can be used to achieve all rate pairs inside the capacity region.

**IV. Conclusion**

We considered the problem of achieving the rates in the capacity region of a discrete memoryless multi-level 3-receiver broadcast channel with degraded message sets through polar coding. We give a new two-level chaining construction to achieve all the points in the capacity region without time-sharing. We also gave a detailed analysis of the probability of decoding error for the constructed coding scheme [9]. We showed the capacity of the broadcast channel does not enlarge, even when receiver-1 requires to recover only its private message.

**V. Acknowledgement**

This work was supported in part by National Science Foundation Grants CCF-1415109 and CCF-1619053.

**References**

[1] E. Ankan, “Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,” IEEE Trans. Inf. Theory, vol. 55, no. 7, pp. 3051–3075, Jul. 2009.

[2] R. A. Chou and M. R. Bloch, “Polar coding for the broadcast channel with confidential messages,” Proc. IEEE Information Theory Workshop (ITW), Jerusalem, 2015, pp. 1–5.

[3] A. El Gamal and Y. H. Kim, Network Information Theory, Cambridge, UK: Cambridge University Press, 2011.

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

[5] J. Honda and H. Yamamoto, “Polar coding without alphabet extension for asymmetric models,” IEEE Trans. Inf. Theory, vol. 59, no. 12, pp. 7829–7838, Dec. 2013.

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

[7] C. Nair and A. El Gamal, “The capacity region of a class of three-receiver broadcast channels with degraded message sets,” IEEE Trans. Inf. Theory, vol. 55, no. 10, pp. 4479–4493, Oct. 2009.

[8] K. Nagarjuna and P. H. Siegel, “Universal polar coding for asymmetric channels,” Proc. IEEE Information Theory Workshop (ITW), Guangzhou, Nov. 2018, pp. 1–5.

[9] K. Nagarjuna Tunuguntla and P. H. Siegel, Polar coding for multilevel 3-receiver broadcast channels, arXiv:2010.07228, 2020.

[10] J. d. Olmo Alsó and J. R. Fonollosa, “Polar coding for common message only wiretap broadcast channel,” Proc. IEEE Int. Symp. Inf. Theory, Paris, France, 2019, pp. 1762–1766.

[11] E. Şasoglu, E. Telatar and E. Ankan, “Polarization for arbitrary discrete memoryless channels,” Proc. IEEE Information Theory Workshop (ITW), Taormina, 2009, pp. 144–148.

[12] E. Şasoglu, “Polarization and polar codes,” Found. Trends Commun. Inf. Theory, vol. 8, no. 4, pp. 259–381, Oct. 2012.