Polar Codes for Broadcast Channels with Receiver Message Side Information and Noncausal State Available at the Encoder

Jin Sima and Wei Chen
Tsinghua National Laboratory for Information Science and Technology (TNList)
Department of Electronic Engineering, Tsinghua University, Beijing, CHINA
Email: smj13@mails.tsinghua.edu.cn, wchen@tsinghua.edu.cn

***Abstract*—In this paper polar codes are proposed for two receiver broadcast channels with receiver message side information (BCSI) and noncausal state available at the encoder, referred to as BCSI with noncausal state for short, where the two receivers know a priori the private messages intended for each other. We establish an achievable rate region for BCSI with noncausal state and show that it is strictly larger than the straightforward extension of the Gelfand-Pinsker result. To achieve the established rate region, we present polar codes for the general Gelfand-Pinsker problem, which adopts chaining construction and utilizes causal information to pre-transmit the frozen bits. It is also shown that causal information is necessary to pre-transmit the frozen bits. Based on the result of Gelfand-Pinsker problem, we then propose polar codes for BCSI with noncausal state. The difficulty is that there are multiple chains sharing common information bit indices. To avoid value assignment conflicts, a nontrivial polarization alignment scheme is presented. It is shown that the proposed rate region is tight for degraded BCSI with noncausal state.**

I. Introduction

In Arikan's pioneering work [1], he introduced polar codes, which constitute a new and promising class of practical capacity achieving codes. In the past years, polar codes have been richly investigated and generalized to various channel/source coding settings. In the work [2], Goela, Abbe, and Gastpar introduced polar codes for realizing superposition strategy and Marton's strategy, which comprise the main coding strategies for broadcast channels. To guarantee the alignment of polarization indices, the coding scheme requires some degradedness conditions with respect to the auxiliary random variables and channel outputs. Such degradedness requirements can be removed by adopting the polarization alignment techniques proposed by Mondelli, Hassani, Sason, and Urbanke [3], where multi-block transmission and block chaining are considered. The work in [4] proposed polar codes for two receiver broadcast channels with receiver message side information (BCSI), where each receiver knows the message intended for the other.

In this paper, we consider polar codes for BCSI with common message and with noncausal state available at the encoder, which is a generalization of Gelfand-Pinsker channel and BCSI. Such channel arises in multi-user cellular communication systems with two-way communication tasks or pairwise message exchange requests. For each pair of users that exchange messages, broadcasting to them in the downlink transmission can be regarded as BCSI with noncausal state, by considering the interference from signals of other users as noncausal state known at the base station. BCSI with noncausal state was studied in a previous work [5], where a coding scheme combining Gelfand-Pinsker binning and network coding was proposed. Its related scenarios, broadcast channels with noncausal state, has received much attention and has been investigated in, e.g., [6]–[8].

Polar codes for Gelfand-Pinsker problems have been presented. Polar codes for binary channels with additive noise and interference was proposed in [9]. Noisy write once memory was considered in [10], where polar codes with polynomial computational and storage complexity were proposed. For general Gelfand-Pinsker settings, the work in [10], [11] proposed polar coding schemes based on the the block chaining method in [3]. The problem of applying the chaining construction to the Gelfand-Pinsker settings is to communicate the state information to the receiver in the first block. This problem was not addressed in [11]. The work in [10] proposed a solution to this problem by using an extra phase to transmit the frozen bits in the first block, where the channel state information is not used by the encoder. As we will show in the next, this solution may not work in some cases. In particular, the state information is needed by the encoder to transmit the frozen bits in the first block.

In this paper, we establish an achievable rate region for BCSI with common message and with noncausal state. Polar coding schemes are presented to achieve the established region. To this end, we first propose polar codes for the general Gelfand-Pinsker problem, based on the chaining construction in [3]. A pre-communication phase that utilizes causal state information is performed to transmit the frozen bits in the first block. It is also shown that the state information is necessary to transmit these frozen bits. We then use the result in Gelfand-Pinsker problem to construct polar codes for BCSI with noncausal state. To overcome the problem that the two chains may overlap and cause value assignment conflicts, the two chains are generated in opposite directions so that the overlapped sets only needs to carry the XOR of the bits contained in the two chains. We present an example to show that it is strictly larger than the achievable rate region in [5]. It

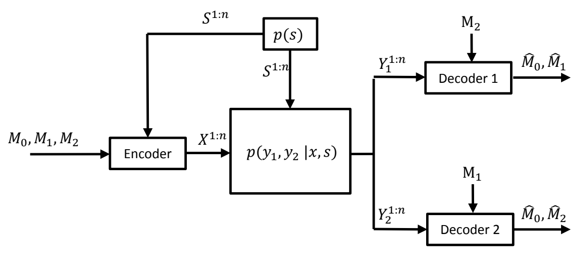

Fig. 1. BCSI with noncausal state

is shown that the established rate region is tight for degraded BCSI with common message and with noncausal state. The extension of the scheme in this paper to higher input alphabet size can be made following the techniques in [12].

II. System Model

Broadcast channels with receiver message side information (BCSI) and with noncausal state available at the encoder (as shown in Fig. 1), which is referred to as BCSI with noncausal state for short, is a two-receiver discrete memoryless broadcast channels (DMBC) with state

$$(\mathcal{X} \times \mathcal{S}, P_{Y_1,Y_2|X,S}(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2), \tag{1}$$

with input alphabet $\mathcal{X}$, state alphabet $\mathcal{S}$, output alphabets $\mathcal{Y}_1, \mathcal{Y}_2$ and conditional distribution $P_{Y_1,Y_2|X,S}$ $(y_1, y_2|x, s)$. The channel state sequence $S^{1:n}$ is a sequence of n i.i.d. random variables with pmf $P_S(s)$ and is noncausally available at the encoder. The sender wishes to send a message tuple $(M_0, M_1, M_2) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ to receivers 1 and 2, where receivers 1 and 2 know side information of messages M_2 and M_1 respectively. M_0 is a common message intended for both receivers.

A $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ code consists of a message set $[1 : 2^{nR_0}] \times [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$, an encoder $\zeta : [1 : 2^{nR_0}] \times [1 : 2^{nR_1}] \times [1 : 2^{nR_2}] \times \mathcal{S}^n \to \mathcal{X}^n$ that maps $(M_0, M_1, M_2, S^{1:n})$ to a codeword $X^{1:n}$, and two decoders $\xi_1 : \mathcal{Y}_1^n \times [1 : 2^{nR_2}] \to [1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$ and $\xi_2 : \mathcal{Y}_2^n \times [1 : 2^{nR_1}] \to [1 : 2^{nR_0}] \times [1 : 2^{nR_2}]$ that map $(Y_1^{1:n}, M_2)$ and $(Y_2^{1:n}, M_1)$ to $(\hat{M}_0, \hat{M}_1)$ and $(\hat{M}_0, \hat{M}_2)$ respectively. Here $Y_i^{1:n}$ is the received sequence of receiver i. A rate tuple (R_0, R_1, R_2) is achievable if there exists a $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ code such that the average error probability of the code

$$\begin{aligned} P_e^{(n)} = P\{&\xi_1(Y_1^{1:n}, M_2) \neq \{M_0, M_1\} \\ &\cup \xi_2(Y_2^{1:n}, M_1) \neq \{M_0, M_2\}\} \end{aligned} \tag{2}$$

tends to zero as n goes to infinity. The capacity region $\mathcal{C}$ is the closure of the set of all achievable rate tuples (R_0, R_1, R_2).

For each random variable U, we shall use the notation $U^{1:n}$ to denote the sequence of n i.i.d. random variables drawn from pmf $P_U(u)$. The i-th element of $U^{1:n}$ is denoted as U^i.

III. BCSI with Common Message and with Noncausal State

In this section a polar coding scheme is proposed for BCSI with common message and with noncausal state (1). It is also shown that the proposed polar coding scheme achieves the capacity region for degraded BCSI with common message and with noncausal state.

The Gelfand-Pinsker capacity for channel with random state noncausally known at the encoder is given by

$$C = \max_{p_{U|S}(u|s), x(u,s)} I(U;Y) - I(U;S). \tag{3}$$

A straightforward extension of the Gelfand-Pinsker capacity for BCSI with noncausal state is given by [5]

$$\begin{aligned} R_0 + R_1 &\leq I(U;Y_1) - I(U;S), \\ R_0 + R_2 &\leq I(U;Y_2) - I(U;S). \end{aligned} \tag{4}$$

We now establish an achievable rate region, which is strictly larger than that characterized by (4), and present polar codes for achieving the region.

Theorem 1. For BCSI with common message and with noncausal state (1), where the input has binary alphabet, there exists a polar code sequence with block length n that achieves (R_0, R_1, R_2) if

$$\begin{aligned} R_1 + R_0 &\leq I(V_1, V_2; Y_1) - I(V_1, V_2; S), \\ R_2 + R_0 &\leq I(V_1; Y_2) - I(V_1; S) \end{aligned} \tag{5}$$

for binary variables V_1, V_2 that satisfy (1) $(V_1, V_2) \to (X, S) \to Y_1$ form a Markov chain, (2) $(V_1, V_2) \to (X, S) \to Y_2$ form a Markov chain, (3) $I(V_2; Y_1|V_1) > I(V_2; S|V_1)$, (4) $I(V_1; Y_1) > I(V_1; S)$, (5) $I(V_1; Y_2) > I(V_1; S)$, and for some function $f(v_1, v_2, s) :\ \{0, 1\}^2 \times \mathcal{S} \to \mathcal{X}$. As n increases, the encoding and decoding complexity is $O(n \log n)$ and the error probability is $O(2^{-n^{\beta}})$ for $0 < \beta < \frac{1}{2}$.

Remark 1. The rate region (5) reduces to (4) when the random variable V_2 remains constant.

Remark 2. Symmetrically, the rate region is achievable if the role of receiver 1 and receiver 2 is reversed.

To give an example where the region (5) is strictly larger than (4), consider a broadcast channels with state $(\mathcal{X} \times \mathcal{S}, P_{Y_1,Y_2|X,S}(y_1, y_2|x, s), \mathcal{Y}_1 \times \mathcal{Y}_2)$, with input alphabet $\mathcal{X} = \{1, 2, 3, 4\}$, and state alphabet $\mathcal{S} = \{0, 1, 2, 3, 4\}$. Such channel can be viewed as memory with stuck faults with 5 states. The state S takes values $s = 1, 2, 3, 4$ with probability $\frac{p}{4}$ respectively. And $S = 0$ with probability $1-p$. The received data $Y_1 = S$ when $S = 1, 2, 3, 4$. And $Y_1 = X$ when $S = 0$. The received data Y_2 is a blurred version of Y_1, where $Y_2 = 0$ when $Y_1 = 1, 2$, and $Y_2 = 1$ when $Y_1 = 3, 4$.

Proposition 1. For the broadcast channels with state described above, the rate region (5) achieves the channel capacity, while the region (4) is strictly smaller than the channel capacity.

Proof: See [?] for details. ■

Now we define the sets for polarization and coding. Let $(V_1^{1:n}, V_2^{1:n})$ be a sequence of n i.i.d. random variables with pmf $P_{V_1,V_2}(v_1, v_2)$. Set the sequences $U_1^{1:n} = V_1^{1:n} G_n$ and

$U_2^{1:n} = V_2^{1:n} G_n$. Define the polarization sets

$$
\begin{aligned}
&\mathcal{H}_{U_1}^{(n)} = \{i \in [n] : Z(U_1^i | U_1^{1:i-1}) \geq 1 - 2^{-n^\beta}\},\\
&\mathcal{L}_{U_1}^{(n)} = \{i \in [n] : Z(U_1^i | U_1^{1:i-1}) \leq 2^{-n^\beta}\},\\
&\mathcal{H}_{U_1|S}^{(n)} = \{i \in [n] : Z(U_1^i | S^{1:n}, U_1^{1:i-1}) \geq 1 - 2^{-n^\beta}\},\\
&\mathcal{L}_{U_1|S}^{(n)} = \{i \in [n] : Z(U_1^i | S^{1:n}, U_1^{1:i-1}) \leq 2^{-n^\beta}\},\\
&\mathcal{H}_{U_1|Y_1}^{(n)} = \{i \in [n] : Z(U_1^i | Y_1^{1:n}, U_1^{1:i-1}) \geq 1 - 2^{-n^\beta}\},\\
&\mathcal{L}_{U_1|Y_1}^{(n)} = \{i \in [n] : Z(U_1^i | Y_1^{1:n}, U_1^{1:i-1}) \leq 2^{-n^\beta}\},\\
&\mathcal{H}_{U_1|Y_2}^{(n)} = \{i \in [n] : Z(U_1^i | Y_2^{1:n}, U_1^{1:i-1}) \geq 1 - 2^{-n^\beta}\},\\
&\mathcal{L}_{U_1|Y_2}^{(n)} = \{i \in [n] : Z(U_1^i | Y_2^{1:n}, U_1^{1:i-1}) \leq 2^{-n^\beta}\},\\
&\mathcal{H}_{U_2|Y_1,U_1}^{(n)} = \{i \in [n] : Z(U_2^i | Y_1^{1:n}, U_1^{1:n}, U_2^{1:i-1}) \geq 1 - 2^{-n^\beta}\},\\
&\mathcal{L}_{U_2|Y_1,U_1}^{(n)} = \{i \in [n] : Z(U_2^i | Y_1^{1:n}, U_1^{1:n}, U_2^{1:i-1}) \leq 2^{-n^\beta}\}.
\end{aligned} \tag{6}
$$

The information sets and the remaining frozen sets for receivers 1 and 2 are defined as follows:

$$
\begin{aligned}
&\mathcal{I}_1 = \mathcal{H}_{U_1|S}^{(n)} \cap \mathcal{L}_{U_1|Y_1}^{(n)},\ \mathcal{F}_{1a} = \mathcal{H}_{U_1|S}^{(n)} \cap \{\mathcal{L}_{U_1|Y_1}^{(n)}\}^c,\\
&\mathcal{F}_{1r} = (\mathcal{H}_{U_1|S}^{(n)})^c \cap \{\mathcal{L}_{U_1|Y_1}^{(n)}\}^c,\ \mathcal{F}_{1f} = (\mathcal{H}_{U_1|S}^{(n)})^c \cap \{\mathcal{L}_{U_1|Y_1}^{(n)}\},\\
&\mathcal{I}_2 = \mathcal{H}_{U_1|S}^{(n)} \cap \mathcal{L}_{U_1|Y_2}^{(n)},\ \mathcal{F}_{2a} = \mathcal{H}_{U_1|S}^{(n)} \cap \{\mathcal{L}_{U_1|Y_2}^{(n)}\}^c,\\
&\mathcal{F}_{2r} = (\mathcal{H}_{U_1|S}^{(n)})^c \cap \{\mathcal{L}_{U_1|Y_2}^{(n)}\}^c,\ \mathcal{F}_{2f} = (\mathcal{H}_{U_1|S}^{(n)})^c \cap \{\mathcal{L}_{U_1|Y_2}^{(n)}\}.
\end{aligned} \tag{7}
$$

A. *Polar Codes for the General Gelfand-Pinsker Problem*

Let us now consider polar codes for realizing the Gelfand-Pinsker binning scheme. Without loss of generality, transmission to receiver 1 is assumed. We use the chaining construction, stated as follows. In block 1, the encoder puts the message information in the bits $u^{\mathcal{I}_1}$, and generates the remaining frozen bits $u^{\mathcal{I}_1^c}$ using randomly chosen maps with randomness shared between the encoder and the decoders. For block $j = 2, \ldots, k$, the encoder chooses a subset of the information set $\mathcal{R}_1 \subseteq \mathcal{I}_1$ and fills the bits $u_1^{\mathcal{R}_1}$ with the information contained in $u^{\mathcal{F}_{1r}}$ of block $j-1$, which is approximately determined by the state sequence S^n and can not be recovered by using the received signal $y_1^{1:n}$. Then the encoder puts information in the bits $u^{\mathcal{I}_1 \backslash \mathcal{R}_1}$ and generates the frozen bits $u^{\mathcal{I}_1^c}$ according to randomly chosen maps. Here the bit sets $u^{\mathcal{R}_1}$ in blocks $j = 1, \ldots, k$ can be regarded as the chain to transmit the frozen bits $u^{\mathcal{F}_{1r}}$ to user 1.

Decoder 1 decodes from block k to block 1. Note that for block $j = k-1, \ldots, 1$, the bits $u_1^{\mathcal{F}_{1r}}$ can be recovered if decoding in block $j+1$ is successful. Since the remaining bits can be recovered either by applying maximum a posteriori rule or by using the randomly chosen maps, decoder 1 is able to decode the sequence $u^{1:n}$ for block $j = k-1, \ldots, 1$ if it decodes $u^{1:n}$ of block $j = k$ successfully. The main difficulty here is the transmission of block k. The work in [10] proposed a scheme to transmit the bits of block k by using an extra transmission phase, where state side information is not used at the encoder. There are counterexamples indicating that the scheme in [10] may not work. Consider a binary symmetric channel with additive interference $Y = X \oplus Z \oplus S$, where $Z \sim Bern(p)$ and $S \sim Bern(\frac{1}{2})$. It is easy to see that the channel capacity when the encoder does not use the state side information is zero, meaning that the extra phase is not capable of transmitting information. However, when the causal state information is utilized at the encoder, the channel capacity becomes $1 - H(p)$, which is nonzero when $0 \leq p < \frac{1}{2}$. Hence the information can be transmitted. The following lemma shows that it is sufficient to pre-communicate the bits $u_1^{\mathcal{F}_{1r}}$ of block k by adopting polar coding with causal side information.

Lemma 1. For a channel with random state $(\mathcal{X} \times \mathcal{S}, P_{Y|X,S}(y|x,s), \mathcal{Y})$, where the state is noncausally known at the encoder, if the channel capacity

$$
C = \max_{p_{U|S}(u|s), f(u,s)} I(U;Y) - I(U;S) \tag{8}
$$

is greater than 0, then $\max_{p_U(u), f(u,s)} I(U;Y) > 0$, i.e., the capacity for channel with causal state known at the encoder is greater than 0.

To pre-transmit the bits $u_1^{\mathcal{F}_{1r}}$ of block k, an extra phase that consists of t blocks is used, where the encoder adopts polar codes for channel with causal state. The encoder first chooses a random variable $(V', f'(v,s)) = \arg\max_{P_V(v), f(v,s)} I(V;Y)$ and sets the sequence $U'^{1:n} = V'^{1:n} G_n$. In each block $j = 1, \ldots, t$, the bits $u_1^{\mathcal{F}_{1r}}$ of block k are put in locations $\mathcal{I}_1' = \mathcal{H}_{U'} \cap \mathcal{L}_{U'|Y_1}$. And the frozen bits $u^{(\mathcal{I}_1')^c}$ are generated using randomly chosen maps as usual. Then the encoder transmits $f'(v', s)$ over the channel. Upon decoding, decoder 1 decodes the sequence $u'^{1:n}$ by applying maximum a posteriori rule and using the randomly chosen maps. Let $C_{causal} = \max_{P_V(v), f(v,s)} I(V;Y)$ be the capacity for channel with state sequence causally available at the encoder. According to Lemma 1, $C_{causal} > 0$. By fixing $t = \left\lceil \frac{|\mathcal{F}_{1r}|}{C_{causal}} \right\rceil$, the pre-communication of bits $u_1^{\mathcal{F}_{1r}}$ of block k can be completed in t blocks. The average message rate is given by

$$
\begin{aligned}
R_1 =& \frac{1}{kn + tn} [k(|\mathcal{I}_1| - |\mathcal{R}_1|) + |\mathcal{I}_1 \backslash \mathcal{R}_1|]\\
=& \frac{1}{kn + 2tn} [k(|\mathcal{H}_U^{(n)} \cap \mathcal{L}_{U|Y_1}^{(n)} \backslash \mathcal{H}_U^{(n)} \cap (\mathcal{H}_{U|S}^{(n)})^c|\\
& - |\mathcal{H}_U^{(n)} \cap (\mathcal{H}_{U|S}^{(n)})^c \backslash \mathcal{H}_U^{(n)} \cap \mathcal{L}_{U|Y_1}^{(n)}|) + |\mathcal{I}_1 \backslash \mathcal{R}_1|]\\
=& \frac{1}{kn + 2tn} [k(|\mathcal{H}_U^{(n)} \cap \mathcal{L}_{U|Y_1}^{(n)}|\\
& - |\mathcal{H}_U^{(n)} \cap (\mathcal{H}_{U|S}^{(n)})^c|) + |\mathcal{I}_1 \backslash \mathcal{R}_1|]\\
=& \frac{k}{k+2t} (I(V;Y_1) - I(V;S)) + \frac{1}{kn+2tn} |\mathcal{I}_1 \backslash \mathcal{R}_1| + o(1).
\end{aligned} \tag{9}
$$

As k increases to infinity, the rate R_1 approaches $I(V;Y_1) - I(V;S)$. Similar to polar codes for BCSI with common message, the coding complexity is $O(n \log n)$ and the error probability is $O(2^{-n^\beta})$ for any $0 < \beta < \frac{1}{2}$.

B. Polar Codes for BCSI with noncausal state

To begin with, split the message M_1 into messages M_{11} and M_{10} at rates R_{11} and R_{10} respectively. The coding scheme for BCSI with noncausal state employs a superposition strategy, where the information of (M_0, M_{10}, M_2) is carried by a sequence $u_1^{1:n}$ and the message M_{11} is put in another sequence $u_2^{1:n}$. The encoder transmits $f(v_1, v_2, s)$, where $v_1^{1:n} = u_1^{1:n} G_n$ and $v_2^{1:n} = u_2^{1:n} G_n$. Let the information rates carried by $u_1^{1:n}$ and $u_2^{1:n}$ be given by

$$\begin{aligned} R_0 + R_{10} &\leq I(V_1; Y_1) - I(V_1; S), \\ R_0 + R_2 &\leq I(V_1 : Y_2) - I(V_1; S), \\ R_{11} &\leq I(V_2; Y_1|V_1) - I(V_2; S|V_1). \end{aligned} \tag{10}$$

Summing the first and the third inequality in (10), we get (5).

Let us first deal with the transmission of the sequence $u_1^{1:n}$, which can be viewed as Gelfand-Pinsker binning simultaneously for the two users. The difficulty here is that the chain construction involves multiple chains. In particular, each decoder m needs a chain to transmit the frozen bits $\mathcal{F}_{mr}$. The two chains must be aligned in a same codeword without conflicts, where a position is assigned with two different values. To tackle the problem that the two chains may overlap and cause conflicts, we first deal with the case when the two chains do not overlap. Then we show that the case when the two chains overlap can be converted to the first case.

Let us assume that $R_{10} \geq R_2$. The arguments will be similar when $R_{10} \leq R_2$. Split the message M_{10} into messages M_{100} and M_{101} at rates R_{100} and R_{101} respectively such that $R_{100} = R_2$. The new equivalent common message is set as $M_0' = (M_{100} \oplus M_2, M_0)$. Then we have $R_1 + R_0 = R_0 + R_{10} + R_{11} = R_0' + R_{101} + R_{11}$, $R_2 + R_0 = R_0'$. Set $R_0' = \frac{|\mathcal{I}_2| - |\mathcal{F}_{2r}|}{n}$ and $R_0' + R_{101} = \frac{|\mathcal{I}_1| - |\mathcal{F}_{1r}|}{n}$. Consider the following two cases: (a) $nR_0' \geq |\mathcal{I}_1 \cap \mathcal{I}_2|$. (b) $nR_0' \leq |\mathcal{I}_1 \cap \mathcal{I}_2|$.

Case (a) : In this case, we can choose a subset $\mathcal{R}_1 \subseteq (\mathcal{I}_1 - \mathcal{I}_2)$ and a subset $\mathcal{R}_2 \subseteq (\mathcal{I}_2 - \mathcal{I}_1)$ such that $|\mathcal{R}_1| = |\mathcal{F}_{1r}|$ and $|\mathcal{R}_2| = |\mathcal{F}_{2r}|$. Similar as in the single user Gelfand-Pinsker case, the subsets $\mathcal{R}_1$ and $\mathcal{R}_2$ act the roles of generating the two chains to transmit the frozen bits $u^{\mathcal{F}_{1r}}$ and $u^{\mathcal{F}_{2r}}$ to the two users respectively. In case (a) the two chains do not overlap. Define the sets

$$\begin{aligned} \mathcal{M}_1 = \mathcal{I}_1 \backslash \mathcal{R}_1, \ \mathcal{M}_2 = \mathcal{I}_2 \backslash \mathcal{R}_2 \\ \mathcal{D}_1 = \mathcal{M}_1 - \mathcal{M}_2, \ \mathcal{D}_2 = \mathcal{M}_2 - \mathcal{M}_1. \end{aligned} \tag{11}$$

Let $\mathcal{D}_{10} \subseteq \mathcal{D}_1$ be a subset of $\mathcal{D}_1$ such that $|\mathcal{D}_{10}| = |\mathcal{D}_2|$. The coding scheme to transmit $u_1^{1:n}$ is presented in Fig.5. The first t blocks $j = 1, \ldots, t$ are used to pre-communicate the bits $u_1^{\mathcal{F}_{2r}}$ of block $j = t+1$. And the last t blocks $j = k+t+1, \ldots, k+2t$ conveys the bits $u_1^{\mathcal{F}_{1r}}$ of block $j = k+t$. In block $j = t+1$, the encoder fills the bits $u_1^{\mathcal{R}_2}$ with the information contained in $u_1^{\mathcal{F}_{2r}}$ of block $j+1$ and puts the M_0' information into bits $u_1^{\mathcal{M}_2}$. In block $j = t+2, \ldots, k+t-1$, the encoder copies the bits $u_1^{\mathcal{F}_{2r}}$ of block $j+1$ and the bits $u_1^{\mathcal{F}_{1r}}$ of block $j-1$ to $u_1^{\mathcal{R}_2}$ and $u_1^{\mathcal{R}_1}$ respectively. The bits $u_1^{\mathcal{D}_{10}}$ are filled with $u_1^{\mathcal{D}_2}$ bits of block $j-1$. The bits $u_1^{\mathcal{D}_1 \backslash \mathcal{D}_{10}}$ and bits $u_1^{\mathcal{M}_2}$ are inserted with M_{101} bits and M_0' bits respectively. In

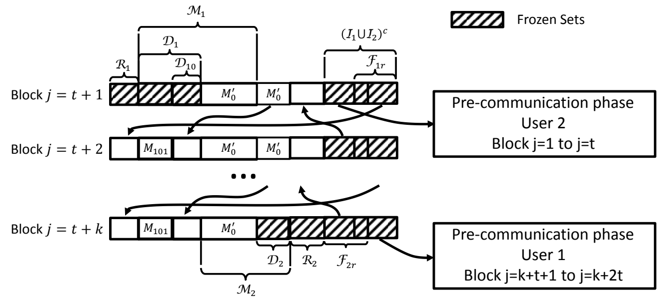

Fig. 2. Polar codes for transmitting $u_1^{1:n}$ in case (a).

block $j = k+t$, the encoder inserts the positions $\mathcal{R}_1$ with the information contained in $u_1^{\mathcal{F}_{1r}}$ of block $j-1$. The bits $u_1^{\mathcal{D}_{10}}$ are filled with $u_1^{\mathcal{D}_2}$ of block $j-1$ and the bits $u_1^{\mathcal{M}_1 \backslash \mathcal{D}_{10}}$ are filled with the information of M_{101}. The remaining bits are frozen and generated using randomized maps and the randomness is shared between the encoder and the decoders.

Upon decoding, user 2 begins by decoding the first t blocks in the pre-communication phase. Then it starts from block $j = t+1$ to block $j = k+t$. For block $t+1$, the bits $u_1^{\mathcal{I}_2 \cup \mathcal{F}_{2f}}$ can be decoded by maximum a posteriori rule and the bits $u_1^{\mathcal{F}_{2a}}$ can be recovered using the shared randomized maps. The bits $u_1^{\mathcal{F}_{2r}}$ are pre-communicated through the first t blocks . For block $j = t+2, \ldots, k+t-1$, The bits $u_1^{\mathcal{F}_{2r}}$, $u_1^{\mathcal{D}_{10}}$, and $u_1^{\mathcal{R}_1}$ can be recovered since the content therein is contained in the bits $u_1^{\mathcal{R}_2}$, $u_1^{\mathcal{D}_2}$, and $u_1^{\mathcal{F}_{1r}}$ respectively decoded in the last block $j-1$. Meanwhile, the bits $u_1^{\mathcal{D}_1 - \mathcal{D}_{10}}$ is available at user 2 as side information. The bits $u_1^{\mathcal{I}_2}$ can be decoded based on the received sequence $y_2^{1:n}$. The remaining frozen bits $u_1^{(\mathcal{I}_1 \cup \mathcal{I}_2)^c}$ can be calculated using the shared randomized maps. Therefore, user 2 decodes successfully. In block $j = k$, the decoding of the bits $u_1^{(\mathcal{R}_2)^c}$ is the same as that in block $j = t+2, \ldots, k+t-1$. The bits $u_1^{\mathcal{R}_2}$ are recovered using the randomly chosen maps. Similarly, user 1 starts from block $k+2t$ to block $t+1$ and is able to decode successfully.

Case (b) : In this case, $|\mathcal{F}_{2r}| > |\mathcal{I}_2 - \mathcal{I}_1|$, which implies that $\mathcal{R}_2 \cap \mathcal{I}_1 \neq \emptyset$ for any subset $\mathcal{R}_2 \in \mathcal{I}_2$ with $|\mathcal{R}_2| = |\mathcal{F}_{2r}|$. Hence in this case the two chains may overlap with each other. To avoid the value assignment conflicts in the overlapped set, the main idea is to let the bits $u_1^{\mathcal{R}_2 \cap \mathcal{I}_1}$ carry the information contained in $u_1^{\mathcal{R}_2}$ and $u_1^{\mathcal{I}_1}$ simultaneously. Let W_1' and W_2' be a subset of information carried in $(M_{101}, u_1^{\mathcal{R}_1})$ and $u_1^{\mathcal{R}_2}$ respectively such that $\log_2 |W_1'| = \log_2 |W_2'| = |\mathcal{I}_1 \cap \mathcal{I}_2| - nR_0'$. Let $M_0'' = (M_0', W_1' \oplus W_2')$, where $W_1' \oplus W_2'$ is the bitwise XOR of W_1' and W_2'. Since $R_0'' = \frac{|\mathcal{I}_1 \cap \mathcal{I}_2|}{n}$, we can adopt the coding scheme of case (a), by regarding M_0'' as the new equivalent common message. Note that in block $j = t+1$, the bits $u_1^{\mathcal{R}_1}$ does not contain information. Hence decoder 2 can recover W_1' and thus the information contained in W_2'. For blocks $j = t+2, \ldots, k+t$, decoder 2 knows the information of $(M_{101}, u_1^{\mathcal{R}_1})$ since $u_1^{\mathcal{R}_1}$ copies the bits $u_1^{\mathcal{F}_{1r}}$ from block $j-1$. Hence decoder 2 can recover the information contained in W_2'. Similarly, decoder 1 can recover the information contained in

W_1'. The message rates (R_0, R_{10}, R_2) are given by

$$\begin{aligned}
R_0 + R_{10} &= \frac{1}{kn+2tn}[(k-1)(|\mathcal{I}_1| - |\mathcal{R}_1|) + |\mathcal{M}_1 \cap \mathcal{M}_2|] \\
&= \frac{k-1}{k+2t}(I(V_1;Y_1) - I(V_1;S)) + \frac{1}{k}|\mathcal{M}_1 \cap \mathcal{M}_2| + o(1) \\
R_0 + R_2 &= \frac{1}{kn+2tn}[(k-1)(|\mathcal{I}_2| - |\mathcal{R}_2|) + |\mathcal{M}_1 \cap \mathcal{M}_2|] \\
&= \frac{k-1}{k+2t}(I(V_1;Y_2) - I(V_1;S)) + \frac{1}{k}|\mathcal{M}_1 \cap \mathcal{M}_2| + o(1)
\end{aligned} \tag{12}$$

The transmission of sequence $u_2^{1:n}$ can be regarded as Gelfand-Pinsker binning for user 1. Define

$$\begin{aligned}
\mathcal{I}_{11} &= \mathcal{H}^{(n)}_{U_2|S,U_1} \cap \mathcal{L}^{(n)}_{U_2|Y_1,U_1}, \\
\mathcal{F}_{11r} &= (\mathcal{H}^{(n)}_{U_2|S,U_1})^c \cap \{\mathcal{L}^{(n)}_{U_2|Y_1,U_1}\}^c,
\end{aligned} \tag{13}$$

The encoder uses t blocks as pre-communication phase and transmits M_{11} through k blocks. Choose a subset $\mathcal{R}_{11} \subseteq \mathcal{I}_{11}$ such that $|\mathcal{R}_{11}| = |\mathcal{F}_{11r}|$. The average rate per symbol R_{11} is given by

$$\begin{aligned}
R_{11} &= \frac{1}{kn+tn}[k(|\mathcal{I}_{11}| - |\mathcal{R}_{11}|) + |\mathcal{I}_{11} \backslash \mathcal{R}_{11}|] \\
&= \frac{k}{kn+tn}[I(V_2;Y_1|V_1) - I(V_2;S|V_1) + |\mathcal{I}_{11} \backslash \mathcal{R}_{11}| + o(1)].
\end{aligned} \tag{14}$$

Let C_{causal} be $C_{causal} = \max\{\max_{P_V(v),x(v,s)} I(V;Y_1), \max_{P_V(v),x(v,s)} I(V;Y_2)\}$. According to Lemma 1, $C_{causal} > 0$. Choose $t = \min\{\left\lceil \frac{|\mathcal{F}_{1r}|}{C_{causal}} \right\rceil, \left\lceil \frac{|\mathcal{F}_{2r}|}{C_{causal}} \right\rceil, \left\lceil \frac{|\mathcal{F}_{11r}|}{C_{causal}} \right\rceil\}$ to be fixed. Then according to (12) and (14), $R_1 + R_0$ and $R_2 + R_0$ approach arbitrarily closed to $I(V_1, V_2; Y_1) - I(V_1, V_2; S)$ and $I(V_1;Y_2) - I(V_1;S)$ respectively, as k grows to infinity. As n goes to infinity, the encoding and decoding complexity for each user is $O(n \log n)$. The error probability is upper bounded by $O(2^{-n^\beta})$ for $0 < \beta < \frac{1}{2}$.

C. Degraded BCSI with Common Message and with Noncausal State

Let us now establish the capacity region for degraded BCSI with common message and with noncausal state. A broadcast channels $P_{Y_1,Y_2|X,S}(y_1,y_2|x,s)$ is physically degraded if

$$P_{Y_2|X,S}(y_2|x,s) = P_{Y_2|Y_1}(y_2|y_1)P_{Y_1|X,S}(y_1|x,s) \tag{15}$$

for some distribution $P_{Y_1|Y_2}(y_1|y_2)$, i.e., $(X,S) \to Y_1 \to Y_2$ form a Markov chain. A broadcast channels $P_{Y_1,Y_2|X,S}(y_1,y_2|x,s)$ is stochastically degraded if

$$P_{Y_2|X,S}(y_2|x,s) = \sum_{y_1 \in \mathcal{Y}_1} P_{Y_2|Y_1}(y_2|y_1)P_{Y_1|X,S}(y_1|x,s) \tag{16}$$

for some distribution $P_{Y_1|Y_2}(y_1|y_2)$. Since the channel capacity depends only on the conditional marginals $P_{Y_1|X,S}(y_1|x,s)$ and $P_{Y_2|X,S}(y_2|x,s)$, the capacity region of a stochastically degraded BC is the same as that of a corresponding physically degraded BC [13]. Hence the notion of physically degraded and stochastically degraded are referred to as degraded, and the degradedness is denoted as $P_{Y_1|X,S}(y_1|x,s) \succ P_{Y_2|X,S}(y_2|x,s)$.

Theorem 2. Let $\mathcal{R}$ be the set of tuples (R_0, R_1, R_2) that satisfy

$$\begin{aligned}
R_1 + R_0 &\le I(V_1, V_2; Y_1) - I(V_1, V_2; S), \\
R_2 + R_0 &\le I(V_1; Y_2) - I(V_1; S)
\end{aligned} \tag{17}$$

for some random variables V_1, V_2 such that (1) $I(V_2;Y_1|V_1) > I(V_2;S|V_1)$, and (2) $(V_1,V_2) \to (S,X) \to Y_1 \to Y_2$ form a Markov chain, and for some function $\phi: \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{S} \to \mathcal{X}$ such that $x = \phi(v_1, v_2, s)$. Then $\mathcal{R}$ is the capacity region of the degraded BCSI with common message and with noncausal state $(\mathcal{X} \times \mathcal{S}, P_{Y_1,Y_2|X,S}(y_1,y_2|x,s), \mathcal{Y}_1 \times \mathcal{Y}_2)$.

IV. Conclusion

In this paper polar coding schemes are proposed for broadcast channels with receiver message side information (BCSI) and with noncausal state available at the encoder. It is proved that polar codes are able to achieve the Gelfand-Pinsker capacity through a two-phase transmission. In the first phase the encoder pre-communicates information through polar coding for channel with causal state. In the second phase the encoder transmits messages using chaining construction of polar codes. The presented polar coding scheme for BCSI with common message and with noncausal state has a superposition coding flavor in the sense that the code sequences are successively generated.

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, 2009.

[2] N. Goela, E. Abbe, and M. Gastpar, "Polar codes for broadcast channels," in *Proc, IEEE Int. Symp. Inform. Theory*. IEEE, 2013, pp. 1127–1131.

[3] M. Mondelli, S. Hassani, I. Sason, and R. Urbanke, "Achieving marton's region for broadcast channels using polar codes," *IEEE Trans. Inform. Theory*, pp. 783–800, 2015.

[4] M. Andersson, R. F. Schaefer, T. J. Oechtering, and M. Skoglund, "Polar coding for bidirectional broadcast channels with common and confidential messages," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1901–1908, 2013.

[5] T. Oechtering and M. Skoglund, "Bidirectional broadcast channel with random states noncausally known at the encoder," *IEEE Trans. Inform. Theory*, vol. 59, pp. 64–75, 2013.

[6] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2867–2877, 2005.

[7] C. Nair, A. E. Gamal, and Y.-K. Chia, "An achievability scheme for the compound channel with state noncausally available at the encoder," *arXiv preprint arXiv:1004.3427*, 2010.

[8] R. Khosravi-Farsani and F. Marvasti, "Capacity bounds for multiuser channels with non-causal channel state information at the transmitters," in *Proc. IEEE Inform. Theo. Workshop*, Oct 2011, pp. 195–199.

[9] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, 2009.

[10] E. E. Gad, Y. Li, J. Kliewer, M. Langberg, A. Jiang, and J. Bruck, "Asymmetric error correction and flash-memory rewriting using polar codes," *arXiv preprint arXiv:1410.3542*, 2014.

[11] D. Burshtein, "Coding for asymmetric side information channels with applications to polar codes," in *Proc, IEEE Int. Symp. Inform. Theory*, 2015, pp. 1527–1531.

[12] R. Mori and T. Tanaka, "Channel polarization on q-ary discrete memoryless channels by arbitrary kernels," in *Proc, IEEE Int. Symp. Inform. Theory*. IEEE, 2010, pp. 894–898.

[13] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.