KTH Electrical Engineering

Polar Codes for Bidirectional Broadcast Channels with Common and Confidential Messages

MATTIAS ANDERSSON, RAFAEL F. WYREMBELSKI,
TOBIAS J. OECHTERING, MIKAEL SKOGLUND

Stockholm 2012

Communication Theory
School of Electrical Engineering
Kungliga Tekniska Hgskolan

Polar Codes for Bidirectional Broadcast Channels with Common and Confidential Messages

Mattias Andersson*, Rafael F. Wyrembelski†, Tobias J. Oechtering*, and Mikael Skoglund*

*School of Electrical Engineering and the ACCESS Linnaeus Center
Royal Institute of Technology (KTH), Stockholm, Sweden

†Lehrstuhl für Theoretische Informationstechnik
Technische Universität München, Germany

***Abstract*—We consider the bidirectional broadcast channel with common and confidential messages. We show that polar codes achieve the capacity of binary input symmetrical bidirectional broadcast channels with confidential messages, if one node's channel is a degraded version of the other node's channel. We also find a new bound on the cardinality of the auxiliary random variable in this setup.**

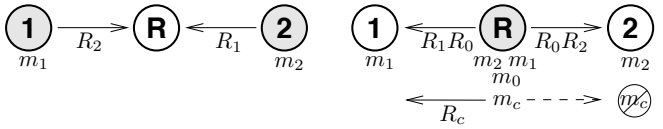
I. Introduction

Recent developments have significantly increased the performance of wireless networks. One research area that is gaining more importance is the efficient implementation of multiple services at the physical layer. For example, in current cellular systems operators establish not only (bidirectional) voice communication, but also offer further multicast or confidential services that are subject to certain secrecy constraints. These should be wisely integrated to increase the spectral efficiency of next generation cellular systems.

Further, it has been shown that the concept of bidirectional relaying improves the performance and coverage in wireless networks. This is mainly based on the fact that it advantageously exploits the property of bidirectional communication to reduce the inherent loss in spectral efficiency induced by half-duplex relays [1,2]. Bidirectional relaying applies to three-node networks, where a half-duplex relay node establishes a bidirectional communication between two other nodes using a two-phase decode-and-forward protocol [3–5]. This is also known as two-way relaying.

Here, we consider physical layer service integration for bidirectional relaying where the relay integrates additional common and confidential messages in the broadcast phase. In addition to the transmission of both individual messages, it has the following tasks as visualized in Figure 1: the transmission of a common message to both nodes and the transmission of a confidential message to one node, which has to be kept secret from the other, non-legitimate node. This necessitates the analysis of the *bidirectional broadcast channel (BBC) with common and confidential messages*. Note that both receiving nodes can use their own message from the previous phase for decoding so that this channel differs from the classical broadcast channel with common and confidential messages.

The secrecy capacity region of the discrete memoryless BBC with common and confidential messages is derived in [6]. The design of practical coding schemes for the BBC is discussed in [7], while [8] addresses the problem of joint network and channel coding in multi-way relay channels.



(a) MAC phase (b) BBC phase

Fig. 1. Physical layer service integration in bidirectional relay networks. In the initial MAC phase, nodes 1 and 2 transmit their messages m_1 and m_2 with rates R_2 and R_1 to the relay node. Then, in the BBC phase, the relay forwards the messages m_1 and m_2 and adds a common message m_0 with rate R_0 to the communication and further a confidential message m_c for node 1 with rate R_c which should be kept secret from node 2.

In this work we consider polar codes for the BBC with common and confidential messages. Polar codes were introduced by Arıkan and were shown to be capacity achieving for a large class of channels in [9,10]. They have been further studied for a large range of multi-user channels in [11–17].

II. Polar Codes

We consider binary polar codes which are block codes of length $N = 2^n$. Let $\mathcal{X}$ be the binary field and let $G = RF^{\otimes n}$, where R is the bit-reversal mapping defined in [9], $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, and $F^{\otimes n}$ denotes the n^{th} Kronecker power of F. Apply the linear transformation G to N bits u_1^N and send the result through N independent copies of a binary input memoryless channel $W(y|x)$. This gives an N-dimensional channel $W_N(y_1^N|u_1^N)$, and Arıkan's observation was that the channels seen by individual bits, defined by

$$W_N^{(i)}(y_1^N, u_1^{i-1}|u_i) = \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{2^{N-1}} W_N(y_1^N|u_1^N), \quad (1)$$

polarize, i.e as N grows $W_N^{(i)}$ approaches either an error-free channel or a completely noisy channel. We refer to the error-free channels as *good* channels, and the idea of polar coding is to send information only over the good channels, while keeping the input to the bad channels fixed, and known both at the destination and the sender.

Given a subset $\mathcal{A} \subset \{1, \ldots, N\}$ and a binary vector $u_{\mathcal{F}}$ of length $N - |\mathcal{A}|$ we define the polar code $P(N, \mathcal{A}, u_{\mathcal{F}})$ of length N as follows. Let $G_{\mathcal{A}}$ be the submatrix of G formed by rows with indices in $\mathcal{A}$, and let $u_{\mathcal{A}}$ be the corresponding subvector of u_1^N. We call $\mathcal{A}^{\mathcal{C}} = \mathcal{F}$ the frozen set, and the (fixed) bits $u_{\mathcal{F}}$ frozen bits. The codewords of $P(N, \mathcal{A}, u_{\mathcal{F}})$ are given by $x^N = u_{\mathcal{A}} G_{\mathcal{A}} \oplus u_{\mathcal{F}} G_{\mathcal{F}}$ and the rate is $|\mathcal{A}|/N$.

The block error probability using the successive cancellation (SC) decoding rule defined by

$$\hat{u}_i = \begin{cases} u_i & i \in \mathcal{F}, \\ 0 & \text{if } \frac{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} | u_i = 0)}{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} | u_i = 1)} \geq 1 \text{ and } i \in \mathcal{A}, \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

can be upper bounded by $\sum_{i \in \mathcal{A}} Z_N^{(i)}$, where $Z_N^{(i)}$ is the Bhattacharyya parameter for the channel $W_N^{(i)}$[9]. It was shown in [18] that for any $\beta < 1/2$,

$$\liminf_{n \to \infty} \frac{1}{N} |\{i : Z_N^{(i)} < 2^{-N^\beta}\}| = I(W), \quad (3)$$

where $I(W)$ is the symmetric capacity of W, which equals the Shannon capacity for symmetric channels. Thus if we let the good channels be given by

$$\mathcal{G}_N = \{i : Z_N^{(i)} < 2^{-N^\beta}\}, \quad (4)$$

the rate of $P(N, \mathcal{G}_N, u_{\mathcal{F}})$ approaches $I(W)$ as N grows. Also the block error probability P_e using SC decoding is upper bounded by

$$P_e \leq N 2^{-N^\beta}. \quad (5)$$

We define the nested polar code $P(N, \mathcal{A}, \mathcal{B}, u_{\mathcal{F}})$ of length N where $\mathcal{B} \subset \mathcal{A}$ as follows. The codewords of $P(N, \mathcal{A}, \mathcal{B}, u_{\mathcal{F}})$ are the same as the codewords for $P(N, \mathcal{A}, u_{\mathcal{F}})$. The nested structure is defined by partitioning $P(N, \mathcal{A}, u_{\mathcal{F}})$ as cosets of $P(N, \mathcal{B}, [0\ u_{\mathcal{F}}])$. Thus codewords in $P(N, \mathcal{A}, \mathcal{B}, u_{\mathcal{F}})$ are given by $x^N = u_{\mathcal{B}} G_{\mathcal{B}} \oplus u_{\mathcal{A} \setminus \mathcal{B}} G_{\mathcal{A} \setminus \mathcal{B}} \oplus u_{\mathcal{F}} G_{\mathcal{F}}$, where $u_{\mathcal{A} \setminus \mathcal{B}}$ determines which coset the codeword lies in. Note that each coset will be a polar code with $\mathcal{B}^{\mathcal{C}}$ as the frozen set. The frozen bits u_i are either given by $u_{\mathcal{F}}$ (if $i \in \mathcal{A}^{\mathcal{C}}$) or they equal the corresponding bits in $u_{\mathcal{A} \setminus \mathcal{B}}$.

For the following analysis we will need two results relating degraded channels and nested polar codes. Let W_1 and W_2 be two symmetric binary input memoryless channels, and let W_2 be degraded with respect to W_1. Denote the polarized channels as defined in (1) by $W_{1,N}^{(i)}$ and $W_{2,N}^{(i)}$, and their Bhattacharyya parameters by $Z_{1,N}^{(i)}$ and $Z_{2,N}^{(i)}$. We use the following lemma:

Lemma 1 ([11, Lemma 4.7]). *If W_2 is degraded with respect to W_1, then $W_{2,N}^{(i)}$ is degraded with respect to $W_{1,N}^{(i)}$ and $Z_{2,N}^{(i)} \geq Z_{1,N}^{(i)}$.*

The following result for degraded wiretap channels [19] was shown in [13–16]:

Theorem 1 ([13–16]). *Let W be a symmetric wiretap channel and denote the marginal channels to the main user and the wiretapper by W_1 and W_2 respectively. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be the corresponding sets given by (4). If W_2 is degraded with respect to W_1, the nested polar code $P(N, \mathcal{G}_1, \mathcal{G}_2, u_{\mathcal{F}})$ achieves the secrecy capacity of the wiretap channel.*

III. Polar Codes for the Bidirectional Broadcast Channel

Let $\mathcal{X}$ and $\mathcal{Y}_k$, $k = 1, 2$, be finite input and output sets. Then for input and output sequences $x^N \in \mathcal{X}^N$ and $y_k^N \in \mathcal{Y}_k^N$, $k = 1, 2$, of length N, the discrete memoryless broadcast channel is given by $W_N(y_1^N, y_2^N | x^N) := \prod_{i=1}^N W(y_{1,i}, y_{2,i} | x_i)$. Since we do not allow any cooperation between the receiving nodes, it is sufficient to consider the marginal transition probabilities $W_{k,N} := \prod_{i=1}^N W_k(y_{k,i} | x_i)$, $k = 1, 2$ only.

We consider the standard model with a block code of arbitrary but fixed length N. The set of individual messages of node k, $k = 1, 2$, is denoted by $\mathcal{M}_k := \{1, ..., M_k^{(N)}\}$. The sets of common and confidential messages of the relay node are denoted by $\mathcal{M}_0 := \{1, ..., M_0^{(N)}\}$ and $\mathcal{M}_c := \{1, ..., M_c^{(N)}\}$, respectively. Further, we use $\mathcal{M} := \mathcal{M}_c \times \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2$.

In the bidirectional broadcast phase, we assume that the relay has successfully decoded both individual messages $m_1 \in \mathcal{M}_1$ and $m_2 \in \mathcal{M}_2$ that nodes 1 and 2 transmitted in the previous multiple access phase. Thus m_k is known at node k and at the relay. Besides both individual messages the relay additionally transmits a common message $m_0 \in \mathcal{M}_0$ to both nodes and a confidential message $m_c \in \mathcal{M}_c$ to node 1, which should be kept secret from node 2, cf. Figure 1.

The ignorance of the non-legitimate node 2 about the confidential message $m_c \in \mathcal{M}_c$ is measured by the concept of equivocation rate. Here, the equivocation rate $\frac{1}{N} H(\mathrm{M}_c | \mathrm{Y}_2^N \mathrm{M}_2)$ characterizes the secrecy level of the confidential message. The higher the equivocation rate, the more ignorant node 2 is about the confidential message. We consider the case of *perfect secrecy* and thus require that the confidential rate R_c fulfills

$$\tfrac{1}{N} H(\mathrm{M}_c | \mathrm{Y}_2^N \mathrm{M}_2) \geq R_c - \delta$$

for some (small) $\delta > 0$. This is often equivalently written as $\frac{1}{N} I(\mathrm{M}_c; \mathrm{Y}_2^N | \mathrm{M}_2) \leq \delta$.

The BBC with common and confidential messages was analyzed in [6] for discrete memoryless channels. Its secrecy capacity region is restated in the following theorem:

Theorem 2 ([6]). *The secrecy capacity region of the BBC with common and confidential messages is the set of rate tuples $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ that satisfy*

$$R_c \leq I(\mathrm{V}; \mathrm{Y}_1 | \mathrm{U}) - I(\mathrm{V}; \mathrm{Y}_2 | \mathrm{U})$$
$$R_0 + R_k \leq I(\mathrm{U}; \mathrm{Y}_k), \quad k = 1, 2$$

for random variables $\mathrm{U} - \mathrm{V} - \mathrm{X} - (\mathrm{Y}_1, \mathrm{Y}_2)$. The cardinalities of the ranges of U and V can be bounded by

$$|\mathcal{U}| \leq |\mathcal{X}| + 3, \qquad |\mathcal{V}| \leq |\mathcal{X}|^2 + 4|\mathcal{X}| + 3.$$

For the following analysis of polar codes we need the case where the marginal channels are degraded, i.e., $\mathrm{X}-\mathrm{Y}_1-\mathrm{Y}_2$.

Corollary 1. *The secrecy capacity region of the degraded BBC with common and confidential messages is the set of rate tuples* $(R_c, R_0, R_1, R_2) \in \mathbb{R}_+^4$ *that satisfy*

$$R_c \le I(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - I(\mathrm{X};\mathrm{Y}_2|\mathrm{U})$$
$$R_0 + R_k \le I(\mathrm{U};\mathrm{Y}_k), \quad k=1,2$$

for random variables $\mathrm{U}-\mathrm{X}-\mathrm{Y}_1-\mathrm{Y}_2$. *The cardinality of the range of* U *can be bounded by*

$$|\mathcal{U}| \le |\mathcal{X}|.$$

Proof: The achievability follows immediately from the non-degraded case in Theorem 2, cf. also [6]. The converse and the bound on the cardinality of $\mathcal{U}$ is devoted to the appendix. ∎

A. Polar Codes for the BBC

First consider a binary input BBC W without common and confidential messages. The capacity region is given by

$$R_1 \le C_1 \tag{6}$$
$$R_2 \le C_2 \tag{7}$$

where C_1 and C_2 are the capacities of W_1 and W_2 respectively.

Theorem 3. *Let W be a BBC with binary input alphabet and symmetric marginal channels W_1 and W_2. Then there exists a polar coding scheme that achieves the rates given by (6) and (7).*

Proof: Fix $0 < \beta < 1/2$. Let $W_{k,N}^{(i)}$ and $Z_{k,N}^{(i)}$ for $k = 1, 2$ denote the polarized marginal channels and their Bhattacharyya parameters. Now define the following sets:

$$\mathcal{G}_{1,N} = \{i : Z_{1,N}^{(i)} < 2^{-N^\beta} \text{ and } Z_{2,N}^{(i)} \ge 2^{-N^\beta}\}$$
$$\mathcal{G}_{2,N} = \{i : Z_{1,N}^{(i)} \ge 2^{-N^\beta} \text{ and } Z_{2,N}^{(i)} < 2^{-N^\beta}\}$$
$$\mathcal{G}_{12,N} = \{i : Z_{1,N}^{(i)} < 2^{-N^\beta} \text{ and } Z_{2,N}^{(i)} < 2^{-N^\beta}\}$$
$$\mathcal{B}_N = \{i : Z_{1,N}^{(i)} \ge 2^{-N^\beta} \text{ and } Z_{2,N}^{(i)} \ge 2^{-N^\beta}\}$$

$\mathcal{G}_{1,N}$ are the channels that are good only for node 1, $\mathcal{G}_{2,N}$ the channels that are good only for node 2, $\mathcal{G}_{12,N}$ are the channels that are good for both nodes, and $\mathcal{B}_N$ are the channels that are bad for both nodes. Consider the polar code $\mathcal{P}(N, \mathcal{G}_{1,N} \cup \mathcal{G}_{2,N} \cup \mathcal{G}_{12,N}, u_{\mathcal{F}})$ with input bits given by

$$u_i = \begin{cases} m_{2i} & \text{if } i \in \mathcal{G}_{1,N}, \\ m_{1i} & \text{if } i \in \mathcal{G}_{2,N}, \\ m_{1i} \oplus m_{2i} & \text{if } i \in \mathcal{G}_{12,N}, \end{cases}$$

where we assume that the messages m_1 and m_2 are binary vectors. Since node 1 knows m_1 he treats the input bits in $\mathcal{G}_{2,N}$ as frozen and decodes the input bits u_i for $i \in \mathcal{G}_{1,N} \cup \mathcal{G}_{12,N}$ using the SC decoder (2). Finally he subtracts the bits of m_1 that appear in bits in $\mathcal{G}_{12,N}$. Thus the rate for node 1 becomes

$$R_{1,N} = \frac{|\mathcal{G}_{1,N}| + |\mathcal{G}_{12,N}|}{N}. \tag{8}$$

Node 2 treats the input bits m_2 in $\mathcal{G}_{1,N}$ as frozen and gets the rate

$$R_{2,N} = \frac{|\mathcal{G}_{2,N}| + |\mathcal{G}_{12,N}|}{N}. \tag{9}$$

By the definition of $\mathcal{G}_{1,N}, \mathcal{G}_{2,N}, \mathcal{G}_{12,N}, \mathcal{B}_N$ and using (3) - (5) we see that the error probability goes to zero as N increases, and that the rates R_1 and R_2 approach the capacities C_1 and C_2. ∎

B. Polar Codes for the BBC with Common Messages

Note that we can use some of the input bits in $\mathcal{G}_{12,N}$ to transmit a common message m_0, unknown at both destinations, by transferring parts of the rates R_1 and R_2 to R_0.

Corollary 2. *Let W be a BBC with binary input alphabet and symmetric marginal channels W_1 and W_2, where W_2 is degraded with respect to W_1. If we consider an additional common message m_0, the scheme in Theorem 3 achieves the following rate triples, which is the capacity region.*

$$R_0 + R_1 \le C_1 \tag{10}$$
$$R_0 + R_2 \le C_2. \tag{11}$$

Proof: It is easy to see that C_1 and C_2 are outer bounds to the capacity region. Since W_2 is degraded with respect to W_1 we have $\mathcal{G}_{2,N} = \emptyset$ by Lemma 1. Thus, by (3),

$$\lim_{N\to\infty} R_{0,N} + R_{1,N} = \lim_{N\to\infty} \frac{|\mathcal{G}_{1,N}| + |\mathcal{G}_{12,N}|}{N} = C_1, \tag{12}$$

and

$$\lim_{N\to\infty} R_{0,N} + R_{2,N} = \lim_{N\to\infty} \frac{|\mathcal{G}_{12,N}|}{N} = C_2. \tag{13}$$

which completes the proof. ∎

C. Polar Codes for the BBC with Confidential Messages

Now we show how to design polar codes for a BBC with a confidential message. For simplicity we consider the case where W_1 and W_2 are binary symmetric channels (BSC) with transition probabilities p_1 and p_2, with $p_2 > p_1$. We call such a channel a binary symmetric BBC. Using the same arguments as in [20, Example 15.6.3] it is easy to show that choosing U to be a Ber(1/2) binary random variable, and $p_{\mathrm{X}|\mathrm{U}}$ to be a BSC(α), with $0 < \alpha < 1/2$ is optimal. In this case the secrecy capacity region in Corollary 1 becomes

$$R_c \le h_2(\alpha \star p_1) - h_2(p_1) - h_2(\alpha \star p_2) + h_2(p_2),$$
$$R_0 + R_k \le 1 - h_2(\alpha \star p_k), \quad k=1,2$$

where $h_2(x) = -x\log x - (1-x)\log(1-x)$ and $\alpha \star \beta = (1-\alpha)\beta + \alpha(1-\beta)$.

Our main result is the following:

Theorem 4. *There exists a polar code $\mathcal{C}_{BBC}$ designed for the binary symmetric BBC, and a polar code $\mathcal{C}_{WT}$ designed for the binary symmetric wiretap channel such that transmitting*

$$\mathrm{X}^N = \mathrm{X}_{BBC}^N \oplus \mathrm{X}_{WT}^N,$$

for $\mathrm{X}_{BBC}^N \in \mathcal{C}_{BBC}$ *and* $\mathrm{X}_{WT}^N \in \mathcal{C}_{WT}$ *achieves the secrecy capacity region for the binary symmetric BBC with common and confidential messages*.

Proof: Fix $0 < \alpha < 1/2$. We design $\mathcal{C}_{BBC}$ for a binary symmetric BBC with transition probabilities $\alpha \star p_1$ and $\alpha \star p_2$. Assume that X_{WT}^N is statistically indistinguishable from an i.i.d. Ber(α) vector. Then, by Corollary 2, $\mathcal{C}_{BBC}$ can achieve all rate triples satisfying

$$R_0 + R_k \leq 1 - h_2(\alpha \star p_k), \quad k = 1, 2.$$

Both nodes can now decode X_{BBC}^N and remove its contribution. Note that since the channels are symmetric the error probabilities do not depend on the values of the frozen bits, and we can choose them to be zero [9]. Also note that since X_{BBC}^N and X_{WT}^N are independent, X_{BBC}^N provides no information about X_{WT}^N. Thus assuming that node 2 decodes X_{BBC}^N does not increase the equivocation of m_c at node 2.

Let $\mathcal{C}_{WT}$ be a polar code with input weight α designed for a binary symmetric wiretap channel with transition probabilities p_1 and p_2 using Theorem 1. To design a polar code with input weight α we augment the binary channel with a virtual q-ary input and then design a polar code for the augmented channel. For details see [10,11]. This construction achieves all rates satisfying

$$R_c \leq h_2(\alpha \star p_1) - h_2(p_1) - h_2(\alpha \star p_2) + h_2(p_2),$$

while keeping the message perfectly secret from node 2.

In order to make the codewords of $\mathcal{C}_{WT}$ statistically indistinguishable from an i.i.d. Ber(α) vector we average over all possible values of the frozen bits of $\mathcal{C}_{WT}$. Let $P_{e,BBC}(u_{\mathcal{F}})$, $P_{e,WT}(u_{\mathcal{F}})$, and $P_e(u_{\mathcal{F}})$ be the average error probabilities of $\mathcal{C}_{BBC}$, $\mathcal{C}_{WT}$, and the overall scheme when using $u_{\mathcal{F}}$ as the frozen bits for $\mathcal{C}_{WT}$. Choosing $u_{\mathcal{F}}$ uniformly at random we can make

$$E_{\mathrm{U}_{\mathcal{F}}}[P_e(\mathrm{U}_{\mathcal{F}})] \leq E_{\mathrm{U}_{\mathcal{F}}}[P_{e,BBC}(\mathrm{U}_{\mathcal{F}}) + P_{e,WT}(\mathrm{U}_{\mathcal{F}})]$$

arbitrarily small if we choose N large enough, since the codewords of $\mathcal{C}_{WT}$ are i.i.d. Ber(α) when we average over $u_{\mathcal{F}}$. Since the average error probability is small there exists at least one $u_{\mathcal{F}}$ such that $P_e(u_{\mathcal{F}})$ is small. ∎

IV. Conclusions

We have given a polar coding scheme that achieves the secrecy capacity region of a binary symmetric bidirectional broadcast channel with common and confidential messages. We have also found a new bound on the cardinality of the auxiliary random variable in this setup.

Appendix

A. Proof of Weak Converse

For any sequence of codes for the degraded BBC with common and confidential messages with error probabilities going to zero, we want to show that there exists random variables $\mathrm{U} - \mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2$ such that

$$\begin{aligned} \tfrac{1}{N} H(\mathrm{M}_c|\mathrm{Y}_2^N \mathrm{M}_2) &\leq I(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - I(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) + o(N^0) \\ \tfrac{1}{N}(H(\mathrm{M}_0) + H(\mathrm{M}_k)) &\leq I(\mathrm{U};\mathrm{Y}_k) + o(N^0), \ k = 1, 2 \end{aligned}$$

We do this by using techniques similar to [21] and the Fano-like inequalities

$$\begin{aligned} H(\mathrm{M}_c \mathrm{M}_0 \mathrm{M}_2 | \mathrm{Y}_1^N \mathrm{M}_1) &\leq N\epsilon_1^{(N)}, \\ H(\mathrm{M}_0 \mathrm{M}_1 | \mathrm{Y}_2^N \mathrm{M}_2) &\leq N\epsilon_2^{(N)}, \end{aligned}$$

from [6]. Let $\mathrm{M}_{012} = (\mathrm{M}_0 \mathrm{M}_1 \mathrm{M}_2)$ and introduce the random variable $\mathrm{U}_i = (\mathrm{M}_{012} \mathrm{Y}_1^{i-1})$.

We first bound $N(R_0 + R_1) \leq H(\mathrm{M}_0) + H(\mathrm{M}_2)$ as

$$\begin{aligned} H(\mathrm{M}_0) + H(\mathrm{M}_2) &\leq I(\mathrm{M}_{012}; \mathrm{Y}_1^N) + N\epsilon_1^{(N)} \\ &= \sum_{i=1}^N I(\mathrm{M}_{012}; \mathrm{Y}_{1i} | \mathrm{Y}_1^{i-1}) + N\epsilon_1^{(N)} \\ &\leq \sum_{i=1}^N I(\mathrm{M}_{012} \mathrm{Y}_1^{i-1}; \mathrm{Y}_{1i}) + N\epsilon_1^{(N)} \\ &= \sum_{i=1}^N I(\mathrm{U}_i; \mathrm{Y}_{1i}) + N\epsilon_1^{(N)}. \end{aligned}$$

Then we bound $N(R_0 + R_2) \leq H(\mathrm{M}_0) + H(\mathrm{M}_1)$ as

$$\begin{aligned} H(\mathrm{M}_0) + H(\mathrm{M}_1) &\leq I(\mathrm{M}_{012}; \mathrm{Y}_2^N) + N\epsilon_2^{(N)} \\ &= \sum_{i=1}^N I(\mathrm{M}_{012}; \mathrm{Y}_{2i} | \mathrm{Y}_2^{i-1}) + N\epsilon_2^{(N)} \\ &\leq \sum_{i=1}^N I(\mathrm{M}_{012} \mathrm{Y}_1^{i-1} \mathrm{Y}_2^{i-1}; \mathrm{Y}_{2i}) + N\epsilon_2^{(N)} \\ &\overset{(*)}{=} \sum_{i=1}^N I(\mathrm{M}_{012} \mathrm{Y}_1^{i-1}; \mathrm{Y}_{2i}) + N\epsilon_2^{(N)} \\ &= \sum_{i=1}^N I(\mathrm{U}_i; \mathrm{Y}_{2i}) + N\epsilon_2^{(N)} \end{aligned}$$

where we use the degradedness $\mathrm{X}_i - \mathrm{Y}_{1i} - \mathrm{Y}_{2i}$ in $(*)$.

Finally we bound $NR_c \leq H(\mathrm{M}_c|\mathrm{Y}_2^N \mathrm{M}_2)$ as

$$\begin{aligned} &H(\mathrm{M}_c|\mathrm{Y}_2^N \mathrm{M}_2) \\ &= H(\mathrm{M}_c|\mathrm{Y}_2^N \mathrm{M}_{012}) + I(\mathrm{M}_c; \mathrm{M}_0 \mathrm{M}_1 | \mathrm{Y}_2^N \mathrm{M}_2) \\ &\leq H(\mathrm{M}_c|\mathrm{Y}_2^N \mathrm{M}_{012}) + N\epsilon_2^{(N)} \\ &= I(\mathrm{M}_c; \mathrm{Y}_1^N | \mathrm{Y}_2^N \mathrm{M}_{012}) + H(\mathrm{M}_c|\mathrm{Y}_2^N \mathrm{M}_{012} \mathrm{Y}_1^N) + N\epsilon_2^{(N)} \\ &\leq I(\mathrm{M}_c; \mathrm{Y}_1^N | \mathrm{Y}_2^N \mathrm{M}_{012}) + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\ &\leq I(\mathrm{M}_c \mathrm{X}^N; \mathrm{Y}_1^N | \mathrm{Y}_2^N \mathrm{M}_{012}) + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\ &= I(\mathrm{X}^N; \mathrm{Y}_1^N | \mathrm{Y}_2^N \mathrm{M}_{012}) + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\ &= H(\mathrm{X}^N|\mathrm{M}_{012} \mathrm{Y}_2^N) - H(\mathrm{X}^N|\mathrm{M}_{012} \mathrm{Y}_2^N \mathrm{Y}_1^N) + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\ &= H(\mathrm{X}^N|\mathrm{M}_{012} \mathrm{Y}_2^N) - H(\mathrm{X}^N|\mathrm{M}_{012} \mathrm{Y}_1^N) + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\ &= I(\mathrm{X}^N; \mathrm{Y}_1^N | \mathrm{M}_{012}) - I(\mathrm{X}^N; \mathrm{Y}_2^N | \mathrm{M}_{012}) + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} I(\mathrm{X}^N; \mathrm{Y}_{1i}|\mathrm{M}_{012}\mathrm{Y}_1^{i-1}) - I(\mathrm{X}^N; \mathrm{Y}_{2i}|\mathrm{M}_{012}\mathrm{Y}_2^{i-1}) \\
&\quad + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\
&= \sum_{i=1}^{N} H(\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}) - H(\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}\mathrm{X}^N) + \\
&\quad - H(\mathrm{Y}_{2i}|\mathrm{Y}_2^{i-1}\mathrm{M}_{012}) + H(\mathrm{Y}_{2i}|\mathrm{Y}_2^{i-1}\mathrm{M}_{012}\mathrm{X}^N) \\
&\quad + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\
&\leq \sum_{i=1}^{N} H(\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}) - H(\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}\mathrm{X}_i) + \\
&\quad - H(\mathrm{Y}_{2i}|\mathrm{Y}_2^{i-1}\mathrm{Y}_1^{i-1}\mathrm{M}_{012}) + H(\mathrm{Y}_{2i}|\mathrm{Y}_2^{i-1}\mathrm{M}_{012}\mathrm{X}_i) \\
&\quad + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\
&= \sum_{i=1}^{N} H(\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}) - H(\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}\mathrm{X}_i) + \\
&\quad - H(\mathrm{Y}_{2i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}) + H(\mathrm{Y}_{2i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}\mathrm{X}_i) \\
&\quad + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\
&= \sum_{i=1}^{N} I(\mathrm{X}_i; \mathrm{Y}_{1i}|\mathrm{U}_i) - I(\mathrm{X}_i; \mathrm{Y}_{2i}|\mathrm{U}_i) + N\epsilon_1^{(N)} + N\epsilon_2^{(N)}.
\end{aligned}
$$

Now we get the desired bounds by letting J be a R.V. uniformly distributed over $\{1, \ldots, N\}$, and choosing $\mathrm{U} = (\mathrm{U}_\mathrm{J}, \mathrm{J}), \mathrm{X} = \mathrm{X}_\mathrm{J}, \mathrm{Y}_1 = \mathrm{Y}_{1\mathrm{J}}$, and $\mathrm{Y}_2 = \mathrm{Y}_{2\mathrm{J}}$.

B. Proof of Bound on Cardinality of $\mathcal{U}$

We follow [22] closely, and use their notation. By [22, Lemma 3] the secrecy capacity region is given by

$$
\begin{aligned}
\{(R_c, R_0, R_1, R_2) \geq 0 : \forall \underline{\lambda} \geq 0, \\
\underline{\lambda}^t (R_c, R_0 + R_1, R_0 + R_2)^t \leq G(\underline{\lambda})\},
\end{aligned}
$$

where $\underline{\lambda} \in \mathbb{R}^3$, and $G(\underline{\lambda})$ is given by

$$
\sup_{\mathrm{U}} \underline{\lambda}^t (I(\mathrm{X}; \mathrm{Y}_1|\mathrm{U}) - I(\mathrm{X}; \mathrm{Y}_2|\mathrm{U}), I(\mathrm{U}; \mathrm{Y}_1), I(\mathrm{U}; \mathrm{Y}_2))^t,
$$

and the supremum is taken over all R.V. U s.t. $P_{\mathrm{U}\mathrm{X}\mathrm{Y}_1\mathrm{Y}_2} = P_\mathrm{U} P_{\mathrm{X}|\mathrm{U}} P_{\mathrm{Y}_1\mathrm{Y}_2|\mathrm{X}}$. Now let $\mathcal{P}$ be the set of probability distributions on $\mathcal{X}$, and let $P_\mathrm{X} \in \mathcal{P}$. We define the following $|\mathcal{X}|$ functions on $\mathcal{P}$:

$$
\begin{aligned}
f_j(P_\mathrm{X}) &= P_\mathrm{X}(j), \quad j = 1, 2, \ldots, |\mathcal{X}| - 1, \\
f_{|\mathcal{X}|}(P_\mathrm{X}) &= \lambda_1 (I_{P_\mathrm{X}}(\mathrm{X}; \mathrm{Y}_1) - I_{P_\mathrm{X}}(\mathrm{X}; \mathrm{Y}_1)) - \lambda_2 H_{P_\mathrm{X}}(\mathrm{Y}_1) \\
&\quad - \lambda_3 H_{P_\mathrm{X}}(\mathrm{Y}_2),
\end{aligned}
$$

where $I_{P_\mathrm{X}}(\mathrm{X}; \mathrm{Y}_i)$ and $H_{P_\mathrm{X}}(\mathrm{Y}_i)$ are the corresponding mutual informations and entropies when the distribution of X is P_X. Each probability distribution P_U defines a measure $\mu(\mathrm{d}P_\mathrm{X})$ on $\mathcal{P}$. Let P_X^* be the probability distribution that achieves $G(\underline{\lambda})$, and let μ^* be the measure that achieves P_X^*. Note that

$$
\begin{aligned}
\int f_j(P_\mathrm{X}) \mu^*(\mathrm{d}P_\mathrm{X}) &= P_\mathrm{X}^*(j), \quad j = 1, 2, \ldots, |\mathcal{X}| - 1, \\
\int f_{|\mathcal{X}|}(P_\mathrm{X}) \mu^*(\mathrm{d}P_\mathrm{X}) &= \lambda_1 (I_{P_\mathrm{X}^*}(\mathrm{X}; \mathrm{Y}_1|\mathrm{U}) - I_{P_\mathrm{X}^*}(\mathrm{X}; \mathrm{Y}_1|\mathrm{U})) \\
&\quad - \lambda_2 H_{P_\mathrm{X}^*}(\mathrm{Y}_1|\mathrm{U}) - \lambda_3 H_{P_\mathrm{X}^*}(\mathrm{Y}_2|\mathrm{U}).
\end{aligned}
$$

From $f_1(P_\mathrm{X}^*), \ldots, f_{|\mathcal{X}|-1}(P_\mathrm{X}^*)$ we can calculate $H_{P_\mathrm{X}^*}(\mathrm{Y}_1)$ and $H_{P_\mathrm{X}^*}(\mathrm{Y}_2)$ and form

$$
\int f_{|\mathcal{X}|}(P_\mathrm{X}) \mu^*(\mathrm{d}P_\mathrm{X}) + \lambda_2 H_{P_\mathrm{X}^*}(\mathrm{Y}_1) + \lambda_3 H_{P_\mathrm{X}^*}(\mathrm{Y}_2) = G(\underline{\lambda}).
$$

Now it follows from [22, Lemma 2] that it is sufficient to consider R.V. U with $|\mathcal{U}| \leq |\mathcal{X}|$.

REFERENCES

[1] B. Rankov and A. Wittneben, "Spectral Efficient Protocols for Half-Duplex Fading Relay Channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[2] P. Larsson, N. Johansson, and K.-E. Sunell, "Coded Bi-directional Relaying," in *Proc. 5th Scandinavian Workshop on Ad Hoc Networks*, Stockholm, Sweden, May 2005, pp. 851–855.

[3] T. J. Oechtering, C. Schnurr, I. Bjelaković, and H. Boche, "Broadcast Capacity Region of Two-Phase Bidirectional Relaying," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 454–458, Jan. 2008.

[4] S. J. Kim, P. Mitran, and V. Tarokh, "Performance Bounds for Bidirectional Coded Cooperation Protocols," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5235–5241, Nov. 2008.

[5] G. Kramer and S. Shamai (Shitz), "Capacity for Classes of Broadcast Channels with Receiver Side Information," in *Proc. IEEE Inf. Theory Workshop*, Tahoe City, CA, USA, Sep. 2007, pp. 313–318.

[6] R. Wyrembelski and H. Boche, "Bidirectional broadcast channels with common and confidential messages," in *Information Theory Workshop (ITW), 2011 IEEE*, oct. 2011, pp. 713 –717.

[7] C. Schnurr, T. J. Oechtering, and S. Stańczak, "On Coding for the Broadcast Phase in the Two-Way Relay Channel," in *Proc. Conf. Inf. Sciences and Systems*, Baltimore, MD, USA, Mar. 2007, pp. 271–276.

[8] O. Iscan, I. Latif, and C. Hausl, "Network Coded Multi-way Relaying with Iterative Decoding," in *Proc. IEEE Int.l Symp. Personal, Indoor and Mobile Radio Commun.*, Istanbul, Turkey, Sep. 2010, pp. 482–487.

[9] E. Arikan, "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 –3073, Jul. 2009.

[10] E. Sasoglu, E. Telatar, and E. Arikan, "Polarization for Arbitrary Discrete Memoryless Channels," in *Proc. IEEE Inf. Theory Workshop*, Taormina, Italy, Oct. 2009, pp. 144 – 148.

[11] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, EPFL, 2009.

[12] S. B. Korada and R. Urbanke, "Polar Codes for Slepian-Wolf, Wyner-Ziv, and Gelfand-Pinsker," in *Proc. IEEE Inf. Theory Workshop*, Cairo, Egypt, Jan. 2010, pp. 1–5.

[13] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested Polar Codes for Wiretap and Relay Channels," *IEEE Commun. Lett.*, vol. 14, no. 8, pp. 752 –754, Aug. 2010.

[14] H. Mahdavifar and A. Vardy, "Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 913 – 917.

[15] E. Hof and S. Shamai, "Secrecy-achieving polar-coding," in *Information Theory Workshop (ITW), 2010 IEEE*, 30 2010-sept. 3 2010, pp. 1 –5.

[16] O. Koyluoglu and H. El Gamal, "Polar coding for secure transmission and key agreement," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, sept. 2010, pp. 2698 –2703.

[17] R. Blasco-Serrano, R. Thobaben, V. Rathi, and M. Skoglund, "Polar Codes for Compress-and-Forward in Binary Relay Channels," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, USA, Nov. 2010, pp. 1743 – 1747.

[18] E. Arikan and E. Telatar, "On the rate of channel polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jul. 2009, pp. 1493 –1495.

[19] A. D. Wyner, "The wire-tap channel," *Bell. Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[20] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley and Sons, 1991.

[21] H. D. Ly, T. Liu, and Y. Liang, "Multiple-Input Multiple-Output Gaussian Broadcast Channels With Common and Confidential Messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5477–5487, Nov. 2010.

[22] M. Salehi, "Cardinality bounds on auxiliary variables in multiple-user theory via the method of Ahlswede and Körner," Dept. Stat., Stanford Univ., Stanford, CA, Tech. Rep. 33, 1978.