Polar Coding for Bidirectional Broadcast Channels with Common and Confidential Messages

Mattias Andersson, *Student Member, IEEE,* Rafael F. Schaefer, *Member, IEEE,* Tobias J. Oechtering, *Senior Member, IEEE,* Mikael Skoglund, *Senior Member, IEEE*

***Abstract*—The integration of multiple services such as the transmission of private, common, and confidential messages at the physical layer is becoming important for future wireless networks in order to increase spectral efficiency. In this paper, bidirectional relay networks are considered, in which a relay node establishes bidirectional communication between two other nodes using a decode-and-forward protocol. In the broadcast phase, the relay transmits additional common and confidential messages, which then requires the study of the *bidirectional broadcast channel (BBC) with common and confidential messages*. This channel generalizes the broadcast channel with receiver side information considered by Kramer and Shamai. Low complexity polar codes are constructed that achieve the capacity region of both the degraded symmetric BBC, and the BBC with common and confidential messages. The use of polar codes allows an intuitive interpretation of how to incorporate receiver side information and secrecy constraints as different sets of frozen bits at the different receivers for an optimal code design. In order to show that the constructed codes achieve capacity, a tighter bound on the cardinality of an auxiliary random variable used in the converse is found using a method by Salehi.**

***Index Terms*—Polar codes, bidirectional broadcast channel, bidirectional relaying, confidential message, physical layer security.**

I. Introduction

RECENT developments such as multiuser MIMO, cooperative multi-point transmission, or relaying have significantly increased the performance of wireless networks. One additional research area that is gaining more importance is the efficient physical layer implementation of multiple services such as the simultaneous transmission of private, common, and confidential messages. For example, in cellular systems, operators establish not only (bidirectional) voice communication, but also offer further services that are either multicast or subject to certain secrecy constraints. Nowadays this is realized by allocating different services on different logical channels and by applying secrecy techniques on higher levels. In general, this is quite inefficient and there is a trend to merge different services directly on the physical layer to increase spectral efficiency.

Manuscript received: September 14, 2012; revised: March 1,2013. Part of this work has appeared as a conference paper in [1]. This work has been supported in part by the Swedish Research Council.

M. Andersson, T. J. Oechtering, and M. Skoglund are with the School of Electrical Engineering and the ACCESS Linnaeus Centre, KTH Royal Institute of Technology, Stockholm, Sweden (e-mail: {amattias, oech}@kth.se, skoglund@ee.kth.se).

R. F. Schaefer is with the Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Germany (e-mail: wyrembelski@tum.de).

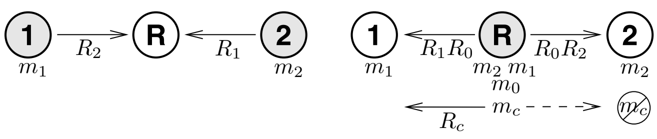
Digital Object Identifier 10.1109/JSAC.2013.130921.

Currently, information is kept secret using cryptographic techniques, which are based on the assumption of insufficient computational capabilities of non-legitimate receivers. With increasing computational power and improved algorithms, these techniques are becoming less and less secure. In this context, the concept of information theoretic security is becoming attractive, since it only uses the properties of the wireless channel in order to establish secrecy. Information theoretic secrecy was initiated by Wyner [2], who introduced the *wiretap channel*, which was later generalized by Csiszár and Körner to the *broadcast channel with confidential messages* [3]. Recently, there has been growing interest in information theoretic secrecy, cf. for instance [4–7] and references therein.

Another key technique to improve the overall performance and coverage for future wireless networks is the concept of *bidirectional relaying*. 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 [8, 9]. Bidirectional relaying applies to three-node networks, in which a half-duplex relay node establishes bidirectional communication between two other nodes using a two-phase decode-and-forward protocol [10–12]. 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 confidential messages to one or both nodes, which have to be kept secret from the other node. This necessitates the analysis of the *bidirectional broadcast channel with common and confidential messages*. The BBC with common and confidential messages is of course not limited to such a two-phase scheme, it is for example a generalization of the broadcast channel with partial side information and degraded message sets considered in [12]. We consider a degraded BBC, where the channel to node 1 is stronger than the channel to node 2. In this setup any message that can be decoded by node 2 can also be decoded by node 1, and thus the only possible receiver of a confidential message is node 1. 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 capacity-equivocation region of the discrete memoryless BBC with common and confidential messages was derived



(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 [13]. The design of practical coding schemes for the BBC was discussed in [14], while [15] addressed the problem of joint network and channel coding in multi-way relay channels.

To pave the way for practical implementation of such concepts, one is interested in finding low complexity coding schemes which achieve capacity. The coding scheme which we consider in this paper are polar codes, which were introduced by Arıkan and were shown to be capacity achieving for a large class of channels in [16, 17]. Polar codes have a natural nested structure [18], which can be used to implement the binning schemes used in multi-user and physical layer security scenarios, and they have been studied for a large range of such setups [19–25]. They generally exhibit weak finite length performance, but recently, polar codes of block length 2048 concatenated with a cyclic redundancy check and decoded with a list decoder were developed, and shown to perform 0.2 dB away from the information theoretical limit over the binary input AWGN channel [26, 27]. This finite block length performance, together with their nested structure and low complexity makes them interesting candidates for practical implementation.

The contributions of this work are the construction of polar codes for the BBC with common and confidential messages and showing that the constructed codes achieve the whole capacity-equivocation region. In order to design polar codes for the BBC with common and confidential messages, we first design capacity achieving schemes for the standard symmetric BBC without confidential messages. We then use superposition coding together with a polar code for the wiretap channel to achieve the capacity-equivocation region. To show that the scheme is capacity achieving, we tighten the outer bound for the capacity-equivocation region obtained in [13] to the degraded setting. Using the methods in [28], we further develop a new bound on the cardinality of the range of the auxiliary random variable U involved in the coding scheme from [13], to show that it is sufficient to consider binary U. This improved bound can simplify the code construction for multi-user scenarios significantly, and the method is not widely used in the literature.

As noted previously, the BBC with common and confidential messages is a generalization of the broadcast channel with partial side information and degraded message sets considered in [12], so our scheme is also capacity achieving for degraded channels of this type. To our knowledge, this is the first work to construct low complexity coding schemes which utilize receiver side information in the multi-user setting.

This paper is structured as follows. In Section II, we review polar codes for binary input symmetric channels, and for Wyner's wiretap channel [2]. In Section III, we introduce the BBC with common and confidential messages and construct polar coding schemes for it. In Section IV, we conclude the paper.

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 [16], $F = \left[\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right]$, 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 an index set $\mathcal{I} \subset \{1, \dots, N\}$ and a binary vector u_1^N, let $G_\mathcal{I}$ be the submatrix formed by the rows of G with indices in $\mathcal{I}$, and let $u_\mathcal{I}$ be the corresponding subvector of u_1^N. Given such an index set $\mathcal{A}$, and a binary vector $u_\mathcal{F}$ of length $N - |\mathcal{A}|$ we define the polar code $\mathcal{C}(N, \mathcal{A}, u_\mathcal{F})$ of length N as follows. We call $\mathcal{A}^\mathcal{C} = \mathcal{F}$ the frozen set, and the (fixed) bits $u_\mathcal{F}$ frozen bits. The codewords of $\mathcal{C}(N, \mathcal{A}, u_\mathcal{F})$ are given by

$$x^N = u_\mathcal{A} G_\mathcal{A} \oplus u_\mathcal{F} G_\mathcal{F},$$

and the rate is given by $|\mathcal{A}|/N$.

Polar codes can be decoded using the successive cancellation (SC) decoding rule defined by

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

where the bits u_i are decoded successively from 1 to N. It was shown in [16] that the block error probability when using SC decoding can be bounded from above by $\sum_{i\in\mathcal{A}} Z_N^{(i)}$, where $Z_N^{(i)}$ is the Bhattacharyya parameter for the channel $W_N^{(i)}$. Further, it was shown in [29] 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 given by $I(\mathrm{X};\mathrm{Y})$ when the input distribution P_X is uniform. This is called the symmetric capacity of W and is equal to the Shannon capacity for symmetric channels. Using (3), we see that if we let the good channels be given by

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

the block error probability P_e using SC decoding is bounded from above by

$$P_e \leq 2^{-N^\beta}, \tag{5}$$

and the rate of $\mathcal{C}(N, \mathcal{G}_N, u_{\mathcal{F}})$ approaches $I(W)$ as N grows.

Arıkan further showed that the encoder and decoder can be implemented with complexity $O(N \log N)$.

We define the nested polar code $\mathcal{C}(N, \mathcal{A}, \mathcal{B}, u_{\mathcal{F}})$ of length N where $\mathcal{B} \subset \mathcal{A}$ as follows. The codewords of $\mathcal{C}(N, \mathcal{A}, \mathcal{B}, u_{\mathcal{F}})$ are the same as the codewords for $\mathcal{C}(N, \mathcal{A}, u_{\mathcal{F}})$. The nested structure is defined by partitioning $\mathcal{C}(N, \mathcal{A}, u_{\mathcal{F}})$ as cosets of $\mathcal{C}(N, \mathcal{B}, u_{\mathcal{F}_\mathcal{B}})$, where the entries of $u_{\mathcal{F}_\mathcal{B}}$ are zero if they correspond to an index in $\mathcal{B} \setminus \mathcal{A}$, and given by the corresponding entry in $u_{\mathcal{F}}$ otherwise. Thus the codewords in $\mathcal{C}(N, \mathcal{A}, \mathcal{B}, u_{\mathcal{F}})$ 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 will use the following lemma:

Lemma 1 ([19, 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 [2] was shown in [21–24]:

Theorem 1 ([21–24])**.** *Let W be a degraded symmetric wiretap channel and denote the marginal channels to the main user and the wiretapper by W_1 and W_2 respectively. Let $\mathcal{G}_{1,N}$ and $\mathcal{G}_{2,N}$ be the corresponding sets given by (4). If W_2 is degraded with respect to W_1, the nested polar code $\mathcal{C}(N, \mathcal{G}_{1,N}, \mathcal{G}_{2,N}, u_{\mathcal{F}})$ achieves the capacity-equivocation region of the wiretap channel.*

The secrecy capacity of the wiretap channel is achieved by transmitting the message m over the channels in $\mathcal{G}_{1,N} \setminus \mathcal{G}_{2,N}$, while sending random bits over the channels in $\mathcal{G}_{2,N}$.

In the next section we introduce the BBC with common and confidential messages, and construct polar coding schemes for the BBC both with and without common and confidential messages.

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. For a rate-equivocation tuple $(R_c, R_e, R_0, R_1, R_2) \in \mathbb{R}_+^5$ to be achievable we require, in addition to the error probabilities of decoding the legitimate messages going to zero, the equivocation rate to fulfill

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

for some (small) $\delta > 0$.

The case where the equivocation rate R_e equals the confidential rate R_c is called *perfect secrecy*. 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 [13] for discrete memoryless channels. Its capacity-equivocation region is restated in the following theorem:

Theorem 2 ([13])**.** *The capacity-equivocation region of the BBC with common and confidential messages is the set of rate-equivocation tuples $(R_c, R_e, R_0, R_1, R_2) \in \mathbb{R}_+^5$ that satisfy*

$$\begin{aligned} 0 \leq R_e &\leq R_c \\ R_e &\leq I(\mathrm{V}; \mathrm{Y}_1 | \mathrm{U}) - I(\mathrm{V}; \mathrm{Y}_2 | \mathrm{U}) \\ R_c + R_0 + R_k &\leq I(\mathrm{V}; \mathrm{Y}_1 | \mathrm{U}) + I(\mathrm{U}; \mathrm{Y}_k), \quad k = 1, 2 \\ R_0 + R_k &\leq I(\mathrm{U}; \mathrm{Y}_k), \quad k = 1, 2 \end{aligned}$$

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 capacity-equivocation region of the degraded BBC with common and confidential messages is the* *set of rate tuples* $(R_c, R_e, R_0, R_1, R_2) \in \mathbb{R}_+^5$ *that satisfy*

$$
\begin{aligned}
0 \leq R_e &\leq R_c \\
R_e &\leq I(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - I(\mathrm{X};\mathrm{Y}_2|\mathrm{U}) \\
R_c + R_0 + R_k &\leq I(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + I(\mathrm{U};\mathrm{Y}_k), \quad k = 1,2 \\
R_0 + R_k &\leq I(\mathrm{U};\mathrm{Y}_k), \quad k = 1,2
\end{aligned}
$$

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

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

Proof: The achievability follows immediately from the non-degraded case in Theorem 2, cf. also [13]. The converse and the bound on the cardinality of $\mathcal{U}$ is devoted to the appendix. ∎

By considering the case of perfect secrecy, i.e. $R_e = R_c$, we obtain the secrecy capacity region.

Corollary 2. *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*

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

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

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

Polar codes that achieve this region were designed in [1].

Remark 1. *The improved bound on the cardinality of* $\mathcal{U}$ *is particularly helpful when designing coding schemes. In the following subsections we will see that it allows us to consider binary input coding schemes when designing codes for a binary input channel, where a looser bound might have required non-binary schemes.*

Remark 2. *Note that by letting* $R_e = 0$ *in Corollary 3 we drop the secrecy constraint on the message* m_c. *In this case the BBC with common and confidential messages specializes to the broadcast channel with partial receiver side information and degraded message sets considered in [12]. Thus the BBC with common and confidential messages is a generalization of the broadcast channel with partial receiver side information and degraded message sets, and any scheme that is capacity achieving for the first is also capacity achieving for the second.*

In the next subsections we design polar coding schemes for the BBC, and then for the BBC with common and confidential messages.

A. Polar Codes for the BBC

First consider a binary input BBC W with marginal channels W_1 and W_2 with no common and confidential messages. The capacity region is given by

$$R_1 \leq C_1 \tag{6}$$

$$R_2 \leq C_2 \tag{7}$$

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

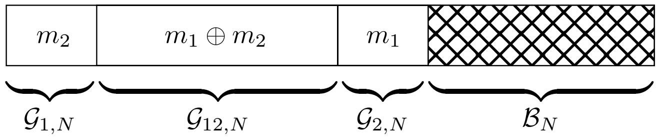


Fig. 2. Frozen sets and encoding for the BBC. A Part of m_1 (m_2) is transmitted over $\mathcal{G}_{1,N}$ ($\mathcal{G}_2, N$), and the remaining part of m_1 and m_2 are transmitted as $m_1 \oplus m_2$ over $\mathcal{G}_{12,N}$.

In the following theorem we present a polar coding scheme for this channel. Note how the values of the frozen bits for the two users correspond to the side information available.

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). The encoders and decoders can be implemented with complexity* $O(N \log N)$.

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:

$$\mathcal{G}_{1,N} = \{i : Z_{1,N}^{(i)} < 2^{-N^\beta} \text{ and } Z_{2,N}^{(i)} \geq 2^{-N^\beta}\}, \tag{8}$$

$$\mathcal{G}_{2,N} = \{i : Z_{1,N}^{(i)} \geq 2^{-N^\beta} \text{ and } Z_{2,N}^{(i)} < 2^{-N^\beta}\}, \tag{9}$$

$$\mathcal{G}_{12,N} = \{i : Z_{1,N}^{(i)} < 2^{-N^\beta} \text{ and } Z_{2,N}^{(i)} < 2^{-N^\beta}\}, \tag{10}$$

$$\mathcal{B}_N = \{i : Z_{1,N}^{(i)} \geq 2^{-N^\beta} \text{ and } Z_{2,N}^{(i)} \geq 2^{-N^\beta}\}, \tag{11}$$

where $\mathcal{G}_{1,N}$ 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 $\mathcal{C}(N, \mathcal{G}_{1,N} \cup \mathcal{G}_{2,N} \cup \mathcal{G}_{12,N}, u_{\mathcal{F}})$ 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. The frozen sets and the encoding is shown in Figure 2. Since node 1 knows m_1 it 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 it subtracts the bits of m_1 that appear in bits in $\mathcal{G}_{12,N}$. Thus the rate for node 1 becomes

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

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

$$R_{2,N} = \frac{|\mathcal{G}_{2,N}| + |\mathcal{G}_{12,N}|}{N}. \tag{13}$$

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. Finally, the complexity of the encoder and the decoder is the same as for the point-to-point channel. ∎

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 3. *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,*

$$R_0 + R_1 \leq C_1 \tag{14}$$
$$R_0 + R_2 \leq C_2. \tag{15}$$

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),

$$\lim_{N\to\infty} R_{0,N} + R_{1,N} = \lim_{N\to\infty} \frac{|\mathcal{G}_{1,N}| + |\mathcal{G}_{12,N}|}{N} = C_1, \tag{16}$$

and

$$\lim_{N\to\infty} R_{0,N} + R_{2,N} = \lim_{N\to\infty} \frac{|\mathcal{G}_{12,N}|}{N} = C_2, \tag{17}$$

which completes the proof. ∎

Remark 3. *Note that the condition that W_2 is degraded with respect to W_1 ensures that $\mathcal{G}_{2,N} = \emptyset$. If W_1 and W_2 are not ordered by degradation, the highest rate for the common message that can be achieved is given by $\liminf_{N\to\infty} |\mathcal{G}_{12,N}|/N$. This quantity is called the compound capacity $C_{P,SC}(W_1, W_2)$ of W_1 and W_2 using polar codes and SC decoding. In general, $C_{P,SC}(W_1, W_2)$ is lower than the minimum of the capacities of W_1 and W_2. Methods to calculate upper and lower bounds on $C_{P,SC}(W_1, W_2)$ were developed in [30].*

In the next subsection we show how to design polar codes for a degraded BBC with common and confidential messages.

B. Polar Codes for the BBC with Confidential Messages

We consider the case where W_1 and W_2 are binary symmetric channels (BSC) with transition probabilities p_1 and p_2, with $p_2 > p_1$.[1] We call such a channel a binary symmetric BBC. Using the upper bound on $|\mathcal{U}|$ from Corollary 1 and the same arguments as in [31, 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 capacity-equivocation region in Corollary 1 becomes

$$\begin{aligned} 0 \leq R_e &\leq R_c \\ R_e &\leq h_2(\alpha \star p_1) - h_2(p_1) - h_2(\alpha \star p_2) + h_2(p_2) \\ R_c + R_0 + R_k &\leq h_2(\alpha \star p_1) - h_2(p_1) + 1 - h_2(\alpha \star p_k), \\ &\quad k = 1, 2 \\ R_0 + R_k &\leq 1 - h_2(\alpha \star p_k), \quad k = 1, 2, \end{aligned}$$

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*

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

for $\mathrm{X}^N_{BBC} \in \mathcal{C}_{BBC}$ and $\mathrm{X}^N_{WT} \in \mathcal{C}_{WT}$ achieves the capacity-equivocation region for the binary symmetric BBC with common and confidential messages. The encoders and decoders can be implemented with complexity $O(N \log N)$.

Proof: Fix $0 < \alpha < 1/2$. We first design $\mathcal{C}_{BBC}$ for a binary symmetric BBC with a common message with transition probabilities $\alpha \star p_1$ and $\alpha \star p_2$. If X^N_{WT} is statistically indistinguishable from an i.i.d. Ber(α) vector, then, by Corollary 3, $\mathcal{C}_{BBC}$ achieves 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^N_{BBC} 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 [16]. Also note that since X^N_{BBC} and X^N_{WT} are independent, X^N_{BBC} provides no information about X^N_{WT}. Thus, assuming that node 2 decodes X^N_{BBC} does not increase the equivocation of m_c at node 2.

Let $\mathcal{C}_{WT}$ be a polar code with input weight $\alpha' \in \mathbb{Q}$ designed for a binary symmetric wiretap channel with transition probabilities p_1 and p_2 using Theorem 1. To design a polar code with rational input weight α', we augment the binary channel with a virtual q-ary input and then design a polar code for this augmented channel. This technique was introduced by Gallager [32], and used for polar codes in [17, 19]. Since any $\alpha \in \mathbb{R}$ can be approximated arbitrarily well by an $\alpha' \in \mathbb{Q}$, such a construction achieves all rate-equivocation pairs satisfying

$$\begin{aligned} R_c &\leq h_2(\alpha \star p_1) - h_2(p_1), \\ R_e &\leq h_2(\alpha \star p_1) - h_2(p_1) - h_2(\alpha \star p_2) + h_2(p_2). \end{aligned}$$

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 respectively, when using $u_{\mathcal{F}}$ as the frozen bits for $\mathcal{C}_{WT}$. Choosing $u_{\mathcal{F}}$ uniformly at random we can make the average error probability

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

arbitrarily small by choosing N large enough, since the codewords of $\mathcal{C}_{WT}$ are i.i.d. Ber(α) when averaged 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, and using this $u_{\mathcal{F}}$ as the frozen bits for $\mathcal{C}_{WT}$ makes the overall error probability small.

Finally, the complexity of the encoders and the decoders are the same as in the point-to-point setting. ∎

Remark 4. *Consider a BBC with non-degraded marginal channels. As in Remark 3, R_0 is bounded from above by $C_{P,SC}(W_1, W_2)$, but more importantly, the analysis of the equivocation rate R_e becomes difficult. It was conjectured in [23] that it is possible to achieve the secrecy capacity of non-degraded wiretap channels using polar codes. A proof of this conjecture would also apply to our scheme.*

[1] This apparent simplification is made to make the exposition clearer. Our results generalize to arbitrary q-ary input BBCs with degraded marginal channels using results from [17].

IV. Conclusions

We have given a polar coding scheme with complexity $O(N \log N)$ for the degraded symmetric bidirectional broadcast channel with common and confidential messages. The different side information available at the different nodes is used in an intuitive way as different values of the frozen bits in the constituent polar codes. In order to show that the coding scheme is optimal, we have specialized the outer bound from [13] to the degraded setting. This outer bound includes an auxiliary random variable $\mathrm{U} \in \mathcal{U}$, and using methods from [28] we have shown that $\mathcal{U}$ can be chosen to have cardinality equal to the cardinality of the input alphabet $\mathcal{X}$. This allowed us to completely characterize the capacity-equivocation region and show that polar codes can achieve the whole region.

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 exist random variables $\mathrm{U} - \mathrm{X} - \mathrm{Y}_1 - \mathrm{Y}_2$ such that

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

We do this by using techniques similar to [33] 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 [13]. 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

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

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

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

where (a) follows from the degradedness $\mathrm{X}_i - \mathrm{Y}_{1i} - \mathrm{Y}_{2i}$.

We bound $H(\mathrm{M}_c)$:

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

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)\\
&\quad + 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)\\
&\quad + 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)\\
&\quad + 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})\\
&\quad + N\epsilon_1^{(N)} + N\epsilon_2^{(N)}\\
&= \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})\\
&\qquad + 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) +\\
&\qquad - 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)\\
&\qquad + 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) +\\
&\qquad - 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)\\
&\qquad + N\epsilon_1^{(N)} + N\epsilon_2^{(N)}
\end{aligned}$$

$$\stackrel{(b)}{=} \sum_{i=1}^{N} H(\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}) - H(\mathrm{Y}_{1i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}\mathrm{X}_i) + \\ - H(\mathrm{Y}_{2i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}) + H(\mathrm{Y}_{2i}|\mathrm{Y}_1^{i-1}\mathrm{M}_{012}\mathrm{X}_i) \\ + N\epsilon_1^{(N)} + N\epsilon_2^{(N)} \\ = \sum_{i=1}^{N} I(\mathrm{X}_i;\mathrm{Y}_{1i}|\mathrm{U}_i) - I(\mathrm{X}_i;\mathrm{Y}_{2i}|\mathrm{U}_i) + N\epsilon_1^{(N)} + N\epsilon_2^{(N)},$$

where (b) follows from the Markov chain $(\mathrm{Y}_1^{i-1}, \mathrm{Y}_2^{i-1}, \mathrm{M}_{012}) - \mathrm{X}_i - \mathrm{Y}_{2i}$, which is due to the channel being memoryless.

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

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

We follow [28] closely, and use their notation. By [28, Lemma 3] the capacity-equivocation region is given by

$$\{(R_e, R_c, R_0, R_1, R_2) \in \mathbb{R}_+^5 : \forall(\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5) \in \mathbb{R}_+^5, \\ \lambda_1 R_e + \lambda_2(R_c + R_0 + R_1) + \lambda_3(R_c + R_0 + R_2) + \\ \lambda_4(R_0 + R_1) + \lambda_5(R_0 + R_2) \le G(\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)\},$$

where $G(\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)$ is given by the supremum of

$$\lambda_1(I(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - I(\mathrm{X};\mathrm{Y}_2|\mathrm{U})) + \lambda_2(I(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + I(\mathrm{U};\mathrm{Y}_1)) + \\ \lambda_3(I(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) + I(\mathrm{U};\mathrm{Y}_2)) + \lambda_4 I(\mathrm{U};\mathrm{Y}_1) + \lambda_5 I(\mathrm{U};\mathrm{Y}_2),$$

taken over all R.V. U s.t. $P_{\mathrm{UXY_1Y_2}} = P_{\mathrm{U}} P_{\mathrm{X|U}} P_{\mathrm{Y_1Y_2|X}}$. 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 $\mathcal{P}$:

$$f_j(P_{\mathrm{X}}) = P_{\mathrm{X}}(j), \quad j = 1,2,\ldots,|\mathcal{X}|-1, \\ f_{|\mathcal{X}|}(P_{\mathrm{X}}) = \lambda_1(I_{P_{\mathrm{X}}}(\mathrm{X};\mathrm{Y}_1) - I_{P_{\mathrm{X}}}(\mathrm{X};\mathrm{Y}_2)) \\ + \lambda_2(I_{P_{\mathrm{X}}}(\mathrm{X};\mathrm{Y}_1) - H_{P_{\mathrm{X}}}(\mathrm{Y}_1)) \\ + \lambda_3(I_{P_{\mathrm{X}}}(\mathrm{X};\mathrm{Y}_1) - H_{P_{\mathrm{X}}}(\mathrm{Y}_2)) \\ - \lambda_4 H_{P_{\mathrm{X}}}(\mathrm{Y}_1) - \lambda_5 H_{P_{\mathrm{X}}}(\mathrm{Y}_2),$$

where $I_{P_{\mathrm{X}}}(\mathrm{X};\mathrm{Y}_i)$ and $H_{P_{\mathrm{X}}}(\mathrm{Y}_i)$ are the corresponding mutual information 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 $G(\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5)$, and let μ^* be the corresponding measure. Note that

$$\int f_j(P_{\mathrm{X}})\mu^*(\mathrm{d}P_{\mathrm{X}}) = P_{\mathrm{X}}^*(j), \quad j = 1,2,\ldots,|\mathcal{X}|-1, \\ \int f_{|\mathcal{X}|}(P_{\mathrm{X}})\mu^*(\mathrm{d}P_{\mathrm{X}}) = \lambda_1(I_{P_{\mathrm{X}}^*}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - I_{P_{\mathrm{X}}^*}(\mathrm{X};\mathrm{Y}_2|\mathrm{U})) \\ + \lambda_2(I_{P_{\mathrm{X}}^*}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - H_{P_{\mathrm{X}}^*}(\mathrm{Y}_1|\mathrm{U})) \\ + \lambda_3(I_{P_{\mathrm{X}}^*}(\mathrm{X};\mathrm{Y}_1|\mathrm{U}) - H_{P_{\mathrm{X}}^*}(\mathrm{Y}_2|\mathrm{U})) \\ - \lambda_4 H_{P_{\mathrm{X}}^*}(\mathrm{Y}_1|\mathrm{U}) - \lambda_5 H_{P_{\mathrm{X}}^*}(\mathrm{Y}_2|\mathrm{U}).$$

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

$$\int f_{|\mathcal{X}|}(P_{\mathrm{X}})\mu^*(\mathrm{d}P_{\mathrm{X}}) + (\lambda_2+\lambda_4)H_{P_{\mathrm{X}}^*}(\mathrm{Y}_1) + \\ (\lambda_3+\lambda_5)H_{P_{\mathrm{X}}^*}(\mathrm{Y}_2) = G(\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5).$$

Now it follows from [28, Lemma 2] that it is sufficient to consider R.V. U with $|\mathcal{U}| \le |\mathcal{X}|$.

REFERENCES

[1] M. Andersson, R. F. Wyrembelski, T. J. Oechtering, and M. Skoglund, “Polar codes for bidirectional broadcast channels with common and confidential messages,” in *Wireless Communication Systems (ISWCS), 2012 International Symposium on*, Aug. 2012, pp. 1014 –1018.

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

[3] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339 – 348, May 1978.

[4] Y. Liang, H. V. Poor, and S. Shamai (Shitz), “Information Theoretic Security,” *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4-5, pp. 355–580, 2009.

[5] E. Jorswieck, A. Wolf, and S. Gerbracht, “Secrecy on the physical layer in wireless networks,” *Trends in Telecommunications Technologies*, pp. 413–435, 2010.

[6] R. Liu and W. Trappe, Eds., *Securing Wireless Communications at the Physical Layer*. Springer, 2010.

[7] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] G. Kramer and S. Shamai, “Capacity for classes of broadcast channels with receiver side information,” in *Information Theory Workshop (ITW), IEEE*, Sep. 2007, pp. 313 –318.

[13] R. F. Wyrembelski and H. Boche, “Bidirectional broadcast channels with common and confidential messages,” in *Information Theory Workshop (ITW), IEEE*, Oct. 2011, pp. 713 –717.

[14] C. Schnurr, T. J. Oechtering, and S. Stańczak, “On coding for the broadcast phase in the two-way relay channel,” in *Information Sciences and Systems (CISS). 41st Annual Conference on*, Mar. 2007, pp. 271–276.

[15] O. Iscan, I. Latif, and C. Hausl, “Network coded multi-way relaying with iterative decoding,” in *Personal Indoor and Mobile Radio Communications (PIMRC), IEEE 21st International Symposium on*, Sep. 2010, pp. 482 –487.

[16] E. Arıkan, “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.

[17] E. Sasoglu, E. Telatar, and E. Arıkan, “Polarization for arbitrary discrete memoryless channels,” in *Information Theory Workshop (ITW), IEEE*, Oct. 2009, pp. 144 –148.

[18] R. Zamir, S. Shamai, and U. Erez, “Nested linear/lattice codes for structured multiterminal binning,” *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun 2002.

[19] S. B. Korada, “Polar codes for channel and source coding,” Ph.D. dissertation, EPFL, 2009.

[20] S. B. Korada and R. Urbanke, “Polar Codes for Slepian-Wolf, Wyner-Ziv, and Gelfand-Pinsker,” in *Information Theory Workshop (ITW), IEEE*, Jan. 2010, pp. 1–5.

[21] 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.

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

[23] E. Hof and S. Shamai, “Secrecy-achieving polar-coding,” in *Information Theory Workshop (ITW), IEEE*, Sep. 2010, pp. 1 –5.

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

[25] R. Blasco-Serrano, R. Thobaben, M. Andersson, V. Rathi, and M. Skoglund, “Polar codes for cooperative relaying,” *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3263 –3273, Nov. 2012.

[26] I. Tal and A. Vardy, “List Decoding of Polar Codes,” *ArXiv e-prints*, May 2012.

[27] B. Li, H. Shen, and D. Tse, "An adaptive successive cancellation list decoder for polar codes with cyclic redundancy check," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 2044–2047, Dec. 2012.
[28] 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.
[29] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jul. 2009, pp. 1493 –1495.
[30] S. Hassani, S. Korada, and R. Urbanke, "The compound capacity of polar codes," in *Communication, Control, and Computing. 47th Annual Allerton Conference on*, Oct. 2009, pp. 16 –21.
[31] T. Cover and J. Thomas, *Elements of Information Theory*. New York, NY, USA: John Wiley & Sons, Inc., 1991.
[32] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: John Wiley & Sons, Inc., 1968.
[33] 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.

Mattias Andersson (S'07) received the M.Sc. degree in Engineering Physics from the KTH Royal Institute of Technology, Stockholm, Sweden in 2007. In 2007 he joined the Communication Theory laboratory of the School of Electrical Engineering at the KTH Royal Institute of Technology, Stockholm, Sweden. His research interests include digital communications and information theory with focus on code design for physical layer security.

Rafael F. Schaefer (formerly Wyrembelski) (S'08-M'12) received the Dipl.-Ing. degree in Electrical Engineering and Computer Science in 2007 from the Technische Universität Berlin, Germany, and the Dr.-Ing. degree in Electrical Engineering in 2012 from the Technische Universität München. Between 2007 and 2010 he worked as a research and teaching assistant at the Heinrich-Hertz-Lehrstuhl für Mobilkommunikation at the Technische Universität Berlin, Germany. Since November 2010 he has been with the Lehrstuhl für Theoretische Informationstechnik at the Technische Universität München, Germany, where he is currently working as a Post-Doctoral researcher.

Tobias J. Oechtering (S'01-M'08-SM'12) received his Dipl.-Ing. degree in Electrical Engineering and Information Technology in 2002 from RWTH Aachen University, Germany, his Dr.-Ing. degree in Electrical Engineering in 2007 from the Technische Universität Berlin, Germany, and his Docent degree in Communication Theory in 2012 from KTH Royal Institute of Technology. Between 2002 and 2008 he has been with Technische Universität Berlin and Fraunhofer German-Sino Lab for Mobile Communications, Berlin, Germany. In 2008 he joined the Communication Theory Lab at the KTH Royal Institute of Technology, Stockholm, Sweden and has been an Associate Professor since May 2013. Presently, he is serving as an editor for IEEE Communications Letters. Dr. Oechtering received the "Förderpreis 2009" from the Vodafone Foundation. His research interests include information theory, distributed detection, and signal processing for wireless communication, as well as communication for networked control.

Mikael Skoglund (S'93-M'97-SM'04) received the Ph.D. degree in 1997 from Chalmers University of Technology, Sweden. In 1997, he joined the KTH Royal Institute of Technology, Stockholm, Sweden, where he was appointed to the Chair in Communication Theory in 2003. At KTH, he heads the Communication Theory Lab and he is the Assistant Dean for Electrical Engineering.

Dr. Skoglund's research interests are in the theoretical aspects of wireless communications. He has worked on problems in source-channel coding, coding and transmission for wireless communications, Shannon theory and statistical signal processing. He has authored and co-authored more than 300 scientific papers in these areas, and he holds six patents.

Dr. Skoglund has served on numerous technical program committees for IEEE conferences. During 2003–08 he was an associate editor with the IEEE Transactions on Communications and he is presently on the editorial board for IEEE Transactions on Information Theory.