ISIT'2013 1569720299

Polar Codes for Broadcast Channels

Naveen Goela[†], Emmanuel Abbe[♯], and Michael Gastpar[§]

***Abstract*—Building on polar code constructions proposed by the authors for deterministic broadcast channels, two theorems are introduced in the present paper for *noisy* two-user broadcast channels. The theorems establish polar code constructions for two important information-theoretic broadcast strategies: (1) Cover's superposition strategy; (2) Marton's construction. One aspect of the polar code constructions is the alignment of polarization indices via constraints placed on the auxiliary and channel-input distributions. The codes achieve capacity-optimal rates for several classes of broadcast channels (e.g., binary-input stochastically degraded channels). Applying Arıkan's original matrix kernel for polarization, it is shown that the average probability of error in decoding two private messages at the broadcast receivers decays as $\mathcal{O}(2^{-n^\beta})$ where $0 < \beta < \frac{1}{2}$ and n is the code length. The encoding and decoding complexities remain $\mathcal{O}(n \log n)$. The error analysis is made possible by defining new polar code ensembles for broadcast channels.**

***Index Terms*—Polar codes, Deterministic broadcast channel, Cover's superposition codes, Marton's construction.**

I. Introduction

Introduced by Cover in 1972, the broadcast problem consists of a single source transmitting m independent private messages to m receivers through a discrete, memoryless, broadcast channel (DM-BC) [1]. The private-message capacity region is known if the channel structure is *deterministic*, *degraded*, *less-noisy*, or *more-capable* (see e.g., the text of [2]). For general classes of DM-BCs, there exist inner bounds such as Marton's inner bound [3] and outer bounds such as the Nair-El-Gamal outer bound [4]. Several codes relying on *random binning*, *superposition*, and *Marton's construction* have been analyzed in the literature [5].

In his celebrated work, Arıkan introduced and proved that polar codes achieve the capacity of binary-input, symmetric, point-to-point channels with low encoding and decoding complexity [6]. A refined rate of polarization was established in [7]. Since then, polarization methods have been applied to several multi-user information theory problems: m-user multiple-access channels [8], [9]; cooperative relaying [10]; wiretap channels [11], [12]; and Gelfand-Pinsker and Wyner-Ziv problems [13]. In addition, polarization methods have been applied for lossless and lossy source compression, Slepian-Wolf distributed and multi-user source coding [14]–[17]. To our knowledge, constructing low-complexity broadcast codes based on polarization of random variables (r.v.'s) is a new area of research which contains several open questions. Initially, our codes rely on polarization theorems for binary r.v.'s. However, the codes may be modified to accommodate larger auxiliary, input, and output alphabets for broadcast channels. In extending to arbitrary alphabet sizes, there exist several papers analyzing generalized constructions [18], and polarization for q-ary sources and channels [16], [19], [20].

A polar code construction for m-user *deterministic* broadcast channels was recently proposed by the authors [21]. The code is a low-complexity implementation of random binning. Following this work, we propose two polar code constructions for *noisy* discrete, memoryless, broadcast channels (DM-BCs). The new codes provide insight into two fundamental strategies for broadcast: (1) Cover's superposition strategy; (2) Marton's strategy. One important contribution is to replace the classical error analysis of random code ensembles with a modern error analysis of polar code ensembles, extending beyond point-to-point channels [6], [22] to multi-user channels.

II. Model

Definition 1 (Discrete, Memoryless, Broadcast Channel): A DM-BC with m broadcast receivers consists of a discrete input alphabet $\mathcal{X}$, discrete output alphabets $\mathcal{Y}_i$ for $i \in [m]$, and a conditional distribution $P_{Y_1 Y_2 \ldots Y_m|X}(y_1, y_2, \ldots, y_m|x)$ where $x \in \mathcal{X}$ and $y_i \in \mathcal{Y}_i$.

Definition 2 (Private Messages): Let $m, n \in \mathbb{Z}_+$ and $R_i \in \mathbb{R}_+$. For a DM-BC with m broadcast receivers, there exist m private messages $\{W_i\}_{i\in[m]}$ such that each message W_i is composed of $\lfloor nR_i \rfloor$ bits where $(W_1, W_2, \ldots, W_m)$ are uniformly distributed over $[2^{\lfloor nR_1 \rfloor}] \times [2^{\lfloor nR_2 \rfloor}] \times \cdots \times [2^{\lfloor nR_m \rfloor}]$.

Definition 3 (Channel Encoding and Decoding): For the DM-BC with private independent messages, let the vector of rates $\vec{R} \triangleq \begin{bmatrix} R_1 & R_2 & \ldots & R_m \end{bmatrix}^T$. An $(\vec{R}, n)$ code for the DM-BC consists of one encoder

$$x^n : [2^{\lfloor nR_1 \rfloor}] \times [2^{\lfloor nR_2 \rfloor}] \times \cdots \times [2^{\lfloor nR_m \rfloor}] \to \mathcal{X}^n,$$

and m decoders specified by $\hat{W}_i : \mathcal{Y}_i^n \to [2^{\lfloor nR_i \rfloor}]$ for $i \in [m]$. Based on received observations $\{Y_i(j)\}_{j\in[n]}$, each decoder outputs a decoded message $\hat{W}_i$.

Definition 4 (Average Probability of Error): The average probability of error $P_e^{(n)}$ for a DM-BC code is defined to be the probability that the decoded message at all receivers is not equal to the transmitted message,

$$P_e^{(n)} = \mathbb{P}\left\{ \bigvee_{i=1}^{m} \hat{W}_i\left(\{Y_i(j)\}_{j\in[n]}\right) \neq W_i \right\}. \tag{1}$$

Definition 5 (Private-Message Capacity Region): If there exists a sequence of $(\vec{R}, n)$ codes with $P_e^{(n)} \to 0$, then the

[†]N. Goela is currently with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, Berkeley, CA 94720-1770 USA (e-mail: ngoela@eecs.berkeley.edu).

[♯]E. Abbe is currently with the School of Engineering and Applied Sciences, Princeton University, Princeton, NJ, 08544 USA (e-mail: eabbe@princeton.edu).

[§]M. C. Gastpar is currently with the School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland (e-mail: michael.gastpar@epfl.ch).

rates $\vec{R} \in \mathbb{R}_+^m$ are achievable. The private-message capacity region is the closure of the set of achievable rates.

Definition 6: A two-user DM-BC $P_{Y_1Y_2|X}(y_1, y_2|x)$ is *stochastically degraded* if its conditional marginal distributions are the same as that of a physically degraded DM-BC, i.e., if there exists a distribution $\tilde{P}_{Y_2|Y_1}(y_2|y_1)$ such that

$$P_{Y_2|X}(y_2|x) = \sum_{y_1 \in \mathcal{Y}_1} P_{Y_1|X}(y_1|x)\tilde{P}_{Y_2|Y_1}(y_2|y_1). \quad (2)$$

If (2) holds for two conditional distributions $P_{Y_1|X}(y_1|x)$ and $P_{Y_2|X}(y_2|x)$ defined over the same input, then the property is denoted as follows: $P_{Y_1|X}(y_1|x) \succ P_{Y_2|X}(y_2|x)$.

Definition 7 (Bhattacharyya Parameter): Let $(T, V) \sim P_{T,V}$ where $T \in \{0, 1\}$ and $V \in \mathcal{V}$ where $\mathcal{V}$ is an arbitrary discrete alphabet. The Bhattacharyya parameter $Z(T|V) \in [0, 1]$ is defined

$$Z(T|V) = 2\sum_{v \in \mathcal{V}} P_V(v)\sqrt{P_{T|V}(0|v)P_{T|V}(1|v)}. \quad (3)$$

III. Polar Superposition Code

Theorem 1: For any two-user DM-BC $P_{Y_1Y_2|X}(y_1, y_2|x)$ with binary input alphabet $\mathcal{X}$ and output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$, there exists a polar code sequence indexed by a code length n which achieves the region

$$\mathfrak{R}(V, X, Y_1, Y_2) \triangleq \Big\{R_1, R_2 \;\Big|\; R_1 \le I(X; Y_1|V), R_2 \le I(V; Y_2)\Big\}, \quad (4)$$

for r.v.'s V, X, Y_1, Y_2 with listed properties:

- V is a binary random variable.
- $P_{Y_1|V}(y_1|v) \succ P_{Y_2|V}(y_2|v)$.
- $P_{VXY_1Y_2} = P_V P_{X|V} P_{Y_1Y_2|X}$.

For this code sequence, $P_e^{(n)} = \mathcal{O}(2^{-n^\beta})$ where $0 < \beta < \frac{1}{2}$. The complexity of encoding and decoding is $\mathcal{O}(n \log n)$.

Remark 1: See [23] for a full proof of Theorem 1. Note that the theorem applies to any two-user DM-BC assuming the listed properties hold. The requirement that auxiliary V and X be binary variables is remedied by applying q-ary polarization theorems. The requirement $P_{Y_1|V}(y_1|v) \succ P_{Y_2|V}(y_2|v)$ guarantees that polarization indices are *aligned* and nested for the coarse message carried by auxiliary V.

Example 1 (Binary Symmetric DM-BC): Let $0 < p_1 < p_2 < \frac{1}{2}$. Consider a DM-BC composed of two binary symmetric channels (*BSC*s), a *BSC* with bit-flip probability p_1 to the first receiver and a *BSC* with bit-flip probability p_2 to the other receiver. For $\alpha \in [0, \frac{1}{2}]$, Cover's information-theoretic superposition inner bound is optimal:

$$\Big\{R_1, R_2 \;\Big|\; R_1 \le h_b(\alpha * p_1) - h_b(p_1), R_2 \le 1 - h_b(\alpha * p_2)\Big\}.$$

In Theorem 1, let V be a uniform Bernoulli r.v. and $X = V \oplus S$ where S is a Bernoulli r.v. with $\mathbb{P}\{S = 1\} = \alpha$. Then Theorem 1 establishes the existence of polar broadcast codes which achieve all rates on the capacity-boundary with low complexity.

IV. Marton's Coding Scheme

Theorem 2: For any two-user DM-BC $P_{Y_1Y_2|X}(y_1, y_2|x)$, there exists a polar code sequence indexed by code length n which achieves the region

$$\mathfrak{R}(V_1, V_2, X, Y_1, Y_2) \triangleq \Big\{R_1, R_2 \;\Big|\; R_1 \le I(V_1; Y_1), R_2 \le I(V_2; Y_2) - I(V_1; V_2)\Big\}, \quad (5)$$

for r.v.'s V_1, V_2, X, Y_1, Y_2 with listed properties:

- V_1 and V_2 are binary r.v.'s.
- $P_{Y_2|V_2}(y_2|v_2) \succ P_{V_1|V_2}(v_1|v_2)$.
- For a deterministic function $\phi : \{0, 1\}^2 \to \mathcal{X}$, the joint distribution of all r.v.'s is given by

$$P_{V_1V_2XY_1Y_2}(v_1, v_2, x, y_1, y_2) = P_{V_1V_2}(v_1, v_2)\mathbb{1}_{[x=\phi(v_1,v_2)]}P_{Y_1Y_2|X}(y_1, y_2|x).$$

For this code sequence, $P_e^{(n)} = \mathcal{O}(2^{-n^\beta})$ where $0 < \beta < \frac{1}{2}$. The complexity of encoding and decoding is $\mathcal{O}(n \log n)$.

Remark 2: See [23] for a full proof of Theorem 2. The theorem is a polarization-based version of Marton's coding scheme which includes a symbol-by-symbol deterministic ϕ-mapping without loss of generality [2, Chapter 8]. The property $P_{Y_2|V_2}(y_2|v_2) \succ P_{V_1|V_2}(v_1|v_2)$ guarantees an *alignment* of polarization indices in the multi-user setting. It is a *natural* restriction since it also implies that $I(Y_2; V_2) > I(V_1; V_2)$ so that $R_2 > 0$. However, certain joint distributions on r.v.'s are not permitted. Thus Theorem 2 differs slightly from Marton's information-theoretic strategy.

V. Sketch of Proof for Theorem 2

A. Polar Transform

Figure 1 provides a block diagram for a polar broadcast code based on Marton's strategy. For a code length of n, consider n independent and identically distributed copies $(V_1^{1:n}, V_2^{1:n}, X^{1:n}, Y_1^{1:n}, Y_2^{1:n})$ of the r.v.'s (V_1, V_2, X, Y_1, Y_2). The r.v.'s have a joint distribution with listed properties as specified in Theorem 2. For example, $P_{X|V_1V_2}(x|v_1, v_2) = \mathbb{1}_{[x=\phi(v_1,v_2)]}$ for ϕ a deterministic mapping.

The polar transform $\mathbf{G}_n$ (see [14]) is applied to the auxiliary r.v.'s: $U_1^{1:n} \triangleq V_1^{1:n}\mathbf{G}_n$; $U_2^{1:n} \triangleq V_2^{1:n}\mathbf{G}_n$. In the transformed domain, the joint distribution of $(U_1^{1:n}, U_2^{1:n})$ is given by $P_{U_1^{1:n}U_2^{1:n}}\big(u_1^{1:n}, u_2^{1:n}\big) \triangleq P_{V_1^{1:n}V_2^{1:n}}\big(u_1^{1:n}\mathbf{G}_n, u_2^{1:n}\mathbf{G}_n\big)$. For polar coding purposes, the joint distribution is also decomposed as follows:

$$P_{U_1^{1:n}U_2^{1:n}}\big(u_1^{1:n}, u_2^{1:n}\big) = P_{U_1^{1:n}}(u_1^{1:n})P_{U_2^{1:n}|U_1^{1:n}}\big(u_2^{1:n}\big|u_1^{1:n}\big) = \prod_{j=1}^{n} P\big(u_1^j\big|u_1^{1:j-1}\big)P\big(u_2^j\big|u_2^{1:j-1}, u_1^{1:n}\big). \quad (6)$$

The conditional probabilities in (6) may be computed efficiently using Arıkan's recursive protocols [14]. Noting that $\mathbf{G}_n^{-1} = \mathbf{G}_n$, the above discussion is consistent with Figure 1. The r.v.'s $(U_1^{1:n}, U_2^{1:n})$ are highly *dependent* r.v.'s.

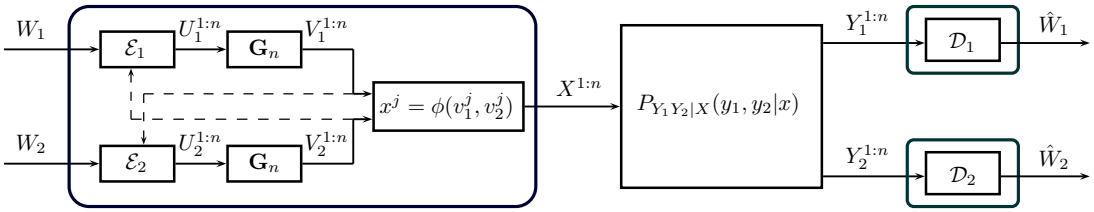

Fig. 1. Block diagram of a polar code based on Marton's strategy for a two-user noisy broadcast channel. In Marton's scheme, the auxiliary r.v.'s $V_1^{1:n}$ and $V_2^{1:n}$ may be *correlated* even though the private messages W_1 and W_2 are independent; this is expressed via the dashed arrows at the encoder. The deterministic function ϕ is a symbol-by-symbol mapping which produces the codeword input to the channel.

B. Effective Channel

An effective channel exists between the polar-transformed auxiliary variables $(U_1^{1:n}, U_2^{1:n})$ and the channel outputs $(Y_1^{1:n}, Y_2^{1:n})$.

$$P^{\phi}_{Y_1^{1:n}Y_2^{1:n}|U_1^{1:n}U_2^{1:n}}\big(y_1^{1:n}, y_2^{1:n} \,\big|\, u_1^{1:n}, u_2^{1:n}\big) \triangleq P^{\phi}_{Y_1^{1:n}Y_2^{1:n}|V_1^{1:n}V_2^{1:n}}\big(y_1^{1:n}, y_2^{1:n} \,\big|\, u_1^{1:n}\mathbf{G}_n, u_2^{1:n}\mathbf{G}_n\big)$$

where $P^{\phi}_{Y_1^{1:n}Y_2^{1:n}|V_1^{1:n}V_2^{1:n}}\big(y_1^{1:n}, y_2^{1:n} \,\big|\, v_1^{1:n}, v_2^{1:n}\big) \triangleq \prod_{j=1}^{n} P_{Y_1Y_2|X}\big(y_1^j, y_2^j \,\big|\, \phi(v_1^j, v_2^j)\big)$.

C. Polarization Theorems

Definition 8 (Polarization Sets): Let r.v.'s $U_1^{1:n} \triangleq V_1^{1:n}\mathbf{G}_n$ and $U_2^{1:n} \triangleq V_2^{1:n}\mathbf{G}_n$ where $(V_1^{1:n}, V_2^{1:n}, X^{1:n}, Y_1^{1:n}, Y_2^{1:n})$ were defined in Section V-A. Selecting $\delta_n = 2^{-n^\beta}$ for $0 < \beta < \frac{1}{2}$, define the following *polarization sets*:

$$\begin{aligned}
\mathcal{H}_{V_1}^{(n)} &\triangleq \Big\{j \in [n] : Z\big(U_1^j \,\big|\, U_1^{1:j-1}\big) \geq 1 - \delta_n\Big\},\\
\mathcal{L}_{V_1|Y_1}^{(n)} &\triangleq \Big\{j \in [n] : Z\big(U_1^j \,\big|\, U_1^{1:j-1}, Y_1^{1:n}\big) \leq \delta_n\Big\},\\
\mathcal{H}_{V_2|V_1}^{(n)} &\triangleq \Big\{j \in [n] : Z\big(U_2^j \,\big|\, U_2^{1:j-1}, V_1^{1:n}\big) \geq 1 - \delta_n\Big\},\\
\mathcal{L}_{V_2|V_1}^{(n)} &\triangleq \Big\{j \in [n] : Z\big(U_2^j \,\big|\, U_2^{1:j-1}, V_1^{1:n}\big) \leq \delta_n\Big\},\\
\mathcal{H}_{V_2|Y_2}^{(n)} &\triangleq \Big\{j \in [n] : Z\big(U_2^j \,\big|\, U_2^{1:j-1}, Y_2^{1:n}\big) \geq 1 - \delta_n\Big\},\\
\mathcal{L}_{V_2|Y_2}^{(n)} &\triangleq \Big\{j \in [n] : Z\big(U_2^j \,\big|\, U_2^{1:j-1}, Y_2^{1:n}\big) \leq \delta_n\Big\}.
\end{aligned}$$

Definition 9 (Message Sets): Given the polarization sets of Definition 8, define the following *message sets*:

$$\begin{aligned}
\mathcal{M}_1^{(n)} &\triangleq \mathcal{H}_{V_1}^{(n)} \cap \mathcal{L}_{V_1|Y_1}^{(n)},\\
\mathcal{M}_2^{(n)} &\triangleq \mathcal{H}_{V_2|V_1}^{(n)} \cap \mathcal{L}_{V_2|Y_2}^{(n)}.
\end{aligned}$$

Proposition 1 (Polarization [6] [7]): Consider the message sets specified in Definition 8 and Definition 9 with $\delta_n = 2^{-n^\beta}$ where $0 < \beta < \frac{1}{2}$. Fix a constant $\tau > 0$. Then there exists an $N_o = N_o(\beta, \tau)$ such that $\forall n > N_o$,

$$\frac{1}{n}\left|\mathcal{M}_1^{(n)}\right| \geq I(V_1; Y_1) - \tau, \tag{7}$$

$$\frac{1}{n}\left|\mathcal{M}_2^{(n)}\right| \geq I(V_2; Y_2) - I(V_1; V_2) - \tau. \tag{8}$$

Lemma 1 (Alignment of Polarization Indices [23]): Consider the polarization sets specified in Definition 8. If $P_{Y_2|V_2}(y_2|v_2) \succ P_{V_1|V_2}(v_1|v_2)$ holds, then $I(V_2; Y_2) \geq I(V_1; V_2)$ and for all $j \in [n]$,

$$Z\big(U_2^j \,\big|\, U_2^{1:j-1}, Y_2^{1:n}\big) \leq Z\big(U_2^j \,\big|\, U_2^{1:j-1}, V_1^{1:n}\big)$$

As a result, $\mathcal{L}_{V_2|V_1}^{(n)} \subseteq \mathcal{L}_{V_2|Y_2}^{(n)}$ and $\mathcal{H}_{V_2|Y_2}^{(n)} \subseteq \mathcal{H}_{V_2|V_1}^{(n)}$.

Remark 3: The alignment of polarization indices characterized by Lemma 1 is diagrammed in Figure 2. The alignment ensures that the cardinality $\left|\mathcal{M}_2^{(n)}\right| > 0$. The indices in $\mathcal{M}_2^{(n)}$ represent those message bits freely set at the broadcast encoder and simultaneously those message bits reliably decoded by the second receiver $\mathcal{D}_2$ given its observations.

D. Partially-Polarized Indices

As shown in Figure 2, for the Marton coding scheme, exact alignment of polarization indices is not possible because there exist sets of partially-polarized indices.

Definition 10 (Sets of Partially-Polarized Indices):

$$\Delta_1 \triangleq [n] \setminus \big(\mathcal{H}_{V_2|V_1}^{(n)} \cup \mathcal{L}_{V_2|V_1}^{(n)}\big), \tag{9}$$

$$\Delta_2 \triangleq [n] \setminus \big(\mathcal{H}_{V_2|Y_2}^{(n)} \cup \mathcal{L}_{V_2|Y_2}^{(n)}\big). \tag{10}$$

The number of partially-polarized indices grows *sub-linearly* in n, constituting a negligible fraction of the total number of indices as $n \to \infty$. For an arbitrarily small $\eta > 0$,

$$\frac{\left|\Delta_1 \cup \Delta_2\right|}{n} \leq \eta, \tag{11}$$

for all n sufficiently large enough. As shown in [23], providing these $o(n)$ bits as "genie-given" bits to the decoders results in a rate penalty. However, the rate penalty can be forced to be arbitrarily small for sufficiently large n.

E. Broadcast Encoding Based on Polarization: $\mathcal{E}_1, \mathcal{E}_2$

As diagrammed in Figure 1, the broadcast encoder must map two independent messages (W_1, W_2) uniformly distributed over $[2^{\lfloor nR_1 \rfloor}] \times [2^{\lfloor nR_2 \rfloor}]$ to one codeword $x^n \in \mathcal{X}^n$. The encoding block $\mathcal{E}_1$ maps the message W_1 to a sequence $u_1^{1:n} \in \{0,1\}^n$. The encoding block $\mathcal{E}_2$ maps the message W_2 to a sequence $u_2^{1:n} \in \{0,1\}^n$. The aim of both $\mathcal{E}_1$ and $\mathcal{E}_2$ is to ensure that $(u_1^{1:n}, u_2^{1:n})$ obey the joint statistics of $(U_1^{1:n}, U_2^{1:n})$ otherwise the encoding contributes to the overall decoding error at the receivers as shown in [23].

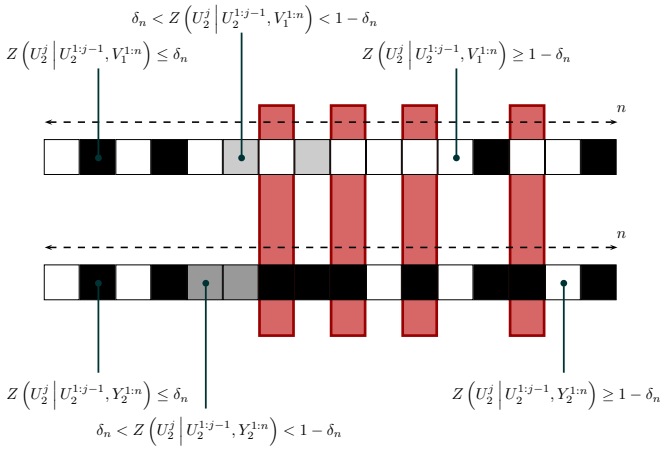

Fig. 2. The alignment of polarization indices for Marton coding over noisy broadcast channels with respect to the second receiver. The message set $\mathcal{M}_2^{(n)}$ is highlighted by the vertical red rectangles and represents those bits transmitted reliably to the second receiver. At finite code length n, exact alignment is not possible due to partially-polarized indices pictured in gray.

1) Encoding Block $\mathcal{E}_1$: More precisely, the block $\mathcal{E}_1$ constructs the sequence $u_1^{1:n}$ bit-by-bit successively by inserting message bits of W_1 at precise indices given by set $\mathcal{M}_1^{(n)}$, and then applies a *randomized map* to fill in the rest of the indices.

$$u_1^j = \begin{cases} W_1 \text{ message bit}, & \textit{if } j \in \mathcal{M}_1^{(n)}, \\ \Psi_1^{(j)}\big(u_1^{1:j-1}\big), & \textit{otherwise.} \end{cases}$$

Taking into account the statistics of $U_1^{1:n}$, each randomized map $\Psi_1^{(j)} : \{0,1\}^{j-1} \to \{0,1\}$ for a fixed $j \in [n]$ is defined:

$$\Psi_1^{(j)}\big(u_1^{1:j-1}\big) \triangleq \begin{cases} 0, & \textit{w.p. } \lambda_0\big(u_1^{1:j-1}\big), \\ 1, & \textit{w.p. } 1-\lambda_0\big(u_1^{1:j-1}\big), \end{cases}$$
$$\text{where } \lambda_0\big(u_1^{1:j-1}\big) \triangleq \mathbb{P}\big(U_1^j = 0 \,\big|\, U_1^{1:j-1} = u_1^{1:j-1}\big).$$

2) Encoding Block $\mathcal{E}_2$: The block $\mathcal{E}_2$ constructs the sequence $u_2^{1:n}$ bit-by-bit successively by inserting message bits of W_2 at precise indices given by set $\mathcal{M}_2^{(n)}$, and then applies appropriate *randomized maps* to fill in the rest of the indices. Encoding block $\mathcal{E}_2$ must take into account the fact that $\mathcal{E}_1$ constructs the sequence $u_1^{1:n}$ first. Equivalently, $\mathcal{E}_2$ has knowledge of $v_1^{1:n} \triangleq u_1^{1:n}\mathbf{G}_n$.

$$u_2^j = \begin{cases} W_2 \text{ message bit}, & \textit{if } j \in \mathcal{M}_2^{(n)}, \\ \Gamma(j), & \textit{if } j \in \mathcal{H}_{V_2|V_1}^{(n)} \setminus \mathcal{M}_2^{(n)}, \\ \Psi_2^{(j)}\big(u_2^{1:j-1}, v_1^{1:n}\big), & \textit{otherwise.} \end{cases}$$

According to the statistics of both $(U_1^{1:n}, U_2^{1:n})$, the randomized map $\Psi_2^{(j)} : \{0,1\}^{n+j-1} \to \{0,1\}$ for $j \in [n]$ is defined:

$$\Psi_2^{(j)}\big(u_2^{1:j-1}, v_1^{1:n}\big) \triangleq \begin{cases} 0, & \textit{w.p. } \lambda_0\big(u_2^{1:j-1}, v_1^{1:n}\big), \\ 1, & \textit{w.p. } 1-\lambda_0\big(u_2^{1:j-1}, v_1^{1:n}\big) \end{cases}$$
$$\text{where } \lambda_0\big(u_2^{1:j-1}, v_1^{1:n}\big) \triangleq \mathbb{P}\big(U_2^j = 0 \,\big|\, U_2^{1:j-1} = u_2^{1:j-1}, V_1^{1:n} = v_1^{1:n}\big).$$

The randomized map $\Gamma : [n] \to \{0,1\}$ is simply equal to 0 or 1 with equal probability for any input.

Remark 4: The random maps $\Psi_1^{(j)}$ and $\Psi_2^{(j)}$ may be characterized as vectors of independent Bernoulli random variables for $j \in [n]$ fixed. Each Bernoulli random variable of the vector is zero or one with a well-defined fixed probability. The random map Γ may be characterized as an n-length vector of Bernoulli$(\frac{1}{2})$ random variables.

3) Achievable Rates: As mentioned, the encoding blocks $\mathcal{E}_1$ and $\mathcal{E}_2$ encode messages W_1 and W_2 into the sequences $(u_1^{1:n}, u_2^{1:n})$. As a result of Proposition 1, the encoding achieves the following rates for a finite code length.

$$R_1 = \frac{1}{n}\left|\mathcal{M}_1^{(n)}\right|.$$
$$R_2 = \frac{1}{n}\left|\mathcal{M}_2^{(n)}\right|.$$

4) Encoding Protocol: After encoding blocks $\mathcal{E}_1$ and $\mathcal{E}_2$ form sequences $(u_1^{1:n}, u_2^{1:n})$, the broadcast encoder computes $v_1^{1:n} \triangleq u_1^{1:n}\mathbf{G}_n$ and $v_2^{1:n} \triangleq u_2^{1:n}\mathbf{G}_n$. The last step in the encoding protocol is to form the codeword $x^{1:n}$ using the deterministic symbol-by-symbol mapping ϕ.

Codeword Construction: $\forall j \in [n],\ x^j = \phi\big(v_1^j, v_2^j\big)$.

For $j \in \Delta_2$, the encoder records the realization of u_2^j. These bits are provided to the second receiver as "genie-given" bits.

F. Broadcast Decoding Based on Polarization: $\mathcal{D}_1, \mathcal{D}_2$

1) Decoder $\mathcal{D}_1$: The first receiver decodes the binary sequence $\hat{u}_1^{1:n}$ using its observations $y_1^{1:n}$. The message W_1 is located at the indices $j \in \mathcal{M}_1^{(n)}$ in the sequence $\hat{u}_1^{1:n}$. More precisely, we define the following deterministic polar decoding function for the j-th bit:

$$\xi_{u_1}^{(j)}\big(u_1^{1:j-1}, y_1^{1:n}\big) \triangleq \underset{u \in \{0,1\}}{\arg\max} \Big\{ \mathbb{P}\big(U_1^j = u \,\big|\, U_1^{1:j-1} = u_1^{1:j-1}, Y_1^{1:n} = y_1^{1:n}\big) \Big\}.$$

Decoder $\mathcal{D}_1$ reconstructs $\hat{u}_1^{1:n}$ bit-by-bit successively as follows using the same identical random mapping $\Psi_1^{(j)}$ at the

encoder:

$$\hat{u}_1^j = \begin{cases} \xi_{u_1}^{(j)}\big(\hat{u}_1^{1:j-1}, y_1^{1:n}\big), & \textit{if } j \in \mathcal{M}_1^{(n)}, \\ \Psi_1^{(j)}\big(\hat{u}_1^{1:j-1}\big), & \textit{otherwise.} \end{cases}$$

Given that all previous bits $\hat{u}_1^{1:j-1}$ have been decoded correctly, $\mathcal{D}_1$ makes a mistake on the j-th bit $\hat{u}_1^j$ only if $j \in \mathcal{M}_1^{(n)}$. For the other indices, $\mathcal{D}_1$ produces the same bit produced at the encoder due to shared random maps.

2) Decoder $\mathcal{D}_2$: The second receiver decodes the binary sequence $\hat{u}_2^{1:n}$ using observations $y_2^{1:n}$. The message W_2 is located at the indices $j \in \mathcal{M}_2^{(n)}$ of the sequence $\hat{u}_2^{1:n}$. Define the following deterministic polar decoding functions

$$\xi_{u_2}^{(j)}\big(u_2^{1:j-1}, y_2^{1:n}\big) \triangleq \arg\max_{u\in\{0,1\}} \Big\{ \mathbb{P}\big(U_2^j = u \,\big|\, U_2^{1:j-1} = u_2^{1:j-1}, Y_2^{1:n} = y_2^{1:n}\big) \Big\}.$$

Decoder $\mathcal{D}_2$ reconstructs the sequence $\hat{u}_2^{1:n}$ bit-by-bit successively as follows using the same identical random mapping Γ used at the encoder. Including all but $o(n)$ of the indices,

$$\hat{u}_2^j = \begin{cases} \xi_{u_2}^{(j)}\big(\hat{u}_2^{1:j-1}, y_2^{1:n}\big), & \textit{if } j \in \mathcal{L}_{V_2|Y_2}^{(n)}, \\ \Gamma(j), & \textit{if } j \in \mathcal{H}_{V_2|Y_2}^{(n)}. \end{cases}$$

For those $o(n)$ indices $j \in \Delta_2$ where Δ_2 is the set of partially-polarized indices defined in (10), the decoder $\mathcal{D}_2$ is provided with "genie-given" bits from the encoder. Thus, all bits are recovered reliably, and $\mathcal{D}_2$ only makes a successive cancelation error for indices $j \in \mathcal{L}_{V_2|Y_2}^{(n)}$.

Remark 5: Note that $\mathcal{D}_2$ does not utilize the random mapping $\Psi_2^{(j)}\big(u_2^{1:j-1}, v_1^{1:n}\big)$ used by the encoding block $\mathcal{E}_2$ because $\mathcal{D}_2$ does not have access to the sequence $v_1^{1:n}$. This is the key to Marton's coding scheme.

G. Error Analysis

The encoding blocks $(\mathcal{E}_1, \mathcal{E}_2)$ and the decoding blocks $(\mathcal{D}_1, \mathcal{D}_2)$ utilize random maps $\Psi_1^{(j)}$, $\Psi_2^{(j)}$, and Γ. Therefore, the average probability of error defined in (1) is a *random quantity:* $P_e^{(n)}\big\{\Psi_1^{(j)}, \Psi_2^{(j)}, \Gamma\big\}$. The expectation of this random quantity decays stretched-exponentially in the code length n.

Lemma 2 (Average Error Probability [23]): Consider the polarization-based Marton code described in Section V-E and Section V-F. Let (R_1, R_2) be the broadcast rates selected according to the Bhattacharyya criterion given in Proposition 1. Then for $0 < \beta < \frac{1}{2}$ and sufficiently large n,

$$\mathbb{E}_{\Psi_1^{(j)}, \Psi_2^{(j)}, \Gamma}\Big[P_e^{(n)}\big\{\Psi_1^{(j)}, \Psi_2^{(j)}, \Gamma\big\}\Big] < 2^{-n^\beta}.$$

Remark 6: If $\mathbb{E}_{\Psi_1^{(j)}, \Psi_2^{(j)}, \Gamma}\Big[P_e^{(n)}\big\{\Psi_1^{(j)}, \Psi_2^{(j)}, \Gamma\big\}\Big] \to 0$ as $n \to \infty$, then there must exist at least one fixed set of maps for which $P_e^{(n)} \to 0$. In the proof of Lemma 2, the error is decomposed into the error incurred at the encoding blocks $(\mathcal{E}_1, \mathcal{E}_2)$ due to the imprecise mapping between messages and sequences, and the error incurred at each of the decoders $(\mathcal{D}_1, \mathcal{D}_2)$ due to channel noise and successive cancelation decoding.

VI. Conclusion

Combined with the authors' contributions in [21], polar code constructions have been designed for several classes of broadcast channels. The codes offer new insight into important information-theoretic broadcast strategies: (1) random binning; (2) Cover's superposition strategy; (3) Marton's construction. Our current aim is to supplement the theory with experimental evidence to analyze in depth the relationship between the broadcast code length n, average probability of error $P_e^{(n)}$, and broadcast rates R_i chosen within the capacity region.

References

[1] T. M. Cover, "Broadcast channels," *IEEE Trans. on Inform. Theory*, vol. 18, pp. 2–14, Jan. 1972.

[2] A. E. Gamal and Y.-H. Kim, *Network Inform. Theory*. New York: Cambridge University Press, 2011.

[3] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. on Inform. Theory*, vol. 25, pp. 306–311, May 1979.

[4] C. Nair and A. E. Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. on Inform. Theory*, vol. 53, pp. 350–355, Jan. 2007.

[5] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. on Inform. Theory*, vol. 44, pp. 2524–2530, Oct. 1998.

[6] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. on Inform. Theory*, vol. 55, pp. 3051–3073, July 2009.

[7] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *Proc. of the IEEE Intern. Symp. on Inform. Theory*, (South Korea), July 2009.

[8] E. Şaşoğlu, E. Yeh, and E. Telatar, "Polar codes for the two-user binary-input multiple-access channel," in *Proc. IEEE Inform. Theory Workshop*, (Cairo, Egypt), Jan. 2010.

[9] E. Abbe and E. Telatar, "Polar codes for the m-user multiple access channel," *IEEE Trans. on Inform. Theory*, vol. 58, pp. 5437–5448, Aug. 2012.

[10] R. Blasco-Serrano, R. Thobaben, M. Andersson, V. Rathi, and M. Skoglund, "Polar codes for cooperative relaying," *IEEE Trans. on Comm.*, vol. 60, pp. 3263–3273, Nov. 2012.

[11] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. on Inform. Theory*, vol. 57, pp. 6428–6443, Oct. 2011.

[12] O. Koyluoglu and H. El Gamal, "Polar coding for secure transmission and key agreement," *IEEE Trans. on Inform. Forensics and Security*, vol. 7, pp. 1472–1483, Oct. 2012.

[13] S. B. Korada, *Polar Codes for Channel and Source Coding*. PhD thesis, EPFL, 2009.

[14] E. Arıkan, "Source polarization," in *Proc. of the IEEE Intern. Symp. on Inform. Theory*, June 2010.

[15] S. B. Korada and R. L. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Trans. on Inform. Theory*, vol. 56, no. 4, pp. 1751–1768, 2010.

[16] E. Abbe, "Randomness and dependencies extraction via polarization," in *Proc. of the Inform. Theory and Applic. Workshop (ITA)*, Feb. 2011.

[17] E. Arıkan, "Polar coding for the slepian-wolf problem based on monotone chain rules," in *Proc. of the IEEE Intern. Symp. on Inform. Theory*, July 2012.

[18] S. Korada, E. Şaşoğlu, and R. Urbanke, "Polar codes: Characterization of exponent, bounds, and constructions," *IEEE Trans. on Inform. Theory*, vol. 56, pp. 6253–6264, Dec. 2010.

[19] E. Şaşoğlu, E. Telatar, and E. Arıkan, "Polarization for arbitrary discrete memoryless channels," *CoRR*, vol. abs/0908.0302, 2009.

[20] W. Park and A. Barg, "Polar codes for q-ary channels, $q = 2^r$," *IEEE Trans. on Inform. Theory*, vol. 59, pp. 955–969, Feb. 2013.

[21] N. Goela, E. Abbe, and M. Gastpar, "Polar codes for the deterministic broadcast channel," in *Proc. Intern. Zurich Seminar on Comm.*, (Switzerland), pp. 51–54, Feb. 2012.

[22] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric channels," in *Proc. of the IEEE Intern. Symp. on Inform. Theory*, (Cambridge, USA), July 2012.

[23] N. Goela, E. Abbe, and M. Gastpar, "Polar codes for broadcast channels," *CoRR*, vol. abs/1301.6150, Jan. 2013.