Strong Coordination with Polar Codes

Matthieu R. Bloch[1], Laura Luzzi[2], and Jörg Kliewer[3]

Abstract— In this paper, we design explicit codes for strong coordination in two-node networks. Specifically, we consider a two-node network in which the action imposed by nature is binary and uniform, and the action to coordinate is obtained via a symmetric discrete memoryless channel. By observing that polar codes are useful for channel resolvability over binary symmetric channels, we prove that nested polar codes achieve a subset of the strong coordination capacity region, and therefore provide a constructive and low complexity solution for strong coordination.

I. Introduction

The characterization of the information-theoretic limits of coordination in networks has recently been investigated in [1]. The coordinated actions of nodes in the network are modeled by joint probability distributions, and the level of coordination is measured in terms of how well these joint distributions approximate a target joint distribution. Two types of coordination have been introduced: empirical coordination, which only requires the empirical distribution of coordinated actions to approach a target distribution, and strong coordination, which requires the total variational distance of coordinated actions to approach a target distribution. The concept of coordination sheds light into the fundamental limits of several problems, such as distributed control or task assignment in a network.

The design of practical and efficient coordination schemes approaching the fundamental limits predicted by information theory has attracted little attention to date. One of the hurdles faced for code design is that the metric to optimize is not a probability of error but a variational distance between distributions. Nevertheless, polar codes [2] have recently been successfully adapted [3] for empirical coordination, with an analysis based on results from lossy source coding with polar codes [4]. In this paper, we construct polar codes that are able to achieve strong coordination in some cases. Unlike the construction in [3], which solely relies on source coding with polar codes, our construction also exploits polar codes for channel resolvability [5]. Channel resolvability characterizes the bit rate required to simulate a process at

[1]M. Bloch is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA and with the GT-CNRS UMI 2958, Metz, France. matthieu.bloch@ece.gatech.edu

[2]L. Luzzi was with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, United Kingdom. She is now with Laboratoire ETIS (ENSEA - Université de Cergy-Pontoise - CNRS), 6 Avenue du Ponceau, 95014 Cergy-Pontoise, France. laura.luzzi@ensea.fr

[3]J. Kliewer is with the Klipsch School of Electrical and Computer Engineering, New Mexico State University, Las Cruces, NM, 88003, USA. jkliewer@nmsu.edu

the output of a channel and plays a key role in the analysis of the common information between random variables [6], secure communication over wiretap channels [7], [8], and coordination [1, Lemma 19]. By remarking that polar codes can be used for channel resolvability, we are able to provide a constructive alternative to the information-theoretic proof in [1].

The remainder of the paper is organized as follows. Section II sets the notation and recalls known results for polar codes. Section III shows that polar codes provide channel resolvability for symmetric channels by leveraging results in [9]. Section IV proves that polar codes achieve strong coordination for simple two-node networks with symmetric actions. Finally, Section V concludes the paper with a discussion of potential improvements and extensions.

II. Notation and Preliminaries

First, a word about notation. Given a length n vector $\mathbf{x} = (x_1, \cdots, x_n)$ and $i \in [\![1, n]\!]$, we use the notation $\mathbf{x}_1^i$ as a shorthand for the row vector $(x_1, \cdots, x_i)$. Similarly, for any set $F \in [\![1, n]\!]$, we denote by $\mathbf{x}_F$ the vector of length $|F|$ containing the indices x_i for $i \in F$. The distributions of different random variables defined on the same alphabet $\mathcal{X}$ are denoted by different symbols, e.g. p_X, q_X. For brevity, the subscripts in the distributions may be dropped if the alphabet is clear from the context or from the argument. We also use $\mathbb{D}$ to denote the Kullback-Leibler divergence between two distributions.

Next, we briefly review the concepts and notation related to polar codes that will be used throughout the paper. The key element in the polar coding construction is the decomposition of $n \triangleq 2^m$ independent copies of a given binary-input discrete-memoryless channel $(\mathcal{X}, W_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y})$ with capacity $C(W_{\mathsf{Y}|\mathsf{X}})$ into n *bit-channels* which are essentially either error-free or pure noise channels. Specifically, consider the transformation $G_n \triangleq G_2^{\otimes n}$ where

$$G_2 \triangleq \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

and $\otimes$ is the Kronecker product. A vector $\mathbf{u} \in \{0,1\}^n$ is transformed into $\mathbf{x} = \mathbf{u}G_n$. The i-th bit channel $(\{0,1\}, W_n^{(i)}, \mathcal{Y}^n \times \{0,1\}^{i-1})$ is a composite channel that combines the transformation G_n and the channel, and is defined by its transition probabilities

$$W_n^{(i)}(\mathbf{y}, \mathbf{u}_1^{i-1}|u_i) \triangleq \frac{1}{2^{n-1}} \sum_{\mathbf{u}_{i+1}^n} W_{\mathsf{Y}^n|\mathsf{X}^n}(\mathbf{y}|\mathbf{u}G_n).$$

For n large enough, the bit channels *polarize*, i.e. they become either completely noisy or noise-free. The exact

measure of the noise level will be specified in subsequent sections.

III. CHANNEL RESOLVABILITY WITH POLAR CODES

A. Channel resolvability

In its simplest formulation, the problem of *channel resolvability* [6], [5] can be stated as follows. Consider a discrete memoryless channel $(\mathcal{X}, W_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y})$ whose input is an i.i.d. source distributed according to q_{X}; the output of the channel is then an i.i.d. process distributed according to q_{Y}. The aim is to construct a sequence of codes $\{\mathcal{C}_n\}_{n \geqslant 1}$ of rate R and increasing block length n, such that the output distribution p_{Y^n} induced by a uniform choice of the codewords in $\mathcal{C}_n$ approaches the distribution $q_{\mathsf{Y}^n} \sim \prod_{i=1}^{n} q_{\mathsf{Y}}$ in variational distance, i.e.

$$\lim_{n \to \infty} \mathbb{V}(p_{\mathsf{Y}^n}, q_{\mathsf{Y}^n}) = 0. \tag{1}$$

In this case, the sequence $\{\mathcal{C}_n\}_{n \geqslant 1}$ is called a sequence of *resolvability codes* achieving *resolution rate* R for $(W_{\mathsf{Y}|\mathsf{X}}, q_{\mathsf{X}})$. The *channel resolvability* of $W_{\mathsf{Y}|\mathsf{X}}$ is then defined as the minimum resolution rate such that resolvability codes exist for any input source.

B. Coding scheme for channel resolvability

In this section, we leverage the results of [9] to construct resolvability codes when $(\mathcal{X}, W_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y})$ is a binary-input symmetric DMC and q_{X} is the uniform distribution on $\{0, 1\}$, i.e. $q_{\mathsf{X}} \sim \mathcal{B}(\frac{1}{2})$; this result will be exploited in Section IV for the problem of coordination. We use the notion of symmetry in [10], according to which there exists a permutation $\pi_1 : \mathcal{Y} \to \mathcal{Y}$ such that $\pi_1 = \pi_1^{-1}$ and

$$\forall y \in \mathcal{Y}, \quad W_{\mathsf{Y}|\mathsf{X}}(y|0) = W_{\mathsf{Y}|\mathsf{X}}(\pi_1(y)|1) \tag{2}$$

In particular, the following property of symmetric channels will be useful.

Lemma 1 ([10]): If $(\mathcal{X}, W_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y})$ is a memoryless symmetric channel and if q_{Y} is the output distribution corresponding to the uniform input distribution q_{X} on $\mathcal{X}$, then

$$\forall x \in \mathcal{X}, \quad C(W_{\mathsf{Y}|\mathsf{X}}) = \mathbb{D}\big(W_{\mathsf{Y}|\mathsf{X}=x}\|q_{\mathsf{Y}}\big),$$

where $W_{\mathsf{Y}|\mathsf{X}=x}$ is the output distribution induced by the fixed symbol x.

Let $W_n^{(i)}$ denote the set of bit channels corresponding to $W_{\mathsf{Y}|\mathsf{X}}$, and define the sets of "good bits" $\mathcal{G}_n$ and "bad bits" $\mathcal{B}_n$ as

$$\mathcal{G}_n \triangleq \left\{ i \in [\![1, n]\!] : C(W_n^{(i)}) \geqslant 2^{-n^\beta} \right\},$$

and $\mathcal{B}_n \triangleq [\![1, n]\!] \setminus \mathcal{G}_n.$

Our strategy to simulate the i.i.d. process distributed according to q_{Y} is to send random uniform bits on the good bits, and fixed bits on the bad bits. Intuitively, the uniform bits will be preserved by the noiseless bit-channels, while the pure noise bit-channels will produce almost-uniform bits for any input. Formally, let $r = |\mathcal{G}_n|$ and consider the polar codes defined in Section II. We will use the $(n, r, \mathcal{G}_n, \mathbf{0}^{n-r})$ *coset*

code $\mathcal{C}_n$ [2] obtained by using $\mathcal{G}_n$ as the set of information bits and $\mathcal{B}_n$ as the set of *frozen* bits.

Proposition 1: If the channel $(\mathcal{X}, W_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y})$ is symmetric and $q_{\mathsf{X}} \sim \mathcal{B}(\frac{1}{2})$, then $\{\mathcal{C}_n\}_{n \geqslant 1}$ is a sequence of resolvability codes of resolution rate $C(W_{\mathsf{Y}|\mathsf{X}})$ for $(W_{\mathsf{Y}|\mathsf{X}}, q_{\mathsf{X}})$.

Proof:
We know from [9, Proposition 20] that

$$\lim_{n \to \infty} \frac{r}{n} = C(W_{\mathsf{Y}|\mathsf{X}}),$$

so that the condition regarding the resolution rate is satisfied. Following [9], given two vectors $\mathbf{x}^r \in \{0, 1\}^r$ and $\mathbf{s}^{n-r} \in \{0, 1\}^{n-r}$, we let $(\mathbf{x}^r, \mathbf{s}^{n-r})$ denote the vector $\mathbf{v}^n \in \{0, 1\}^n$ such that $\mathbf{v}_{|\mathcal{G}_n} = \mathbf{x}^r$ and $v_{|\mathcal{B}_n} = \mathbf{s}^{n-r}$. We then define a composite channel $(\{0, 1\}^{n-r}, W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}, \mathcal{Y}^n)$, which includes the polar code and the random bits sent on the good bits $\mathcal{G}_n$, so that

$$W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}(\mathbf{y}^n|\mathbf{s}^{n-r})$$
$$\triangleq \frac{1}{2^r} \sum_{\mathbf{x}^r \in \{0,1\}^r} W_{\mathsf{Y}^n|\mathsf{X}^n}\left(\mathbf{y}^n \middle| (\mathbf{x}^r, \mathbf{s}^{n-r})G_n\right).$$

It is shown in [9, Proposition 13] that $W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}$ is symmetric and that

$$C(W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}) \leqslant \sum_{i \in \mathcal{B}_n} C(W_n^{(i)}) \leqslant (n - r)2^{-n^\beta}.$$

We now show that this last inequality implies that $\{\mathcal{C}_n\}_{n \geqslant 1}$ form a sequence of resolvability codes.

By the definition of coset codes [2], the output distribution p_{Y^n} induced by the code $\mathcal{C}_n$ coincides with the output distribution $W_{\mathsf{Y}^n|\mathsf{S}^{n-r}=\mathbf{0}^{n-r}}$ of the constant input $\mathbf{0}^{n-r}$ through $W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}$. Moreover, since $W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}$ is symmetric and G_n is full-rank, the output of the channel $W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}$ to a uniformly distributed input on $\{0, 1\}^{n-r}$ has the desired output distribution q_{Y^n}. Hence, applying Lemma 1 to the channel $W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}$, we find that

$$\mathbb{D}(p_{\mathsf{Y}^n}\|q_{\mathsf{Y}^n}) = \mathbb{D}\big(W_{\mathsf{Y}^n|\mathsf{S}^{n-r}=\mathbf{0}^{n-r}}\|q_{\mathsf{Y}^n}\big) = C(W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}),$$

so that $\lim_{n \to \infty} \mathbb{D}(p_{\mathsf{Y}^n}\|q_{\mathsf{Y}^n}) = 0$. Pinsker's inequality then ensures that

$$\lim_{n \to \infty} \mathbb{V}(p_{\mathsf{Y}^n}, q_{\mathsf{Y}^n}) = 0$$

∎

Remark 1: The choice of frozen bits set at $\mathbf{0}^{n-r}$ is arbitrary. The choice of a different coset code characterized by $\mathbf{u}_F$ in place of $\mathbf{0}^{n-r}$ does not alter the reasoning. In particular, the symmetry of the channel $W_{\mathsf{Y}^n|\mathsf{S}^{n-r}}$ and Lemma 1 still hold.

IV. STRONG COORDINATION WITH POLAR CODES

A. Strong coordination for a two-node network

The problem of strong coordination for the two-node network [11] is illustrated in Figure 1. Node X with actions distributed according to $q_{\mathsf{X}^n} \sim \prod_{i=1}^{n} q_{\mathsf{X}}$ and given by nature wishes to coordinate with node Y to obtain the

Fig. 1. Coordination for two-node network

Fig. 2. Illustration of partition sets F_1, F_2, F_3 (after reordering of indices).
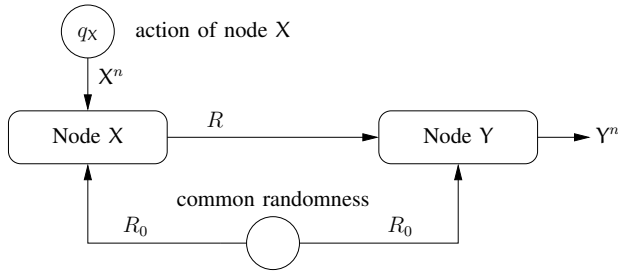
joint distribution of actions $q_{X^nY^n} \sim \prod_{i=1}^n q_{XY}$. Nodes X and Y have access to an independent source of common randomness, which provides uniform random numbers in $[\![1, 2^{nR_0}]\!]$, and node X transmits messages in $[\![1, 2^{nR}]\!]$ to node Y. Specifically, a $(2^{nR}, 2^{nR_0}, n)$ coordination code $\mathcal{C}_n$ for this network consists of a stochastic encoding function

$$f : \mathcal{X}^n \times [\![1, 2^{nR_0}]\!] \to [\![1, 2^{nR}]\!]$$

and of a stochastic decoding function

$$g : [\![1, 2^{nR}]\!] \times [\![1, 2^{nR_0}]\!] \to \mathcal{Y}^n.$$

We let $U_0 \in [\![1, 2^{nR_0}]\!]$ denote the common randomness and $p_{X^n g(f(X^n, U_0))}$ be the distribution induced by the co-ordination code. A coordination q_{XY} is achievable with rates (R, R_0) if there exists a sequence of $(2^{nR}, 2^{nR_0}, n)$ coordination codes $\{\mathcal{C}_n\}_{n \geqslant 1}$ such that

$$\lim_{n \to \infty} \mathbb{V}\big(p_{X^n g(f(X^n), U_0)}, q_{X^n Y^n}\big) = 0$$

Let X and Y be the random variables with joint distribution q_{XY}. It is shown in [11] that the set of achievable rates (R, R_0) is the following.

Theorem 1 ([11, Theorem 3.1]): The set of achievable rates (R, R_0) for coordination q_{XY} is

$$\bigcup_{X \to V \to Y} \left\{ (R, R_0) : \begin{array}{l} R + R_0 \geqslant \mathbb{I}(XY; V) \\ R \geqslant \mathbb{I}(X; V) \end{array} \right\}$$

In the sequel, we restrict our attention to the case where $\mathcal{X} = \{0, 1\}$, $q_X \sim \mathcal{B}(1/2)$, and the conditional distribution of actions $q_{Y|X}$ is symmetric.

B. Coding scheme for strong coordination

In this section, we describe the proposed scheme to achieve strong coordination. Let X and Y be the random variables with joint distribution q_{XY}, and let $V \in \{0, 1\}$ be a binary random variable satisfying the following conditions.

C1: $X \to V \to Y$ forms a Markov chain;
C2: the transition probability $W_{X|V}$ corresponds to a binary symmetric channel;
C3: the transition probability $W_{Y|V}$ is symmetric.

By assumption, such a random variable V exists and is distributed according to $\mathcal{B}(1/2)$.

We first construct polar codes of length $n \triangleq 2^m$ for the channel with transition probabilities $W_{YX|V}$ as follows.
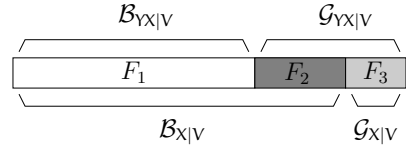
- For the symmetric channel $W_{YX|V}$, and for $i \in [\![1, n]\!]$, we let $\overline{W}_n^{(i)}$ be the corresponding set of bit channels. We define the sets

$$\mathcal{G}_{YX|V} \triangleq \left\{ i \in [\![1, n]\!] : C(\overline{W}_n^{(i)}) \geqslant 2^{-n^\beta} \right\},$$
$$\mathcal{B}_{YX|V} \triangleq [\![1, n]\!] \setminus \mathcal{G}_{YX|V}. \qquad (3)$$

- For the symmetric channel $W_{X|V}$, and for $i \in [\![1, n]\!]$, we let $\widetilde{W}_n^{(i)}$ be the corresponding set of bit channels. We define the sets

$$\mathcal{G}_{X|V} \triangleq \left\{ i \in [\![1, n]\!] : C(\widetilde{W}_n^{(i)}) \geqslant 2^{-n^\beta} \right\},$$
$$\mathcal{B}_{X|V} \triangleq [\![1, n]\!] \setminus \mathcal{G}_{X|V}. \qquad (4)$$

The sets defined in Eq. (3) and (4) satisfy the following property.

Lemma 2: $\mathcal{G}_{X|V} \subset \mathcal{G}_{YX|V}$ and $\mathcal{B}_{YX|V} \subset \mathcal{B}_{X|V}$.

Proof: The channel $W_{X|V}$ is physically degraded with respect to the channel $W_{YX|V}$. Therefore, [4, Lemma 21] guarantees that, for all $i \in [\![1, n]\!]$, $\widetilde{W}_n^{(i)}$ is degraded with respect to $\overline{W}_n^{(i)}$, so that

$$C(\widetilde{W}_n^{(i)}) \leqslant C(\overline{W}_n^{(i)}). \qquad \blacksquare$$

Consequently, the sets F_1, F_2 and F_3 defined as

$$F_1 \triangleq \mathcal{B}_{YX|V},$$
$$F_2 \triangleq \mathcal{G}_{YX|V} \cap \mathcal{B}_{X|V},$$
$$F_3 \triangleq \mathcal{G}_{YX|V} \cap \mathcal{G}_{X|V}.$$

form a partition of $[\![1, n]\!]$, which is illustrated in Figure 2.

We now exploit these sets to construct a coordination code. The bits in positions F_1 are frozen bits with values $\mathbf{u}_{F_1} = \mathbf{0}_{F_1}$ fixed at all times. The encoding and decoding procedures are then the following.

Operation at node X. To encode a sequence of binary actions $\mathbf{x} \in \mathcal{X}^n$ provided by nature, node X performs successive-cancellation (SC) encoding to determine the value of the bits $\mathbf{u}_{F_3}$ in F_3, using the bits $\mathbf{u}_{F_2}$ from the common randomness in positions F_2 and the frozen bits $\mathbf{u}_{F_1}$ in position F_1. Specifically, the probability of obtaining a bit

u_i during SC encoding is the following [4].

$$\tilde{p}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) = \begin{cases} 1 & i \in F_1, u_i = (\mathbf{u}_{F_1})_i \\ 0 & i \in F_1, u_i \neq (\mathbf{u}_{F_1})_i \\ \frac{1}{2} & i \in F_2 \\ \frac{L_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1})}{1 + L_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1})} & i \in F_3, u_i = 0 \\ \frac{1}{1 + L_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1})} & i \in F_3, u_i = 1 \end{cases}$$
(5)

where

$$L_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1}) \triangleq \frac{\widetilde{W}_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1}|0)}{\widetilde{W}_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1}|1)}.$$

The bits in F_3 are then transmitted to node Y. Note that the encoding complexity is that of SC encoding, which is $O(n \log n)$.

Operation at node Y. To create a sequence of coordinated actions $\mathbf{y} \in \mathcal{Y}^n$, node Y creates a vector $\mathbf{u}$ with frozen bits $\mathbf{u}_{F_1}$, common randomness bits $\mathbf{u}_{F_2}$, and received bits $\mathbf{u}_{F_3}$ in positions F_1, F_2, F_3, respectively. It then computes the vector $\mathbf{u}G_n$, and simulates its transmission over a memoryless channel with transition probabilities $W_{\mathsf{Y}|\mathsf{V}}$. The resulting vector $\mathbf{y}$ is used as the sequence of coordinated actions. The encoding complexity is again $O(n \log n)$.

Remark 2: Nodes X and Y require randomness to perform either SC encoding or simulate a memoryless channel. The evaluation of encoding complexity implicitly assumes that the cost of generating randomness bit-wise is $O(n)$ in both cases.

The constructed scheme operates at rate $R \triangleq \frac{|F_3|}{n}$ between nodes X and Y and requires a rate $R_0 \triangleq \frac{|F_2|}{n}$ of common randomness. Our main result, which we establish in Section IV-C, is the following.

Proposition 2: For any random variable V satisfying the conditions **C1**, **C2**, and **C3**, the coordination q_{XY} is achievable with any rates (R, R_0) such that

$$R + R_0 > C(W_{\mathsf{YX}|\mathsf{V}}) \quad \text{and} \quad R > C(W_{\mathsf{X}|\mathsf{V}})$$

C. Proof of Proposition 2

The proof is a constructive counterpart of the information-theoretic proof in [11]. We first define the distribution $\tilde{p}$ induced by the encoding/decoding procedures described in Section IV-B. By definition,

$$\tilde{p}(\mathbf{u}_{F_2}, \mathbf{u}_{F_3}, \mathbf{x}, \mathbf{y})$$
$$\triangleq \frac{1}{2^{|F_2|}} \prod_{i \in F_3} \tilde{p}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) q_{\mathsf{X}^n}(\mathbf{x}) W_{\mathsf{Y}^n|\mathsf{V}^n}(\mathbf{y}|\mathbf{u}G_n),$$

where the vector $\mathbf{u}$ is such that $\mathbf{u}_{F_1} = \mathbf{0}_{F_1}$. We also define the distribution $\hat{p}$ induced by the nested polar code with uniform inputs transmitted over the symmetric channel $W_{\mathsf{YX}|\mathsf{V}}$ (see Section III-B); we have

$$\hat{p}(\mathbf{u}_{F_2}, \mathbf{u}_{F_3}, \mathbf{x}, \mathbf{y})$$
$$\triangleq \frac{1}{2^{|F_2|}} \frac{1}{2^{|F_3|}} W_{\mathsf{X}^n|\mathsf{V}^n}(\mathbf{x}|\mathbf{u}G_n) W_{\mathsf{Y}^n|\mathsf{V}^n}(\mathbf{y}|\mathbf{u}G_n).$$

By applying the triangle inequality repeatedly, we upper bound the variational distance between the induced distribution $\tilde{p}(\mathbf{x}, \mathbf{y})$ and the target coordination $q(\mathbf{x}, \mathbf{y})$ as follows.

$$\sum_{\mathbf{x}, \mathbf{y}} |\tilde{p}(\mathbf{x}, \mathbf{y}) - q(\mathbf{x}, \mathbf{y})|$$

$$\leqslant \sum_{\mathbf{x}, \mathbf{y}} |\tilde{p}(\mathbf{x}, \mathbf{y}) - \hat{p}(\mathbf{x}, \mathbf{y})| + \sum_{\mathbf{x}, \mathbf{y}} |\hat{p}(\mathbf{x}, \mathbf{y}) - q(\mathbf{x}, \mathbf{y})|$$

$$\leqslant \sum_{\mathbf{x}, \mathbf{y}, \mathbf{u}_{F_2}, \mathbf{u}_{F_3}} |\tilde{p}(\mathbf{u}_{F_2}, \mathbf{u}_{F_3}, \mathbf{x}, \mathbf{y}) - \hat{p}(\mathbf{u}_{F_2}, \mathbf{u}_{F_3}, \mathbf{x}, \mathbf{y})|$$

$$+ \sum_{\mathbf{x}, \mathbf{y}} |\hat{p}(\mathbf{x}, \mathbf{y}) - q(\mathbf{x}, \mathbf{y})|$$

$$\overset{(a)}{=} \sum_{\mathbf{x}, \mathbf{u}_{F_2}, \mathbf{u}_{F_3}} |\tilde{p}(\mathbf{u}_{F_2}, \mathbf{u}_{F_3}, \mathbf{x}) - \hat{p}(\mathbf{u}_{F_2}, \mathbf{u}_{F_3}, \mathbf{x})|$$

$$+ \sum_{\mathbf{x}, \mathbf{y}} |\hat{p}(\mathbf{x}, \mathbf{y}) - q(\mathbf{x}, \mathbf{y})|$$

$$\triangleq \mathbb{V}\left(\tilde{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n}, \hat{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n}\right) + \mathbb{V}(\hat{p}_{\mathsf{X}^n\mathsf{Y}^n}, q_{\mathsf{X}^n\mathsf{Y}^n}), \quad (6)$$

where equality (a) follows from the definition of $\hat{p}$ and $\tilde{p}$. We first establish that, as n goes to infinity, the coding scheme operates at the sum rate in Proposition 2 and $\mathbb{V}(\hat{p}_{\mathsf{X}^n\mathsf{Y}^n}, q_{\mathsf{X}^n\mathsf{Y}^n})$ vanishes.

Lemma 3: The sequence of coding schemes satisfies

$$\lim_{n \to \infty} R_0 + R = C(W_{\mathsf{YX}|\mathsf{V}}), \quad (7)$$

$$\text{and} \quad \lim_{n \to \infty} \mathbb{V}(\hat{p}_{\mathsf{X}^n\mathsf{Y}^n}, q_{\mathsf{X}^n\mathsf{Y}^n}) = 0. \quad (8)$$

Proof: By recalling that $F_2 \cup F_3 \triangleq \mathcal{G}_{\mathsf{XY}|\mathsf{V}}$ and that the bits in positions F_2 and F_3 are i.i.d $\mathcal{B}(1/2)$ random bits, Proposition 1 guarantees that the coding scheme is a resolvability code for $(W_{\mathsf{YX}|\mathsf{V}}, q_{\mathsf{X}})$ with a resolution rate satisfying $\lim_{n \to \infty} \frac{1}{n}|\mathcal{G}_{\mathsf{YX}|\mathsf{V}}| = C(W_{\mathsf{YX}|\mathsf{V}})$. ∎

Next, we show that, as n goes to infinity, the coding scheme achieves the communication rate R of Proposition 2 and the *average* over all possible choices of frozen bits $\mathbf{u}_{F_1}$ of $\mathbb{V}\left(\tilde{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n}, \hat{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n}\right)$ vanishes, as well.

Lemma 4: The sequence of coding schemes satisfies

$$\lim_{n \to \infty} R = C(W_{\mathsf{X}|\mathsf{V}}), \quad (9)$$

$$\text{and} \quad \lim_{n \to \infty} \mathbb{E}_{\mathsf{U}_{F_1}}\left(\mathbb{V}\left(\tilde{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n}, \hat{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n}\right)\right) = 0. \quad (10)$$

Proof: Since $F_3 \triangleq \mathcal{G}_{\mathsf{X}|\mathsf{V}}$ and since the bits in position F_3 are i.i.d. $\mathcal{B}(1/2)$ random bits, Proposition 1 and Remark 1 ensure that $\lim_{n \to \infty} \frac{1}{n}|\mathcal{G}_{\mathsf{X}|\mathsf{V}}| = C(W_{\mathsf{X}|\mathsf{V}})$.

We now define two new distributions on $\mathcal{U}^n \times \mathcal{X}^n$ as follows.

$$\hat{P}(\mathbf{u}, \mathbf{x}) \triangleq \frac{1}{2^{|F_1|}} \frac{1}{2^{|F_2|}} \frac{1}{2^{|F_3|}} W_{\mathsf{X}^n|\mathsf{V}^n}(\mathbf{x}|\mathbf{u}G_n), \quad (11)$$

$$\tilde{P}(\mathbf{u}, \mathbf{x}) \triangleq q_{\mathsf{X}^n}(\mathbf{x}) \prod_{i=1}^{n} \tilde{P}(u_i|\mathbf{x}\mathbf{u}_1^{i-1}), \quad (12)$$

where

$$\tilde{P}(u_i|\mathbf{x}\mathbf{u}_1^{i-1}) \triangleq \begin{cases} \frac{1}{2} & \text{for } i \in F_1 \cup F_2 \\ \tilde{p}(u_i|\mathbf{x}\mathbf{u}_1^{i-1}) & \text{for } i \in F_3. \end{cases}$$

Remark 3: An important property shown in [4] is that, for $i \in F_3$, we have

$$\tilde{P}(u_i|\mathbf{x}\mathbf{u}_1^{i-1}) \triangleq \tilde{p}(u_i|\mathbf{x}\mathbf{u}_1^{i-1}) = \hat{P}(u_i|\mathbf{x}\mathbf{u}_1^{i-1}).$$

One can check that

$$\mathbb{E}_{\mathsf{U}_{F_1}}\big(\mathbb{V}\big(\tilde{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n}, \hat{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n}\big)\big) = \mathbb{V}\big(\tilde{P}_{\mathsf{U}^n\mathsf{X}^n}, \hat{P}_{\mathsf{U}^n\mathsf{X}^n}\big).$$

We now develop an upper bound for $\mathbb{V}\big(\tilde{P}_{\mathsf{U}^n\mathsf{X}^n}, \hat{P}_{\mathsf{U}^n\mathsf{X}^n}\big)$. Note that

$$\mathbb{V}\big(\tilde{P}_{\mathsf{U}^n\mathsf{X}^n}, \hat{P}_{\mathsf{U}^n\mathsf{X}^n}\big)$$

$$= \sum_{\mathbf{u},\mathbf{x}} \left| q(\mathbf{x}) \prod_{i=1}^n \tilde{P}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) - \hat{P}(\mathbf{x}) \prod_{i=1}^n \hat{P}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) \right|$$

$$\leqslant \underbrace{\sum_{\mathbf{u},\mathbf{x}} \left| q(\mathbf{x}) - \hat{P}(\mathbf{x}) \right| \prod_{i=1}^n \tilde{P}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1})}_{\triangleq A_n}$$

$$+ \underbrace{\sum_{\mathbf{u},\mathbf{x}} \hat{P}(\mathbf{x}) \left| \prod_{i=1}^n \tilde{P}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) - \prod_{i=1}^n \hat{P}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) \right|}_{\triangleq B_n}$$

$$\tag{13}$$

Since $\hat{P}_{\mathsf{U}^n}$ is uniform on $\mathcal{U}^n$ and since G_n defines a bijective map from $\mathcal{U}^n$ to $\mathcal{X}^n$, the distribution $\hat{P}_{\mathsf{X}^n}$ is also uniform and the term A_n on the right-hand side of Eq. (13) is zero. By applying a telescoping equality to the term B_n as in the proof of [4, Lemma 4], and by recalling that $\forall i \in F_3$, $\hat{P}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) = \tilde{P}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1})$, and that $\forall i \in F_1 \cup F_2$, $\hat{P}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) = \frac{1}{2}$, we obtain

$$B_n \leqslant \sum_{i \in F_1 \cup F_2} \sum_{u_i, \mathbf{u}_1^{i-1}, \mathbf{x}} \hat{P}(\mathbf{x}) \hat{P}(\mathbf{u}_1^{i-1}|\mathbf{x}) \left| \frac{1}{2} - \hat{P}(u_i|\mathbf{u}_1^{i-1}\mathbf{x}) \right|$$

$$= \sum_{i \in F_1 \cup F_2} \sum_{u_i, \mathbf{u}_1^{i-1}, \mathbf{x}} \left| \frac{1}{2}\hat{P}(\mathbf{u}_1^{i-1}, \mathbf{x}) - \hat{P}(u_i, \mathbf{u}_1^{i-1}\mathbf{x}) \right|$$

$$= \frac{1}{2} \sum_{i \in F_1 \cup F_2} \sum_{u_i, \mathbf{u}_1^{i-1}, \mathbf{x}} \left| \hat{P}(\mathbf{u}_1^{i-1}, \mathbf{x}) - \widetilde{W}_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1}|u_i) \right|$$

$$\triangleq \frac{1}{2} \sum_{i \in F_1 \cup F_2} \sum_{u_i} \mathbb{V}\big(\hat{P}_{\mathsf{X}^n\mathsf{U}_1^{i-1}}, \widetilde{W}_{\mathsf{X}^n\mathsf{U}_1^{i-1}|\mathsf{U}_i=u_i}\big) \tag{14}$$

By noting that

$$\hat{P}(\mathbf{u}_1^{i-1}, \mathbf{x}) = \frac{1}{2} \sum_{u \in \{0,1\}} \widetilde{W}_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1}|u_i = u),$$

and since the bit-channels $\widetilde{W}_n^{(i)}$ are symmetric [2, Proposition 13], we can argue as in the proof of Proposition 1 that for any $u \in \{0,1\}$ and $i \in F_1 \cup F_2$,

$$\mathbb{D}\big(\widetilde{W}_{\mathsf{X}^n\mathsf{U}_1^{i-1}|\mathsf{U}_i=u} \| \hat{P}_{\mathsf{X}^n\mathsf{U}_1^{i-1}}\big) = C(\widetilde{W}_n^{(i)}) \leqslant 2^{-n^\beta}.$$

Using Pinsker's inequality, we obtain

$$\mathbb{V}\big(\hat{P}_{\mathsf{X}^n\mathsf{U}_1^{i-1}}, \widetilde{W}_{\mathsf{X}^n\mathsf{U}_1^{i-1}|\mathsf{U}_i=u}\big) \leqslant 2^{-\frac{1}{2}n^\beta}\sqrt{2\ln 2},$$

and we conclude that $B_n \leqslant n 2^{-\frac{1}{2}n^\beta}\sqrt{2\ln 2}$. Therefore,

$$\lim_{n\to\infty} \mathbb{V}\big(\tilde{P}_{\mathsf{U}^n\mathsf{X}^n}, \hat{P}_{\mathsf{U}^n\mathsf{X}^n}\big) \leqslant \lim_{n\to\infty} B_n = 0.$$

$\blacksquare$

Finally we show that $\mathbb{V}\big(\tilde{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n\mathsf{Y}^n}, \hat{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n\mathsf{Y}^n}\big)$ is independent of the value of the frozen bits $\mathbf{u}_{F_1}$.

Lemma 5:

$$\mathbb{E}_{\mathsf{U}_{F_1}}\big(\mathbb{V}\big(\tilde{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n\mathsf{Y}^n}, \hat{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n\mathsf{Y}^n}\big)\big)$$

$$= \mathbb{V}\big(\tilde{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n\mathsf{Y}^n}, \hat{p}_{\mathsf{U}_{F_2}\mathsf{U}_{F_3}\mathsf{X}^n\mathsf{Y}^n}\big). \tag{15}$$

Proof: Since the channel $W_{\mathsf{X}|\mathsf{V}}$ is symmetric, there exists a permutation $\pi_1 : \mathcal{X} \to \mathcal{X}$ such that $\pi_1 = \pi_1^{-1}$ and

$$\forall x \in \mathcal{X}, \quad W_{\mathsf{X}|\mathsf{V}}(x|0) = W_{\mathsf{X}|\mathsf{V}}(\pi_1(x)|1). \tag{16}$$

Defining the identity $\pi_0 : \mathcal{X} \to \mathcal{X}$ allows us to define an action $\{0,1\} \times \mathcal{X} \to \mathcal{X}$ given by

$$v \cdot x = \pi_v(x).$$

This can be extended component-wise to an action $\{0,1\}^n \times \mathcal{X}^n \to \mathcal{X}^n$ as

$$(v_1, \ldots, v_n) \cdot (x_1, \ldots, x_n) = (\pi_{v_1}(x_1), \ldots, \pi_{v_n}(x_n)).$$

Therefore, we have $\forall \mathbf{v}, \mathbf{w} \in \{0,1\}^n, \forall \mathbf{x} \in \mathcal{X}^n$,

$$W_{\mathsf{X}^n|\mathsf{V}^n}(\mathbf{x}|\mathbf{v}) = W_{\mathsf{X}^n|\mathsf{V}^n}(\mathbf{w} \cdot \mathbf{x}|\mathbf{v} \oplus \mathbf{w}) \tag{17}$$

Lemma 8 in [4] shows that $\forall i \in [\![1,n]\!]$

$$L_n^{(i)}(\mathbf{w} \cdot \mathbf{x}, (\mathbf{w}G_n^{-1})_1^{i-1} \oplus \mathbf{u}_1^{i-1})$$

$$= \begin{cases} L_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1}) & \text{if } (\mathbf{w}G_n^{-1})_i = 0 \\ (L_n^{(i)}(\mathbf{x}, \mathbf{u}_1^{i-1}))^{-1} & \text{if } (\mathbf{w}G_n^{-1})_i = 1 \end{cases} \tag{18}$$

Now, consider the two SC encodings corresponding to two values of the frozen bits $\breve{\mathbf{u}}_{F_1}$ and $\bar{\mathbf{u}}_{F_1}$.

$$\breve{p}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) = \begin{cases} 1 & i \in F_1, u_i = (\breve{\mathbf{u}}_{F_1})_i \\ 0 & i \in F_1, u_i \neq (\breve{\mathbf{u}}_{F_1})_i \\ \frac{1}{2} & i \in F_2 \\ \frac{L_n^{(i)}(\mathbf{x},\mathbf{u}_1^{i-1})}{1+L_n^{(i)}(\mathbf{x},\mathbf{u}_1^{i-1})} & i \in F_3, u_i = 0 \\ \frac{1}{1+L_n^{(i)}(\mathbf{x},\mathbf{u}_1^{i-1})} & i \in F_3, u_i = 1 \end{cases}$$

$$\tag{19}$$

$$\bar{p}(u_i|\mathbf{x}, \mathbf{u}_1^{i-1}) = \begin{cases} 1 & i \in F_1, u_i = (\bar{\mathbf{u}}_{F_1})_i \\ 0 & i \in F_1, u_i \neq (\bar{\mathbf{u}}_{F_1})_i \\ \frac{1}{2} & i \in F_2 \\ \frac{L_n^{(i)}(\mathbf{x},\mathbf{u}_1^{i-1})}{1+L_n^{(i)}(\mathbf{x},\mathbf{u}_1^{i-1})} & i \in F_3, u_i = 0 \\ \frac{1)}{1+L_n^{(i)}(\mathbf{x},\mathbf{u}_1^{i-1})} & i \in F_3, u_i = 1 \end{cases}$$

$$\tag{20}$$

Using (18) and following the proof of Lemma 9 in [4], one can show by induction that if $\mathbf{w} \in \{0,1\}^n$ is such that

$$\bar{\mathbf{u}}_{F_1} \oplus \breve{\mathbf{u}}_{F_1} = (\mathbf{w}G_n^{-1})_{F_1}, \tag{21}$$

then $\forall i \in [\![1, n]\!]$, $\forall u_i \in \{0, 1\}$, we have

$$\breve{p}(u_i | \mathbf{x}, \mathbf{u}_1^{i-1})$$
$$= \bar{p}(u_i \oplus (\mathbf{w} G_n^{-1})_i | \mathbf{w} \cdot \mathbf{x}, (\mathbf{w} G_n^{-1})_1^{i-1} \oplus \mathbf{u}_1^{i-1}). \quad (22)$$

Note that, given $\breve{\mathbf{u}}_{F_1}$ and $\bar{\mathbf{u}}_{F_1}$, a $\mathbf{w}$ satisfying (21) always exists since G_n is one-to-one.

Similarly to Lemma 10 in [4], where it is shown that the average distortion is independent of the choice of frozen bits, we prove that the variational distance is independent of the choice of the frozen bits $\mathbf{u}_{F_1}$. Consider two resolvability codes for the channel $W_{X|V}$ obtained by transmitting i.i.d. $\mathcal{B}(1/2)$ random bits on F_2 and F_3 and by freezing the bits in F_1 to $\bar{\mathbf{u}}_{F_1}$ and $\breve{\mathbf{u}}_{F_1}$, respectively. Denote the induced distribution by $\hat{p}_{X^n \cup^n | \cup_{F_1} = \bar{\mathbf{u}}_{F_1}}$ and $\hat{p}_{X^n \cup^n | \cup_{F_1} = \breve{\mathbf{u}}_{F_1}}$, respectively. Our goal is to show that

$$\mathbb{V}\left(\bar{p}_{X^n \cup^n}, \hat{p}_{X^n \cup^n | \cup_{F_1} = \bar{\mathbf{u}}_{F_1}}\right) = \mathbb{V}\left(\breve{p}_{X^n \cup^n}, \hat{p}_{X^n \cup^n | \cup_{F_1} = \breve{\mathbf{u}}_{F_1}}\right).$$

In fact, we have

$$\mathbb{V}\left(\bar{p}_{X^n \cup^n}, \hat{p}_{X^n \cup^n | \cup_{F_1} = \bar{\mathbf{u}}_{F_1}}\right)$$
$$= \sum_{\mathbf{x}, \mathbf{u}} \frac{1}{2^{|F_2|}} \mathbb{1}_{\{\mathbf{u}_{F_1} = \bar{\mathbf{u}}_{F_1}\}} \left| q(\mathbf{x}) \prod_{i \in F_3} \bar{p}(u_i | \mathbf{x}, \mathbf{u}_1^{i-1}) \right.$$
$$\left. - \frac{1}{2^{|F_3|}} W_{X^n | V^n}(\mathbf{x} | \mathbf{u} G_n) \right|. \quad (23)$$

Consider the change of variables $\mathbf{u} = \mathbf{v} \oplus \mathbf{w} G_n^{-1}$ and $\mathbf{x} = \mathbf{w} \cdot \mathbf{z}$, where $\mathbf{w}$ satisfies (21). Equation (23) becomes

$$\sum_{\mathbf{z}, \mathbf{v}} \frac{1}{2^{|F_2|}} \mathbb{1}_{\{\mathbf{v}_{F_1} = \breve{\mathbf{u}}_{F_1}\}}$$
$$\left| q(\mathbf{w} \cdot \mathbf{z}) \prod_{i \in F_3} \bar{p}(v_i \oplus (\mathbf{w} G_n^{-1})_i | \mathbf{w} \cdot \mathbf{z}, \mathbf{v}_1^{i-1} \oplus (\mathbf{w} G_n^{-1})_1^{i-1}) \right.$$
$$\left. - \frac{1}{2^{|F_3|}} W_{X^n | V^n}(\mathbf{w} \cdot \mathbf{z} | \mathbf{v} G_n \oplus \mathbf{w}) \right|$$

Using Eq. (22) and Eq. (17), this further simplifies as

$$\sum_{\mathbf{z}, \mathbf{v}} \frac{1}{2^{|F_2|}} \mathbb{1}_{\{\mathbf{v}_{F_1} = \breve{\mathbf{u}}_{F_1}\}} \left| q(\mathbf{w} \cdot \mathbf{z}) \prod_{i \in F_3} \breve{p}(v_i | \mathbf{z}, \mathbf{v}_1^{i-1}) \right.$$
$$\left. - \frac{1}{2^{|F_3|}} W_{X^n | V^n}(\mathbf{z} | \mathbf{v} G_n) \right|$$
$$= \mathbb{V}\left(\breve{p}_{X^n \cup^n}, \hat{p}_{X^n \cup^n | \cup_{F_1} = \breve{\mathbf{u}}_{F_1}}\right),$$

where the last inequality follows because q_{X^n} is the uniform distribution on $\mathcal{X}^n$. ∎

Combining the results of Lemma 3, Lemma 4, and Lemma 5 with Eq. (6), we conclude that the proposed coding scheme is a resolvability code.

V. DISCUSSION

In general, the achievable coordination region with polar codes given in Proposition 2 is strictly smaller than the coordination capacity region given in Theorem 1 because
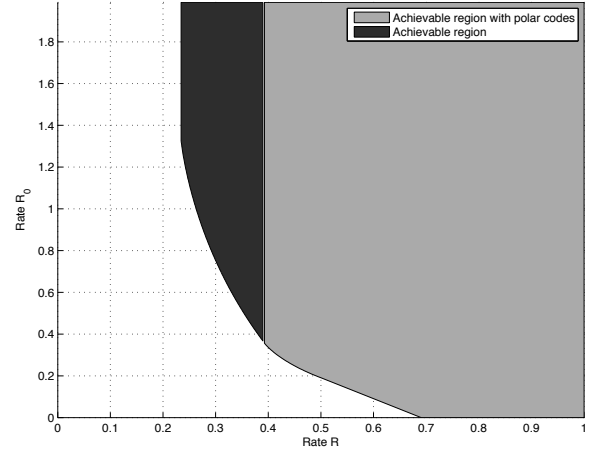

Fig. 3. Example of achievable rates for coordination with and without polar codes.

of the constraints **C1**, **C2**, and **C3**, on the random variable V (see Section IV-B).

As a first illustration, consider the situation in which $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{0, ?, 1\}$, $q_X \sim \mathcal{B}(\frac{1}{2})$ and $q_{Y|X}$ corresponds to the concatenation of a binary symmetric channel with crossover probability p with a binary erasure channel with erasure probability ϵ. In other words, the transition probability matrix corresponding to $q_{Y|X}$ is

$$\begin{pmatrix} (1-p)(1-\epsilon) & \epsilon & p(1-\epsilon) \\ p(1-\epsilon) & \epsilon & (1-p)(1-\epsilon) \end{pmatrix}$$

Because of condition **C2** (see Section IV-B), one can show that the boundary of the region of achievable rates (R, R_0) is characterized by

$$R \geqslant 1 - \mathbb{H}_b(q)$$
$$R_0 + R \geqslant (1 - \epsilon)\left(\mathbb{H}_b(p) - \mathbb{H}_b\left(\frac{p-q}{1-2q}\right)\right) + 1 - \mathbb{H}_b(q)$$

for $q \in [0, \min(\frac{1}{2}, p)]$. On the other hand, it is not difficult to show that the following rates are also admissible by Theorem 1.

$$R \geqslant (1 - \nu)(1 - \mathbb{H}_b(p))$$
$$R_0 + R \geqslant \mathbb{H}_b(\epsilon) + (1 - \epsilon)\mathbb{H}_b(p) + (1 - \nu)\mathbb{H}_b\left(\frac{\epsilon - \nu}{1 - \nu}\right)$$
$$+ (1 - \nu)(1 - \mathbb{H}_b(p))$$

for $\nu \in [0, \min(1, \epsilon)]$. The regions achievable with and without polar codes are illustrated in Figure 3, for the case $\epsilon = 0.4$ and $p = 0.15$.

As a second illustration, consider the situation in which $\mathcal{X} = \{0, 1\}$, $\mathcal{Y} = \{0, ?, 1\}$, and $q_{Y|X}$ corresponds to binary erasure channel with erasure probability ϵ. The coordination capacity region for this model is characterized in [11], and it is shown that the optimal choice of V such that $X \to V \to Y$ forms a Markov chain is a ternary random variable. In contrast, one can check that the only possible choice of such a V satisfying the constraints **C1**, **C2**, and **C3** is $V =$
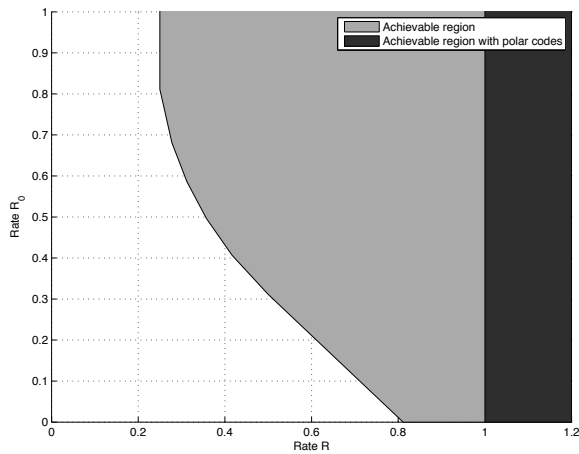
Fig. 4. Example of achievable rates for coordination with and without polar codes.

X. Consequently, the achievable coordination rate with polar codes is the trivial region $\{(R, R_0) : R_0 \geqslant 0, R \geqslant 1$, which is achievable without any coding. The regions are illustrated in Figure 4.

The generalization of the results beyond binary actions at node X can be carried out by leveraging known results about non-binary polar codes. However, the generalization to non-uniform actions and asymmetric channels seems much more challenging, since the proofs used in this paper heavily rely on the symmetry properties and uniformity of the actions to coordinate. Finding an explicit coordination scheme in a more general case remains an open problem and will be the topic of future research.

ACKNOWLEDGEMENT

REFERENCES

[1] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination Capacity," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4181–4206, 2010.
[2] E. Arikan, "Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.
[3] R. Blasco-Serrano, R. Thobaben, and M. Skoglund, "Polar Codes for Coordination in Cascade Networks," in *Proc. of International Zurich Seminar on Communications*, Zurich, Switzerland, March 2012, pp. 55–58.
[4] S. B. Korada and R. L. Urbanke, "Polar Codes are Optimal for Lossy Source Coding," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1751–1768, April 2010.
[5] T. Han and S. Verdú, "Approximation Theory of Output Statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.
[6] A. Wyner, "The Common Information of Two Dependent Random Variables," *IEEE Trans. Inf. Theory*, vol. 21, no. 2, pp. 163–179, March 1975.
[7] M. Hayashi, "General Nonasymptotic and Asymptotic Formulas in Channel Resolvability and Identification Capacity and their Application to the Wiretap Channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, April 2006.
[8] M. R. Bloch, "Achieving Secrecy: Capacity vs. Resolvability," in *Proc. of IEEE International Symposium on Information Theory*, Saint Petersburg, Russia, August 2011, pp. 632–636.
[9] 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, 2011.
[10] R. G. Gallager, *Information Theory and Reliable Communication*. Wyley, 1968.
[11] P. Cuff, "Communication requirements for generating correlated random variables," in *Proc. IEEE International Symposium on Information Theory ISIT 2008*, 2008, pp. 1393–1397.