# Polar Coding for Empirical and Strong Coordination via Distribution Approximation

Rémi A. Chou and Matthieu R. Bloch
School of Electrical and Computer Engineering
Georgia Institute of Technology

Jörg Kliewer
Department of Electrical and Computer Engineering
New Jersey Institute of Technology

***Abstract*—We design low-complexity polar codes for empirical and strong coordination in two-node network. Our constructions hinge on the observation that polar codes may be used to approximate distribution; which we leverage to prove that nested polar codes achieve the capacity region of empirical coordination and strong coordination.**

## I. Introduction

The information-theoretic limits of coordination in networks have 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 a sequence of coordinated actions to approach a target distribution, and strong coordination, which requires the total variational distance of a sequence 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.

In this paper, we demonstrate the ability of polar codes to provide a constructive alternative to the information-theoretic proof in [1] for two-node network. Specifically, the contributions of this paper are the following.

- We design a polar coding scheme that achieves the entire empirical coordination capacity region when common randomness, whose rate vanishes to zero as the block-length grows, is available at the nodes. This construction extends [2], which only deals with uniform distributions and requires a non negligible rate of common randomness available at the nodes,
- We construct a polar coding scheme that achieves the entire strong coordination capacity region. It generalizes [3], which only deals with uniform distributions and actions obtained via a symmetric channel.

The research was supported in part by NSF grants CCF-1320304 and CCF-1440014.

A key characteristic of our coding schemes is to rely on distribution approximation and to not require a reconstruction algorithm, such as the successive-cancellation decoder for source coding or channel coding schemes.

The remainder of the paper is organized as follows. Section II sets the notation. Section III formally introduces the problem. For a two-node network, Sections IV and V provide polar coding schemes that achieve the empirical coordination capacity region and the strong coordination capacity region, respectively.

## II. Notation

We define the integer interval $[\![a,b]\!]$, as the set of integers between $\lfloor a \rfloor$ and $\lceil b \rceil$. For $n \in \mathbb{N}$, we let $G_n \triangleq \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{\otimes n}$ be the source polarization transform defined in [4]. We note the components of a vector, $X^{1:N}$, of size $N$, with superscripts, i.e., $X^{1:N} \triangleq (X^1, X^2, \dots, X^N)$. For any set $\mathcal{A} \subset [\![1,N]\!]$, we note $X^{1:N}[\mathcal{A}]$ the components of $X^{1:N}$ whose indices are in $\mathcal{A}$. We note $\mathbb{V}(\cdot,\cdot)$ and $\mathbb{D}(\cdot||\cdot)$ the variational distance and the divergence, respectively, between two distributions. We note the conditional divergence w.r.t. $p_X$, $\mathbb{D}(\cdot||\cdot|p_X)$. Finally, we note the indicator function $\mathbb{1}\{\omega\}$, which is equal to 1 if the predicate $\omega$ is true and 0 otherwise.

## III. Problem Statement

Consider a source $(\mathcal{X}\mathcal{Y}, q_{XY})$ with $\mathcal{Y} \triangleq \{0,1\}$ and $\mathcal{X}$ a countable alphabet. Consider two nodes, Node 1 and Node 2.

**Definition 1.** *A $(2^{NR}, 2^{NR_0}, N)$ coordination code $\mathcal{C}_N$ for a fixed joint distribution $q_{XY}$ consists of*

- *common randomness $C$ with rate $R_0$;*
- *an encoding function $f_N : \mathcal{X}^N \times [\![1, 2^{nR_0}]\!] \to [\![1, 2^{NR}]\!]$;*
- *a decoding function $g_N : [\![1, 2^{NR}]\!] \times [\![1, 2^{nR_0}]\!] \to \mathcal{Y}^N$,*

*and operates as follows*

- *Node 1 observes $X^{1:N}$, $N$ i.i.d. realizations of $(\mathcal{X}, q_X)$;*
- *Node 1 transmits $f_N(X^{1:N}, C)$ to Node 2;*
- *Node 2 forms $\widetilde{Y}^{1:N} \triangleq g_N(f_N(X^{1:N}, C), C)$, whose joint distribution with $X^{1:N}$ is denoted $\widetilde{p}_{X^{1:N}Y^{1:N}}$.*

**Definition 2.** *A rate pair $(R, R_0)$ for a fixed joint distribution $q_{XY}$ is achievable for empirical coordination if there exists a sequence of $(2^{NR}, 2^{NR_0}, N)$ coordination codes, $\{\mathcal{C}_N\}_{N \geqslant 1}$ such that for $\epsilon > 0$*

$$\lim_{N \to \infty} \mathbb{P}[\mathbb{V}(q_{XY}, T_{X^{1:N}\widetilde{Y}^{1:N}}) > \epsilon] = 0,$$

*where for a sequence* $(x^{1:N}, \widetilde{y}^{1:N})$ *generated at Nodes 1, 2, and for* $(x,y) \in \mathcal{X} \times \mathcal{Y}$,

$$T_{x^{1:N}\widetilde{y}^{1:N}}(x,y) \triangleq \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\{(x^i, \widetilde{y}^i) = (x,y)\}.$$

*The closure of the set of achievable rates is called the empirical coordination capacity region.*

**Definition 3.** *A rate pair* $(R, R_0)$ *for a fixed joint distribution* $q_{XY}$ *is achievable for strong coordination if there exists a sequence of* $(2^{NR}, 2^{NR_0}, N)$ *coordination codes,* $\{\mathcal{C}_N\}_{N\geqslant 1}$ *such that for* $\epsilon > 0$

$$\lim_{N\to\infty} \mathbb{V}(\widetilde{p}_{X^{1:N}Y^{1:N}}, q_{X^{1:N}Y^{1:N}}) = 0.$$

*The closure of the set of achievable rate pairs is called the strong coordination capacity region.*

The coordination capacity regions have been fully characterized using random coding arguments.

**Theorem 1** ([1])**.** *The empirical coordination capacity region is*

$$\mathcal{R}_{\text{EC}}(q_{XY}) \triangleq \{(R,R_0) : R \geqslant I(X;Y),\ R_0 = 0\}.$$

**Theorem 2** ([1])**.** *The strong coordination capacity region is*

$$\mathcal{R}_{\text{SC}}(q_{XY}) \triangleq \bigcup_{X\to V\to Y} \{(R,R_0) : R{+}R_0 \geqslant I(XY;V),\ R \geqslant I(X;V)\}.$$

## IV. Polar Coding for Empirical Coordination

The high-level idea of the coding scheme can be summarized as follows. From $X^{1:N}$ and some randomness $C_1$ shared with Node 2, Node 1 constructs a random variable $\widetilde{Y}^{1:N}$ whose joint probability distribution with $X^{1:N}$ is close to the target distribution $q_{X^{1:N}Y^{1:N}}$. Moreover, Node 1 can construct a message $M$ with rate close to $I(X;Y)$ such that Node 2 can reconstruct $\widetilde{Y}^{1:N}$ with $M$ and $C_1$. Finally, by performing encoding over $k \in \mathbb{N}^*$ blocks with the same randomness $C_1$, the overall rate of shared randomness vanishes to zero as the number of blocks increases.

Note that the coding scheme for each block is similar to lossy source coding schemes [5], [6], as suggested by the optimal communication rate described in Theorem 1. However, the performance metric of interest is totally different.

### A. Coding Scheme

Consider the random variables $X$, $Y$ distributed according to the fixed p.m.f. $q_{XY}$. Let $N \triangleq 2^n$, $n \in \mathbb{N}^*$. Define $U^{1:N} \triangleq Y^{1:N} G_n$ and define for $\beta < 1/2$, $\delta_N \triangleq 2^{-N^\beta}$, the sets

$$\begin{aligned}
\mathcal{V}_Y &\triangleq \left\{ i \in [\![1,N]\!] : H(U^i | U^{1:i-1}) > 1 - \delta_N \right\}, \\
\mathcal{H}_Y &\triangleq \left\{ i \in [\![1,N]\!] : H(U^i | U^{1:i-1}) > \delta_N \right\}, \\
\mathcal{V}_{Y|X} &\triangleq \left\{ i \in [\![1,N]\!] : H(U^i | U^{1:i-1} X^{1:N}) > 1 - \delta_N \right\}, \\
\mathcal{H}_{Y|X} &\triangleq \left\{ i \in [\![1,N]\!] : H(U^i | U^{1:i-1} X^{1:N}) > \delta_N \right\}.
\end{aligned}$$

Encoding is performed over $k \in \mathbb{N}^*$ blocks of length $N$. We use the subscript $i \in [\![1,k]\!]$ to denote random variables associated to encoding Block $i$. The encoding and decoding procedures are described in Algorithms 1, 2, respectively.

**Algorithm 1** Encoding algorithm at Node 1 for empirical coordination

**Require:** A vector $C_1$ of $|\mathcal{V}_{Y|X}|$ uniformly distributed bits shared with Node 2 and $X_{1:k}^{1:N}$.

**for** Block $i = 1$ to $k$ **do**
  $C_i \leftarrow C_1$
  $\widetilde{U}_i^{1:N}[\mathcal{V}_{Y|X}] \leftarrow C_i$
  Given $X_i^{1:N}$, successively draw the remaining bits of $\widetilde{U}_i^{1:N}$ according to $\widetilde{p}_{U_i^{1:N}X_i^{1:N}}$ defined by

$$\widetilde{p}_{U_i^j|U_i^{1:j-1}X_i^{1:N}}(u_i^j|\widetilde{U}_i^{1:j-1}X_i^{1:N}) \triangleq \begin{cases} q_{U^j|U^{1:j-1}X^{1:N}}(u_i^j|\widetilde{U}_i^{1:j-1}X_i^{1:N}) & \text{if } j \in \mathcal{H}_Y \backslash \mathcal{V}_{Y|X} \\ q_{U^j|U^{1:j-1}}(u_i^j|\widetilde{U}_i^{1:j-1}) & \text{if } j \in \mathcal{H}_Y^c \end{cases} \quad (1)$$

  Transmit $M_i \triangleq \widetilde{U}_i^{1:N}[\mathcal{H}_Y \backslash \mathcal{V}_{Y|X}]$ and $\widetilde{C}_i$, the randomness necessary to draw $\widetilde{U}_i^{1:N}[\mathcal{H}_Y^c]$, to Node 2.
**end for**

**Algorithm 2** Decoding algorithm at Node 2 for empirical coordination

**Require:** The vector $C_1$ used in Algorithm 1 and $M_{1:k}$.

**for** Block $i = 1$ to $k$ **do**
  $C_i \leftarrow C_1$
  $\widetilde{U}_i^{1:N}[\mathcal{V}_{Y|X}] \leftarrow C_i$
  $\widetilde{U}_i^{1:N}[\mathcal{H}_Y \backslash \mathcal{V}_{Y|X}] \leftarrow M_i$
  Using $\widetilde{C}_i$, successively draw the remaining bits of $\widetilde{U}_i^{1:N}$ according to $q_{U^j|U^{1:j-1}}$
**end for**

### B. Scheme Analysis

The following lemma shows that $\widetilde{p}_{Y^{1:N}X^{1:N}}$, defined by $\widetilde{p}_{X^{1:N}} \triangleq q_{X^{1:N}}$ and Equation (1), approximates $q_{X^{1:N}Y^{1:N}}$.

**Lemma 1.** *For any* $i \in [\![1,k]\!]$,

$$\mathbb{V}(q_{X^{1:N}Y^{1:N}}, \widetilde{p}_{X_i^{1:N}Y_i^{1:N}}) \leqslant \sqrt{2\log 2}\sqrt{N\delta_N}.$$

The proof of Lemma 1 is similar to the proof of Lemma 4 and is thus omitted. The following lemma shows that empirical coordination holds for each block.

**Lemma 2.** *For* $i \in [\![1,k]\!]$, *we have*

$$\lim_{N\to\infty} \mathbb{P}[\mathbb{V}(q_{XY}, T_{X_i^{1:N}\widetilde{Y}_i^{1:N}}) > \epsilon] = 0.$$

*Proof.* For $\epsilon > 0$, define $\mathcal{T}_\epsilon(q_{XY}) \triangleq \{(x^{1:N}, y^{1:N}) : \mathbb{V}(q_{XY}, T_{x^{1:N}y^{1:N}}) \leqslant \epsilon\}$. We note for a joint distribution $q$ over $(\mathcal{X} \times \mathcal{Y})$,

$$\mathbb{P}_q[(X^{1:N}, Y^{1:N}) \in \mathcal{T}_\epsilon(q_{XY})] \triangleq \sum_{x^{1:N}, y^{1:N}} q_{X^{1:N}Y^{1:N}}(x^{1:N}, y^{1:N}) \mathbb{1}\{(x^{1:N}, y^{1:N}) \in \mathcal{T}_\epsilon(q_{XY})\}.$$

Let $i \in [\![1,k]\!]$, we have

$$\mathbb{P}_{\widetilde{p}}[\mathbb{V}(q_{XY}, T_{X_i^{1:N}\widetilde{Y}_i^{1:N}}) > \epsilon]$$

$$= \sum_{x^{1:N},y^{1:N}} \Big[\widetilde{p}_{X_i^{1:N}Y_i^{1:N}}(x^{1:N},y^{1:N}) - q_{X^{1:N}Y^{1:N}}(x^{1:N},y^{1:N}) + q_{X^{1:N}Y^{1:N}}(x^{1:N},y^{1:N})\Big] \mathbb{1}\{(x^{1:N},y^{1:N}) \notin \mathcal{T}_\epsilon(q_{XY})\}$$
$$\leqslant \mathbb{V}(\widetilde{p}_{X_i^{1:N}Y_i^{1:N}}, q_{X^{1:N}Y^{1:N}}) + \mathbb{P}_q[(X^{1:N},Y^{1:N}) \notin \mathcal{T}_\epsilon(q_{XY})]$$
$$\xrightarrow{N\to\infty} 0,$$

where we have used Lemma 1 and the AEP for strongly typical set [7]. ∎

We now show that empirical coordination holds for all blocks jointly.

**Lemma 3.** *We have*

$$\lim_{N\to\infty} \mathbb{P}[\mathbb{V}(q_{XY}, T_{X_{1:k}^{1:N}\widetilde{Y}_{1:k}^{1:N}}) > \epsilon] = 0.$$

*Proof.* We have

$$\mathbb{V}(q_{XY}, T_{x_{1:k}^{1:N}\widetilde{y}_{1:k}^{1:N}})$$
$$= \sum_{x,y} \left| q_{XY}(x,y) - \frac{1}{kN}\sum_{j=1}^{k}\sum_{i=1}^{N} \mathbb{1}\{(x_j^i,\widetilde{y}_j^i) = (x,y)\} \right|$$
$$= \sum_{x,y} \left| \sum_{j=1}^{k} \left( \frac{1}{k} q_{XY}(x,y) - \frac{1}{kN}\sum_{i=1}^{N} \mathbb{1}\{(x_j^i,\widetilde{y}_j^i) = (x,y)\} \right) \right|$$
$$\leqslant \frac{1}{k}\sum_{j=1}^{k}\sum_{x,y} \left| q_{XY}(x,y) - \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\{(x_j^i,\widetilde{y}_j^i) = (x,y)\} \right|$$
$$\leqslant \frac{1}{k}\sum_{j=1}^{k} \mathbb{V}(q_{XY}, T_{x_j^{1:N}\widetilde{y}_j^{1:N}}),$$

hence,

$$\mathbb{E}_{\widetilde{p}_{X_{1:k}^{1:N},Y_{1:k}^{1:N}}}[\mathbb{V}(q_{XY}, T_{x_{1:k}^{1:N}\widetilde{y}_{1:k}^{1:N}})]$$
$$\leqslant \frac{1}{k}\sum_{j=1}^{k} \mathbb{E}_{\widetilde{p}_{X_{1:k}^{1:N},Y_{1:k}^{1:N}}}[\mathbb{V}(q_{XY}, T_{x_j^{1:N}\widetilde{y}_j^{1:N}})] \xrightarrow{N\to\infty} 0,$$

where the limit holds by Lemma 2 and because convergence in probability and uniform integrability implies convergence in the mean. We then obtain the result because convergence in the mean implies convergence in probability. ∎

**Theorem 3.** *The coding scheme described in Algorithms 1, 2 achieves the empirical coordination capacity region.*

*Proof.* The communication rate is

$$\frac{k|\mathcal{H}_Y\backslash\mathcal{V}_{Y|X}|}{kN} \leqslant \frac{|\mathcal{V}_Y\backslash\mathcal{V}_{Y|X}|+|\mathcal{H}_Y\backslash\mathcal{V}_Y|}{N}$$
$$= \frac{|\mathcal{V}_Y|-|\mathcal{V}_{Y|X}|+|\mathcal{H}_Y|-|\mathcal{V}_Y|}{N} \xrightarrow{N\to\infty} I(X;Y),$$

where we have used [4] and [8, Lemma 1].

Node 1 also communicates randomness to reconstruct $U^{1:N}[\mathcal{H}_Y^c]$, but its rate is $o(N)$ since

$$\lim_{N\to\infty} \frac{1}{N}\sum_{j\in\mathcal{H}_Y^c} H(\widetilde{U}^j|\widetilde{U}^{1:j-1}) = 0,$$

which can be shown using Lemma 1 similarly to [9, Lemma 9]. Then, the common randomness rate is

$$\frac{|\mathcal{V}_{Y|X}|}{kN} \xrightarrow{N\to\infty} \frac{H(Y|X)}{k} \xrightarrow{k\to\infty} 0,$$

where we have used [8, Lemma 1]. Finally, we conclude with Lemma 3. ∎

## V. Polar Coding for Strong Coordination

The principle of the strong coordination coding scheme can be summarized as follows. From $X^{1:N}$ and some randomness $(C_1,\bar{C}_1)$ shared with Node 2, Node 1 constructs a random variable $\widetilde{V}^{1:N}$ whose joint probability distribution with $X^{1:N}$ is close to the target distribution $q_{X^{1:N}V^{1:N}}$. Moreover, Node 1 can construct a message $M$ with rate close to $I(X;V)$ such that Node 2 can reconstruct $\widetilde{V}^{1:N}$ with $M$ and $(C_1,\bar{C}_1)$. Then, Node 2 performs channel prefixing on $\widetilde{V}^{1:N}$ to form $\widetilde{Y}^{1:N}$ whose joint distribution with $X^{1:N}$ is close to $q_{X^{1:N}Y^{1:N}}$. Finally, by performing encoding over $k\in\mathbb{N}^*$ blocks with the same randomness $\bar{C}_1$, the overall rate of shared randomness becomes the rate of $C_1$, which can be chosen on the order of $I(V;Y|X)$.

### A. Coding Scheme

Consider the random variables $X$, $Y$, $V$ distributed according to the fixed p.m.f. $q_{XYV}$ such that $X\to V\to Y$ and $|\mathcal{X}|=|\mathcal{Y}|=|\mathcal{V}|=2$. Let $N\triangleq 2^n$, $n\in\mathbb{N}^*$. Define $U^{1:N}\triangleq V^{1:N}G_n$, $T^{1:N}\triangleq Y^{1:N}G_n$, and define for $\beta<1/2$ and $\delta_N\triangleq 2^{-N^\beta}$ the sets

$$\mathcal{H}_V \triangleq \left\{i\in[\![1,N]\!] : H(U^i|U^{1:i-1}) > \delta_N\right\},$$
$$\mathcal{V}_{V|X} \triangleq \left\{i\in[\![1,N]\!] : H(U^i|U^{1:i-1}X^{1:N}) > 1-\delta_N\right\},$$
$$\mathcal{V}_{V|XY} \triangleq \left\{i\in[\![1,N]\!] : H(U^i|U^{1:i-1}X^{1:N}Y^{1:N}) > 1-\delta_N\right\},$$
$$\mathcal{V}_{Y|V} \triangleq \left\{i\in[\![1,N]\!] : H(T^i|T^{1:i-1}V^{1:N}) > 1-\delta_N\right\}.$$

Note that $\mathcal{V}_{V|XY}\subset\mathcal{V}_{V|X}\subset\mathcal{V}_V$.

We define $\mathcal{F}_1\triangleq\mathcal{H}_V^c$, $\mathcal{F}_2\triangleq\mathcal{V}_{V|XY}$, $\mathcal{F}_3\triangleq\mathcal{V}_{V|X}\backslash\mathcal{V}_{V|XY}$, and $\mathcal{F}_4\triangleq\mathcal{H}_V\backslash\mathcal{V}_{V|X}$. Observe that $(\mathcal{F}_1,\mathcal{F}_2,\mathcal{F}_3,\mathcal{F}_4)$ forms a partition of $[\![1,N]\!]$. Encoding is performed over $k\in\mathbb{N}^*$ blocks of length $N$. We use the subscript $i\in[\![1,k]\!]$ to denote random variables associated to encoding Block $i$. The encoding and decoding procedures are described in Algorithms 3, 4, respectively.

### B. Scheme Analysis

The following lemma shows that $\widetilde{p}_{V^{1:N}X^{1:N}}$, defined by $\widetilde{p}_{X^{1:N}}\triangleq q_{X^{1:N}}$ and Equation (2), approximates $q_{V^{1:N}X^{1:N}}$.

**Lemma 4.** *For any* $i\in[\![1,k]\!]$,

$$\mathbb{V}(q_{V^{1:N}X^{1:N}}, \widetilde{p}_{V_i^{1:N}X_i^{1:N}}) \leqslant \delta_N^{(A)},$$

*where* $\delta_N^{(A)}\triangleq\sqrt{2\log 2}\sqrt{N\delta_N}$.

*Proof.* We have

$$\mathbb{D}(q_{V^{1:N}X^{1:N}}, \widetilde{p}_{V_i^{1:N}X_i^{1:N}})$$
$$\overset{(a)}{=} \mathbb{D}(q_{U^{1:N}X^{1:N}}||\widetilde{p}_{U_i^{1:N}X_i^{1:N}})$$

**Algorithm 3** Encoding algorithm at Node 1 for strong coordination

**Require:** A vector $C_{1:k}$ of $k|\mathcal{F}_3|$ uniformly distributed bits shared with Node 2. A vector $\bar{C}_1$ of $|\mathcal{F}_2|$ uniformly distributed bits shared with Node 2 and $X_{1:k}^{1:N}$.

1: **for** Block $i = 1$ to $k$ **do**

2: $\bar{C}_i \leftarrow \bar{C}_1$

3: $\widetilde{U}_i^{1:N}[\mathcal{F}_2] \leftarrow \bar{C}_i$

4: $\widetilde{U}_i^{1:N}[\mathcal{F}_3] \leftarrow C_i$

5: Given $X_i^{1:N}$, successively draw the remaining bits of $\widetilde{U}_i^{1:N}$ according to $\widetilde{p}_{U_i^{1:N} X_i^{1:N}}$ defined by

$$\widetilde{p}_{U_i^j | U_i^{1:j-1} X_i^{1:N}}(u_i^j | \widetilde{U}_i^{1:j-1} X_i^{1:N}) \triangleq \begin{cases} q_{U^j|U^{1:j-1}}(u_i^j|\widetilde{U}_i^{1:j-1}) & \text{if } j \in \mathcal{F}_1 \\ q_{U^j|U^{1:j-1}X^{1:N}}(u_i^j|\widetilde{U}_i^{1:j-1} X_i^{1:N}) & \text{if } j \in \mathcal{F}_4 \end{cases} \tag{2}$$

6: Transmit $M_i \triangleq \widetilde{U}_j^{1:N}[\mathcal{F}_4]$ and $C_i'$, the randomness necessary to draw $\widetilde{U}_i^{1:N}[\mathcal{F}_1]$, to Node 2.

7: **end for**

**Algorithm 4** Decoding algorithm at Node 2 for strong coordination

**Require:** The vectors $C_{1:k}$ and $\bar{C}_1$ used in Algorithm 3 and $M_{1:k}$.

1: **for** Block $i = 1$ to $k$ **do**

2: $\bar{C}_i \leftarrow \bar{C}_1$

3: $\widetilde{U}_i^{1:N}[\mathcal{F}_2] \leftarrow \bar{C}_i$

4: $\widetilde{U}_i^{1:N}[\mathcal{F}_3] \leftarrow C_i$

5: $\widetilde{U}_i^{1:N}[\mathcal{F}_4] \leftarrow M_i$

6: Using $C_i'$, successively draw the remaining bits of $\widetilde{U}_i^{1:N}$ according to $q_{U^j|U^{1:j-1}}$

7: $\widetilde{V}_i^{1:N} \leftarrow \widetilde{U}_i^{1:N} G_n$

8: Channel prefixing: Given $\widetilde{V}_i^{1:N}$, successively draw the bits of $\widetilde{T}_i^{1:N}$ according to $\widetilde{p}_{T_i^{1:N} V_i^{1:N}}$ defined by

$$\widetilde{p}_{T_i^j | T_i^{1:j-1} V_i^{1:N}}(t_i^j | \widetilde{T}_i^{1:j-1} \widetilde{V}_i^{1:N}) \triangleq \begin{cases} 1/2 & \text{if } j \in \mathcal{V}_{Y|V} \\ q_{T^j|T^{1:j-1}V^{1:N}}(t_i^j|\widetilde{T}_i^{1:j-1} \widetilde{V}_i^{1:N}) & \text{if } j \in \mathcal{V}_{Y|V}^c \end{cases}$$

9: $\widetilde{Y}_i^{1:N} \leftarrow \widetilde{T}_i^{1:N} G_n$

10: **end for**

$$\begin{aligned}
&\overset{(b)}{=} \mathbb{D}(q_{U^{1:N}|X^{1:N}} || \widetilde{p}_{U_i^{1:N}|X_i^{1:N}} | q_{X^{1:N}}) \\
&\overset{(c)}{=} \sum_{j=1}^{N} \mathbb{D}(q_{U^j|U^{1:j-1}X^{1:N}} || \widetilde{p}_{U_i^j|U_i^{1:j-1}X_i^{1:N}} | q_{U^{1:j-1}X^{1:N}}) \\
&\overset{(d)}{=} \sum_{j \in \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3} \mathbb{D}(q_{U^j|U^{1:j-1}X^{1:N}} || \widetilde{p}_{U_i^j|U_i^{1:j-1}X_i^{1:N}} | q_{U^{1:j-1}X^{1:N}}) \\
&\overset{(e)}{=} \sum_{j \in \mathcal{V}_{V|X}} (1 - H(U^j|U^{1:j-1}X^{1:N})) \\
&\quad + \sum_{j \in \mathcal{H}_V^c} (H(U^j|U^{1:j-1}) - H(U^j|U^{1:j-1}X^{1:N})) \\
&\leqslant |\mathcal{V}_{V|X}|\delta_N + |\mathcal{H}_V^c|\delta_N \leqslant N\delta_N,
\end{aligned}$$

where $(a)$ holds by invertibility of $G_n$, $(b)$ and $(c)$ hold by the chain rule for divergence [7], $(d)$ and $(e)$ hold by (2). Finally, we conclude with Pinsker's inequality. ■

We now show that strong coordination holds for each block in the following lemma.

**Lemma 5.** *For $i \in [\![1, k]\!]$, we have*

$$\mathbb{V}(\widetilde{p}_{X_i^{1:N} Y_i^{1:N}}, q_{X^{1:N}Y^{1:N}}) \leqslant \mathbb{V}(\widetilde{p}_{V_i^{1:N} X_i^{1:N} Y_i^{1:N}}, q_{V^{1:N}X^{1:N}Y^{1:N}}) \leqslant \delta_N^{(B)},$$

*where $\delta_N^{(B)} \triangleq (\sqrt{2}+2)\sqrt{2\log 2}\sqrt{N\delta_N}$.*

*Proof.* We have

$$\begin{aligned}
&\mathbb{V}(\widetilde{p}_{V_i^{1:N} X_i^{1:N} Y_i^{1:N}}, q_{V^{1:N}X^{1:N}Y^{1:N}}) \\
&= \mathbb{V}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}X_i^{1:N}} \widetilde{p}_{V_i^{1:N}X_i^{1:N}}, q_{Y^{1:N}|V^{1:N}X^{1:N}} q_{V^{1:N}X^{1:N}}) \\
&= \mathbb{V}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}} \widetilde{p}_{V_i^{1:N}X_i^{1:N}}, q_{Y^{1:N}|V^{1:N}} q_{V^{1:N}X^{1:N}}) \\
&\overset{(a)}{\leqslant} \mathbb{V}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}} \widetilde{p}_{V_i^{1:N}X_i^{1:N}}, \widetilde{p}_{Y_i^{1:N}|V_i^{1:N}} q_{V^{1:N}X^{1:N}}) \\
&\quad + \mathbb{V}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}} q_{V^{1:N}X^{1:N}}, q_{Y^{1:N}|V^{1:N}} q_{V^{1:N}X^{1:N}}) \\
&= \mathbb{V}(\widetilde{p}_{V_i^{1:N}X_i^{1:N}}, q_{V^{1:N}X^{1:N}}) \\
&\quad + \mathbb{V}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}} q_{V^{1:N}}, q_{Y^{1:N}V^{1:N}}) \\
&\overset{(b)}{\leqslant} \delta_N^{(A)} + \mathbb{V}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}} q_{V^{1:N}}, q_{Y^{1:N}V^{1:N}}) \\
&\overset{(c)}{\leqslant} \delta_N^{(A)} + \mathbb{V}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}} q_{V^{1:N}}, \widetilde{p}_{Y_i^{1:N}V_i^{1:N}}) \\
&\quad + \mathbb{V}(\widetilde{p}_{Y_i^{1:N}V_i^{1:N}}, q_{Y^{1:N}V^{1:N}}) \\
&= \delta_N^{(A)} + \mathbb{V}(q_{V^{1:N}}, \widetilde{p}_{V_i^{1:N}}) + \mathbb{V}(\widetilde{p}_{Y_i^{1:N}V_i^{1:N}}, q_{Y^{1:N}V^{1:N}}) \\
&\overset{(d)}{\leqslant} 2\delta_N^{(A)} + \mathbb{V}(\widetilde{p}_{Y_i^{1:N}V_i^{1:N}}, q_{Y^{1:N}V^{1:N}}),
\end{aligned} \tag{3}$$

where $(a)$ and $(c)$ hold by the triangle inequality, $(b)$ and $(d)$ hold by Lemma 4. We bound the r.h.s. of (3) via its associated divergence by analyzing Step 8. of Algorithm 4 as follows.

$$\begin{aligned}
&\mathbb{D}(q_{Y^{1:N}V^{1:N}} || \widetilde{p}_{Y_i^{1:N}V_i^{1:N}}) \\
&\overset{(e)}{=} \mathbb{D}(q_{T^{1:N}|V^{1:N}} || \widetilde{p}_{T_i^{1:N}|V_i^{1:N}} | q_{V^{1:N}}) + \mathbb{D}(q_{U^{1:N}} || \widetilde{p}_{U_i^{1:N}}) \\
&\overset{(f)}{\leqslant} \mathbb{D}(q_{T^{1:N}|V^{1:N}} || \widetilde{p}_{T_i^{1:N}|V_i^{1:N}} | q_{V^{1:N}}) \\
&\quad + \mathbb{D}(q_{U^{1:N}X^{1:N}} || \widetilde{p}_{U_i^{1:N}X_i^{1:N}}) \\
&\overset{(g)}{\leqslant} N\delta_N + \mathbb{D}(q_{T^{1:N}|V^{1:N}} || \widetilde{p}_{T_i^{1:N}|V_i^{1:N}} | q_{V^{1:N}}) \\
&\overset{(h)}{=} N\delta_N + \sum_{j=1}^{N} \mathbb{D}(q_{T^j|T^{1:j-1}V^{1:N}} || \widetilde{p}_{T_i^j|T_i^{1:j-1}V_i^{1:N}} | q_{T^{1:j-1}V^{1:N}}) \\
&= N\delta_N + \sum_{j \in \mathcal{V}_{Y|V}} \mathbb{D}(q_{T^j|T^{1:j-1}V^{1:N}} || \widetilde{p}_{T_i^j|T_i^{1:j-1}V_i^{1:N}} | q_{T^{1:j-1}V^{1:N}})
\end{aligned}$$

$$= N\delta_N + \sum_{j \in \mathcal{V}_{Y|V}} (1 - H(T^j | T^{1:j-1} V^{1:N}))$$

$$\overset{(i)}{\leqslant} N\delta_N + |\mathcal{V}_{Y|V}| \delta_N \leqslant 2N\delta_N, \tag{4}$$

where $(e)$ holds by invertibility of $G_n$ and the chain rule for divergence, $(f)$ holds by the chain rule for divergence and positivity of the divergence, $(g)$ holds by the proof of Lemma 4, $(h)$ holds by the chain rule for divergence, $(i)$ holds by definition of $\mathcal{V}_{Y|V}$. Finally, combining (3), (4), and Pinsker's inequality yields the result. ∎

Using Lemma 5, we show in the following lemma that an asymptotic independence result holds for two consecutive blocks.

**Lemma 6.** *For $i \in [\![2, k]\!]$, we have,*

$$\mathbb{V}\left(\widetilde{p}_{X^{1:N}_{i-1:i} Y^{1:N}_{i-1:i} \bar{C}_1}, \widetilde{p}_{Y^{1:N}_{i-1} X^{1:N}_{i-1}} \widetilde{p}_{X^{1:N}_i Y^{1:N}_i \bar{C}_1}\right) \leqslant \delta_N^{(C)},$$

*where* $\delta_N^{(C)} \triangleq \sqrt{2\log 2}\sqrt{N\delta_N + 2\delta_N^{(B)}\left(3N - \log \delta_N^{(B)}\right)}$.

*Proof.* Let $i \in [\![1, k]\!]$. We have for $N$ large enough

$$\left| H(\widetilde{U}_i^{1:N}[\mathcal{V}_{V|XY}] | \widetilde{X}_i^{1:N} \widetilde{Y}_i^{1:N}) - H(U^{1:N}[\mathcal{V}_{V|XY}] | X^{1:N} Y^{1:N}) \right|$$
$$\overset{(a)}{\leqslant} \left| H(\widetilde{U}_i^{1:N}[\mathcal{V}_{V|XY}] \widetilde{X}_i^{1:N} \widetilde{Y}_i^{1:N}) - H(U^{1:N}[\mathcal{V}_{V|XY}] X^{1:N} Y^{1:N}) \right| + \left| H(\widetilde{X}_i^{1:N} \widetilde{Y}_i^{1:N}) - H(X^{1:N} Y^{1:N}) \right|$$
$$\overset{(b)}{\leqslant} \mathbb{V}(\widetilde{p}_{U_i^{1:N}[\mathcal{V}_{V|XY}] X_i^{1:N} Y_i^{1:N}}, q_{U^{1:N}[\mathcal{V}_{V|XY}] X^{1:N} Y^{1:N}}) \times \log \frac{2^{3N}}{\mathbb{V}(\widetilde{p}_{U_i^{1:N}[\mathcal{V}_{V|XY}] X_i^{1:N} Y_i^{1:N}}, q_{U^{1:N}[\mathcal{V}_{V|XY}] X^{1:N} Y^{1:N}})} + \left| H(\widetilde{X}_i^{1:N} \widetilde{Y}_i^{1:N}) - H(X^{1:N} Y^{1:N}) \right|$$
$$\overset{(c)}{\leqslant} \delta_N^{(B)}(3N - \log \delta_N^{(B)}) + \left| H(\widetilde{X}_i^{1:N} \widetilde{Y}_i^{1:N}) - H(X^{1:N} Y^{1:N}) \right|$$
$$\leqslant 2\delta_N^{(B)}\left(3N - \log \delta_N^{(B)}\right) \triangleq \delta_N^{(VXY)},$$

where $(a)$ holds by the triangle inequality, $(b)$ follows from [10, Lemma 2.7], $(c)$ holds by Lemma 5, invertibility of $G_n$, and because $x \mapsto x \log x$ is decreasing for $x > 0$ small enough.

Hence,

$$I(\widetilde{X}_i^{1:N} \widetilde{Y}_i^{1:N}; \widetilde{U}_i^{1:N}[\mathcal{V}_{V|XY}])$$
$$\overset{(d)}{=} |\mathcal{V}_{V|XY}| - H(\widetilde{U}_i^{1:N}[\mathcal{V}_{V|XY}] | \widetilde{X}_i^{1:N} \widetilde{Y}_i^{1:N})$$
$$\leqslant |\mathcal{V}_{V|XY}| - H(U^{1:N}[\mathcal{V}_{V|XY}] | X^{1:N} Y^{1:N}) + \delta_N^{(VXY)}$$
$$\overset{(e)}{\leqslant} |\mathcal{V}_{V|XY}| - \sum_{j \in \mathcal{V}_{V|XY}} H(U^j | U^{1:j-1} X^{1:N} Y^{1:N}) + \delta_N^{(VXY)}$$
$$\overset{(f)}{\leqslant} |\mathcal{V}_{V|XY}| - |\mathcal{V}_{V|XY}|(1 - \delta_N) + \delta_N^{(VXY)} \leqslant N\delta_N + \delta_N^{(VXY)},$$

where $(d)$ holds by uniformity of $\widetilde{U}_i^{1:N}[\mathcal{V}_{V|XY}]$, $(e)$ holds because conditioning reduces entropy, $(f)$ holds by definition of $\mathcal{V}_{V|XY}$.

Then, for $i \in [\![2, k]\!]$, we can show

$$\mathbb{D}\left(\widetilde{p}_{X^{1:N}_{i-1:i} Y^{1:N}_{i-1:i} \bar{C}_1} || \widetilde{p}_{Y^{1:N}_{i-1} X^{1:N}_{i-1}} \widetilde{p}_{X^{1:N}_i Y^{1:N}_i \bar{C}_1}\right)$$
$$= I(\widetilde{X}_{i-1}^{1:N} \widetilde{Y}_{i-1}^{1:N}; \bar{C}_1 \widetilde{X}_i^{1:N} \widetilde{Y}_i^{1:N})$$
$$= I(\widetilde{X}_{i-1}^{1:N} \widetilde{Y}_{i-1}^{1:N}; \widetilde{U}_{i-1}^{1:N}[\mathcal{V}_{V|XY}]) \leqslant N\delta_N + \delta_N^{(VXY)}.$$

∎

Using Lemma 6 we can show an asymptotical independence result for all blocks as stated in the following lemma.

**Lemma 7.** *We have*

$$\mathbb{V}\left(\widetilde{p}_{X^{1:N}_{1:k} Y^{1:N}_{1:k}}, \prod_{i=1}^{k} \widetilde{p}_{X^{1:N}_i Y^{1:N}_i}\right) \leqslant k\delta_N^{(C)}.$$

*where $\delta_N^{(C)}$ is defined in Lemma 6.*

Using Lemmas 5, 7, we can show that strong coordination holds over all blocks as stated in the following lemma.

**Lemma 8.** *We have*

$$\mathbb{V}\left(\widetilde{p}_{X^{1:N}_{1:k} Y^{1:N}_{1:k}}, q_{X^{1:kN} Y^{1:kN}}\right) \leqslant k\delta_N^{(D)}.$$

*where* $\delta_N^{(C)} \triangleq \delta_N^{(B)} + \delta_N^{(C)}$.

Finally, using [4], [8, Lemma 1], and using Lemma 5 as in the proof of [9, Lemma 9], we can prove that the communication rate and the common randomness rate are optimal. Our final result is stated as follows.

**Theorem 4.** *The coding scheme described in Algorithms 3, 4 achieves strong coordination for the rate pair $(R, R_0)$ such that $R + R_0 = I(XY; V)$, $R = I(X; V)$ for any p.m.f. $q_{XYV} = q_{XV} q_{Y|V}$ with $|\mathcal{X}| = |\mathcal{Y}| = |\mathcal{V}| = 2$.*

Note that our analysis for a binary alphabet $\mathcal{V}$ can be extended to alphabets with prime size to achieve $\mathcal{R}_{\text{SC}}(q_{XY})$. Justification of this extension and the proofs of Lemmas 7, 8 and Theorem 4 are omitted due to space constraints.

## References


[1] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4181–4206, 2010.

[2] R. Blasco-Serrano, R. Thobaben, and M. Skoglund, "Polar codes for coordination in cascade networks," in *Int. Zurich Seminar on Communications, Zurich, Switzerland proceedings*, 2012, pp. 55–58.

[3] M. Bloch, L. Luzzi, and J. Kliewer, "Strong Coordination with Polar Codes," in *Proc. of the Annual Allerton Conf. on Communication Control and Computing*, 2012.

[4] E. Arikan, "Source Polarization," in *IEEE Int. Symp. Info. Theory*, 2010, pp. 899–903.

[5] S. Korada and R. Urbanke, "Polar Codes are Optimal for Lossy Source Coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751–1768, 2010.

[6] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric models," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, 2013.

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

[8] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *arXiv preprint arXiv:1305.4746v3*, 2013.

[9] R. A. Chou and M. R. Bloch, "Polar coding for the broadcast channel with confidential messages and constrained randomization," *arXiv preprint arXiv:1411.0281*, 2014.

[10] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge Univ Pr, 1981.