Empirical and Strong Coordination via Soft Covering With Polar Codes

Rémi A. Chou, *Member, IEEE*, Matthieu R. Bloch, *Senior Member, IEEE*, and Jörg Kliewer, *Senior Member, IEEE*

***Abstract*— We design polar codes for empirical coordination and strong coordination in two-node networks. Our constructions hinge on the fact that polar codes enable explicit low-complexity schemes for soft covering. We leverage this property to propose *explicit* and low-complexity coding schemes that achieve the capacity regions of both empirical coordination and strong coordination for sequences of actions taking value in an alphabet of prime cardinality. Our results improve previously known polar coding schemes, which (i) were restricted to uniform distributions and to actions obtained via binary symmetric channels for strong coordination, (ii) required a non-negligible amount of common randomness for empirical coordination, and (iii) assumed that the simulation of discrete memoryless channels could be perfectly implemented. As a by-product of our results, we obtain a polar coding scheme that achieves channel resolvability for an arbitrary discrete memoryless channel whose input alphabet has prime cardinality.**

***Index Terms*— Strong Coordination, empirical coordination, resolvability, channel resolvability, soft covering, polar codes.**

I. Introduction

THE characterization of the information-theoretic limits of coordination in networks has recently been investigated, for instance, in [2]–[5]. The coordinated actions of nodes in a 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 with respect to the variational distance. Two types of coordination have been introduced: empirical coordination, which requires the one-dimensional empirical distribution of a sequence of n actions to approach a target distribution, and strong coordination, which requires the n-dimensional distribution of a sequence of n 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. Several extensions and applications have built upon the results of [3], including channel simulation [4], [5], multiterminal settings for empirical coordination [6] or strong coordination [7]–[9], empirical coordination for joint source-channel coding [10], and coordination for power control [11]. Strong coordination also finds its origin in quantum information theory with "visible compression of mixed states" [2], [12], [13], [14, Sec. 10.6].

The design of practical and efficient coordination schemes approaching the fundamental limits predicted by information theory has, however, 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. Notable exceptions are [15] and [16], which have proposed coding schemes based on polar codes [17], [18] for a small subset of all two-node network coordination problems.

In this paper, we demonstrate how to solve the issue of coding for soft covering with polar codes. Building upon this result, we extend the constructions in [15] and [16] to provide an explicit and low-complexity alternative to the information-theoretic proof in [3] for two-node networks. More specifically, the contributions of this paper are as follows:

- We propose an explicit polar coding scheme to achieve the channel resolvability of an arbitrary memoryless channel whose input alphabet has prime cardinality; low-complexity coding schemes have previously been proposed with polar codes in [16], invertible extractors in [19], and injective group homomorphisms in [20] but are all restricted to symmetric channels. Although [21] has proposed low-complexity linear coding schemes for arbitrary memoryless channels, the construction therein is non-explicit in the sense that only existence results are proved.
- We propose an explicit polar coding scheme that achieves the empirical coordination capacity region for actions from an alphabet of prime cardinality, when common randomness, whose rate vanishes to zero as the block-length grows, is available at the nodes. This construction extends [15], which only deals with uniform distributions and requires a non negligible rate of common randomness available at the nodes.
- We propose an explicit polar coding scheme that achieves the strong coordination capacity region for actions from

Manuscript received August 30, 2016; revised October 18, 2017; accepted January 10, 2018. Date of publication March 20, 2018; date of current version June 20, 2018. This work was supported by NSF under Grant CCF-1320304 and Grant CCF-1440014. Part of this paper was presented at the 2015 IEEE International Symposium on Information Theory [1].

R. A. Chou is with the Department of Electrical Engineering and Computer Science, Wichita State University, Wichita, KS 67260 USA (e-mail: remi.chou@wichita.edu).

M. R. Bloch is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA, and also with GT-CNRS UMI 2958, 57070 Metz, France (e-mail: matthieu.bloch@ece.gatech.edu).

J. Kliewer is with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: jkliewer@njit.edu).

Communicated by A. Tchamkerten, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2018.2817519

an alphabet with prime cardinality. This generalizes [16], which only considers uniform distributions of actions obtained via a binary symmetric channel, and assumes that the simulation of discrete memoryless channels can be perfectly implemented.

Our proposed constructions are *explicit* and handle asymmetric settings through block-Markov encoding instead of relying, as in [22] and related works, on the existence of some maps or a non-negligible amount of shared randomness; we provide further discussion contrasting the present work with [22] in Remark 1.

The coding mechanism underlying our coding schemes is "soft covering," which refers to the approximation of output statistics using codebooks. In particular, soft covering with random codebooks has been used to study problems in information theory such as Wyner's common information [23], the resolvability of a channel [24], secrecy over wiretap channels [20], [25]–[28], secrecy over quantum wiretap channels [29], strong coordination [3], channel synthesis [4], [5], covert and stealth communication [30], [31]. In our coding schemes, soft covering with polar codes is obtained via the special case of soft covering over noiseless channels, together with an appropriate block-Markov encoding to "recycle" common randomness.

Remark 1: [22, Th. 3] provides a polar coding scheme for asymmetric channels for which reliability holds on average over a random choice of the sequence of "frozen bits." This proves the existence of a specific sequence of "frozen bits" that ensures reliability. [22, Sec. III-A] also provides an explicit *construction, which, however, requires that encoder and decoder share a non-negligible amount of randomness as the blocklength grows. To circumvent these issues, we use instead the technique of block-Markov encoding, which has been successfully applied to universal channel coding in [32], to channel coding in [33], and to Wyner-Ziv coding in [34] and [35].*

The idea of randomness recycling is closely related to recursive constructions of seeded extractors in the computer science literature, see for instance [36]. Block-Markov encoding in polar coding schemes has first been used in [32] and [33], for problems involving reliability constraints and requiring the reconstruction of some random variables. Unlike problems that only involve reliability constraints, an additional difficulty of block-Markov encoding for coordination is to ensure approximation of the target distribution jointly over all encoding blocks, despite potential inter-block dependencies.

The remainder of the paper is organized as follows. Section II provides the notation, and Section III reviews the notion of resolvability and coordination. Section IV demonstrates the ability of polar codes to achieve channel resolvability. For a two-node network, Sections V and VI provide polar coding schemes that achieve the empirical coordination capacity region and the strong coordination capacity region, respectively. Finally, Section VII provides concluding remarks.

II. Notation

We let $[\![a, b]\!]$ be the set of integers between $\lfloor a \rfloor$ and $\lceil b \rceil$. We denote the set of strictly positive natural numbers by $\mathbb{N}^*$. 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 [18]. The components of a vector $X^{1:N}$ of size N are denoted with superscripts, i.e., $X^{1:N} \triangleq (X^1, X^2, \ldots, X^N)$. For any set $\mathcal{A} \subset [\![1, N]\!]$, we let $X^{1:N}[\mathcal{A}]$ be the components of $X^{1:N}$ whose indices are in $\mathcal{A}$. For two probability distributions p and q defined over the same alphabet $\mathcal{X}$, we define the variational distance between p and q as

$$\mathbb{V}(p_X, q_X) \triangleq \sum_{x \in \mathcal{X}} |p(x) - q(x)|.$$

We let $\mathbb{D}(\cdot||\cdot)$ denote the Kullback-Leibler divergence between two distributions. If p, q are two distributions over the finite alphabet $\mathcal{X}$, similar to [37], we use the convention $D(p||q) = +\infty$ if there exists $x \in \mathcal{X}$ such that $q(x) = 0$ and $p(x) > 0$. For joint probability distributions p_{XY} and q_{XY} defined over $\mathcal{X} \times \mathcal{Y}$, we write the conditional Kullback-Leibler divergence as

$$\mathbb{E}_{p_X}\left[\mathbb{D}(p_{Y|X}||q_{Y|X})\right] \triangleq \sum_{x \in \mathcal{X}} p_X(x) \mathbb{D}(p_{Y|X=x}||q_{Y|X=x}).$$

All logs are taken with respect to base 2. Unless specified otherwise, random variables are written in upper case letters, and particular realizations of a random variable are written in corresponding lower case letters. Finally, the indicator function is denoted by $\mathbb{1}\{\omega\}$, which is equal to 1 if the predicate ω is true and 0 otherwise.

III. Review of Resolvability and Coordination

We first review the notion of channel resolvability [24], which plays a key role in the analysis of strong coordination. We then review the related notion of resolvability, which forms a building block of our proposed coding schemes. Finally, we review the definition of empirical and strong coordination in two-node networks as introduced in [3]. A common characteristic of these three problems is that they do not require a reconstruction algorithm to ensure a reliability condition, as, for instance, for source or channel coding, but require, instead a "good" approximation of given probability distributions.

A. Soft Covering and Channel Resolvability

Soft covering is concerned with the approximation theory of output statistics [24] and first appeared in the analysis of the common information between random variables [23].

Consider a discrete memoryless channel $(\mathcal{X}, q_{Y|X}, \mathcal{Y})$ and a memoryless source $(\mathcal{X}, q_X)$ with $\mathcal{X}$ and $\mathcal{Y}$ representing finite alphabets. Define the target distribution q_Y as the output of the channel when the input distribution is q_X as

$$\forall y \in \mathcal{Y}, q_Y(y) \triangleq \sum_{x \in \mathcal{X}} q_{Y|X}(y|x) q_X(x). \tag{1}$$

As depicted in Figure 1, the encoder wishes to form a sequence with the least amount of randomness such that the sequence sent over the channel $(\mathcal{X}, q_{Y|X}, \mathcal{Y})$ produces an output $\widetilde{Y}^{1:N}$, whose distribution is close to $q_{Y^{1:N}} \triangleq \prod_{i=1}^{N} q_Y$. A formal definition is given as follows.

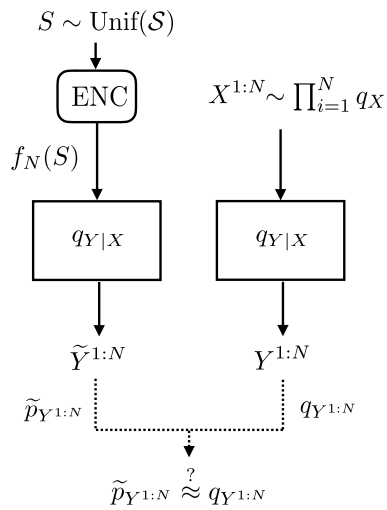

Fig. 1. Description of the channel resolvability problem. Unif($\mathcal{S}$) denotes the uniform distribution over $\mathcal{S}$.

Definition 2: Consider a discrete memoryless channel $(\mathcal{X}, q_{Y|X}, \mathcal{Y})$. *A* $(2^{NR}, N)$ *soft covering code* $\mathcal{C}_N$ *consists of*

- *a randomization sequence* S *uniformly distributed over* $\mathcal{S} \triangleq [\![1, 2^{NR}]\!]$;
- *an encoding function* $f_N : \mathcal{S} \to \mathcal{X}^N$;

and operates as follows:

- *the encoder forms* $f_N(S)$;
- *the encoder transmits* $f_N(S)$ *over the channel* $q_{Y^{1:N}|X^{1:N}} = \prod_{i=1}^{N} q_{Y|X}$.

Definition 3: R is an achievable soft covering rate for an input distribution q_X *if there exists a sequence of* $(2^{NR}, N)$ *soft covering codes,* $\{\mathcal{C}_N\}_{N \geqslant 1}$, *such that*

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

where $q_{Y^{1:N}} = \prod_{i=1}^{N} q_Y$ *with* q_Y *defined in* (1) *for the input distribution* q_X *and* $\forall y^{1:N} \in \mathcal{Y}^N$,

$$\widetilde{p}_{Y^{1:N}}(y^{1:N}) \triangleq \sum_{s\in\mathcal{S}} q_{Y^{1:N}|X^{1:N}}\left(y^{1:N} | f_N(s)\right) \frac{1}{|\mathcal{S}|}.$$

As summarized in Theorem 4, the infimum of achievable rates for any input distribution q_X is called channel resolvability and has been characterized in [4], [20], [24], and [30] for various metrics.

Theorem 4 (Channel Resolvability): Consider a discrete memoryless channel $W \triangleq (\mathcal{X}, q_{Y|X}, \mathcal{Y})$. *The channel resolvability of* W *is* $\max_{q_X} I(X;Y)$, *where* X *and* Y *have joint distribution* $q_{Y|X} q_X$.

B. Resolvability and Conditional Resolvability

Resolvability corresponds to channel resolvability over a noiseless channel and thus characterizes the minimum rate of a uniformly distributed sequence required to simulate a source with given statistics. We review here the slightly more general notion of *conditional resolvability* [38].

Consider a discrete memoryless source $(\mathcal{X} \times \mathcal{Y}, q_{XY})$ with $\mathcal{X}$ and $\mathcal{Y}$ finite alphabets. As depicted in Figure 2, given

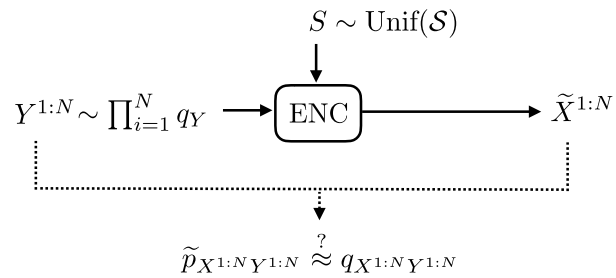

Fig. 2. Illustration of the conditional resolvability problem. Unif($\mathcal{S}$) denotes the uniform distribution over $\mathcal{S}$.

N realizations of the memoryless source $(\mathcal{Y}, q_Y)$, the encoder wishes to form $\widetilde{X}^{1:N}$ with the minimal amount of randomness such that the joint distribution of $(\widetilde{X}^{1:N}, Y^{1:N})$, denoted by $\widetilde{p}_{X^{1:N}Y^{1:N}}$, is close to $q_{X^{1:N}Y^{1:N}} \triangleq \prod_{i=1}^{N} q_{XY}$. We refer to this setting as the conditional resolvability problem. Note that the traditional definition of resolvability [24] corresponds to $\mathcal{Y} = \emptyset$. A formal definition is as follows.

Definition 5: A $(2^{NR}, N)$ *code* $\mathcal{C}_N$ *for a discrete memoryless source* $(\mathcal{X} \times \mathcal{Y}, q_{XY})$ *consists of*

- *a randomization sequence* S *uniformly distributed over* $\mathcal{S} \triangleq [\![1, 2^{NR}]\!]$;
- *an encoding function* $f_N : \mathcal{S} \times \mathcal{Y}^N \to \mathcal{X}^N$;

and operates as follows:

- *the encoder observes* N *realizations* $Y^{1:N}$ *of the memoryless source* $(\mathcal{Y}, q_Y)$;
- *the encoder forms* $\widetilde{X}^{1:N} \triangleq f_N(S, Y^{1:N})$.

The joint distribution of $(\widetilde{X}^{1:N}, Y^{1:N})$ *is denoted by* $\widetilde{p}_{X^{1:N}Y^{1:N}}$.

Definition 6: R is an achievable conditional resolution rate for a discrete memoryless source $(\mathcal{X} \times \mathcal{Y}, q_{XY})$ *if there exists a sequence of* $(2^{NR}, N)$ *codes,* $\{\mathcal{C}_N\}_{N \geqslant 1}$ *such that*

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

with $\widetilde{p}_{X^{1:N}Y^{1:N}}$ *the joint probability distribution of the encoder output* $\widetilde{X}^{1:N}$, *and where* $Y^{1:N}$ *is available at the encoder.*

The infimum of such achievable rates is called the conditional resolvability and is characterized as follows [4], [38].

Theorem 7 (Conditional Resolvability): The conditional resolvability of a discrete memoryless source $(\mathcal{X} \times \mathcal{Y}, q_{XY})$ *is* $H(X|Y)$.

Remark 8: In [38], the conditional resolvability is described as the minimum randomness required to approximate a target conditional distribution representing a channel given a fixed input process. We prefer to approach conditional resolvability as an extension of resolvability since the corresponding interpretation in terms of random number generation [24] and the special case $\mathcal{Y} = \emptyset$ *seem more natural in the context of our proofs.*

The operation described in Definitions 5 and 6 may be viewed as performing soft covering over noiseless channels. Codes achieving conditional resolvability will be the main building block of our coding schemes to emulate soft covering over noisy channels.

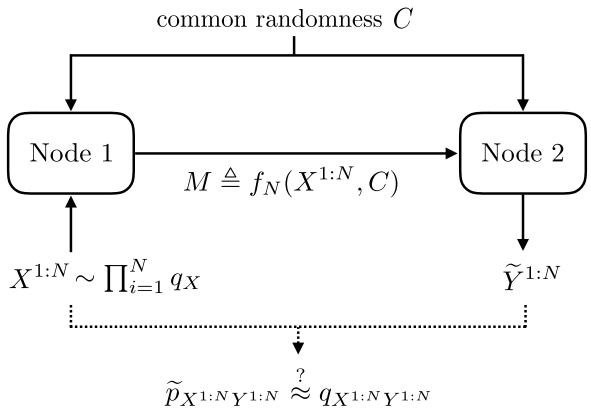

Fig. 3. Coordination setup for a two-node network.

C. Coordination

Consider a memoryless source $(\mathcal{X}\mathcal{Y}, q_{XY})$ with $\mathcal{X}$ and $\mathcal{Y}$ finite alphabets, and two nodes, Node 1 and Node 2. As depicted in Figure 3, Node 1 observes a sequence of actions $X^{1:N}$ and sends a message M over a noiseless channel to Node 2. M must be constructed such that from M and some randomness C, pre-shared with Node 1, Node 2 can produce $\widetilde{Y}^{1:N}$ such that the joint distribution of $(X^{1:N}, \widetilde{Y}^{1:N})$, denoted by $\widetilde{p}_{X^{1:N}Y^{1:N}}$, is close to $q_{Y^{1:N}} \triangleq \prod_{i=1}^N q_{XY}$. A formal definition is as follows.

Definition 9: 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 shared by Node 1 and Node 2;*
- *an encoding function $f_N : \mathcal{X}^N \times [\![1, 2^{NR_0}]\!] \to [\![1, 2^{NR}]\!]$ at Node 1;*
- *a decoding function $g_N : [\![1, 2^{NR}]\!] \times [\![1, 2^{NR_0}]\!] \to \mathcal{Y}^N$ at Node 2,*

and operates as follows:

- *Node 1 observes $X^{1:N}$, N independent 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 by $\widetilde{p}_{X^{1:N}Y^{1:N}}$.*

The notion of empirical and strong coordination are then defined as follows.

Definition 10: 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 $\epsilon > 0$

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

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

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

is the joint histogram of $(x^{1:N}, \widetilde{y}^{1:N})$. The closure of the set of achievable rates is called the empirical coordination capacity region.

Definition 11: 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

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

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

The capacity regions for empirical coordination and strong coordination have been fully characterized in [3].

Theorem 12 ([3]): The empirical coordination capacity region is

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

Theorem 13 ([3]): The strong coordination capacity region is

$$\mathcal{R}_{\mathrm{SC}}(q_{XY}) \triangleq \bigcup_{\substack{X \to V \to Y \\ |\mathcal{V}| \leqslant |\mathcal{X}||\mathcal{Y}|+1}} \{(R, R_0) : R + R_0 \geqslant I(XY; V), \quad R \geqslant I(X; V)\}.$$

IV. Polar Coding for Channel Resolvability

We first develop an explicit and low-complexity coding scheme to achieve channel resolvability. The key ideas that will be reused in our coding scheme for empirical and strong coordination are (i) resolvability-achieving random number generation and (ii) randomness recycling through block-Markov encoding.

Informally, our coding scheme operates over $k \in \mathbb{N}^*$ encoding blocks of length $N \triangleq 2^n$, $n \in \mathbb{N}^*$ as follows. In the first block, using a rate $H(X)$ of randomness, we generate a random variable whose distribution is close to $q_{X^{1:N}}$. When the produced random variable is sent over the channel $q_{Y|X}$, the channel output distribution is close to $q_{Y^{1:N}}$. The amount of randomness used is non-optimal, since we are approximately "wasting" a fraction $H(X|Y)$ of randomness by Theorem 4. For the next encoding blocks, we proceed as in the first block except that part of the randomness is now recycled from the previous block. More specifically, we recycle the bits of randomness used at the input of the channel in the previous block that are almost independent from the channel output. The rate of those bits can be shown to approach $H(X|Y)$. The main difficulty is to ensure that the target distribution at the output of the channel is jointly approximated over all blocks despite the randomness reuse from one block to another.

We provide a formal description of the coding scheme in Section IV-A, and present its analysis in Section IV-B. Part of the analysis for channel resolvability will be directly reused for the problem of strong coordination in Section VI.

A. Coding Scheme

Fix a joint probability distribution q_{XY} over $\mathcal{X} \times \mathcal{Y}$, where $|\mathcal{X}|$ is a prime number. Define $U^{1:N} \triangleq X^{1:N} G_n$, where G_n is

defined in Section II, and define for $\beta < 1/2$, $\delta_N \triangleq 2^{-N^\beta}$ and the sets

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

Note that the sets $\mathcal{V}_X$ and $\mathcal{V}_{X|Y}$ are defined with respect to q_{XY}. Intuitively, $U^{1:N}[\mathcal{V}_{X|Y}]$ corresponds to the components of $U^{1:N}$ that are almost independent from Y (see [39] for an interpretation of $\mathcal{V}_X$ and $\mathcal{V}_{X|Y}$ in terms of randomness extraction). Note also that

$$\lim_{N\to\infty} |\mathcal{V}_X|/N = H(X),$$
$$\lim_{N\to\infty} |\mathcal{V}_{X|Y}|/N = H(X|Y)$$

by [39, Lemma 7], which relies on a proof technique used in [22] and a result from [40].

We use the subscript $i \in [\![1, k]\!]$ to denote random variables associated with the encoding of Block i, and we use the notation $X_{i:j} \triangleq (X_l)_{l \in [\![i,j]\!]}$, when $i < j$. The encoding process is described in Algorithm 1. The functional dependence graph of the coding scheme is depicted in Figure 4 for the reader's convenience.

Algorithm 1 Encoding Algorithm for Channel Resolvability

Require: A vector $\bar{C}_1$ of $|\mathcal{V}_{X|Y}|$ uniformly distributed symbols shared by the encoder and decoder, and k vectors $C_{1:k}$ of $|\mathcal{V}_X \backslash \mathcal{V}_{X|Y}|$ uniformly distributed symbols.

1: **for** Block $i = 1$ to k **do**
2: $\bar{C}_i \leftarrow \bar{C}_1$
3: $\widetilde{U}_i^{1:N}[\mathcal{V}_{X|Y}] \leftarrow \bar{C}_i$
4: $\widetilde{U}_i^{1:N}[\mathcal{V}_X \backslash \mathcal{V}_{X|Y}] \leftarrow C_i$
5: Successively draw the remaining components $\widetilde{U}_i^{1:N}[\mathcal{V}_X^c]$, according to

$$\widetilde{p}_{U_i^j | U_i^{1:j-1}}(u_i^j | \widetilde{U}_i^{1:j-1}) \triangleq q_{U^j|U^{1:j-1}}(u_i^j | \widetilde{U}_i^{1:j-1}) \text{ if } j \in \mathcal{V}_X^c. \quad (2)$$

6: Transmit $\widetilde{X}_i^{1:N} \triangleq \widetilde{U}_i^{1:N} G_n$ over the channel $q_{Y|X}$. We denote $\widetilde{Y}_i^{1:N}$ the corresponding channel output.
7: **end for**

In essence, the protocol described in Algorithm 1 performs a resolvability-achieving random number generation [41, Definition 2.2.2] for each encoding block, and recycles randomness $\bar{C}_1$ over all blocks.

Remark 14: The randomizations described in (2) *could be replaced by deterministic decisions for $j \in \mathcal{H}_X^c$, i.e., randomized decisions are only needed for $j \in \mathcal{V}_X^c \backslash \mathcal{H}_X^c$, as shown in [35].*

B. Scheme Analysis

Our analysis of Algorithm 1 exploits simple relations satisfied by the Kullback-Leibler divergence presented in Appendix A. We denote the distribution induced by the coding scheme, i.e., the joint distribution of $\widetilde{X}_i^{1:N}$ and $\widetilde{Y}_i^{1:N}$, by $\widetilde{p}_{X_i^{1:N} Y_i^{1:N}} = q_{Y^{1:N}|X^{1:N}} \widetilde{p}_{X_i^{1:N}}$, $i \in [\![1, k]\!]$. We start with

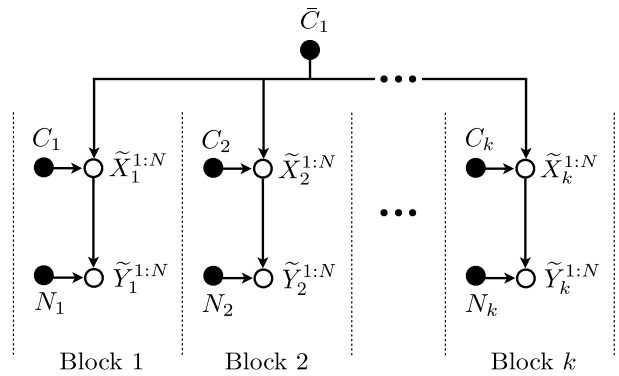

Fig. 4. Functional dependence graph of the block encoding scheme for channel resolvability. N_i, $i \in [\![1, k]\!]$, is the channel noise corresponding to the transmission over Block i. For Block i, $(C_i, \bar{C}_{i-1})$ is the randomness used at the encoder to form $\widetilde{X}_i$, where $\forall i \in [\![2, k]\!]$, $\bar{C}_i = \bar{C}_{i-1}$ and C_i is only used in Block i.

Lemma 15 that helps us show in Remark 16 that each block individually performs soft covering.

Lemma 15: For block $i \in [\![1, k]\!]$, we have

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

where $\delta_N^{(1)} \triangleq N\delta_N$.

Proof: For $i \in [\![1, k]\!]$, we have

$$\begin{aligned}
&\mathbb{D}(q_{X^{1:N}Y^{1:N}} || \widetilde{p}_{X_i^{1:N} Y_i^{1:N}}) \\
&= \mathbb{E}_{q_{X^{1:N}}} \left[\mathbb{D}(q_{Y^{1:N}|X^{1:N}} || \widetilde{p}_{Y_i^{1:N}|X_i^{1:N}}) \right] + \mathbb{D}(q_{X^{1:N}} || \widetilde{p}_{X_i^{1:N}}) \\
&\overset{(a)}{=} \mathbb{D}(q_{X^{1:N}} || \widetilde{p}_{X_i^{1:N}}) \\
&\overset{(b)}{=} \mathbb{D}(q_{U^{1:N}} || \widetilde{p}_{U_i^{1:N}}) \\
&\overset{(c)}{=} \sum_{j=1}^{N} \mathbb{E}_{q_{U^{1:j-1}}} \left[\mathbb{D}(q_{U^j|U^{1:j-1}} || \widetilde{p}_{U_i^j|U_i^{1:j-1}}) \right] \\
&\overset{(d)}{=} \sum_{j \in \mathcal{V}_X} \mathbb{E}_{q_{U^{1:j-1}}} \left[\mathbb{D}(q_{U^j|U^{1:j-1}} || \widetilde{p}_{U_i^j|U_i^{1:j-1}}) \right] \\
&\overset{(e)}{=} \sum_{j \in \mathcal{V}_X} (\log|\mathcal{X}| - H(U^j | U^{1:j-1})) \\
&\overset{(f)}{\leqslant} |\mathcal{V}_X| \delta_N \\
&\leqslant N\delta_N,
\end{aligned}$$

where (a) holds by definition of $\widetilde{p}_{X_i^{1:N} Y_i^{1:N}}$, (b) holds by invertibility of G_n, (c) holds by the chain rule for divergence [42], (d) holds by (2), (e) holds by uniformity of the symbols in positions $\mathcal{V}_X$, (f) holds by definition of $\mathcal{V}_X$. ■

Remark 16: The encoding algorithm performs a resolvability-achieving random number generation in each block. A formal proof can be found in Appendix B and relies on Lemma 15 and an asymptotic symmetry result for the Kullback-Leibler divergence.

We now establish the asymptotic independence of consecutive blocks with the following two lemmas.

Lemma 17: For $i \in [\![2, k]\!]$, the outputs of two consecutive blocks are asymptotically independent; specifically,

$$\mathbb{D}\left(\widetilde{p}_{Y_{i-1:i}^{1:N} \bar{C}_1} || \widetilde{p}_{Y_{i-1}^{1:N} \bar{C}_1} \widetilde{p}_{Y_i^{1:N}} \right) \leqslant \delta_N^{(2)},$$

where $\delta_N^{(2)} = O\left(N^2 \delta_N^{1/2}\right)$.

Proof: Let $i \in [\![1, k]\!]$. We have

$$\begin{aligned}
&H(U^{1:N}[\mathcal{V}_{X|Y}]|Y^{1:N}) - H(\widetilde{U}_i^{1:N}[\mathcal{V}_{X|Y}]|\widetilde{Y}_i^{1:N}) \\
&\quad = H(U^{1:N}[\mathcal{V}_{X|Y}]Y^{1:N}) - H(\widetilde{U}_i^{1:N}[\mathcal{V}_{X|Y}]\widetilde{Y}_i^{1:N}) \\
&\qquad + H(\widetilde{Y}_i^{1:N}) - H(Y^{1:N}) \\
&\quad \overset{(a)}{\leqslant} ND_1 \log(|\mathcal{Y}|/D_1) + ND_2 \log(|\mathcal{X}||\mathcal{Y}|/D_2) \\
&\quad \overset{(b)}{\leqslant} ND_1 \log(|\mathcal{Y}|/D_1) + ND_3 \log(|\mathcal{X}||\mathcal{Y}|/D_3) \\
&\quad \overset{(c)}{\leqslant} 2N\sqrt{2\ln 2}\sqrt{\delta_N^{(1)}} \log\left(|\mathcal{X}||\mathcal{Y}|/\left[\sqrt{2\ln 2}\sqrt{\delta_N^{(1)}}\right]\right) \\
&\quad \triangleq \delta_N^{(UY)},
\end{aligned} \tag{3}$$

where (a) holds for N large enough by Lemma 39 in Appendix A with $D_1 \triangleq \sqrt{2\ln 2}\sqrt{\mathbb{D}(q_{U^{1:N}[\mathcal{V}_{X|Y}]Y^{1:N}} || \widetilde{p}_{U_i^{1:N}[\mathcal{V}_{X|Y}]Y_i^{1:N}})}$ and $D_2 \triangleq \sqrt{2\ln 2}\sqrt{\mathbb{D}(q_{Y^{1:N}} || \widetilde{p}_{Y_i^{1:N}})}$, (b) holds for N large enough because $D_2 \leqslant D_3$ by the chain rule for relative Kullback-Leibler divergence and invertibility of G_n with $D_3 \triangleq \sqrt{2\ln 2}\sqrt{\mathbb{D}(q_{X^{1:N}Y^{1:N}} || \widetilde{p}_{X_i^{1:N}Y_i^{1:N}})}$, (c) holds for N large enough by Lemma 15 and because $D_1 \leqslant D_3$.

Hence, for $i \in [\![2, k]\!]$,

$$\begin{aligned}
&\mathbb{D}\left(\widetilde{p}_{Y_{i-1:i}^{1:N}\bar{C}_1} || \widetilde{p}_{Y_{i-1}^{1:N}\bar{C}_1}\widetilde{p}_{Y_i^{1:N}}\right) \\
&\quad = I(\widetilde{Y}_{i-1}^{1:N}\bar{C}_1; \widetilde{Y}_i^{1:N}) \\
&\quad = I(\widetilde{Y}_i^{1:N}; \bar{C}_1) + I(\widetilde{Y}_{i-1}^{1:N}; \widetilde{Y}_i^{1:N}|\bar{C}_1) \\
&\quad \overset{(d)}{=} I(\widetilde{Y}_i^{1:N}; \bar{C}_1) \\
&\quad = I(\widetilde{Y}_i^{1:N}; \widetilde{U}_i^{1:N}[\mathcal{V}_{X|Y}]) \\
&\quad \overset{(e)}{=} |\mathcal{V}_{X|Y}| \log|\mathcal{X}| - H(\widetilde{U}_i^{1:N}[\mathcal{V}_{X|Y}]|\widetilde{Y}_i^{1:N}) \\
&\quad \overset{(f)}{\leqslant} |\mathcal{V}_{X|Y}| \log|\mathcal{X}| - H(U^{1:N}[\mathcal{V}_{X|Y}]|Y^{1:N}) + \delta_N^{(UY)} \\
&\quad \overset{(g)}{\leqslant} |\mathcal{V}_{X|Y}| \log|\mathcal{X}| - \sum_{j\in\mathcal{V}_{X|Y}} H(U^j|U^{1:j-1}Y^{1:N}) + \delta_N^{(UY)} \\
&\quad \leqslant |\mathcal{V}_{X|Y}| \log|\mathcal{X}| - |\mathcal{V}_{X|Y}|(\log|\mathcal{X}| - \delta_N) + \delta_N^{(UY)} \\
&\quad \leqslant N\delta_N + \delta_N^{(UY)},
\end{aligned}$$

where (d) holds because $\widetilde{Y}_{i-1}^{1:N} - \bar{C}_1 - \widetilde{Y}_i^{1:N}$, as seen in Figure 4, (e) holds by uniformity of $\widetilde{U}_i^{1:N}[\mathcal{V}_{X|Y}]$, (f) holds by (3), (g) holds because conditioning reduces entropy. ■

Lemma 18: The outputs of all the blocks are asymptotically independent; specifically,

$$\mathbb{D}\left(\widetilde{p}_{Y_{1:k}^{1:N}} \,\middle\|\, \prod_{i=1}^{k} \widetilde{p}_{Y_i^{1:N}}\right) \leqslant (k-1)\delta_N^{(2)}.$$

Proof: We have

$$\begin{aligned}
&\mathbb{D}\left(\widetilde{p}_{Y_{1:k}^{1:N}} \,\middle\|\, \prod_{i=1}^{k} \widetilde{p}_{Y_i^{1:N}}\right) \\
&\quad \overset{(a)}{=} \sum_{i=2}^{k} I(\widetilde{Y}_i^{1:N}; \widetilde{Y}_{1:i-1}^{1:N}) \\
&\quad \leqslant \sum_{i=2}^{k} I(\widetilde{Y}_i^{1:N}; \widetilde{Y}_{1:i-1}^{1:N}\bar{C}_1) \\
&\quad = \sum_{i=2}^{k} \left(I(\widetilde{Y}_i^{1:N}; \widetilde{Y}_{i-1}^{1:N}\bar{C}_1) + I(\widetilde{Y}_i^{1:N}; \widetilde{Y}_{1:i-2}^{1:N}|\bar{C}_1\widetilde{Y}_{i-1}^{1:N})\right) \\
&\quad \overset{(b)}{=} \sum_{i=2}^{k} I(\widetilde{Y}_i^{1:N}; \widetilde{Y}_{i-1}^{1:N}\bar{C}_1) \\
&\quad = \sum_{i=2}^{k} \mathbb{D}\left(\widetilde{p}_{Y_{i-1:i}^{1:N}\bar{C}_1}, \widetilde{p}_{Y_{i-1}^{1:N}\bar{C}_1}\widetilde{p}_{Y_i^{1:N}}\right) \\
&\quad \overset{(c)}{\leqslant} \sum_{i=2}^{k} \delta_N^{(2)} \\
&\quad = (k-1)\delta_N^{(2)},
\end{aligned}$$

where (a) holds by Lemma 37 in Appendix A, (b) holds because for any $i \in [\![3, k]\!]$, the Markov chain $\widetilde{Y}_{1:i-2}^{1:N} - \bar{C}_1\widetilde{Y}_{i-1}^{1:N} - \widetilde{Y}_i^{1:N}$ holds as seen in Figure 4, (c) holds by Lemma 17. ■

We are now ready to show that the target output distribution is jointly approximated over all blocks.

Lemma 19: We have,

$$\mathbb{D}\left(\widetilde{p}_{Y_{1:k}^{1:N}} \,\middle\|\, q_{Y^{1:kN}}\right) \leqslant \delta_N^{(3)},$$

where $\delta_N^{(3)} = O\left(k^{3/2}N^2\delta_N^{1/4}\right)$.

Proof: First, observe that

$$\begin{aligned}
\mathbb{D}\left(q_{Y^{1:kN}} \,\middle\|\, \prod_{i=1}^{k} \widetilde{p}_{Y_i^{1:N}}\right) &= \mathbb{D}\left(\prod_{i=1}^{k} q_{Y^{1:N}} \,\middle\|\, \prod_{i=1}^{k} \widetilde{p}_{Y_i^{1:N}}\right) \\
&= \sum_{i=1}^{k} \mathbb{D}\left(q_{Y^{1:N}} \,\middle\|\, \widetilde{p}_{Y_i^{1:N}}\right) \\
&\leqslant k\delta_N^{(1)},
\end{aligned} \tag{4}$$

where the inequality holds by Lemma 15. Then, we have for N large enough,

$$\begin{aligned}
&\mathbb{D}\left(\widetilde{p}_{Y_{1:k}^{1:N}} \| q_{Y^{1:kN}}\right) \\
&\overset{(a)}{\leqslant} \log\left(\frac{1}{\mu_{q_Y}^{kN}}\right)\sqrt{2\ln 2}\left[\sqrt{\mathbb{D}\left(\widetilde{p}_{Y_{1:k}^{1:N}} \,\middle\|\, \prod_{i=1}^{k} \widetilde{p}_{Y_i^{1:N}}\right)}\right. \\
&\qquad\qquad \left. + \sqrt{\mathbb{D}\left(q_{Y^{1:kN}} \,\middle\|\, \prod_{i=1}^{k} \widetilde{p}_{Y_i^{1:N}}\right)}\right] \\
&\overset{(b)}{\leqslant} kN \log\left(\frac{1}{\mu_{q_Y}}\right)\sqrt{2\ln 2}\left[\sqrt{k\delta_N^{(1)}} + \sqrt{(k-1)\delta_N^{(2)}}\right],
\end{aligned}$$

where (a) holds by Lemma 38 in Appendix A and because $\mu_{q_{Y^{1:kN}}} = \mu_{q_Y}^{kN}$, (b) holds by Lemma 18 and by (4). ■

Our encoding scheme exploits randomness to draw symbols according to (2), whose rate is for any $i \in [\![1, k]\!]$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j\in\mathcal{V}_X^c} H(\widetilde{U}_i^j | \widetilde{U}_i^{1:j-1}).$$

We quantify this rate in the following lemma.

Lemma 20:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j\in\mathcal{V}_X^c} H(\widetilde{U}_i^j | \widetilde{U}_i^{1:j-1}) = 0. \tag{5}$$

Proof: We have for $i \in [\![1, k]\!]$, for $j \in \mathcal{V}_X^c$,

$$\begin{aligned}
&H(\widetilde{U}_i^j | \widetilde{U}_i^{1:j-1}) - H(U^j | U^{1:j-1}) \\
&\quad = H(\widetilde{U}_i^{1:j}) - H(U^{1:j}) + H(U^{1:j-1}) - H(\widetilde{U}_i^{1:j-1}) \\
&\quad \overset{(a)}{\leqslant} 2ND\log(|\mathcal{X}|/D) \\
&\quad \overset{(b)}{\leqslant} 2N\sqrt{2\ln 2}\sqrt{\delta_N^{(1)}}\log\left(|\mathcal{X}|/\left[\sqrt{2\ln 2}\sqrt{\delta_N^{(1)}}\right]\right) \\
&\quad \triangleq \delta_N^{(U)},
\end{aligned} \tag{6}$$

where (a) holds by Lemma 39 in Appendix A for N large enough and similar to the proof of Lemma 17 with $D \triangleq \sqrt{2\ln 2}\sqrt{\mathbb{D}(q_{X^{1:N}} || \widetilde{p}_{X_i^{1:N}})}$, (b) holds by Lemma 15 for N large enough.

Defining $\mathcal{H}_X \triangleq \{i \in [\![1, N]\!] : H(U^i | U^{1:i-1}) > \delta_N\}$, we thus obtain

$$\begin{aligned}
&\sum_{j\in\mathcal{V}_X^c} H(\widetilde{U}_i^j | \widetilde{U}_i^{1:j-1}) \\
&\quad = \sum_{j\in\mathcal{H}_X^c\cup(\mathcal{H}_X\setminus\mathcal{V}_X)} H(\widetilde{U}_i^j | \widetilde{U}_i^{1:j-1}) \\
&\quad \leqslant |\mathcal{H}_X\setminus\mathcal{V}_X|\log|\mathcal{X}| + \sum_{j\in\mathcal{H}_X^c} H(\widetilde{U}_i^j | \widetilde{U}_i^{1:j-1}) \\
&\quad = (|\mathcal{H}_X| - |\mathcal{V}_X|)\log|\mathcal{X}| + \sum_{j\in\mathcal{H}_X^c} H(\widetilde{U}_i^j | \widetilde{U}_i^{1:j-1}) \\
&\quad \overset{(a)}{\leqslant} (|\mathcal{H}_X| - |\mathcal{V}_X|)\log|\mathcal{X}| + \sum_{j\in\mathcal{H}_X^c} (H(U^j | U^{1:j-1}) + \delta_N^{(U)}) \\
&\quad \overset{(b)}{\leqslant} (|\mathcal{H}_X| - |\mathcal{V}_X|)\log|\mathcal{X}| + |\mathcal{H}_X^c|(\delta_N + \delta_N^{(U)}) \\
&\quad \leqslant (|\mathcal{H}_X| - |\mathcal{V}_X|)\log|\mathcal{X}| + N(\delta_N + \delta_N^{(U)}),
\end{aligned} \tag{7}$$

where (a) holds by (6), (b) holds by definition of $\mathcal{H}_X$. Hence, (7) yields (5) by [39, Lemmas 6,7]. ■

Finally, combining the previous lemmas we obtain the following theorem.

Theorem 21: The coding scheme of Section IV-A, which operates over k blocks of length N, achieves channel resolvability with respect to the Kullback-Leibler divergence over the discrete memoryless channel $(\mathcal{X}, q_{Y|X}, \mathcal{Y})$, where $|\mathcal{X}|$ is a prime number. More specifically, the channel output distribution $\widetilde{p}_{Y_{1:k}^{1:N}}$ approaches the target distribution $q_{Y^{1:kN}}$ with convergence rate

$$\mathbb{D}\left(\widetilde{p}_{Y_{1:k}^{1:N}} \| q_{Y^{1:kN}}\right) = O\left(k^{3/2}N^2 2^{-\frac{N^{-\beta}}{4}}\right), \beta \in]0, 1/2[,$$

with a rate of input of randomness equal to $I(X;Y) + \frac{H(X|Y)}{k}$, as $N \to \infty$, which approaches $I(X;Y)$ as $k \to \infty$. It thus provides an explicit coding scheme with complexity in $O(kN\log N)$ for Theorem 4.

Proof: By Lemma 20, the overall rate of uniform symbols required is only

$$\begin{aligned}
\frac{|\bar{C}_1| + |C_{1:k}|}{kN} &= \frac{|\mathcal{V}_{X|Y}| + k|\mathcal{V}_X\setminus\mathcal{V}_{X|Y}|}{kN} \\
&= \frac{|\mathcal{V}_{X|Y}|}{kN} + \frac{|\mathcal{V}_X| - |\mathcal{V}_{X|Y}|}{N} \\
&\xrightarrow{N\to\infty} I(X;Y) + \frac{H(X|Y)}{k} \\
&\xrightarrow{k\to\infty} I(X;Y),
\end{aligned}$$

where we have used [39, Lemma 7]. Finally, we conclude that the optimal rate $I(X;Y)$ is achieved with Lemma 19. ■

V. Polar Coding for Empirical Coordination

We now develop an explicit and low-complexity coding scheme for empirical coordination that achieves the capacity region when the actions of Node 2 are from an alphabet of prime cardinality. The idea is to perform (i) a *conditional* resolvability-achieving random number generation and (ii) randomness recycling through block-Markov encoding. However, the coding scheme is simpler than in Section IV as the common randomness recycling can be performed and studied more directly. In particular, we will see that the encoding blocks may be treated independently of each other since the approximation of the target distribution is concerned with a one dimensional probability distribution, as opposed to a kN dimensional probability distribution as in Section IV.

More specifically, the coding scheme can be informally summarized as follows. From $X^{1:N}$ and some common randomness of rate close to $H(Y|X)$ 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 constructs a message with rate close to $I(X;Y)$ such that Node 2 reconstructs $\widetilde{Y}^{1:N}$ with the message and the common randomness. Finally, encoding is performed over $k \in \mathbb{N}^*$ blocks by recycling the same common randomness, so that the overall rate of shared randomness vanishes as the number of blocks increases. We formally describe the coding scheme in Section V-A, and present its analysis in Section V-B.

A. Coding Scheme

In the following, we redefine the following notation to simplify discussion. Consider the random variables X, Y distributed according to q_{XY} over $\mathcal{X} \times \mathcal{Y}$, where $|\mathcal{Y}|$ is a prime number. Let $N \triangleq 2^n$, $n \in \mathbb{N}^*$. Define $U^{1:N} \triangleq Y^{1:N}G_n$,

where G_n is defined in Section II, and define for $\beta < 1/2$, $\delta_N \triangleq 2^{-N^\beta}$ as well as the sets

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

Note that the sets $\mathcal{V}_Y$, $\mathcal{H}_Y$, $\mathcal{V}_{Y|X}$, and $\mathcal{H}_{Y|X}$ are defined with respect to q_{XY}. Note also that

$$\lim_{N\to\infty} |\mathcal{V}_Y|/N = H(Y) = \lim_{N\to\infty} |\mathcal{V}_Y|/N,$$
$$\lim_{N\to\infty} |\mathcal{V}_{X|Y}|/N = H(X|Y) = \lim_{N\to\infty} |\mathcal{H}_{X|Y}|/N,$$

by [39, Lemmas 6,7], where [39, Lemma 6] follows from [43, Th. 3.2]. An interpretation of $\mathcal{H}_{Y|X}$ in terms of source coding with side information is provided in [18], and an interpretation of $\mathcal{V}_{Y|X}$ in terms of privacy amplification is provided in [34] and [39]. 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 with encoding Block i. The encoding and decoding procedures are described in Algorithms 2 and 3, respectively.

Remark 22: The coding scheme for each block is similar to lossy source coding schemes [22], [44], as suggested by the optimal communication rate described in Theorem 12. However, the performance metric of interest is different.

Algorithm 2 Encoding algorithm at Node 1 for empirical coordination

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

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

$$\widetilde{p}_{U_i^j|U_i^{1:j-1}X_i^{1:N}}(u_i^j|\widetilde{U}_i^{1:j-1}X_i^{1:N}) \triangleq \begin{cases} q_{U^j|U^{1:j-1}X^{1:N}}(u_i^j|\widetilde{U}_i^{1:j-1}X_i^{1:N}) \text{ if } j \in \mathcal{H}_Y\backslash\mathcal{V}_{Y|X} \\ q_{U^j|U^{1:j-1}}(u_i^j|\widetilde{U}_i^{1:j-1}) \text{ if } j \in \mathcal{H}_Y^c \end{cases} \tag{8}$$

5: 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 $\widetilde{U}_i^{1:N}[\mathcal{H}_Y^c]$
6: **end for**

B. Scheme Analysis

The following lemma shows that $\widetilde{p}_{X^{1:N}Y^{1:N}}$, defined by $\widetilde{p}_{X^{1:N}} \triangleq q_{X^{1:N}}$ and Equation (8), approximates $q_{X^{1:N}Y^{1:N}}$.

Lemma 23: For any $i \in [\![1, k]\!]$,

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

Algorithm 3 Decoding algorithm at Node 2 for empirical coordination

Require: The vectors C_1, $\widetilde{C}_{1:k}$, used in Algorithm 2, and $M_{1:k}$.

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

Remark 24: The encoder in each block performs a conditional resolvability-achieving random number generation by Lemma 23 and because $|C_1|/N = |\mathcal{V}_{Y|X}|/N \xrightarrow{N\to\infty} H(Y|X)$. The proof of Lemma 23 is similar to the proof of Lemma 28 in Section VI-B and is thus omitted. The following lemma shows that empirical coordination holds for each block.

Lemma 25: Let $\epsilon > 0$. For $i \in [\![1, k]\!]$, we have

$$\mathbb{P}\left[\mathbb{V}\left(q_{XY}, T_{X_i^{1:N}\widetilde{Y}_i^{1:N}}\right) > \epsilon\right] \leqslant \delta^{(A)}(N),$$

where $\delta^{(A)}(N) = O\left(\sqrt{N\delta_N}\right)$.

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

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

Note that $\lim_{N\to\infty} \mathbb{P}_q[(X^{1:N}, Y^{1:N}) \notin \mathcal{T}_\epsilon(q_{XY})] = 0$ by the AEP [42], and we can precise the convergence rate as follows.

$$\begin{aligned}
&\mathbb{P}_q[(X^{1:N}, Y^{1:N}) \notin \mathcal{T}_\epsilon(q_{XY})] \\
&= \mathbb{P}_q\left[\mathbb{V}\left(q_{XY}, T_{X^{1:N}Y^{1:N}}\right) > \epsilon\right] \\
&= \mathbb{P}_q\left[\sum_{x,y} |q_{XY}(x, y) - T_{X^{1:N}Y^{1:N}}(x, y)| > \epsilon\right] \\
&\leqslant \mathbb{P}_q\left[\exists(x, y), |q_{XY}(x, y) - T_{X^{1:N}Y^{1:N}}(x, y)| \geqslant \frac{\epsilon}{|\mathcal{X}||\mathcal{Y}|}\right] \\
&\leqslant \sum_{x,y} \mathbb{P}_q\left[|q_{XY}(x, y) - T_{X^{1:N}Y^{1:N}}(x, y)| \geqslant \frac{\epsilon}{|\mathcal{X}||\mathcal{Y}|}\right] \\
&= \sum_{x,y} \mathbb{P}_q\left[\left|\frac{1}{N}\sum_{j=1}^{N}\left[q_{XY}(x, y) - \mathbb{1}\{(X^j, Y^j) = (x, y)\}\right]\right| \geqslant \frac{\epsilon}{|\mathcal{X}||\mathcal{Y}|}\right] \\
&\leqslant \sum_{x,y} 2\exp\left(-\frac{N\epsilon^2}{2|\mathcal{X}|^2|\mathcal{Y}|^2}\right) \\
&= 2|\mathcal{X}||\mathcal{Y}|\exp\left(-\frac{N\epsilon^2}{2|\mathcal{X}|^2|\mathcal{Y}|^2}\right),
\end{aligned} \tag{9}$$

where the second inequality holds by Hoeffding's inequality applied for each pair (x, y) to the

independent and zero-mean random variables $Z_j(x,y) \triangleq \left(q_{XY}(x,y) - \mathbb{1}\{(X^j, Y^j) = (x,y)\}\right) \in [-1,1]$, $j \in [\![1, N]\!]$. Next, let $i \in [\![1,k]\!]$. we have

$$\begin{aligned}
&\mathbb{P}_{\widetilde{p}}\left[\mathbb{V}\left(q_{XY}, T_{X_i^{1:N}\widetilde{Y}_i^{1:N}}\right) > \epsilon\right] \\
&= \sum_{x^{1:N}, y^{1:N}} \widetilde{p}_{X_i^{1:N}Y_i^{1:N}}(x^{1:N}, y^{1:N}) \mathbb{1}\{(x^{1:N}, y^{1:N}) \notin \mathcal{T}_\epsilon(q_{XY})\} \\
&= \sum_{x^{1:N}, y^{1:N}} \Big[\widetilde{p}_{X_i^{1:N}Y_i^{1:N}}(x^{1:N}, y^{1:N}) - q_{X^{1:N}Y^{1:N}}(x^{1:N}, y^{1:N}) \\
&\quad + q_{X^{1:N}Y^{1:N}}(x^{1:N}, y^{1:N})\Big] \mathbb{1}\{(x^{1:N}, y^{1:N}) \notin \mathcal{T}_\epsilon(q_{XY})\} \\
&\leqslant \mathbb{V}(\widetilde{p}_{X_i^{1:N}Y_i^{1:N}}, q_{X^{1:N}Y^{1:N}}) + \mathbb{P}_q[(X^{1:N}, Y^{1:N}) \notin \mathcal{T}_\epsilon(q_{XY})] \\
&\leqslant \sqrt{2\log 2}\sqrt{N\delta_N} + 2|\mathcal{X}||\mathcal{Y}| \exp\left(-\frac{N\epsilon^2}{2|\mathcal{X}|^2|\mathcal{Y}|^2}\right),
\end{aligned}$$

where we have used Lemma 23 and (9). ■

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

Lemma 26: Let $\epsilon > 0$ and $\alpha \in]0, 1/2[$. We have

$$\mathbb{P}\left[\mathbb{V}\left(q_{XY}, T_{X_{1:k}^{1:N}\widetilde{Y}_{1:k}^{1:N}}\right) > \epsilon\right] \leqslant \delta^{(B)}(N),$$

where $\delta^{(B)}(N) = O\left(k\sqrt{N\delta_N}\right)$.

Proof: We have

$$\begin{aligned}
&\mathbb{V}\left(q_{XY}, T_{x_{1:k}^{1:N}\widetilde{y}_{1:k}^{1:N}}\right) \\
&= \sum_{x,y}\left|q_{XY}(x,y) - \frac{1}{kN}\sum_{j=1}^{k}\sum_{i=1}^{N}\mathbb{1}\{(x_j^i, \widetilde{y}_j^i) = (x,y)\}\right| \\
&= \sum_{x,y}\left|\sum_{j=1}^{k}\left(\frac{1}{k}q_{XY}(x,y) - \frac{1}{kN}\sum_{i=1}^{N}\mathbb{1}\{(x_j^i, \widetilde{y}_j^i) = (x,y)\}\right)\right| \\
&\leqslant \frac{1}{k}\sum_{j=1}^{k}\sum_{x,y}\left|q_{XY}(x,y) - \frac{1}{N}\sum_{i=1}^{N}\mathbb{1}\{(x_j^i, \widetilde{y}_j^i) = (x,y)\}\right| \\
&\leqslant \frac{1}{k}\sum_{j=1}^{k}\mathbb{V}\left(q_{XY}, T_{x_j^{1:N}\widetilde{y}_j^{1:N}}\right), \qquad (10)
\end{aligned}$$

hence,

$$\begin{aligned}
&\mathbb{P}\left[\mathbb{V}\left(q_{XY}, T_{X_{1:k}^{1:N}\widetilde{Y}_{1:k}^{1:N}}\right) > \epsilon\right] \\
&\overset{(a)}{\leqslant} \mathbb{P}\left[\frac{1}{k}\sum_{j=1}^{k}\mathbb{V}\left(q_{XY}, T_{x_j^{1:N}\widetilde{y}_j^{1:N}}\right) > \epsilon\right] \\
&\leqslant \mathbb{P}\left[\exists j \in [\![1,k]\!], \mathbb{V}\left(q_{XY}, T_{x_j^{1:N}\widetilde{y}_j^{1:N}}\right) > \epsilon\right] \\
&\leqslant \sum_{j=1}^{k}\mathbb{P}\left[\mathbb{V}\left(q_{XY}, T_{x_j^{1:N}\widetilde{y}_j^{1:N}}\right) > \epsilon\right] \\
&\overset{(b)}{\leqslant} k\delta_N^{(A)},
\end{aligned}$$

where (a) holds by (10), (b) holds by Lemma 25. ■

Theorem 27: The coding scheme described in Algorithms 2 and 3, which operates over k blocks of length N, achieves the two-node network empirical coordination capacity region of Theorem 12 for an arbitrary target distribution q_{XY} over $\mathcal{X} \times \mathcal{Y}$, where $|\mathcal{Y}|$ is a prime number. The coding scheme is explicit with complexity in $O(kN\log N)$.

Proof: The communication rate is

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

where we have used [39, Lemmas 6,7] for the limit. Node 1 also communicates the randomness $\widetilde{C}_{1:k}$ to allow Node 2 to form $\widetilde{U}_{1:k}^{1:N}[\mathcal{H}_Y^c]$, 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 23 similar to the proof of Theorem 21. 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 the limit holds by [39, Lemma 7]. Finally, we conclude that the region described in Theorem 12 is achieved with Lemma 26. ■

VI. Polar Coding for Strong Coordination

We finally design an explicit and low-complexity coding scheme for strong coordination that achieves the capacity region when the actions of Node 2 are from an alphabet of prime cardinality. The idea is again to perform (i) conditional resolvability-achieving random number generation and (ii) common randomness recycling with block-Markov encoding as in Section IV. In addition, we also simulate discrete memoryless channels with polar codes as opposed to assuming that this operation can be perfectly realized.

An informal description of the strong coordination coding scheme is as follows. From $X^{1:N}$ and some common randomness of rate close to $H(V|X)$ 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 constructs a message with rate close to $I(X;V)$ such that Node 2 reconstructs $\widetilde{V}^{1:N}$ with the message and the common randomness. Then, Node 2 simulates a discrete memoryless channel with input $\widetilde{V}^{1:N}$ to form $\widetilde{Y}^{1:N}$ whose joint distribution with $X^{1:N}$ is close to $q_{X^{1:N}Y^{1:N}}$. Finally, *part* of the common randomness is recycled by encoding over $k \in \mathbb{N}^*$ blocks, so that the overall rate of shared randomness becomes on the order of $I(V;Y|X)$. As in Section IV for channel resolvability, the main difficulty is to ensure that the joint probability distributions of the actions approach the target distribution over all blocks jointly, despite reusing part of the common randomness over all blocks and despite an imperfect simulation of discrete memoryless channels. The coding scheme is formally described in Section VI-A, and its analysis is presented in Section VI-B.

A. Coding Scheme

We first redefine the following notation. Consider the random variables X, Y, V distributed according to q_{XYV} over $\mathcal{X}\times\mathcal{Y}\times\mathcal{V}$ such that $X \to V \to Y$. Moreover, assume that $|\mathcal{Y}|$ and $|\mathcal{V}|$ are prime numbers. By Theorem 13, one can choose $|\mathcal{V}|$ as the smallest prime number greater than or equal to $(|\mathcal{X}||\mathcal{Y}|+1)$. Let $N \triangleq 2^n$, $n \in \mathbb{N}^*$. Define $U^{1:N} \triangleq V^{1:N}G_n$, $T^{1:N} \triangleq Y^{1:N}G_n$, where G_n is defined in Section II, and define for $\beta < 1/2$ and $\delta_N \triangleq 2^{-N^\beta}$ the sets

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

Note that the sets $\mathcal{H}_V$, $\mathcal{V}_{V|X}$, $\mathcal{V}_{V|XY}$, and $\mathcal{V}_{Y|V}$ are defined with respect to q_{XYV}. Similar to the previous sections, we have

$$\begin{aligned}
\lim_{N\to\infty} |\mathcal{H}_V|/N &= H(V),\\
\lim_{N\to\infty} |\mathcal{V}_{V|X}|/N &= H(V|X),\\
\lim_{N\to\infty} |\mathcal{V}_{V|XY}|/N &= H(V|XY),\\
\lim_{N\to\infty} |\mathcal{V}_{Y|V}|/N &= H(Y|V).
\end{aligned}$$

Note also 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 $\mathcal{F}_4 \triangleq \mathcal{H}_V\backslash\mathcal{V}_{V|X}$ such 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 4 and 5, respectively. The functional dependence graph of the coding scheme is depicted in Figure 5.

B. Scheme Analysis

Although the coordination metric defined in Section III-C is the variational distance, our analysis will be performed with the Kullback-Leibler divergence to highlight similarities with the analysis of channel resolvability in Section IV-B. Reverting back to the variational distance is directly obtained with Pinsker's inequality.

The following lemma shows that $\widetilde{p}_{V_i^{1:N}X_i^{1:N}}$, defined by $\widetilde{p}_{X_i^{1:N}} \triangleq q_{X^{1:N}}$ and Equation (11), approximates $q_{V^{1:N}X^{1:N}}$.

Lemma 28: For any $i \in [\![1,k]\!]$,

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

where $\delta_N^{(A)} \triangleq N\delta_N$.

Proof: We have

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

Algorithm 4 Encoding algorithm at Node 1 for strong coordination

Require: A vector $C_{1:k}$ of $k|\mathcal{F}_3|$ uniformly distributed symbols over $[\![1,|\mathcal{V}|]\!]$ shared with Node 2. A vector $\bar{C}_1$ of $|\mathcal{F}_2|$ uniformly distributed symbols over $[\![1,|\mathcal{V}|]\!]$ shared with Node 2 and $X_{1:k}^{1:N}$.

for Block $i = 1$ to k **do**
- $\bar{C}_i \leftarrow \bar{C}_1$
- $\widetilde{U}_i^{1:N}[\mathcal{F}_2] \leftarrow \bar{C}_i$
- $\widetilde{U}_i^{1:N}[\mathcal{F}_3] \leftarrow C_i$
- Given $X_i^{1:N}$, successively draw the remaining components of $\widetilde{U}_i^{1:N}$ according to $\widetilde{p}_{U_i^{1:N}X_i^{1:N}}$ defined by

$$\widetilde{p}_{U_i^j|U_i^{1:j-1}X_i^{1:N}}(u_i^j|\widetilde{U}_i^{1:j-1}X_i^{1:N}) \triangleq \begin{cases} q_{U^j|U^{1:j-1}}(u_i^j|\widetilde{U}_i^{1:j-1}) & \text{if } j \in \mathcal{F}_1,\\ q_{U^j|U^{1:j-1}X^{1:N}}(u_i^j|\widetilde{U}_i^{1:j-1}X_i^{1:N}) & \text{if } j \in \mathcal{F}_4. \end{cases} \tag{11}$$

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

end for

Algorithm 5 Decoding algorithm at Node 2 for strong coordination

Require: The vectors $C_{1:k}$, $C_{1:k}'$, and $\bar{C}_1$ used in Algorithm 4 and $M_{1:k}$.

for Block $i = 1$ to k **do**
- $\bar{C}_i \leftarrow \bar{C}_1$
- $\widetilde{U}_i^{1:N}[\mathcal{F}_2] \leftarrow \bar{C}_i$
- $\widetilde{U}_i^{1:N}[\mathcal{F}_3] \leftarrow C_i$
- $\widetilde{U}_i^{1:N}[\mathcal{F}_4] \leftarrow M_i$
- Using C_i', successively draw the remaining components of $\widetilde{U}_i^{1:N}$ according to $q_{U^j|U^{1:j-1}}$
- $\widetilde{V}_i^{1:N} \leftarrow \widetilde{U}_i^{1:N}G_n$
- Channel simulation: Given $\widetilde{V}_i^{1:N}$, successively draw the components of $\widetilde{T}_i^{1:N}$ according to $\widetilde{p}_{T_i^{1:N}V_i^{1:N}}$ defined by

$$\widetilde{p}_{T_i^j|T_i^{1:j-1}V_i^{1:N}}(t_i^j|\widetilde{T}_i^{1:j-1}\widetilde{V}_i^{1:N}) \triangleq \begin{cases} 1/|\mathcal{Y}| & \text{if } j \in \mathcal{V}_{Y|V},\\ q_{T^j|T^{1:j-1}V^{1:N}}(t_i^j|\widetilde{T}_i^{1:j-1}\widetilde{V}_i^{1:N}) & \text{if } j \in \mathcal{V}_{Y|V}^c. \end{cases} \tag{12}$$

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

end for

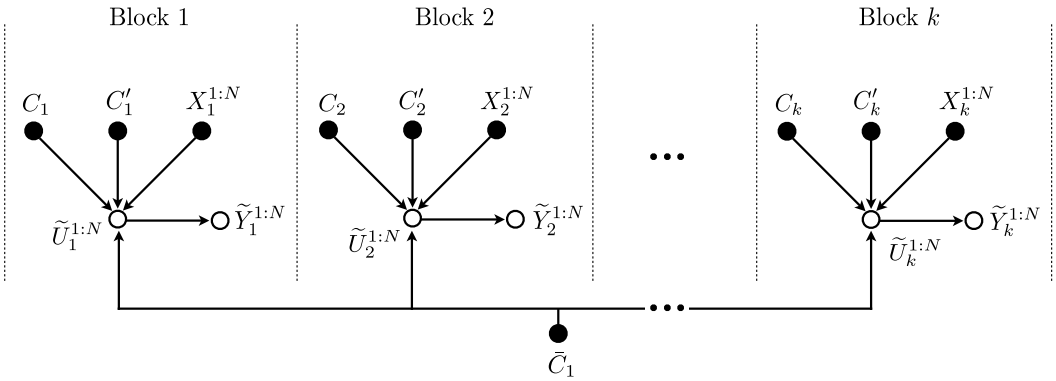

Fig. 5. Functional dependence graph of the block encoding scheme for strong coordination. For Block i, $(C_i, C_i', \bar{C}_i)$ is the common randomness shared by Node 1 and 2 used at the encoder to form $\widetilde{U}_i$, where $\bar{C}_i = \bar{C}_1$ is reused over all blocks.

where (a) holds by invertibility of G_n, (b) and (c) hold by the chain rule for divergence [42], (d) holds by (11), (e) holds by uniformity of $\widetilde{U}_i^{1:N}[\mathcal{F}_2 \cup \mathcal{F}_3] = \widetilde{U}_i^{1:N}[\mathcal{V}_{V|X}]$, and by definition of $\widetilde{p}_{U_i^j|U_i^{1:j-1}X_i^{1:N}}$ in (11). ■

Remark 29: By Lemma 28 and because $|\mathcal{F}_2 \cup \mathcal{F}_3|/N = |\mathcal{V}_{V|X}|/N \xrightarrow{N\to\infty} H(V|X)$, *note that the encoding algorithm of Section VI-A performs a conditional resolvability-achieving random number generation in each block.*

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

Lemma 30: For $i \in [\![1, k]\!]$, *we have*

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

where $\delta_N^{(B)} = O\left(N^{3/2}\delta_N^{1/2}\right)$.

Proof: We have

$$\begin{aligned}&\mathbb{D}(\widetilde{p}_{V_i^{1:N}X_i^{1:N}Y_i^{1:N}}||q_{V^{1:N}X^{1:N}Y^{1:N}}) \\ &= \mathbb{D}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}X_i^{1:N}}\widetilde{p}_{V_i^{1:N}X_i^{1:N}}||q_{Y^{1:N}|V^{1:N}X^{1:N}}q_{V^{1:N}X^{1:N}}) \\ &= \mathbb{D}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}}\widetilde{p}_{V_i^{1:N}X_i^{1:N}}||q_{Y^{1:N}|V^{1:N}}q_{V^{1:N}X^{1:N}}) \\ &\overset{(a)}{\leqslant} N\log\left(\frac{1}{\mu_{VXY}}\right)\sqrt{2\ln 2} \\ &\quad\times\Bigg[\sqrt{\mathbb{D}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}}q_{V^{1:N}X^{1:N}}||\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}}\widetilde{p}_{V_i^{1:N}X_i^{1:N}})} \\ &\qquad+\sqrt{\mathbb{D}(\widetilde{p}_{Y_i^{1:N}|V_i^{1:N}}q_{V^{1:N}X^{1:N}}||\widetilde{p}_{Y_i^{1:N}V_i^{1:N}}q_{X^{1:N}|V^{1:N}})} \\ &\qquad+\sqrt{\mathbb{D}(q_{Y^{1:N}|V^{1:N}}q_{V^{1:N}X^{1:N}}||\widetilde{p}_{Y_i^{1:N}V_i^{1:N}}q_{X^{1:N}|V^{1:N}})}\Bigg] \\ &= N\log\left(\frac{1}{\mu_{VXY}}\right)\sqrt{2\ln 2}\Bigg[\sqrt{\mathbb{D}(q_{V^{1:N}X^{1:N}}||\widetilde{p}_{V_i^{1:N}X_i^{1:N}})} \\ &\quad+\sqrt{\mathbb{D}(q_{V^{1:N}}||\widetilde{p}_{V_i^{1:N}})}+\sqrt{\mathbb{D}(q_{Y^{1:N}V^{1:N}}||\widetilde{p}_{Y_i^{1:N}V_i^{1:N}})}\Bigg] \\ &\overset{(b)}{\leqslant} N\log\left(\frac{1}{\mu_{VXY}}\right)\sqrt{2\ln 2}\left[2\sqrt{\delta_N^{(A)}}+\sqrt{\mathbb{D}(q_{Y^{1:N}V^{1:N}}||\widetilde{p}_{Y_i^{1:N}V_i^{1:N}})}\right],\end{aligned} \tag{13}$$

where (a) holds similar to the proof of Lemma 38 in Appendix A, (b) holds by Lemma 28. We bound the right-hand side of (13) by analyzing Step 8 of Algorithm 5 as follows:

$$\begin{aligned}&\mathbb{D}(q_{Y^{1:N}V^{1:N}}||\widetilde{p}_{Y_i^{1:N}V_i^{1:N}}) \\ &\overset{(a)}{=} \mathbb{D}(q_{T^{1:N}V^{1:N}}||\widetilde{p}_{T_i^{1:N}V_i^{1:N}}) \\ &\overset{(b)}{=} \mathbb{E}_{q_{V^{1:N}}}\left[\mathbb{D}(q_{T^{1:N}|V^{1:N}}||\widetilde{p}_{T_i^{1:N}|V_i^{1:N}})\right]+\mathbb{D}(q_{V^{1:N}}||\widetilde{p}_{V_i^{1:N}}) \\ &\overset{(c)}{=} \mathbb{E}_{q_{V^{1:N}}}\left[\mathbb{D}(q_{T^{1:N}|V^{1:N}}||\widetilde{p}_{T_i^{1:N}|V_i^{1:N}})\right]+\mathbb{D}(q_{U^{1:N}}||\widetilde{p}_{U_i^{1:N}}) \\ &\overset{(d)}{\leqslant} N\delta_N+\mathbb{E}_{q_{V^{1:N}}}\left[\mathbb{D}(q_{T^{1:N}|V^{1:N}}||\widetilde{p}_{T_i^{1:N}|V_i^{1:N}})\right] \\ &\overset{(e)}{=} N\delta_N+\sum_{j=1}^{N}\mathbb{E}_{q_{T^{1:j-1}V^{1:N}}}\left[\mathbb{D}(q_{T^j|T^{1:j-1}V^{1:N}}||\widetilde{p}_{T_i^j|T_i^{1:j-1}V_i^{1:N}})\right] \\ &\overset{(f)}{=} N\delta_N+\sum_{j\in\mathcal{V}_{Y|V}}\mathbb{E}_{q_{T^{1:j-1}V^{1:N}}}\left[\mathbb{D}(q_{T^j|T^{1:j-1}V^{1:N}}||\widetilde{p}_{T_i^j|T_i^{1:j-1}V_i^{1:N}})\right] \\ &\overset{(g)}{=} N\delta_N+\sum_{j\in\mathcal{V}_{Y|V}}(\log|\mathcal{Y}|-H(T^j|T^{1:j-1}V^{1:N})) \\ &\overset{(h)}{\leqslant} N\delta_N+|\mathcal{V}_{Y|V}|\delta_N \\ &\leqslant 2N\delta_N,\end{aligned} \tag{14}$$

where (a) and (c) hold by invertibility of G_n and the data processing inequality, (b) holds by the chain rule for divergence [42, Th. 2.5.3], (d) holds because $\mathbb{D}(q_{U^{1:N}}||\widetilde{p}_{U_i^{1:N}}) \leqslant \mathbb{D}(q_{U^{1:N}X^{1:N}}||\widetilde{p}_{U_i^{1:N}X_i^{1:N}})$ (by the chain rule for divergence and positivity of the divergence) and because $\mathbb{D}(q_{U^{1:N}X^{1:N}}||\widetilde{p}_{U_i^{1:N}X_i^{1:N}}) \leqslant N\delta_N$ by the proof of Lemma 28, (e) holds by the chain rule for divergence, (f) and (g) hold by (12), (h) holds by definition of $\mathcal{V}_{Y|V}$. Finally, combining (13), (14) yields the claim. ■

Using Lemma 30, we now establish the asymptotic independence of consecutive blocks.

Lemma 31: For $i \in [\![2, k]\!]$, *we have,*

$$\mathbb{D}\left(\widetilde{p}_{X_{i-1:i}^{1:N}Y_{i-1:i}^{1:N}\bar{C}_1}||\widetilde{p}_{Y_{i-1}^{1:N}X_{i-1}^{1:N}}\widetilde{p}_{X_i^{1:N}Y_i^{1:N}\bar{C}_1}\right) \leqslant \delta_N^{(C)},$$

where $\delta_N^{(C)} = O\left(N^{9/4}\delta_N^{1/4}\right)$.

Proof: We reuse the proof of Lemma 17 with the substitutions $q_{U^{1:N}} \leftarrow q_{V^{1:N}}$, $q_{Y^{1:N}} \leftarrow q_{X^{1:N}Y^{1:N}}$, $\widetilde{p}_{U_i^{1:N}} \leftarrow$

$\widetilde{p}_{Y_i^{1:N}}, \widetilde{p}_{Y_i^{1:N}} \leftarrow \widetilde{p}_{X_i^{1:N}Y_i^{1:N}}$. Note that the Markov condition $X_{i-1}^{1:N}\widetilde{Y}_{i-1}^{1:N} - \bar{C}_1 - X_i^{1:N}\widetilde{Y}_i^{1:N}$ holds as seen in Figure 5. ■

Using Lemma 31 we next show the asymptotic independence across all blocks.

Lemma 32: We have

$$\mathbb{D}\left(\widetilde{p}_{X_{1:k}^{1:N}Y_{1:k}^{1:N}} || \prod_{i=1}^{k} \widetilde{p}_{X_i^{1:N}Y_i^{1:N}}\right) \leqslant (k-1)\delta_N^{(C)}.$$

where $\delta_N^{(C)}$ is defined in Lemma 31.

Proof: We reuse the proof of Lemma 18 with the substitutions $\widetilde{p}_{Y_i^{1:N}} \leftarrow \widetilde{p}_{X_i^{1:N}Y_i^{1:N}}$. Note that the Markov condition $X_{1:i-2}^{1:N}\widetilde{Y}_{1:i-2}^{1:N} - \bar{C}_1\widetilde{X}_{i-1}^{1:N}\widetilde{Y}_{i-1}^{1:N} - X_i^{1:N}\widetilde{Y}_i^{1:N}$ holds, as seen in Figure 5. ■

Using Lemmas 30 and 32, we now show that strong coordination holds over all blocks.

Lemma 33: We have

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

where $\delta_N^{(D)} = O\left(k^{3/2}N^{17/8}\delta_N^{1/8}\right)$.

Proof: We reuse the proof of Lemma 19 with the substitutions $q_{Y^{1:N}} \leftarrow q_{X^{1:N}Y^{1:N}}$, $\widetilde{p}_{Y_i^{1:N}} \leftarrow \widetilde{p}_{X_i^{1:N}Y_i^{1:N}}$. ■

We can now state our final result as follows.

Theorem 34: The coding scheme described in Algorithms 4 and 5, which operate over k blocks of length N, achieves the two-node network strong coordination capacity region of Theorem 13 for an arbitrary target distribution q_{XY} over $\mathcal{X} \times \mathcal{Y}$, where $|\mathcal{Y}|$ is a prime number. The coding scheme is explicit with complexity in $O(kN \log N)$.

Proof: We prove that the communication rate and the common randomness rate are optimal. The common randomness rate is

$$\begin{aligned}\frac{|\bar{C}_1| + |C_{1:k}|}{kN} &= \frac{|\mathcal{V}_{V|XY}| + k|\mathcal{V}_{V|X} \backslash \mathcal{V}_{V|XY}|}{kN} \\ &= \frac{|\mathcal{V}_{V|XY}|}{kN} + \frac{|\mathcal{V}_{V|X}| - |\mathcal{V}_{V|XY}|}{N} \\ &\xrightarrow{N\to\infty} I(V;Y|X) + \frac{H(V|XY)}{k} \\ &\xrightarrow{k\to\infty} I(V;Y|X), \end{aligned} \tag{15}$$

where we have used [39, Lemma 7].

Next, we determine the communication rate. Observe first that for any $i \in [\![1, k]\!]$,

$$\lim_{N\to\infty} \frac{|C_i'|}{N} = \lim_{N\to\infty} \frac{1}{N} \sum_{j \in \mathcal{H}_V^c} H(\widetilde{V}_i^j | \widetilde{V}_i^{1:j-1}) = 0,$$

which can be proved using Lemma 30 similar to the proof of Theorem 21. Hence, the communication rate is

$$\begin{aligned}\frac{|\widetilde{U}_{1:k}^{1:N}[\mathcal{F}_4]|}{kN} &= \frac{k|\mathcal{F}_4|}{kN} \\ &= \frac{|\mathcal{V}_V| - |\mathcal{V}_{V|X}|}{N} \\ &\xrightarrow{N\to\infty} I(V;X), \end{aligned} \tag{16}$$

where the limit holds by [39, Lemma 7].

We also have that the communication rate and the common randomness rate sum to

$$I(V;X) + I(V;Y|X) = I(V;XY),$$

which together with (15) and (16) recovers the bounds of Theorem 13. This result along with Lemma 33 allow us to conclude. ■

Remark 35: The rate of local randomness $H(Y|V)$ used at Node 2 is optimal. It is possible to extend the converse proof of [3] and include the local randomness rate R_L at Node 2 to show that the strong coordination capacity region is

$$\begin{aligned}&\mathcal{R}_{\mathrm{SC}}(q_{XY}) \\ &\triangleq \bigcup_{\substack{X\to V\to Y \\ |\mathcal{V}|\leqslant|\mathcal{X}||\mathcal{Y}|+1}} \{(R, R_0, R_L) : R + R_L \geqslant I(X;V) + H(Y|V), \\ &\qquad R + R_0 + R_L \geqslant I(XY;V) + H(Y|V)\}.\end{aligned}$$

VII. Concluding remarks

We have demonstrated the ability of polar codes to provide solutions to problems related to soft covering. Specifically, we have proposed an explicit and low-complexity coding scheme for channel resolvability by relying on (i) conditional resolvability-achieving random number generation and (ii) randomness recycling through block-Markov encoding. As discussed in the introduction, our coding scheme generalizes previous explicit coding schemes that achieve channel resolvability but were restricted to uniform distributions or symmetric channels.

Furthermore, by leveraging the coding scheme for channel resolvability, we have proposed explicit and low-complexity polar coding schemes that achieve the capacity regions of empirical coordination and strong coordination in two-node networks.

Note that all our coding schemes require that the cardinality of the alphabet of actions at Node 2 (for Sections V, VI) or of the channel input alphabet (for Section IV) be a prime number. This assumption could be removed if [39, Lemma 6,7] used throughout our proofs to establish various rates could be extended to alphabet with arbitrary cardinalities. Such an extension of [39, Lemma 6] is provided in [43], however, the problem of obtaining a similar extension for [39, Lemma 7] remains open.

Appendix A
Simple Relations for the Kullback-Leibler Divergence

We provide in this appendix simple relations satisfied by the Kullback-Leibler divergence. More specifically, Lemma 36 allows us to characterize the symmetry of the Kullback-Leibler divergence around zero. Lemma 37 allows us to express independence in terms of mutual informations. Lemma 38 describes a relation similar to the triangle inequality for small values of the Kullback-Leibler divergence. Finally, Lemma 39 provides an upper-bound on the difference of the entropy of two random variables defined over the same alphabet. In the proofs of the following lemmas, we only need to consider the case where all the Kullback-Leibler divergences are finite, since the lemmas trivially hold when any of the Kullback-Leibler divergences appearing in the lemmas is infinite (see Section II for the convention used in the definition of the Kullback-Leibler divergence).

Lemma 36: Let p, q be two distributions over the finite alphabet $\mathcal{X}$. We have

$$\mathbb{D}(p\|q) \leqslant \log\left(\frac{1}{\mu_q}\right)\sqrt{2\ln 2}\sqrt{\mathbb{D}(q||p)},$$

where $\mu_q \triangleq \min_{x\in\mathcal{X}} q(x)$.

Proof: The result follows from Pinsker's inequality and the following inequality [45, Eq. (323)].

$$\mathbb{D}(p\|q) \leqslant \log\left(\frac{1}{\mu_q}\right)\mathbb{V}(p,q). \tag{17}$$

■

Lemma 37: Let $(X_i)_{i\in[\![1,k]\!]}$ be arbitrary discrete random variables with joint probability $p_{X_{1:k}}$. We have

$$\mathbb{D}\left(p_{X_{1:k}}\|\prod_{i=1}^{k} p_{X_i}\right) = \sum_{i=2}^{k} I(X_i; X_{1:i-1}).$$

Proof: We have by the chain rule for relative Kullback-Leibler divergence [42]

$$\begin{aligned}\mathbb{D}\left(p_{X_{1:k}}\|\prod_{i=1}^{k} p_{X_i}\right) &= \sum_{i=1}^{k}\mathbb{E}_{X_{1:i-1}}\left[\mathbb{D}\left(p_{X_i|X_{1:i-1}}\|p_{X_i}\right)\right]\\ &= \sum_{i=1}^{k}\mathbb{D}\left(p_{X_{1:i}}\|p_{X_{1:i-1}}p_{X_i}\right)\\ &= \sum_{i=2}^{k} I(X_i; X_{1:i-1}).\end{aligned}$$

■

Lemma 38: Let p, q, and r be distributions over the finite alphabet $\mathcal{X}$. We have

$$\begin{aligned}\mathbb{D}(p\|q) \leqslant \log\left(\frac{1}{\mu_q}\right)\sqrt{2\ln 2}\Big[&\sqrt{\min(\mathbb{D}(p\|r), \mathbb{D}(r\|p))}\\ &+\sqrt{\min(\mathbb{D}(q\|r), \mathbb{D}(r\|q))}\Big],\end{aligned}$$

where $\mu_q \triangleq \min_{x\in\mathcal{X}} q(x)$.

Proof: The result follows from (17), the triangle inequality for the variational distance, and Pinsker's inequality. ■

Lemma 39: Let p, q be distributions over the finite alphabet $\mathcal{X}$. Let $H(p)$ and $H(q)$ denote the Shannon entropy associated with p and q, respectively. Provided that $\mathbb{D}(p||q)$ or $\mathbb{D}(q||p)$ is small enough, we have

$$|H(q) - H(p)| \leqslant D\log\left(\frac{|\mathcal{X}|}{D}\right),$$

where $D \triangleq \sqrt{2\ln 2}\sqrt{\min(\mathbb{D}(p||q), \mathbb{D}(q||p))}$.

Proof: The result follows by [46, Lemma 2.7], Pinsker's inequality, and because $x \mapsto -x\log x$ is increasing over $]0, e^{-1}]$. ■

Appendix B
Proof of Remark 16

We first observe that $|\mathcal{V}_X|/N \xrightarrow{N\to\infty} H(X)$. Then, for $i \in [\![1,k]\!]$,

$$\begin{aligned}&\mathbb{D}(\widetilde{p}_{Y_i^{1:N}}||q_{Y^{1:N}})\\ &\quad\overset{(a)}{\leqslant} N\log\left(\frac{1}{\mu_{q_Y}}\right)\sqrt{2\ln 2}\sqrt{\mathbb{D}\left(q_{Y^{1:N}}||\widetilde{p}_{Y_i^{1:N}}\right)}\\ &\quad\overset{(b)}{\leqslant} N\log\left(\frac{1}{\mu_{q_Y}}\right)\sqrt{2\ln 2}\sqrt{\delta_N^{(1)}},\end{aligned}$$

where (a) holds by Lemma 36 in Appendix A with $\mu_{q_{Y^{1:N}}} = \mu_{q_Y}^N$, (b) holds by the chain rule and Lemma 15.

References

[1] R. A. Chou, M. R. Bloch, and J. Kliewer, "Polar coding for empirical and strong coordination via distribution approximation," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 1512–1516.

[2] E. Soljanin, "Compressing quantum mixed-state sources by sending classical information," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2263–2275, Aug. 2002.

[3] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4181–4206, Sep. 2010.

[4] P. Cuff, "Distributed channel synthesis," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7071–7096, Nov. 2013.

[5] M. H. Yassaee, A. Gohari, and M. R. Aref, "Channel simulation via interactive communications," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 2964–2982, Jun. 2015.

[6] A. Bereyhi, M. Bahrami, M. Mirmohseni, and M. R. Aref, "Empirical coordination in a triangular multiterminal network," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2149–2153.

[7] F. Haddadpour, M. H. Yassaee, A. Gohari, and M. R. Aref, "Coordination via a relay," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, Jul. 2012, pp. 3048–3052.

[8] M. R. Bloch and J. Kliewer, "Strong coordination over a three-terminal relay network," in *Proc. IEEE Inf. Theory Workshop*, Hobart, TAS, Australia, Nov. 2014, pp. 646–650.

[9] B. N. Vellambi, J. Kliewer, and M. R. Bloch, "Strong coordination over multi-hop line networks using channel resolvability codebooks," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1132–1162, Feb. 2018.

[10] M. L. Treust, "Joint empirical coordination of source and channel," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 5087–5114, Aug. 2017.

[11] B. Larrousse, S. Lasaulce, and M. R. Bloch, "Coordination in distributed networks via coded actions with application to power control," *IEEE Trans. Inf. Theory*, to be published.

[12] W. Dur, G. Vidal, and J. I. Cirac, "Visible compression of commuting mixed states," *Phys. Rev. A, Gen. Phys.*, vol. 64, no. 2, pp. 022308-1–022308-7, 2001.

[13] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett.*, vol. 83, no. 15, p. 3081, 1999.

[14] M. Hayashi, *Quantum Information: An Introduction*. Berlin, Germany: Springer-Verlag, 2006.

[15] R. Blasco-Serrano, R. Thobaben, and M. Skoglund, "Polar codes for coordination in cascade networks," in *Proc. Int. Zurich Seminar Commun.*, Zürich, Switzerland, 2012, pp. 55–58.

[16] M. R. Bloch, L. Luzzi, and J. Kliewer, "Strong coordination with polar codes," in *Proc. Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Oct. 2012, pp. 565–571.

[17] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[18] E. Arikan, "Source polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 899–903.

[19] R. Chou, M. R. Bloch, and J. Kliewer, "Low-complexity channel resolvability codes for the symmetric multiple-access channel," in *Proc. IEEE Inf. Theory Workshop*, Hobart, TAS, Australia, Nov. 2014, pp. 466–470.

[20] M. Hayashi and R. Matsumoto, "Secure multiplex coding with dependent and non-uniform multiple messages," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2355–2409, May 2016.
[21] R. A. Amjad and G. Kramer, "Channel resolvability codes based on concatenation and sparse linear encoding," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 2111–2115.
[22] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric models," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, Dec. 2013.
[23] A. D. Wyner, "The common information of two dependent random variables," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 163–179, Mar. 1975.
[24] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.
[25] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, Apr. 2006.
[26] I. Csiszár, "Almost independence and secrecy capacity," *Problems Inf. Transmiss.*, vol. 32, no. 1, pp. 40–47, 1996.
[27] M. Bloch and J. Kliewer, "Strong coordination over a line network," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2319–2323.
[28] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Semantic-security capacity for wiretap channels of type II," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3863–3879, Jul. 2016.
[29] M. Hayashi, "Quantum wiretap channel with non-uniform random number and its exponent and equivocation rate of leaked information," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5595–5622, Oct. 2015.
[30] J. Hou and G. Kramer, "Effective secrecy: Reliability, confusion and stealth," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun. 2014, pp. 601–605.
[31] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.
[32] S. H. Hassani and R. Urbanke, "Universal polar codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun. 2014, pp. 1451–1455.
[33] M. Mondelli, S. Hassani, and R. Urbanke, "How to achieve the capacity of asymmetric channels," in *Proc. Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Sep. 2014, pp. 789–796.
[34] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6213–6237, Nov. 2015.
[35] R. A. Chou and M. R. Bloch, "Using deterministic decisions for low-entropy bits in the encoding and decoding of polar codes," in *Proc. Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, Sep. 2015, pp. 1380–1385.
[36] S. Vadhan, "Pseudorandomness," *Found. Trends Theor. Comput. Sci.*, vol. 7, nos. 1–3, pp. 1–336, 2012.
[37] G. Kramer, "Topics in multi-user information theory," *Found. Trends Commun. Inf. Theory*, vol. 4, nos. 4–5, pp. 265–444, 2007.
[38] Y. Steinberg and S. Verdú, "Simulation of random processes and rate-distortion theory," *IEEE Trans. Inf. Theory*, vol. 42, no. 1, pp. 63–86, Jan. 1996.
[39] R. A. Chou and M. R. Bloch, "Polar coding for the broadcast channel with confidential messages: A random binning analogy," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2410–2429, May 2016.
[40] M. Karzand and I. Telatar, "Polar codes for q-ary source coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 909–912.
[41] T. S. Han, *Information-Spectrum Methods in Information Theory*. vol. 50. Berlin, Germany: Springer-Verlag, 2003.
[42] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2001.
[43] E. Sasoglu, "Polar coding theorems for discrete systems," Ph.D. thesis, School Comput. Commun. Sci., École Polytechn. Fédérale Lausanne, Lausanne, Switzerland, 2011.
[44] S. B. Korada and R. L. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751–1768, Apr. 2010.
[45] I. Sason and S. Verdú, "f-divergence inequalities," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 5973–6006, Nov. 2016.
[46] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge, U.K.: Cambridge Univ. Press, 1981.

Rémi A. Chou (M'17) is an Assistant Professor in the Electrical Engineering and Computer Science Department at Wichita State University. He received the Engineering degree from Supélec, Gif-sur-Yvette, France, in 2011, and the Ph.D. degree in Electrical Engineering from the Georgia Institute of Technology, Atlanta, in 2015. From 2015 to 2017, he was a Postdoctoral Scholar at The Pennsylvania State University, University Park.

Matthieu R. Bloch (M'08–SM'16) is an Associate Professor in the School of Electrical and Computer Engineering. He received the Engineering degree from Supélec, Gif-sur-Yvette, France, the M.S. degree in Electrical Engineering from the Georgia Institute of Technology, Atlanta, in 2003, the Ph.D. degree in Engineering Science from the Université de Franche-Comté, Besançon, France, in 2006, and the Ph.D. degree in Electrical Engineering from the Georgia Institute of Technology in 2008. In 2008-2009, he was a postdoctoral research associate at the University of Notre Dame, South Bend, IN. Since July 2009, Dr. Bloch has been on the faculty of the School of Electrical and Computer Engineering, and from 2009 to 2013 Dr. Bloch was based at Georgia Tech Lorraine. His research interests are in the areas of information theory, error-control coding, wireless communications, and cryptography. Dr. Bloch has served on the organizing committee of several international conferences; he was the chair of the Online Committee of the IEEE Information Theory Society from 2011 to 2014, and he has been on the Board of Governors of the IEEE Information Theory Society and an Associate Editor for the *IEEE Transactions on Information* since 2016. He is the co-recipient of the IEEE Communications Society and IEEE Information Theory Society 2011 Joint Paper Award and the co-author of the textbook *Physical-Layer Security: From Information Theory to Security Engineering* published by Cambridge University Press.

Jörg Kliewer (S'97–M'99–SM'04) received the Dipl.-Ing. (M.Sc.) degree in electrical engineering from Hamburg University of Technology, Hamburg, Germany, in 1993 and the Dr.-Ing. degree (Ph.D.) in electrical engineering from the University of Kiel, Germany, in 1999, respectively.

From 1993 to 1998, he was a research assistant at the University of Kiel, and from 1999 to 2004, he was a senior researcher and lecturer with the same institution. In 2004, he visited the University of Southampton, U.K., for one year, and from 2005 until 2007, he was with the University of Notre Dame, IN, as a Visiting assistant professor. From 2007 until 2013 he was with New Mexico State University, Las Cruces, NM, most recently as an associate professor. He is now with the New Jersey Institute of Technology, Newark, NJ, as an associate professor. His research interests span information and coding theory, graphical models, and statistical algorithms, which includes applications to networked communication and security, data storage, and biology.

Dr. Kliewer was the recipient of a Leverhulme Trust Award and a German Research Foundation Fellowship Award in 2003 and 2004, respectively. He was an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS from 2008 until 2014, and since 2015 serves as an Area Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. He is also an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY since 2017 and a member of the editorial board of the IEEE Information Theory Newsletter since 2012.