# Empirical Coordination Subject to a Fidelity Criterion

Michail Mylonakis Division of Inf. Science & Eng. KTH Royal Institute of Technology mmyl@kth.se Photios A. Stavrou Division of Inf. Science & Eng. KTH Royal Institute of Technology fstavrou@kth.se Mikael Skoglund Division of Inf. Science & Eng. KTH Royal Institute of Technology skoglund@kth.se

Abstract—We study the problem of empirical coordination subject to a fidelity criterion for a general set-up. We prove a result which indicates a strong connection between our framework and the framework of empirical coordination developed in [1]. It turns out that when we design codes that achieve empirical coordination according to a given distribution and subject to the fidelity criterion, it is sufficient to consider codes that produce actions of the same joint type for a class of types which is close enough to our desired distribution is some sense.

### I. INTRODUCTION

Communication is one of the most important and expensive resources in a network with nodes who desire to establish cooperative behavior. When relevant information is known only at some nodes, finding the minimum communication requirements to coordinate actions can be posed as a network information theory problem. Specifically, we consider the communication needed to establish coordination summarized by a joint probability distribution of behavior among all nodes in the network (see Fig. 1).

Cuff et al. in [1] introduced two different notions of coordination, empirical and strong. According to [1], empirical coordination is established if the joint type of the actions in the network is close to the desired distribution. This kind of coordination has been studied in various set-ups (see e.g., [2]-[4]) and it has been combined with ideas from other fields like game theory (see e.g., [5]). Strong coordination instead deals with the joint probability distribution of the actions. If the actions in the network are generated randomly so that a statistician cannot reliably distinguish (as measured by total variation) between the constructed *n*-length sequence of actions and random samples from the desired distribution, then strong coordination is achieved. The literature related to strong coordination is vast and includes more complex set-ups, such as extensions in networks with noisy channels, applications in power control etc, (see, e.g., [6]-[8]). It should be noted that [1] establishes a fundamental difference between empirical and strong coordination regarding the impact that common randomness can have to the accomplishment of the coordinated behaviour. Specifically, it turns out that common randomness does not play a necessary role in achieving em-



Fig. 1. An illustrative example of the considered framework. The nodes in this network have rate-limited communication links between them. In this example, actions  $X_1, X_2$ , and  $X_3$  are assigned according to the joint distribution  $p_0(x_1, x_2, x_3)$ . Then, using the communication that is available to all nodes, the actions  $Y_1, Y_2$ , and  $Y_3$  are produced. We ask, the amount of the communication that is required such that the average distance between the joint type of the actions and the desired distribution to be smaller than a certain level  $\Delta$ .

pirical coordination but it is a valuable resource for achieving strong coordination. For more details on the foundation of common randomness, see [9].

In this paper, we deal with empirical coordination. However, instead of asking the joint type of the actions to be close to the desired distribution in probability, we require something less restrictive. We demand the average distance (as measured by total variation) between the joint type of the actions and the desired distribution to be smaller than a certain level  $\Delta$  for long enough *n*-length sequences. To distinguish between the two kinds of empirical coordination, i.e., ours and the one in [1], we will call the one established in [1] *perfect* empirical coordination. This kind of imperfect empirical coordination was first introduced and studied for a specific set-up (including a more general class of fidelity metrics) by Kramer and Savari in [10]. Here,



Fig. 2. To simplify things, we define formally our problems and we state our general result in the context of the cascade set-up of the figure.

we generalize this framework (in the case of the total variation fidelity metric) for a general network setting. We establish formally a coordination result which shows that when our task is to design a good coordination code that achieves imperfect empirical coordination according to some distribution, it is optimal to restrict our search exactly to the class of codes that achieve perfect empirical coordination according to any distribution from a class of distributions which is close enough to our desired distribution with respect to the total variation distance.

#### **II. DEFINITIONS**

In this section, we state the definitions of perfect and imperfect empirical coordination in the context of the cascade network of Fig 2. These definitions have obvious generalizations to other networks. We begin with some basic mathematical concepts and the definition of the  $\Delta$ -neighborhood, a concept which will help us in the statement of our results.

Definition 1 (Joint type): The joint type  $P_{x^n,y^n,z^n}$  of a tuple of sequences  $(x^n, y^n, z^n)$  is the empirical probability mass function, given by

$$P_{x^{n},y^{n},z^{n}}(x,y,z) \triangleq \frac{1}{n} \sum_{i=1}^{n} \mathbf{1} ((x_{i},y_{i},z_{i}) = (x,y,z)),$$

for all  $(x, y, z) \in \mathbb{X} \times \mathbb{Y} \times \mathbb{Z}$ , where 1 is the indicator function.

*Definition 2 (Total variation):* The total variation between two probability mass functions is given by

$$\|p(x, y, z) - q(x, y, z)\|_{TV} \triangleq \frac{1}{2} \sum_{x, y, z} |p(x, y, z) - q(x, y, z)|.$$

Definition 3 ( $\Delta$ -neighborhood): The  $\Delta$ -neighborhood of a distribution p(x, y, z) is defined as

$$N_{\Delta}(p(x, y, z)) \\ \triangleq \{q(x, y, z) : \|p(x, y, z) - q(x, y, z)\|_{TV} \le \Delta\}.$$

A  $(2^{nR_1}, 2^{nR_2}, n)$  coordination code is the protocol which is used to coordinate the actions in the network for a block of n time periods. The coordination code and the distribution of the random actions,  $X_n$ , induce a joint distribution on the actions in the network. Definition 4 (Coordination code [1]): The  $(2^{nR_1}, 2^{nR_2}, n)$  coordination code for our set-up consist of four functions-an encoding function

$$i: \mathbb{X}^n \to \{1, \ldots, 2^{nR_1}\},\$$

a recoding function

$$j: \{1, \dots, 2^{nR_1}\} \to \{1, \dots, 2^{nR_2}\},\$$

and two decoding functions

$$y^{n}: \{1, \dots, 2^{nR_{1}}\} \to \mathbb{Y}^{n},$$
  
 $z^{n}: \{1, \dots, 2^{nR_{2}}\} \to \mathbb{Z}^{n}.$ 

Definition 5 (Induced distribution [1]): The induced distribution  $\tilde{p}(x^n, y^n, z^n)$  is the resulting joint distribution of the actions in the network  $X^n, Y^n$ , and  $Z^n$  when a  $(2^{nR_1}, 2^{nR_2}, n)$  coordination code is used.

In our set-up, the actions  $X^n$  are chosen by nature to be i.i.d according to  $p_0(x)$ . Thus,  $X^n$  is distributed according to a product distribution

$$X^n \sim \prod_{i=1}^n p_0\left(x_i\right).$$

The actions  $Y^n$  and  $Z^n$  are functions of  $X^n$  given by implementing the coordination code as

$$Y^{n} = y^{n} \Big( i \left( X^{n} \right) \Big),$$
$$Z^{n} = z^{n} \Big( j \big( i \left( X^{n} \right) \big) \Big)$$

Perfect empirical coordination is achieved if the joint type of the actions in the network tends to the desired distribution with high probability.

Definition 6 (Achievability for perfect coordination [1]): A desired distribution  $p_{X,Y,Z}(x,y,z) \triangleq p_0(x) p_{YZ|X}(y,z|x)$  is achievable for empirical coordination with the rate-pair  $(R_1, R_2)$  if there exists a sequence of  $(2^{nR_1}, 2^{nR_2}, n)$  coordination codes such that

$$||P_{x^{n},y^{n},z^{n}}(x,y,z) - p_{0}(x) p_{Y,Z|X}(y,z|x) ||_{TV} \to 0,$$

in probability.

Definition 7 (Coordination capacity region [1]): The coordination capacity region  $C_{p_0}^P$  for the source distribution  $p_0(x)$  is the closure of the set of rate-coordination tuples  $(R_1, R_2, p_{Y,Z|X}(y, z|x))$  that are achievable:

$$C_{p_0}^{P} \triangleq \mathbf{Cl} \left\{ \begin{array}{l} \left( R_1, R_2, p_{Y,Z|X} \left( y, z | x \right) \right) :\\ p_0 \left( x \right) p_{Y,Z|X} \left( y, z | x \right) \\ \text{is achievable at rates } \left( R_1, R_2 \right) \end{array} \right\}.$$

Imperfect empirical coordination is achieved if the average distance between the joint type of the actions in the network and the desired distribution is under a certain level  $\Delta$ . We will call this kind of coordination  $\Delta$ -empirical coordination.

Definition 8 (Achievability for  $\Delta$ -empirical coordination): A desired distribution  $p_{X,Y,Z}(x, y, z) \triangleq p_0(x) p_{Y,Z|X}(y, z|x)$ 

is achievable for  $\Delta$ -empirical coordination with the rate-pair  $(R_1, R_2)$  if there is an N such that for all n > N, there exists a coordination code  $(2^{nR_1}, 2^{nR_2}, n)$  such that

$$\mathbb{E}\{\|P_{x^{n},y^{n},z^{n}}(x,y,z)-p_{0}(x)p_{YZ|X}(y,z|x)\|_{TV}\}\leq\Delta$$

Definition 9 (Rate-distortion-coordination region): The ratedistortion-coordination region  $R_{p_0}^I$  for the source distribution  $p_0(x)$  and for a fixed conditional distribution  $p_{Y,Z|X}(y,z|x)$  is defined as:

$$\begin{split} R_{p_0}^{I} & \left( p_{Y,Z|X} \left( y, z|x \right) \right) \\ & \triangleq \mathbf{Cl} \left\{ \begin{array}{l} \left( R_1, R_2, \Delta \right) : \\ p_0 \left( x \right) p_{Y,Z|X} \left( y, z|x \right) \text{ is achievable} \\ \text{ for } \Delta \text{-empirical coordination at rates } \left( R_1, R_2 \right) \end{array} \right\}. \end{split}$$

#### III. MAIN RESULT

In this section, we state and prove the main result. For simplicity, we state and prove the result in the case of the cascade set-up of Fig. 2. Nevertheless, the results herein are general and they can be extended to many other set-ups.

Our achievability part states that every good coordination code designed achieving for perfect empirical coordination according to any distribution in  $N_{\Delta}(p_0(x) p_{YZ|X}(y, z|x)))$ achieves  $\Delta$ -empirical coordination according to  $p_0(x) p_{YZ|X}(y, z|x)$  (see Fig. 3.) This means that, if  $\left(R_1, R_2, q_{\hat{Y}, \hat{Z}|X}\right) \in C^P_{p_0}$  for some  $p_{0}\left(x\right)q_{\hat{Y}\hat{Z}|X}\left(y,z|x\right) \hspace{0.1in} \in \hspace{0.1in} N_{\Delta}\left(p_{0}\left(x\right)p_{YZ|X}\left(y,z|x\right)\right), \hspace{0.1in} \text{then,}$  $(R_1, R_2, \Delta) \in R_{p_0}^I (p_{YZ|X}(y, z|x)).$ 

Our main result makes a stronger statement. According to the converse part, there is no more efficient way of satisfying the  $\Delta$  coordination-distortion limit than by using a coordination code that produces actions of the same joint type which belongs to  $N_{\Delta}(p_0(x) p_{Y,Z|X}(y,z|x))$ . Clearly, a good coordination code, designed for imperfect empirical coordination would produce a variety of different joint types, satisfying on the average the distortion limit. However, given such a coordination code, repeated uses will produce a longer coordination code with the same rates that achieves perfect empirical coordination according to the expected joint type. This expected joint type can be shown to belong to  $N_{\Delta}(p_0(x) p_{Y,Z|X}(y,z|x))$  and therefore according to the achievability part this new code achieves the distortion limit. Since the two codes have the same rates, we do not lose in rate-efficiency if we substitute the old by the new one (see Fig. 4). This further means that, if  $(R_1, R_2, \Delta) \in$  $R_{p_0}^{I}(p_{Y,Z|X}(y,z|x))$ , then,  $(R_1, R_2, q_{\hat{Y},\hat{Z}|X}) \in C_{p_0}^{P}$  for some  $p_0(x) q_{\hat{Y},\hat{Z}|X}(y,z|x)$  in  $N_{\Delta}(p_0(x) p_{Y,Z|X}(y,z|x))$ .

As a consequence, when we are interested in designing a good coordination code that achieves  $\Delta$ -empirical coordination according to some distribution, we can restrict our search, without loss of optimality, exactly to the class of coordination codes that achieve perfect empirical coordination according

to any distribution in the  $\Delta$ -neighborhood of this desired distribution.

The previous discussion is formalized in the next theorem.



Fig. 3. Achievability part: Every good coordination code designed for achieving perfect empirical coordination according to some distribution  $p_0(x) q_{\hat{Y},\hat{Z}|X}(y,z|x) \in N_{\Delta}(p_0(x) p_{Y,Z|X}(y,z|x))$  achieves  $\Delta$ empirical coordination according to  $p_0(x) p_{Y,Z|X}(y,z|x)$ .



Fig. 4. Converse part: For every coordination code that achieves  $\Delta$ -empirical coordination according to  $p_0(x) p_{Y,Z|X}(y,z|x)$ , there is a coordination code with the same rates which achieves perfect empirical coordination according to some distribution in  $N_{\Delta}(p_0(x) p_{Y,Z|X}(y,z|x))$  (red).

Theorem 1: For every  $p_{Y,Z|X}(y,z|x)$ ,

$$R_{p_{0}}^{I}\left(p_{Y,Z|X}\left(y,z|x\right)\right) = \begin{cases} \left(R_{1},R_{2},\Delta\right):\\ \left(R_{1},R_{2},q_{\hat{Y},\hat{Z}|X}\right)\in C_{p_{0}}^{P} \\ \text{for some }\left(\hat{Y},\hat{Z}\right) \text{ which satisfy} \\ p_{0}\left(x\right)q_{\hat{Y},\hat{Z}|X}\left(y,z|x\right)\in N_{\Delta}\left(p_{0}\left(x\right)p_{Y,Z|X}\left(y,z|x\right)\right) \end{cases} \end{cases}$$

*Proof:* Achievability. When a rate-coordination tuple  $(R_1, R_2, q_{\hat{Y}, \hat{Z}|X})$  is in the interior of  $C_{p_0}^P$  for some  $q_{\hat{Y}, \hat{Z}|X}(y, z|x)$  such that  $p_0(x) q_{\hat{Y}, \hat{Z}|X}(y, z|x) \in N_{\Delta}(p_0(x) p_{Y,Z|X}(y, z|x))$ , we are assured (see Definition 6) the existence of a coordination code such that for every  $\epsilon > 0$  and for all large enough n we have

$$\Pr\left\{\|P_{x^{n},y^{n},z^{n}}(x,y,z)-p_{0}(x)q_{\hat{Y},\hat{Z}|X}(y,z|x)\|_{TV} > \epsilon\right\} < \epsilon$$

Applying the triangle inequality on the total variation and because  $p_0(x) q_{\hat{Y},\hat{Z}|X}(y,z|x) \in N_{\Delta}(p_0(x) p_{Y,Z|X}(y,z|x))$ , we obtain

$$\begin{aligned} \|P_{x^{n},y^{n},z^{n}}(x,y,z) - p_{0}(x) p_{Y,Z|X}(y,z|x) \|_{TV} \\ &\leq \|P_{x^{n},y^{n},z^{n}}(x,y,z) - p_{0}(x) q_{\hat{Y},\hat{Z}|X}(y,z|x) \|_{TV} \\ &+ \|p_{0}(x) q_{\hat{Y},\hat{Z}|X}(y,z|x) - p_{0}(x) p_{Y,Z|X}(y,z|x) \|_{TV} \\ &\leq \|P_{x^{n},y^{n},z^{n}}(x,y,z) - p_{0}(x) q_{\hat{Y},\hat{Z}|X}(y,z|x) \|_{TV} + \Delta \end{aligned}$$

Thus, for every  $\epsilon > 0$  and for all large enough n, this coordination code achieves

$$\Pr\left\{\|P_{x^{n},y^{n},z^{n}}\left(x,y,z\right)-p_{0}\left(x\right)p_{Y,Z|X}\left(y,z|x\right)\|_{TV}\right.$$
  
$$>\Delta+\epsilon\right\}<\epsilon,$$

which gives

By choosing  $\epsilon$  arbitrarily small and n large enough, we conclude that this coordination code achieves also  $\Delta$ -empirical coordination according to  $p_{Y,Z|X}(y,z|x)$  and so  $(R_1, R_2, \Delta) \in R_{p_0}^I(p_{Y,Z|X}(y,z|x))$ .

**Converse.** Suppose that  $(R_1, R_2, \Delta)$  is in the interior of  $R_{p_0}^I(p_{Y,Z|X}(y, z|x))$ , i.e., there exists a coordination code with blocklength n large enough which achieves  $\Delta$ -empirical coordination according to  $p_{Y,Z|X}(y, z|x)$ , at rates  $(R_1, R_2)$  such that

$$\mathbb{E}\{\|P_{x^{n},y^{n},z^{n}}(x,y,z)-p_{0}(x)p_{Y,Z|X}(y,z|x)\|_{TV}\}\leq\Delta.$$

This coordination code induces a distribution  $\tilde{p}(x^n, y^n, z^n) = p_{X^n}(x^n) q_{\hat{Y}^n, \hat{Z}^n|X^n}(y^n, z^n|x^n)$  where  $p_{X^n}(x^n) = \prod_{i=1}^n p_0(x_i)$ . We repeat the use of the scheme over k blocks of length n each and, as a result, we induce a joint distribution on  $\left(X^{kn}, \hat{Y}^{kn}, \hat{Z}^{kn}\right)$  that consists of blocks  $\left(X_1^n, \hat{Y}_1^n, \hat{Z}_1^n\right), \dots, \left(X_{kn-n+1}^{kn}, \hat{Y}_{kn-n+1}^{kn}, \hat{Z}_{kn-n+1}^{kn}\right)$  denoted as  $\left(X^{(1)n}, \hat{Y}^{(1)n}, \hat{Z}^{(1)n}\right), \dots, \left(X^{(k)n}, \hat{Y}^{(k)n}, \hat{Z}^{(k)n}\right)$ . The new coding scheme has rates  $R'_i = \frac{\log\left((2^{nR_i})^k\right)}{kn} = R_i$ , for i = 1, 2. By the law of large numbers, we get

$$P_{x^{kn},y^{kn},z^{kn}} = \frac{1}{k} \sum_{i=1}^{k} P_{x^{(i)n},y^{(i)n},z^{(i)n}} \to \mathbb{E}\left\{P_{x^{n},y^{n},z^{n}}\right\},$$

in probability. Point-wise convergence in probability further implies that as k grows

$$||P_{x^{kn},y^{kn},z^{kn}}(x,y,z) - \mathbb{E}\{P_{x^n,y^n,z^n}\}||_{TV} \to 0,$$

in probability. However,

$$\mathbb{E}\left\{P_{x^n,y^n,z^n}\right\} = p_{\overline{X},\overline{\hat{Y}},\overline{\hat{Z}}}(x,y,z),$$

and thus

$$\|P_{x^{kn},y^{kn},z^{kn}}\left(x,y,z\right)-p_{\overline{X},\overline{\hat{Y}},\overline{\hat{Z}}}(x,y,z)\|_{TV}\to 0,$$

in probability, where

$$p_{\overline{X},\overline{\hat{Y}},\overline{\hat{Z}}}(x,y,z) \triangleq \sum_{k=1}^{n} \frac{p_{0k}(x) q_{\hat{Y}_{k},\hat{Z}_{k}|X_{k}}(y,z|x)}{n}$$
$$= \sum_{k=1}^{n} \frac{p_{0}(x) q_{\hat{Y}_{k}\hat{Z}_{k}|X_{k}}(y,z|x)}{n}$$
$$= p_{0}(x) \underbrace{\sum_{k=1}^{n} \frac{q_{\hat{Y}_{k},\hat{Z}_{k}|X_{k}}(y,z|x)}{n}}_{\stackrel{\triangleq}{=} p_{\overline{Y},\overline{\hat{Z}}|\overline{X}}},$$

(see Lemma 5 in the Appendix). Moreover,

$$\begin{split} \Delta &= \mathbb{E} \big\{ \| P_{x^{n},y^{n},z^{n}}\left(x,y,z\right) - p_{0}\left(x\right)p_{Y,Z|X}\left(y,z|x\right)\|_{TV} \big\} \\ &\stackrel{(a)}{\geq} \| \mathbb{E} \left\{ P_{x^{n},y^{n},z^{n}} \right\} - p_{0}\left(x\right)p_{Y,Z|X}\left(y,z|x\right)\|_{TV} \\ &= \| p_{\overline{X}\overline{Y}\overline{Z}}\left(x,y,z\right) - p_{0}\left(x\right)p_{Y,Z|X}\left(y,z|x\right)\|_{TV} \\ &= \| p_{0}\left(x\right)p_{\overline{Y},\overline{Z}|\overline{X}}\left(y,z|x\right) - p_{0}\left(x\right)p_{Y,Z|X}\left(y,z|x\right)\|_{TV}, \end{split}$$

where (a) follows from Jensen's inequality since total variation is convex, i.e., for every  $\lambda \in [0, 1]$  we have

$$\lambda \| p_1(x,y) - q(x,y) \|_{TV} + (1-\lambda) \| p_2(x,y) - q(x,y) \|_{TV} \\\geq \| (\lambda p_1(x,y) + (1-\lambda) p_2(x,y)) - q(x,y) \|_{TV}.$$

Thus, we have constructed a sequence of coordination codes with rates  $(R_1, R_2)$  that achieves perfect empirical coordination according to  $p_{\overline{X},\overline{\hat{Y}},\overline{\hat{Z}}} = p_0(x) p_{\overline{\hat{Y}},\overline{\hat{Z}}|\overline{X}}(y,z|x)$ , i.e.,  $(R_1, R_2, p_{\overline{\hat{Y}},\overline{\hat{Z}}|\overline{X}}) \in C_{p_0}^P$  and additionally  $p_0(x) p_{\overline{\hat{Y}},\overline{\hat{Z}}|\overline{X}}(y,z|x) \in N_{\Delta}(p_0(x) p_{Y,Z|X}(y,z|x))$ . This completes the proof.

#### **IV. EXAMPLES**

In this section, we apply Theorem 1 to get the ratedistortion-coordination region in the two node-network illustrated in Fig. 5 and in the cascade network illustrated in Fig. 6.



Fig. 5. Two node-network.

Lemma 1: Consider the setup of Fig. 5. Then,

$$C_{p_0}^P = \{ (R, p_{Y|X}(y|x)) : R \ge I(X;Y) \}.$$



Fig. 6. Cascade network.

*Proof:* See [1, Theorem 3]. *Lemma 2:* Consider the setup of Fig. 5. Then,

 $R_{p_{0}}^{I}(p_{Y|X}(y|x)) = \begin{cases} (R, \Delta) : \\ R \ge \min_{\substack{q_{\hat{Y}|X}(y|x): \\ p_{0}(x)q_{\hat{Y}|X}(y|x) \in N_{\Delta}(p_{0}(x)p_{Y|X}(y|x))}} I(X; \hat{Y}) \\ \end{cases},$ 

for every  $p_{Y|X}(y|x)$ .

Proof: From Theorem 1, we obtain

$$\begin{split} R_{p_0}^{I}\left(p_{Y|X}\left(y|x\right)\right) \\ &= \begin{cases} \left(R,\Delta\right):\\ \left(R,q_{\hat{Y}|X}\right) \in C_{p_0}^{P} \\ \text{for some } \hat{Y} \text{ which satisfy} \\ p_0\left(x\right)q_{\hat{Y}|X}\left(y|x\right) \in N_{\Delta}\left(p_0\left(x\right)p_{Y|X}\left(y|x\right)\right) \end{cases} \end{split}$$

Using Lemma 1, the latter can be reformulated as

$$\begin{aligned} R_{p_0}^{I}\left(p_{Y|X}\left(y|x\right)\right) \\ &= \begin{cases} (R,\Delta):\\ R \ge I\left(X;\hat{Y}\right) \\ \text{for some } \hat{Y} \text{ which satisfy} \\ p_0\left(x\right)q_{\hat{Y}|X}\left(y|x\right) \in N_{\Delta}\left(p_0\left(x\right)p_{Y|X}\left(y|x\right)\right) \end{cases} \end{cases} , \end{aligned}$$

and by taking into account the non-emptiness and the closeness of  $N_{\Delta}(p_0(x) p_{Y|X}(y|x))$ , we obtain the characterization of the lemma.

Lemma 3: Consider the set-up of Fig. 6. Then,

$$C_{p_0}^{P} = \left\{ \begin{array}{l} \left( R_1, R_2, p_{Y,Z|X} \left( y, z|x \right) \right) : \\ R_1 \ge I \left( X; Y, Z \right), R_2 \ge I \left( X; Z \right) \end{array} \right\}.$$

*Proof:* See [1, Theorem 5]. *Lemma 4:* Consider the set-up of Fig. 6. Then,

$$\begin{split} & = \begin{cases} R_{p_0}^{I} \left( p_{Y,Z|X} \left( y, z|x \right) \right) \\ & = \begin{cases} \left( R_1, R_2, \Delta \right) : \\ R_1 \ge I \left( X; \hat{Y}, \hat{Z} \right), R_2 \ge I \left( X; \hat{Z} \right) \\ \text{for some } \left( \hat{Y}, \hat{Z} \right) \text{ which satisfy} \\ p_0 \left( x \right) q_{\hat{Y}, \hat{Z}|X} \left( y, z|x \right) \in N_\Delta \big( p_0 \left( x \right) p_{Y,Z|X} \left( y, z|x \right) \big) \end{cases} \end{split}$$

for every  $p_{Y,Z|X}(y, z|x)$ . *Proof:* From Theorem 1, we obtain

$$R_{p_{0}}^{I}\left(p_{Y,Z|X}\left(y,z|x\right)\right) = \begin{cases} \left(R_{1}, R_{2}, \Delta\right): \\ \left(R_{1}, R_{2}, q_{\hat{Y}, \hat{Z}|X}\right) \in C_{p_{0}}^{P} \\ \text{for some } \left(\hat{Y}, \hat{Z}\right) \text{ which satisfy} \\ p_{0}\left(x\right) q_{\hat{Y}, \hat{Z}|X}\left(y, z|x\right) \in N_{\Delta}\left(p_{0}\left(x\right) p_{Y,Z|X}\left(y, z|x\right)\right) \end{cases} \end{cases}$$

and by applying Lemma 3, we obtain the characterization of the lemma.

## V. APPENDIX

*Lemma 5:* For a given probability distribution  $p_{X^n}(x^n)$ ,

$$\mathbb{E}\left\{P_{x^{n}}\left(x\right)\right\} = \frac{1}{n}\sum_{k=1}^{n}\left(p_{X_{k}}\left(x\right)\right).$$

Proof:

$$\mathbb{E}\{P_{x^{n}}(x)\} = \sum_{x^{n}} p_{X^{n}}(x^{n}) \frac{n_{x^{n}}(x)}{n}$$
$$= \sum_{x^{n}} \left( p_{X^{n}}(x^{n}) \frac{1}{n} \sum_{k=1}^{n} \mathbf{1} (x_{k} = x) \right)$$
$$= \sum_{k=1}^{n} \sum_{x^{n}} \left( p_{X^{n}}(x^{n}) \frac{1}{n} \mathbf{1} (x_{k} = x) \right)$$
$$= \frac{1}{n} \sum_{k=1}^{n} \left( p_{X_{k}}(x) \right).$$

#### REFERENCES

- 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.
- [2] A. Bereyhi, M. Bahrami, M. Mirmohseni, and M. R. Aref, "Empirical coordination in a triangular multiterminal network," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2013, pp. 2149–2153.
- [3] M. Le Treust, "Empirical coordination with two-sided state information and correlated source and state," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2015, pp. 466–470.
- [4] —, "Empirical coordination with channel feedback and strictly causal or causal encoding," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2015, pp. 471–475.
- [5] M. Le Treust and T. Tomala, "Information design for strategic coordination of autonomous devices with non-aligned utilities," in *Proc. Annual Allerton Conf. on Comm., Control, and Comput.*, Sep. 2016, pp. 233– 242.
- [6] 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.
- [7] G. Cervia, L. Luzzi, M. Le Treust, and M. R. Bloch, "Strong coordination of signals and actions over noisy channels," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2017, pp. 2835–2839.
- [8] B. Larrousse, S. Lasaulce, and M. R. Bloch, "Coordination in distributed networks via coded actions with application to power control," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3633–3654, May 2018.
- [9] A. Wyner, "The common information of two dependent random variables," *IEEE Trans. Inf. Theory*, vol. 21, no. 2, pp. 163–179, March 1975.
- [10] G. Kramer and S. A. Savari, "Communicating probability distributions," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 518–525, Feb 2007.