Strong Coordination of Signals and Actions over Noisy Channels

Giulia Cervia, Laura Luzzi, Mael Le Treust, Matthieu R Bloch

▶ **To cite this version:**

Giulia Cervia, Laura Luzzi, Mael Le Treust, Matthieu R Bloch. Strong Coordination of Signals and Actions over Noisy Channels. IEEE International Symposium on Information Theory, Jun 2017, Aachen, Germany. hal-01522450

HAL Id: hal-01522450
https://hal.science/hal-01522450

Submitted on 16 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Strong Coordination of Signals and Actions over Noisy Channels

Giulia Cervia*, Laura Luzzi*, Maël Le Treust* and Matthieu R. Bloch‡
* ETIS UMR 8051, Université Paris Seine, Université Cergy-Pontoise, ENSEA, CNRS, Cergy, France.
Email: {giulia.cervia, laura.luzzi, mael.le-treust}@ensea.fr
‡School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia
Email: matthieu.bloch@ece.gatech.edu

***Abstract*—We develop a random binning scheme for strong coordination in a network of two nodes separated by a noisy channel, in which the input and output signals have to be coordinated with the source and its reconstruction. In the case of non-causal encoding and decoding, we propose a joint source-channel coding scheme and develop inner and outer bounds for the strong coordination region. While the set of achievable target distributions is the same as for empirical coordination, we characterize the rate of common randomness required for strong coordination.**

I. Introduction

The 5G standard envisions direct device-to-device communication, which is likely to be a key enabler of the Internet of Things. In this decentralized network of connected objects, such as wireless sensors, medical and wearable devices, smart energy meters, home appliances, and self-driving cars, devices will communicate with each other while sensing or acting on their environment. It is essential that these devices, considered as autonomous decision-makers, cooperate and coordinate their actions.

From an information theory perspective, two different metrics have been proposed to measure the level of coordination: *empirical coordination*, which requires the joint histogram of the actions to approach a target distribution, and *strong coordination*, which requires the total variation distance of the distribution of sequences of actions to converge to an i.i.d. target distribution [1]. While empirical coordination investigates the average behavior over time, strong coordination is to be preferred from a security standpoint, since it guarantees that the sequence of actions will be unpredictable to an outside observer. This is a consequence of the fact that statistical tests will produce identically distributed outcomes for distributions that are close in total variation.

Strong coordination with error free links has been studied in [1] and the case in which only the source and the reconstruction have to be coordinated has been considered in [2]. However, in a realistic scenario where the communication links are noisy, the signals that are transmitted and received over the physical channel become a part of what can be observed. One may therefore wish to coordinate both behaviors and communication [3]. In this setting, strong coordination is desirable since the synthesized sequences would appear to be i.i.d. even from the perspective of a malicious eavesdropper who can observe the signals sent over the communication channel [4].

In this paper, we address this problem in a two-node network comprised of an information source and a noisy channel, in which both nodes have access to a common source of randomness. An inner bound for the empirical coordination region has already been established in [3] and we focus here on the problem of achieving strong coordination for the same setting. This scenario presents two conflicting goals: the encoder needs to convey a message to the decoder to coordinate the reconstructed version of the source, while simultaneously coordinating the signals coding the message. We derive an inner and an outer bound for the strong coordination region by developing a joint source-channel scheme in which an auxiliary codebook allows us to satisfy both goals. Since the two bounds do not match, the optimality of our scheme remains an open question. While the set of achievable target distributions is the same as for empirical coordination, we show that a positive rate of common randomness is required for strong coordination.

The remainder of the paper is organized as follows. Section II introduces the notation, Section III describes the model under investigation and states the main result. Section IV proves an inner bound by proposing a random binning scheme and a random coding scheme that have the same statistics. Finally, Section V proves an outer bound.

II. Preliminaries

We define the integer interval $[a, b]$ as the set of integers between a and b. Given a random vector $X^n := (X_1, \dots, X_n)$, we note X^i the first i components of X^n. We note $\mathbb{V}(\cdot, \cdot)$ the variational distance between two distributions.

We now recall some useful results that we use later.

Lemma 1 (Source coding with side information at the decoder): Consider an encoder that observes a sequence X^n and transmits a message $M \in [1, 2^{nR}]$ to a decoder that has access to side information Y^n, where (X^n, Y^n) is a discrete memoryless source. If the encoding rate $R > H(X|Y)$, the decoder can recover X^n from M and Y^n with arbitrarily small error probability.

The work of M.R. Bloch was supported in part by NSF under grant CIF 1320304. The work of M. Le Treust was supported by INS2I CNRS through projects JCJC CoReDe 2015 and PEPS StrategicCoo 2016. This work was conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01).

Lemma 1 is a consequence of the Slepian-Wolf Theorem [5, Theorem 10.1].

Lemma 2: Given a discrete memoryless source (A^n, B^n) and $K = \varphi(B^n)$ a binning of B^n with 2^{nR} values chosen independently and uniformly at random, if $R < H(B|A)$, then we have

$$\lim_{n\to\infty} \mathbb{E}_\varphi \left[\mathbb{V}\left(P^{\varphi}_{A^n K}, Q_K P_{A^n}\right)\right] = 0,$$

where $\mathbb{E}_\varphi$ denotes the average over the random binnings, P^φ is the distribution corrisponding to a fixed realization of the binning and Q_K is the uniform distribution in $[1, 2^{nR}]$.

Lemma 2 is a consequence of [6, Lemma 3.1] and [7, Theorem 1].

Remark 1: We have,

$$\mathbb{V}(P_A, \widehat{P}_A) \le \mathbb{V}(P_{AB}, \widehat{P}_{AB}), \tag{1}$$

$$\mathbb{V}(P_A, \widehat{P}_A) = \mathbb{V}(P_A P_{B|A}, \widehat{P}_A P_{B|A}), \tag{2}$$

where (1) and (2) have been proven in [8, Lemma 16] and [8, Lemma 17] respectively.

III. System model and main result

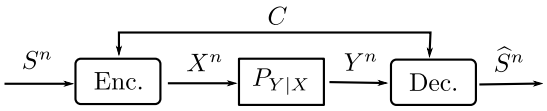

Figure 1. Coordination of signals and actions for a two-node network with a noisy channel.

Consider the model depicted in Figure 1 in which two agents, the encoder and the decoder, have access to a shared source of uniform randomness $C \in [1, 2^{nR_0}]$. The encoder observes an i.i.d. source $S^n \in \mathcal{S}^n$ with distribution $\bar{P}_S$. The encoder then selects a signal $X^n = f_n(S^n, C)$, $f_n : \mathcal{S}^n \times [1, 2^{nR_0}] \to \mathcal{X}^n$. The signal X^n is transmitted over a discrete memoryless channel parametrized by the conditional distribution $\bar{P}_{Y|X}$. Upon observing Y^n and C, the stochastic decoder selects an action $\widehat{S}^n = g_n(Y^n, C)$, $g_n : \mathcal{Y}^n \times [1, 2^{nR_0}] \to \widehat{\mathcal{S}}^n$. For block length n, the pair (f_n, g_n) constitutes a code. We recall the notions of achievability and the strong coordination region [8].

Definition 1: A pair $(\bar{P}_{SXY\hat{S}}, R_0)$ is *achievable* if there exists a sequence (f_n, g_n) of encoders-decoders with rate of common randomness R_0, such that the induced joint distribution $P_{S^n X^n Y^n \widehat{S}^n}$ is nearly indistinguishable from the i.i.d. distribution $\bar{P}_{SXY\hat{S}}$, in total variational distance:

$$\lim_{n\to\infty} \mathbb{V}\left(P_{S^n X^n Y^n \widehat{S}^n}, \bar{P}^{\otimes n}_{SXY\hat{S}}\right) = 0.$$

The *strong coordination region* $\mathcal{R}$ is the set of achievable pairs $(\bar{P}_{SXY\hat{S}}, R_0)$.

In the case of non-causal encoder and decoder, the problem of characterizing the strong coordination region is still open, but we establish the following inner and outer bounds.

Theorem 1: Let $\bar{P}_S$ and $\bar{P}_{Y|X}$ be the given source and channel parameters, then $\mathcal{R}_1 \subseteq \mathcal{R} \subseteq \mathcal{R}_2$ where:

$$\mathcal{R}_1 := \left\{ \begin{array}{c} (\bar{P}_{SXY\widehat{S}}, R_0) \ : \\ \bar{P}_{SXY\widehat{S}} = \bar{P}_S \bar{P}_{X|S} \bar{P}_{Y|X} \bar{P}_{\widehat{S}|SXY} \\ \exists\ U \text{ taking values in } \mathcal{U} \\ \bar{P}_{SXYU\widehat{S}} = \bar{P}_S \bar{P}_{U|S} \bar{P}_{X|US} \bar{P}_{Y|X} \bar{P}_{\widehat{S}|UY} \\ I(U;S) < I(U;Y) \\ R_0 > I(U; SX\hat{S}|Y) \\ |\mathcal{U}| \le |\mathcal{S}||\mathcal{X}||\mathcal{Y}||\widehat{\mathcal{S}}| + 1 \end{array} \right\} \tag{3}$$

$$\mathcal{R}_2 := \left\{ \begin{array}{c} (\bar{P}_{SXY\widehat{S}}, R_0) \ : \\ \bar{P}_{SXY\widehat{S}} = \bar{P}_S \bar{P}_{X|S} \bar{P}_{Y|X} \bar{P}_{\widehat{S}|SXY} \\ \exists\ U \text{ taking values in } \mathcal{U} \\ \bar{P}_{SXYU\widehat{S}} = \bar{P}_S \bar{P}_{U|S} \bar{P}_{X|US} \bar{P}_{Y|X} \bar{P}_{\widehat{S}|UY} \\ I(U;S) \le I(X;Y) \\ R_0 \ge I(U; SX\hat{S}|Y) \\ |\mathcal{U}| \le |\mathcal{S}||\mathcal{X}||\mathcal{Y}||\widehat{\mathcal{S}}| + 1 \end{array} \right\}. \tag{4}$$

Remark 2: Even for empirical coordination, the problem of characterizing the coordination region is still open [3]. The information constraint $I(U;S) \le I(U;Y)$ for empirical coordination [3, Theorem 1] is very similar to ours, as well as the decomposition of the joint probability distribution $\bar{P}_S \bar{P}_{U|S} \bar{P}_{X|US} \bar{P}_{Y|X} \bar{P}_{\widehat{S}|UY}$. The main difference is that strong coordination requires a positive rate of common randomness $R_0 \ge I(U; SX\hat{S}|Y)$.

Remark 3: Our inner bound is a generalization of the one in [2] and the proof follows the same strategy inspired by [7].

IV. Proof of Theorem 1: inner bound

First, we define two random schemes each of which induces a joint distribution.

A. Random binning scheme

Assume that the sequences S^n , X^n, U^n, Y^n and $\widehat{S}^n$ are jointly i.i.d. with distribution $\bar{P}_{S^n} \bar{P}_{U^n|S^n} \bar{P}_{X^n|U^n S^n} \bar{P}_{Y^n|X^n} \bar{P}_{\widehat{S}^n|U^n Y^n}$. We consider two uniform random binnings for U^n:

- first binning $C = \varphi_1(U^n)$, where $\varphi_1 : \mathcal{U}^n \to [1, 2^{nR_0}]$ maps each sequence of $\mathcal{U}^n$ uniformly and independently to the set $[1, 2^{nR_0}]$;
- second binning $F = \varphi_2(U^n)$, $\varphi_2 : \mathcal{U}^n \to [1, 2^{n\tilde{R}}]$.

Note that if $\tilde{R} + R_0 > H(U|Y)$, by Lemma 1, it is possible to recover U^n from Y^n and (C, F) with high probability using a Slepian-Wolf decoder via the conditional distribution

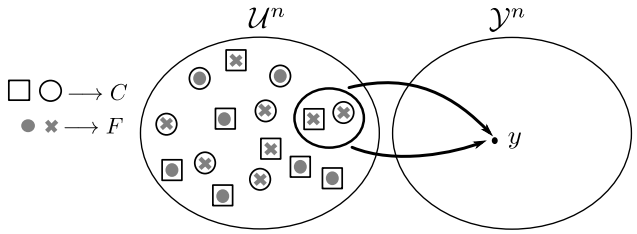

Figure 2. The square and the circle represent the outputs of the first binning C and the dot and the cross the outputs of the second binning F. Given $\mathbf{y}$ and the realizations of C and F, it is possible to recover $\mathbf{u}$.

$P^{SW}_{\widehat{U}^n|CFY^n}$ as depicted in Figure 2. This defines a joint distribution:

$$\bar{P}_{S^nU^n\widehat{U}^nX^nY^nCF\widehat{S}^n} = \bar{P}_{S^n}\bar{P}_{U^n|S^n}\bar{P}_{X^n|U^nS^n}\bar{P}_{C|U^n}\bar{P}_{F|U^n}\bar{P}_{Y^n|X^n}\bar{P}_{\widehat{S}^n|U^nY^n}P^{SW}_{\widehat{U}^n|CFY^n}.$$

In particular, $\bar{P}_{U^n|CFS^n}$ is well defined.

B. Random coding scheme

In this section we follow the approach in [7, Section IV.E] and [2]. Suppose that the encoder and decoder have access not only to common randomness C but also to extra randomness F, where C is generated uniformly at random in $[1, 2^{nR_0}]$ with distribution Q_C and F is generated uniformly at random in $[1, 2^{n\tilde{R}}]$ with distribution Q_F independently of C. Then the encoder generates U^n according to $\bar{P}_{U^n|CFS^n}$ defined in Section IV-A and X^n according to $\bar{P}_{X^n|S^nU^n}$. The encoder sends X^n through the channel. The decoder gets Y^n and (C, F) and reconstructs U^n via the conditional distribution $P^{SW}_{\widehat{U}^n|CFY^n}$. The decoder then generates $\widehat{S}^n$ letter by letter according to the distribution $P_{\widehat{S}^n|\widehat{U}^nY^n}$ (more precisely $\bar{P}_{\widehat{S}^n|U^nY^n}(\widehat{\mathbf{s}}|\widehat{\mathbf{u}}, \mathbf{y})$, where $\widehat{\mathbf{u}}$ is the output of the Slepian-Wolf decoder). This defines a joint distribution:

$$P_{S^nU^n\widehat{U}^nX^nY^nCF\widehat{S}^n} = Q_CQ_FP_{S^n}\bar{P}_{U^n|CFS^n}\bar{P}_{X^n|U^nS^n}\bar{P}_{Y^n|X^n}P^{SW}_{\widehat{U}^n|CFY^n}P_{\widehat{S}^n|\widehat{U}^nY^n}.$$

We want to show that the distribution $\bar{P}$ is achievable for strong coordination:

$$\lim_{n\to\infty} \mathbb{V}\left(\bar{P}_{S^nX^nU^n\widehat{U}^nY^n\widehat{S}^n}, P_{S^nX^nU^n\widehat{U}^nY^n\widehat{S}^n}\right) = 0. \quad (5)$$

We prove that the random coding scheme possesses all the properties of the initial source coding scheme stated in Section IV-A. Note that

$$\begin{aligned}
&\mathbb{V}(\bar{P}_{S^nU^n\widehat{U}^nX^nY^nCF}, P_{S^nU^n\widehat{U}^nX^nY^nCF}) \quad (6)\\
&= \mathbb{V}(\bar{P}_{S^n}\bar{P}_{U^n|S^n}\bar{P}_{X^n|U^nS^n}\bar{P}_{C|U^n}\bar{P}_{F|U^n}\bar{P}_{Y^n|X^n}P^{SW}_{\widehat{U}^n|CFY^n},\\
&\qquad Q_CQ_FP_{S^n}\bar{P}_{U^n|CFS^n}\bar{P}_{X^n|U^nS^n}\bar{P}_{Y^n|X^n}P^{SW}_{\widehat{U}^n|CFY^n})\\
&\overset{(a)}{=}\mathbb{V}(\bar{P}_{S^n}\bar{P}_{U^n|S^n}\bar{P}_{C|U^n}\bar{P}_{F|U^n}, Q_CQ_FP_{S^n}\bar{P}_{U^n|CFS^n})\\
&\overset{(b)}{=}\mathbb{V}(\bar{P}_{S^nCF}, P_{S^n}Q_CQ_F)
\end{aligned}$$

where (a) and (b) come from (2). Then if $R_0 + \widetilde{R} < H(U|S)$, we can apply Lemma 2 where $B^n = U^n$, $K = (C, F)$, $\varphi = (\varphi_1, \varphi_2)$, $A^n = S^n$ and find that

$$\lim_{n\to\infty} \mathbb{E}_\varphi\left[\mathbb{V}\left(\bar{P}^{\varphi}_{S^nCF}, Q_CQ_F\bar{P}_{S^n}\right)\right] = 0.$$

Therefore there exists a fixed binning φ' such that, if we denote with $\bar{P}^{\varphi'}$ and $P^{\varphi'}$ the distributions $\bar{P}$ and P with respect to the choice of a binning φ', we have

$$\lim_{n\to\infty} \mathbb{V}\left(\bar{P}^{\varphi'}_{S^nCF}, P_{S^n}Q_CQ_F\right) = 0$$

which by (6) implies

$$\lim_{n\to\infty} \mathbb{V}(\bar{P}^{\varphi'}_{S^nU^n\widehat{U}^nX^nY^nCF}, P^{\varphi'}_{S^nU^n\widehat{U}^nX^nY^nCF}) = 0. \quad (7)$$

From now on, we will omit φ' to simplify the notation.

Now we would like to show that we have strong coordination for $\widehat{S}^n$ as well, but in the second scheme $\widehat{S}^n$ is generated using $\widehat{U}^n$ and not U^n as in the first scheme. Because of Lemma 1, the inequality $\widetilde{R} + R_0 > H(U|Y)$ implies that $\widehat{U}^n$ is equal to U^n with high probability and we will use this fact to show that the distributions are close in total variational distance. First, we need to establish a technical lemma, whose proof can be found in the Appendix.

Lemma 3: Let V^n and $\widehat{V}^n$ such that $\mathbb{P}\{\widehat{V}^n \neq V^n\} \to 0$ when $n \to \infty$. Then for any random variable W^n and for any joint distribution $P_{W^nV^n\widehat{V}^n}$ we have:

$$\lim_{n\to\infty} \mathbb{V}(P_{W^nV^n\widehat{V}^n}, P_{W^nV^n}\mathbb{1}_{\widehat{V}^n|V^n}) = 0$$

$$\text{where } \mathbb{1}_{\widehat{V}^n|V^n}(\mathbf{v}|\mathbf{v}') = \begin{cases} 1 & \text{if } \mathbf{v} = \mathbf{v}' \\ 0 & \text{if } \mathbf{v} \neq \mathbf{v}' \end{cases}.$$

Since $\widehat{U}^n$ is equal to U^n with high probability, we can apply Lemma 3 and if we denote $Z^n := S^nX^nCF$ we find:

$$\lim_{n\to\infty} \mathbb{V}(\bar{P}_{Z^nY^nU^n\widehat{U}^n}, \bar{P}_{Z^nY^nU^n}\mathbb{1}_{\widehat{U}^n|U^n}) = 0, \quad (8)$$

$$\lim_{n\to\infty} \mathbb{V}(P_{Z^nY^nU^n\widehat{U}^n}, P_{Z^nY^nU^n}\mathbb{1}_{\widehat{U}^n|U^n}) = 0. \quad (9)$$

Then using the triangle inequality, we find that

$$\begin{aligned}
&\mathbb{V}(\bar{P}_{Z^nY^nU^n\widehat{U}^n\widehat{S}^n}, P_{Z^nY^nU^n\widehat{U}^n\widehat{S}^n})\\
&= \mathbb{V}(\bar{P}_{Z^nY^nU^n\widehat{U}^n}\bar{P}_{\widehat{S}^n|U^nY^n}, P_{Z^nY^nU^n\widehat{U}^n}P_{\widehat{S}^n|\widehat{U}^nY^n}) \quad (10)\\
&\leq \mathbb{V}(\bar{P}_{Z^nY^nU^n\widehat{U}^n}\bar{P}_{\widehat{S}^n|U^nY^n}, \bar{P}_{Z^nY^nU^n}\mathbb{1}_{\widehat{U}^n|U^n}\bar{P}_{\widehat{S}^n|U^nY^n})\\
&+\mathbb{V}(\bar{P}_{Z^nY^nU^n}\mathbb{1}_{\widehat{U}^n|U^n}\bar{P}_{\widehat{S}^n|U^nY^n}, P_{Z^nY^nU^n}\mathbb{1}_{\widehat{U}^n|U^n}P_{\widehat{S}^n|\widehat{U}^nY^n})\\
&+\mathbb{V}(P_{Z^nY^nU^n}\mathbb{1}_{\widehat{U}^n|U^n}P_{\widehat{S}^n|\widehat{U}^nY^n}, P_{Z^nY^nU^n\widehat{U}^n}P_{\widehat{S}^n|\widehat{U}^nY^n}).
\end{aligned}$$

The first and the third term go to zero by applying (2) to (8) and (9) respectively. Now observe that $\mathbb{1}_{\widehat{U}^n|U^n}\bar{P}_{\widehat{S}^n|U^nY^n} = \mathbb{1}_{\widehat{U}^n|U^n}P_{\widehat{S}^n|\widehat{U}^nY^n}$ by definition of $P_{\widehat{S}^n|\widehat{U}^nY^n}$. Then by using (2) again the second term is equal to $\mathbb{V}\left(\bar{P}_{Z^nY^nU^n}, P_{Z^nY^nU^n}\right)$ that goes to zero by (7) and (1). Hence we have

$$\lim_{n\to\infty} \mathbb{V}(\bar{P}_{Z^nU^n\widehat{U}^nY^n\widehat{S}^n}, P_{Z^nU^n\widehat{U}^nY^n\widehat{S}^n}) = 0. \quad (11)$$

Then by using (1) we have proved (5).

C. Remove the extra randomness F

Even though the extra common randomness F is required to coordinate $(S^n, X^n, Y^n, \widehat{S}^n, U^n)$, we will show that we do not need it in order to coordinate only $(S^n, X^n, Y^n, \widehat{S}^n)$. Observe that by applying (1), equation (11) implies that

$$\lim_{n\to\infty} \mathbb{V}(\bar{P}_{S^nX^nY^n\widehat{S}^nF}, P_{S^nX^nY^n\widehat{S}^nF}) = 0. \quad (12)$$

As in [7], we would like to reduce the amount of common randomness by having the two nodes to agree on an instance $F = f$. To do so, we apply Lemma 2 again where $B^n = U^n$, $K = F$, $\varphi = \varphi''_2$ and $A^n = S^nX^nY^n\widehat{S}^n$. If $\tilde{R} < H(U|SXY\widehat{S})$, there exists a fixed binning such that

$$\lim_{n\to\infty} \mathbb{V}\left(\bar{P}_{S^nX^nY^n\widehat{S}^nF}, Q_F\bar{P}_{S^nX^nY^n\widehat{S}^n}\right) = 0. \quad (13)$$

Remark 4: Note that in Section IV-B we had already chosen a specific binning φ'_2. In the Appendix we prove that there exists a binning which works for both conditions.

Because of (12), (13) implies

$$\lim_{n\to\infty} \mathbb{V}\left(P_{S^nX^nY^n\widehat{S}^nF}, Q_F\bar{P}_{S^nX^nY^n\widehat{S}^n}\right) = 0. \tag{14}$$

Hence, we can fix $f \in F$ such that $(S^n, X^n, Y^n, \hat{S}^n)$ is almost independent of F according to P. To conclude, we need the following result proved in [7, Lemma 4].

Lemma 4: If $\lim_{n\to\infty} \mathbb{V}\left(P_{Y^n}P_{X^n|Y^n}, P'_{Y^n}P'_{X^n|Y^n}\right) = 0$ then there exists $\mathbf{y} \in Y^n$ such that

$$\lim_{n\to\infty} \mathbb{V}\left(P_{X^n|Y^n=\mathbf{y}}, P'_{X^n|Y^n=\mathbf{y}}\right) = 0.$$

If $f \in F$ is fixed, the distribution $P_{S^nX^nY^n\widehat{S}^n}$ changes to $P_{S^nX^nY^n\widehat{S}^n|F=f}$ and by Lemma 4 we have

$$\lim_{n\to\infty} \mathbb{V}(\bar{P}_{S^nX^nY^n\widehat{S}^n|F=f}, P_{S^nX^nY^n\widehat{S}^n|F=f}) = 0.$$

Since $\bar{P}_{S^nX^nY^n\widehat{S}^n|F=f}$ is close to $\bar{P}_{S^nX^nY^n\widehat{S}^n}$ because of (13), we have

$$\lim_{n\to\infty} \mathbb{V}(\bar{P}_{S^nX^nY^n\widehat{S}^n}, P_{S^nX^nY^n\widehat{S}^n}) = 0.$$

D. Rate constraints

We have imposed the following rate constraints:

$$H(U|Y) < \widetilde{R} + R_0 < H(U|S)$$
$$\widetilde{R} < H(U|SXY\widehat{S}).$$

Therefore we obtain:

$$R_0 > H(U|Y) - H(U|SXY\widehat{S}) = I(U; SX\hat{S}|Y)$$
$$I(U;S) < I(U;Y). \qquad \square$$

V. Proof of Theorem 1: outer bound

Consider a code (f_n, g_n) that induces a distribution $P_{S^nX^nY^n\widehat{S}^n}$ that is ε-close in L^1 distance to the i.i.d. distribution $P^{\otimes n}_{SXY\widehat{S}}$. Let the random variable T be uniformly distributed over the set $[1,n]$ and independent of the induced joint distribution $P_{S^nX^nY^n\widehat{S}^nC}$. The variable T will serve as a random time index. The variable S_T is independent of T because S^n is an i.i.d. source sequence [1]. Then we have

$$\begin{aligned}
0 &\overset{(a)}{\leq} I(X^n;Y^n) - I(C,S^n;Y^n)\\
&\leq I(C,X^n;Y^n) - I(C,S^n;Y^n)\\
&= I(X^n;Y^n|C) - I(S^n;Y^n|C) + I(C;Y^n) - I(C;Y^n)\\
&= H(Y^n|C) - H(Y^n|X^nC) + H(S^n|Y^nC) - H(S^n|C)\\
&\overset{(b)}{\leq} \sum_{t=1}^{n} \big(H(Y_t) - H(Y_t|X_t) + H(S_t|S^{t-1}Y_tY_{\sim t}C) - H(S_t)\big)\\
&\overset{(c)}{\leq} \sum_{t=1}^{n} \left(H(Y_t) - H(Y_t|X_t) + H(S_t|Y_{\sim t}C) - H(S_t)\right)\\
&\overset{(d)}{\leq} nH(Y_T) - nH(Y_T|X_T,T) + nH(S_T|Y_{\sim T}CT) - nH(S_T|T)\\
&\overset{(e)}{=} nH(Y_T) - nH(Y_T|X_T) + nH(S_T|Y_{\sim T}CT) - nH(S_T)\\
&= nI(X_T;Y_T) - nI(S_T;Y_{\sim T},C,T)
\end{aligned}$$

where (a) comes from the Markov chain $Y^n - X^n - (C,S^n)$ and (b) comes from the following facts: conditioning doesn't increase entropy, $\bar{P}_{Y|X}$ is a memoryless channel, the chain rule for the conditional entropy and S^n is an i.i.d. source independent of C. Recall that we note $Y_{\sim t}$ the vector $(Y_i)_{i\neq t}$, $i \in [1,n]$, where the component Y_t has been removed. The inequalities (c) and (d) come from the fact that $H(Y_T|T)$ is smaller or equal to $H(Y_T)$ since conditioning doesn't increase entropy and (e) from the memoryless channel $\bar{P}_{Y|X}$ and the i.i.d. source $\bar{P}_S$.

For the second part of the converse, we need to establish a technical result first. The proof is in the Appendix.

Lemma 5: Let P_{X^n} such that $\mathbb{V}\left(P_{X^n}, \bar{P}_X^{\otimes n}\right) \leq \varepsilon$, then we have

$$\sum_{t=1}^{n} I(X_t; X_{\sim t}) \leq nf(\varepsilon)$$

where $f(\varepsilon)$ goes to zero as ε does.

Then we have

$$\begin{aligned}
&nR_0 \geq H(C) \geq H(C|Y^n) \geq I(S^nX^n\widehat{S}^n; C|Y^n)\\
&= \sum_{t=1}^{n} I(S_tX_t\widehat{S}_t; C|S^{t-1}X^{t-1}\widehat{S}^{t-1}Y_{\sim t}Y_t)\\
&= \sum_{t=1}^{n} I(S_tX_t\widehat{S}_t; CS^{t-1}X^{t-1}\widehat{S}^{t-1}Y_{\sim t}|Y_t)\\
&- \sum_{t=1}^{n} I(S_tX_t\widehat{S}_t; S^{t-1}X^{t-1}\widehat{S}^{t-1}Y_{\sim t}|Y_t)\\
&\geq \sum_{t=1}^{n} I(S_tX_t\widehat{S}_t; CY_{\sim t}|Y_t)\\
&- \sum_{t=1}^{n} I(S_tX_t\widehat{S}_t; S^{t-1}X^{t-1}\widehat{S}^{t-1}Y_{\sim t}|Y_t)\\
&\overset{(a)}{\geq} \sum_{t=1}^{n} I(S_tX_t\widehat{S}_t; CY_{\sim t}|Y_t) - nf(\varepsilon)\\
&= nI(S_TX_T\widehat{S}_T; CY_{\sim T}|Y_TT) - nf(\varepsilon)\\
&= nI(S_TX_T\widehat{S}_T; CY_{\sim T}T|Y_T) - nI(S_T, X_T, \widehat{S}_T; T|Y_T) - nf(\varepsilon)\\
&\geq nI(S_TX_T\widehat{S}_T; CY_{\sim T}T|Y_T) - nI(S_T, X_T, \widehat{S}_T, Y_T; T) - nf(\varepsilon)\\
&\overset{(b)}{\geq} nI(S_TX_T\widehat{S}_T; CY_{\sim T}T|Y_T) - 2nf(\varepsilon)
\end{aligned}$$

where (a) follows from the following chain of inequalities

$$\begin{aligned}
&\sum_{t=1}^{n} I(S_tX_t\widehat{S}_t; S^{t-1}X^{t-1}\widehat{S}^{t-1}Y_{\sim t}|Y_t)\\
&\leq \sum_{t=1}^{n} I(S_tX_t\widehat{S}_t; S_{\sim t}X_{\sim t}\widehat{S}_{\sim t}Y_{\sim t}|Y_t)\\
&\leq \sum_{t=1}^{n} I(S_tX_t\widehat{S}_tY_t; S_{\sim t}X_{\sim t}\widehat{S}_{\sim t}Y_{\sim t}) \leq nf(\varepsilon)
\end{aligned}$$

and $f(\varepsilon)$ is defined in Lemma 5. Finally, the proof of (b) comes from [9, Lemma VI.3].

We identify the auxiliary random variables U_t with $(C, Y_{\sim t})$

for each $t \in [1,n]$ and U with $(C, Y_{\sim T}, T)$. For each $t \in [1,n]$ the following two Markov chains hold: $(S_t, X_t) - (C, Y_{\sim t}, Y_t) - \widehat{S}_t$ and $Y_t - X_t - (C, Y_{\sim t}, S_t)$. Since $U = U_t$ when $T = t$, we also have $(S, X) - (U, Y) - \widehat{S}$ and $Y - X - (U, S)$. The cardinality bound comes from [9, Lemma VI.1].

APPENDIX

A. Proof of Lemma 3

We denote the event that $\widehat{V}^n$ is equal to V^n with $\mathcal{A} := \{V^n = \widehat{V}^n\}$. We know that $\mathbb{P}\{\mathcal{A}\}$ tends to 1. We can write the joint distribution $P_{W^n V^n \widehat{V}^n}$ as

$$\mathbb{P}\{\mathcal{A}\} P_{W^n V^n \widehat{V}^n | \mathcal{A}} + \mathbb{P}\{\mathcal{A}^c\} P_{W^n V^n \widehat{V}^n | \mathcal{A}^c}.$$

Hence, we have

$$\mathbb{V}(P_{W^n V^n \widehat{V}^n}, P_{W^n V^n} \mathbb{1}_{\widehat{V}^n | V^n}) \leq \mathbb{P}\{\mathcal{A}^c\} \left\| P_{W^n V^n \widehat{V}^n | \mathcal{A}^c} \right\|_{L^1} + \left\| \mathbb{P}\{\mathcal{A}\} P_{W^n V^n \widehat{V}^n | \mathcal{A}} - P_{W^n V^n} \mathbb{1}_{\widehat{V}^n | V^n} \right\|_{L^1}$$

where the first term is equal to $(1 - \mathbb{P}\{\mathcal{A}\}) P_{W^n V^n} \mathbb{1}_{\widehat{V}^n | V^n}$ and goes to 0 since $\mathbb{P}\{\mathcal{A}\}$ tends to 1 and the second term goes to 0 since $\mathbb{P}\{\mathcal{A}^c\}$ does.

B. Proof of Remark 4

We want to prove that there exists a binning which works for both the conditions in Section IV-B and Section IV-C. If we denote with $\mathbb{E}_{\varphi_1 \varphi_2}$ and $\mathbb{E}_{\varphi_2}$ the expected value with respect to the random binnings, for all ε, there exists $\bar{n}$ such that $\forall n \geq \bar{n}$

$$\mathbb{E}_{\varphi_1 \varphi_2}\left[\mathbb{V}\left(\bar{P}^{\varphi_1 \varphi_2}_{S^n FC}, Q_F Q_C \bar{P}_{S^n}\right)\right] < \frac{\varepsilon}{2}$$
$$\mathbb{E}_{\varphi_2}\left[\mathbb{V}\left(\bar{P}^{\varphi_2}_{S^n X^n Y^n \widehat{S}^n F}, Q_F \bar{P}_{S^n X^n Y^n \widehat{S}^n}\right)\right] < \frac{\varepsilon}{2}$$

which implies by Markov's inequality

$$\mathbb{P}_{\varphi_1 \varphi_2}\left\{\mathbb{V}\left(\bar{P}^{\varphi_1 \varphi_2}_{S^n FC}, Q_F Q_C \bar{P}_{S^n}\right) < \varepsilon\right\} > \frac{1}{2} \tag{15}$$
$$\mathbb{P}_{\varphi_2}\left\{\mathbb{V}\left(\bar{P}^{\varphi_2}_{S^n X^n Y^n \widehat{S}^n F}, Q_F \bar{P}_{S^n X^n Y^n \widehat{S}^n}\right) < \varepsilon\right\} > \frac{1}{2}.$$

In Section IV-B and IV-C we have chosen the binnings (φ_1', φ_2') and φ_2'' respectively such that

$$\lim_{n \to \infty} \mathbb{V}\left(\bar{P}^{\varphi_1' \varphi_2'}_{S^n FC}, Q_F Q_C \bar{P}_{S^n}\right) = 0$$
$$\lim_{n \to \infty} \mathbb{V}\left(\bar{P}^{\varphi_2''}_{S^n X^n Y^n \widehat{S}^n F}, Q_F \bar{P}_{S^n X^n Y^n \widehat{S}^n}\right) = 0.$$

It follows from (15) that the intersection of the two sets is non-empty, therefore there exists a binning φ_2^* that satisfies both conditions.

C. Proof of Lemma 5

The following result has already been proved in [10, Lemma 2.7].

Lemma 6: Let P and Q two distributions on $\mathcal{X}$ such that $\mathbb{V}(P, Q) = \varepsilon$ and $\varepsilon \leq 1/2$, then

$$|H(P) - H(Q)| \leq \varepsilon \log \frac{|\mathcal{X}|}{\varepsilon}.$$

We also need this lemma proved in [7, Lemma 3.2'].

Lemma 7: If $\mathbb{V}(P_X P_{Y|X}, Q_X Q_{Y|X}) \leq \varepsilon$ then $\mathbb{P}\{x \in \mathcal{X} | \mathbb{V}(P_{Y|X=x}, Q_{Y|X=x}) \leq \sqrt{\varepsilon}\} \geq 1 - 2\sqrt{\varepsilon}$.

Now, consider the set

$$\mathcal{B} := \{\mathbf{x} \in \mathcal{X}^{n-1} | \mathbb{V}(P_{X_t | X_{\sim t} = \mathbf{x}}, \bar{P}_X) \leq \varepsilon\}.$$

By Lemma 7, $\mathbb{P}\{\mathcal{B}\} \geq 1 - 2\sqrt{\varepsilon}$. Observe that

$$\begin{aligned}
&H(X) - H(X_t | X_{\sim t}) \\
&= H(X) - \sum_{\mathbf{x} \in \mathcal{X}^{n-1}} P_{X_{\sim t}}(\mathbf{x}) H(X_t | X_{\sim t} = \mathbf{x}) \\
&\leq \sum_{\mathbf{x} \in \mathcal{X}^{n-1}} \left(P_{X_{\sim t}}(\mathbf{x}) H(X) - P_{X_{\sim t}}(\mathbf{x}) H(X_t | X_{\sim t} = \mathbf{x})\right) \\
&= \sum_{\mathbf{x} \in \mathcal{B}} \left(P_{X_{\sim t}}(\mathbf{x}) H(X) - P_{X_{\sim t}}(\mathbf{x}) | H(X_t | X_{\sim t} = \mathbf{x})\right) \\
&+ \sum_{\mathbf{x} \in \mathcal{B}^c} \left(P_{X_{\sim t}}(\mathbf{x}) H(X) - P_{X_{\sim t}}(\mathbf{x}) | H(X_t | X_{\sim t} = \mathbf{x})\right).
\end{aligned}$$

Hence by Lemma 6

$$|H(X_t | X_{\sim t} = \mathbf{x}) - H(X)| \leq \varepsilon \log \frac{|\mathcal{X}|}{\varepsilon}.$$

Let $\delta := \varepsilon \log \frac{|\mathcal{X}|}{\varepsilon}$, then the first term is bounded by

$$\sum_{\mathbf{x} \in \mathcal{B}} P_{X_{\sim t}}(\mathbf{x}) \delta \leq \delta,$$

while the second term is smaller than

$$\mathbb{P}\{\mathcal{B}^c\} \left(H(X_t) + H(X)\right) \leq 2\sqrt{\varepsilon} \left(2H(X) + \delta\right).$$

Again, by Lemma 6, we have

$$|H(X_t) - H(X)| \leq \delta.$$

Finally, $I(X_t; X_{\sim t}) = H(X_t) - H(X) + H(X) - H(X_t | X_{\sim t})$ is smaller than $f(\varepsilon) = 2\sqrt{\varepsilon}(2H(X) + \delta) + 2\delta$.

REFERENCES

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

[2] F. Haddadpour, M. H. Yassaee, M. R. Aref, and A. Gohari, "When is it possible to simulate a DMC channel from another?", in *Proc. of IEEE Information Theory Workshop (ITW),*, 2013, pp. 1-5.

[3] P. Cuff and C. Schieler, "Hybrid codes needed for coordination over the point-to-point channel," in *Proc. of Allerton Conference on Communication, Control and Computing,* 2011, pp. 235-239.

[4] S. Satpathy and P. Cuff, "Secure cascade channel synthesis," *IEEE Transactions on Information Theory,* vol. 62, no. 11, pp. 6081-6094, 2016.

[5] A. El Gamal and Y. H. Kim, *Network information theory.* Cambridge University Press, 2011.

[6] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography - Part II: CR capacity," *IEEE Transactions on Information Theory,* vol. 44, no. 1, pp. 225-240, 1998.

[7] M. H. Yassaee, M. R. Aref, and A. Gohari, "Achievability proof via output statistics of random binning," *IEEE Transactions on Information Theory,* vol. 60, no. 11, pp. 6760-6786, 2014.

[8] P. Cuff, "Communication in networks for coordinating behavior," Ph.D. dissertation, Stanford University, 2009.

[9] P. Cuff, "Distributed channel synthesis," *IEEE Transactions on Information Theory,* vol. 59, no. 11, pp. 7071-7096, 2013.

[10] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems.* Cambridge University Press, 2011