Joint Strong Coordination and Sensing

Mohammad Mahdi Assaf*, Giulia Cervia†, Michèle Wigger‡

*Télécom Paris, Institut Polytechnique de Paris, LTCI, 91120 Palaiseau, France

†IMT Nord Europe, Institut Mines-Télécom, Centre for Digital Systems, F-59000 Lille, France

‡Université Paris-Saclay, CNRS, CentraleSupélec, L2S, 91190 Gif-sur-Yvette, France

Emails: mohammad.assaf@telecom-paris.fr, giulia.cervia@imt-nord-europe.fr, michele.wigger@centralesupelec.fr

***Abstract*—We study a joint strong coordination and sensing scenario: A transmitter observes an independent and identically distributed (i.i.d.) P_U-source sequence and communicates causally to a receiver, whose goal is to synthesize an action sequence V^N whose joint law with the source sequence U^N is close to an i.i.d. target law Q_{UV} (with marginal P_U). Communication is over a state-dependent memoryless channel with generalized feedback (modelling backscattered signals) to the transmitter, who exploits these signals to sense the hidden channel state S^N. That means, it synthesizes an estimation sequence $\hat{S}^N$ whose joint law with S^N should be close to an i.i.d. target distribution $Q_{S\hat{S}}$. In this paper, we propose a block-Markov coding scheme based on the output statistics of random binning framework and implicit randomness extraction at the two terminals, and we derive the corresponding sufficient conditions for pairs of joint pmfs $(Q_{UV}, Q_{S\hat{S}})$ that can be strongly coordinated this way, in function of the available rate of external common randomness shared between transmitter and receiver. We also provide a converse proof which identifies necessary conditions for pairs of joint pmfs $(Q_{UV}, Q_{S\hat{S}})$ that can be strongly coordinated for some rate of external randomness. A key new step here is to prove optimality of a memoryless estimation kernel at the transmitter side to produce $\hat{S}^N$. Our sufficient and necessary conditions coincide for sufficiently high rates of available common randomness, thus establishing the set of all pairs of joint pmfs $(Q_{UV}, Q_{S\hat{S}})$ that can be strongly coordinated under this assumption.**

I. Introduction

Strong coordination [1]–[5] studies communication systems whose goal is that distributed agents produce sequences approximately following a prescribed joint law.

Recent years saw an emergence of integrated sensing and communication (ISAC) systems, in which the same resources are utilized for both communication and sensing tasks. Specifically, a transmitter communicates to a receiver, and from the backscattered signals (modeled as generalized feedback), the transmitter attempts to estimate parameters of the environment. Theoretic studies on this subject include the works in [6]–[11].

In this work, we take a *strong coordination* approach to this problem. That means, we consider the setup in Figure 1 and assume that the transmitter causally observes an i.i.d. U-sequence. The receiver produces a V-sequence on its side, in a way that the two sequences follow an i.i.d. target joint law Q_{UV}, i.e,. so that the joint UV-sequence is strongly coordinated according to Q_{UV}. From the backscattered signal, the transmitter produces a state-estimation sequence $\hat{S}^N$ that should be strongly coordinated according to an i.i.d. target law $Q_{S\hat{S}}$ with the state sequence S^N of the channel.

Our work is related to the coordination results in [2], [12], which however do not consider sensing. Moreover [12] has no feedback and [2] considers weak and not strong coordination.

In this work, we present necessary and sufficient conditions for target laws Q_{UV} and $Q_{S\hat{S}}$ to allow for strong coordination in our setup. These necessary and sufficient conditions coincide when the transmitter and receiver can share common randomness of sufficiently high rate.

Throughout this article, random variables are denoted by uppercase letters, their realizations by lowercase letters, and their alphabets by calligraphic letters. For a sequence $X^n = (X_1, \ldots, X_n)$, we write $X^{i-1} = (X_1, \ldots, X_{i-1})$, $X_{i+1}^n = (X_{i+1}, \ldots, X_n)$, and $X_{\sim i} = (X^{i-1}, X_{i+1}^n)$. In the block-Markov proof, X_b^n denotes the length-n sequence in block b. We abbreviate *probability mass function* by *pmf* and *independent and identically distributed* by *i.i.d.*. The total variation (TV) distance between two pmfs P and Q is denoted $\|P - Q\|_{TV} = \sum_x |P(x) - Q(x)|/2$.

II. System Model and Problem Formulation

Consider the setup of Fig. 1. Node 1 observes a source process $\{U_i\}_{i\geq 1}$, which is i.i.d. according to a given $\bar{P}_U$, and communicates over a state-dependent memoryless channel (SDMC) to a Node 2, which then produces a coordinated sequence $\{V_i\}$ at its end. Node 1 observes a feedback signal from the SDMC, based on which it produces an estimate $\{\hat{S}_i\}$ of an internal state sequence $\{S_i\}$.

The SDMC is described by the transition law $\bar{P}_{YZ|XS}(y, z|x, s)$, meaning that at each time i and given channel input $X_i = x$ and internal state $S_i = s$, Node 2 observes channel output $Y_i \sim P_{Y|XS}(\cdot|x, s)$ and Node 1 feedback output $Z_i \sim P_{Z|XS}(\cdot|x, s)$. The goal is that Node 2 coordinates its N outputs V^N with the N source symbols U^N according to a product pmf $Q_{UV}^{\otimes N}$ and that Node 1 coordinates its N estimates $\hat{S}^N$ with the first N internal state symbols S^N according to the product pmfs $Q_{S\hat{S}}^{\otimes N}$. We shall allow that the two nodes employ the SDMC slightly more than N times, more precisely, we allow for $N' = (1 + \alpha)N$ channel uses, where α can be an arbitrarily small but positive.

We assume causal encoding and define the following notion of a code.

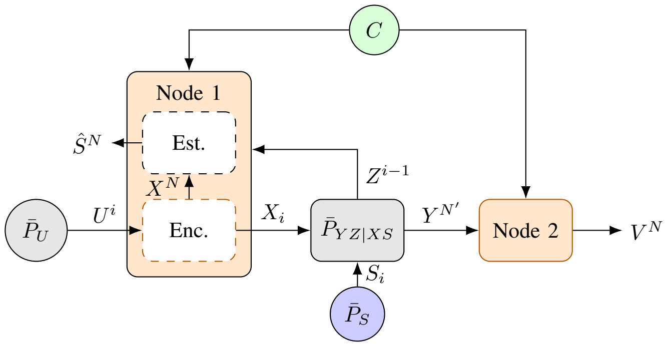

Fig. 1. Joint strong coordination and sensing setup.

Definition 1. An (N, α, R_C) strong coordination–sensing code consists of:

1) A common-randomness variable C, independent of $(U^N, S^{N'})$ and uniformly distributed over $[1 : 2^{NR_C}]$, available noncausally to both nodes;
2) A sequence of stochastic encoding kernels

$$\phi_i = \begin{cases} \mathcal{U}^i \times \mathcal{Z}^{i-1} \times [1:2^{NR_C}] \to \mathcal{P}(\mathcal{X}), & i \in [1:N], \\ \mathcal{U}^N \times \mathcal{Z}^{i-1} \times [1:2^{NR_C}] \to \mathcal{P}(\mathcal{X}), & i \in [N+1:N'], \end{cases}$$

mapping the past and the current source symbols U^i, the previous feedback outputs Z^{i-1}, and the common randomness C to the i-th channel input X_i.
3) A stochastic decoding kernel

$$g\colon \mathcal{Y}^{N'} \times [1 : 2^{NR_C}] \to \mathcal{P}(\mathcal{V}^N),$$

mapping all channel outputs $Y^{N'}$ and C to a sequence V^N.
4) A stochastic estimator kernel

$$h\colon \mathcal{U}^N \times [1 : 2^{NR_C}] \times \mathcal{X}^{N'} \times \mathcal{Z}^{N'} \to \mathcal{P}(\hat{\mathcal{S}}^N),$$

mapping all channel inputs and feedback outputs to the estimated sequence $\hat{S}^N$.

Our goal is to characterize the set of pmfs Q_{UV} and $Q_{S\hat{S}}$ that can be coordinated with a given rate of common randomness R_C and for given channel pmfs $\bar{P}_{YZ|XS}$ and $\bar{P}_S$.

Definition 2. A pair $(Q_{UV}, Q_{S\hat{S}})$ is R_C-achievable if for any $\alpha > 0$ there exists a sequence (in N) of (N, α, R_C) strong coordination–sensing codes such that the induced joint pmfs on (U^N, V^N) and $(S^N, \hat{S}^N)$, denoted $P_{U^N V^N}$ and $P_{S^N \hat{S}^N}$ respectively, satisfy

$$\lim_{N\to\infty} \left\| P_{U^N V^N} - Q_{UV}^{\otimes N} \right\|_{\mathrm{TV}} = 0, \tag{1}$$

$$\lim_{N\to\infty} \left\| P_{S^N \hat{S}^N} - Q_{S\hat{S}}^{\otimes N} \right\|_{\mathrm{TV}} = 0. \tag{2}$$

III. Main Results

Theorem 1 (Sufficient Condition)**.** *A pair of pmfs $(Q_{UV}, Q_{S\hat{S}})$ is R_C-achievable if there exists a pmf Q_{W_1} and kernels $Q_{X|UW_1}$, $Q_{W_2|UZW_1}$, $Q_{K|XZ}$, $Q_{V|YW_1W_2}$, and $Q_{\hat{S}|XZ}$ such that the joint pmf*

$$Q_{UVSXYZ\hat{S}W_1W_2K} = \bar{P}_U \bar{P}_S Q_{W_1} Q_{X|UW_1} Q_{K|XZ} Q_{W_2|UZW_1} \cdot \bar{P}_{YZ|XS} Q_{V|YW_1W_2} Q_{\hat{S}|XZ} \tag{3}$$

satisfies the following three conditions:

1) *Its (U,V)- and $(S,\hat{S})$-marginals are Q_{UV} and $Q_{S\hat{S}}$.*
2) *The following information constraint is satisfied:*

$$I(U,Z;W_2 \mid W_1, Y) + I(X,Z;K \mid W_1, W_2, Y) < I(W_1;Y). \tag{4}$$

3) *The rate of common randomness satisfies:*

$$R_C > \max \left\{ \begin{array}{c} H(W_1|Y) + H(W_2|W_1,Y) \\ \quad - H(W_2|U,V) + \Delta_K^{UV}, \\ H(W_1|Y) + H(W_2|W_1,Y) \\ \quad - H(W_2|S,\hat{S}) + \Delta_K^{S\hat{S}} \end{array} \right\}, \tag{5}$$

where

$$\Delta_K^{UV} \triangleq H(K|W_1,W_2,Y) - H(K|W_2,U,V), \tag{6}$$

$$\Delta_K^{S\hat{S}} \triangleq H(K|W_1,W_2,Y) - H(K|W_2,S,\hat{S}). \tag{7}$$

This sufficient condition is obtained with a block-Markov coding scheme, as described in Section IV-B.

Intuitively, in each block the W_1 sequence is used to describe the current source sequence U to Node 2, while W_2 is used to refine the information from the previous blocks using their feedback outputs Z. The K-sequence is used to extract common randomness from the channel at the two nodes, which can be used in subsequent blocks. The estimate sequence $\hat{S}^N$ is produced by applying the memoryless kernel $Q_{\hat{S}|XZ}$ symbolwise to the input–feedback pairs (X^N, Z^N).

We next provide a necessary condition for R_C-achievability.

Theorem 2 (Necessary Condition)**.** *If a pair of pmfs $(Q_{UV}, Q_{S\hat{S}})$ is R_C-achievable, then there exists a pmf Q_{W_1} and kernels $Q_{X|UW_1}$, $Q_{W_2|UZW_1}$, $Q_{V|YW_1W_2}$, and $Q_{\hat{S}|XZ}$ such that the joint pmf in* (3) *(with K constant) satisfies Conditions 1) and 2) in Theorem 1.*

Having a non-degenerate K can decrease the required amount of common randomness R_C in the achievability result, but not the set of pmfs $(Q_{UV}, Q_{S\hat{S}})$ that are achievable *for some* R_C. This follows also from the next corollary to above two theorems.

Corollary 1. *When the rate of common randomness R_C is sufficiently large (depending on the pmfs $\bar{P}_S$, $\bar{P}_U$, and $P_{YZ|XS}$), then the pair $(Q_{UV}, Q_{S\hat{S}})$ is R_C-achievable if, and only if, the conditions in Theorem 2 are satisfied.*

Remark 1. Consider arbitrary source and state distributions $\bar{P}_U$ and $\bar{P}_S$ and a channel of the form

$$X = \begin{pmatrix} X_{\mathrm{s}} \\ X_{\mathrm{r}} \end{pmatrix} \to (Y, Z) = \begin{pmatrix} Y_{\mathrm{s}}, Z_{\mathrm{s}} \\ (B, B) \end{pmatrix} \tag{8}$$

for B an independent randomness.

Consider any choice of the auxiliaries W_1, W_2, V independent of B, and observe then that the choice $K = B$ independent of $W_1, W_2, U, V, S, \hat{S}$, leads to the same information constraint

$$I(U, Z; W_2|W_1, Y) < I(W_1; Y) \tag{9}$$

as when choosing K a constant. However, it allows to reduce rate of common randomness by $H(K)$. Having K non degenerate thus allows to strictly decrease the common randomness rate, whenever it is positive for K=const.

IV. Proof of Theorem 1

The proof of our sufficient condition is based on a block-Markov coding scheme. To this end, let $\alpha > 0$ and B sufficiently large so that $\alpha \geq \frac{1}{B}$. Then, split the blocklength N' into $B+1$ blocks, each of length $n \triangleq \lfloor \frac{N}{B} \rfloor$. Note that since $N' = (1+\alpha)N \geq (\frac{B+1}{B})N$, there might be a few channel uses left at the end, which we will not be using.

The transmitter describes each source block U_b^n during channel blocks b and $b+1$, for $b = 1, \ldots, B$, and the receiver produces the reconstruction sequence V_b^n based on output blocks Y_b^n and Y_{b+1}^n. The detailed encodings and decodings will be described shortly, see Subsections IV-A and IV-B. At the very end, the transmitter produces the estimation sequence $\hat{S}^N$ by employing a memoryless stochastic kernel $Q_{\hat{S}|XZ}$ to the input and feedback sequences (X^N, Z^N).

Analysis is based on the fact that the block-Markov procedure imposes a joint law of the following form, for some "hidden" random variables $\{H_{1,b}^n, H_{2,b}^n\}_{b=1}^B$ explained later:

$$\begin{aligned} &P_{U^N,V^N,S^N,\hat{S}^N,H_{1,1}^n,\ldots,H_{1,B}^n,H_{2,1}^n,\ldots,H_{2,B}^n} \\ &= P_{U_1^n,H_{1,1}^n} \prod_{b=2}^{B} P_{V_{b-1}^n,S_{b-1}^n,\hat{S}_{b-1}^n,H_{2,b-1}^n,U_b^n,H_{1,b}^n|U_{b-1}^n,H_{1,b-1}^n} \\ &\quad \cdot P_{V_B^n,S_B^n,\hat{S}_B^n,H_{2,B}^n|U_B^n,H_{1,B}^n}. \end{aligned} \tag{10}$$

Using a sliding-window analysis, we will show that

$$P_{U_1^n,H_{1,1}^n} = Q_{U,H_1}^{\otimes n}, \tag{11a}$$

and for $b = 2, \ldots, B$:

$$\begin{aligned} \mathbb{E}_{Q^{\otimes n}_{U',H_1'}}\Big[\Big\| &P_{V_{b-1}^n,S_{b-1}^n,\hat{S}_{b-1}^n,H_{2,b-1}^n,U_b^n,H_{1,b}^n|U_{b-1}^n,H_{1,b-1}^n}(\cdot|U'^n,H_1'^n) \\ &-\Gamma^{\otimes n}_{V,S,\hat{S},H_2,U,H_1|U',H_1'}(\cdot|U'^n,H_1'^n)\Big\|_{\mathrm{TV}}\Big] \leq \eta_n. \end{aligned} \tag{11b}$$

for a sequence $\eta_n \to 0$ as $n \to \infty$ and some pmfs Q_{U,H_1} and $Q_{V,S,\hat{S},H_2|U',H_1'}$ and

$$\Gamma_{V,S,\hat{S},H_2,U,H_1|U',H_1'} \triangleq Q_{V,S,\hat{S},H_2|U',H_1'} Q_{U,H_1}$$

so that the resulting per-block pmf

$$Q_{U,V,S,\hat{S},H_1,H_2} = Q_{U,H_1} Q_{V,S,\hat{S},H_2|U,H_1},$$

has marginals $Q_{U,V}$ and $Q_{S,\hat{S}}$.

The same bound, after marginalizing the newly generated $(U_b^n, H_{1,b}^n)$, applies to the termination block that produces $(V_B^n, S_B^n, \hat{S}_B^n, H_{2,B}^n)$.

By the triangle inequality for TV-distance (for details, see Appendix A-A) and the fact that marginalizing out can only reduce TV-distance [13, Lemma 16], we then obtain over all blocks, with $N = Bn$:

$$\left\| P_{U^N,V^N} - Q_{U,V}^{\otimes N} \right\|_{\mathrm{TV}} \leq (B+1)\eta_n, \tag{12a}$$

$$\left\| P_{S^N,\hat{S}^N} - Q_{S,\hat{S}}^{\otimes N} \right\|_{\mathrm{TV}} \leq (B+1)\eta_n. \tag{12b}$$

Therefore, for any fixed B the TV distance vanishes. Letting $B \to \infty$, we conclude the proof for all values of α.

It remains to explain the detailed encodings and decodings in each block and to prove Inequalities (11). We employ the output statistics of random binning approach (OSRB) to obtain the per-block encodings and decodings. In Subsection IV-A we describe a random binning scheme (sketched in Fig. 2), which implicitly defines two conditional distributions used as per-block encoders and decoders in our block-Markov scheme as depicted in Fig. 3, see Subsection IV-B.

A. Per-Block Random Binning Scheme

Consider the random-binning scheme in Figure 2, where we defined the averaged channel

$$\bar{P}_{YZ|X}(y,z|x) \triangleq \sum_{s\in\mathcal{S}} \bar{P}_S(s)\bar{P}_{YZ|XS}(y,z|x,s) \tag{13}$$

and choose Q_{W_1} and the kernels $Q_{X|UW_1}$, $Q_{W_2|UZW_1}$, $Q_{K|XZ}$, $Q_{V|YW_1W_2}$, and $Q_{\hat{S}|XZ}$.

We then define the target distribution

$$\begin{aligned} Q_{UW_1XYZW_2KV} = \bar{P}_U Q_{W_1} Q_{X|UW_1} \bar{P}_{YZ|X} \\ \cdot Q_{W_2|UZW_1} Q_{K|XZ} Q_{V|YW_1W_2}. \end{aligned} \tag{14}$$

For block 1, draw $W_{1,1}^n \sim Q_{W_1}^{\otimes n}$ and sequences $X_1^n \sim Q_{X|UW_1}^{\otimes n}(\cdot|U_1^n, W_{1,1}^n)$ and $(Y_1^n, Z_1^n) \sim \bar{P}_{YZ|X}^{\otimes n}(\cdot,\cdot|X_1^n)$.

For blocks $b = 2, \ldots, B$, the tuple from the previous block

$$\Sigma_{b-1}^n \triangleq (U_{b-1}^n, W_{1,b-1}^n, X_{b-1}^n, Y_{b-1}^n, Z_{b-1}^n), \tag{15}$$

and $(U_b^n, W_{1,b}^n)$ are fed to the chosen kernels to generate

$$\begin{cases} W_{2,b-1}^n & \sim Q_{W_2|UZW_1}^{\otimes n}(\cdot|U_{b-1}^n, Z_{b-1}^n, W_{1,b-1}^n), \\ K_{b-1}^n & \sim Q_{K|XZ}^{\otimes n}(\cdot|X_{b-1}^n, Z_{b-1}^n), \\ W_{1,b}^n & \sim Q_{W_1}^{\otimes n}, \\ X_b^n & \sim Q_{X|UW_1}^{\otimes n}(\cdot|U_b^n, W_{1,b}^n), \\ (Y_b^n, Z_b^n) & \sim \bar{P}_{YZ|X}^{\otimes n}(\cdot,\cdot|X_b^n). \end{cases}$$

For the termination block $b = B+1$, no new source block is coordinated. Thus draw

$$X_{B+1}^n \sim Q_{X|W_1}^{\otimes n}(\cdot|W_{1,B+1}^n) \tag{16}$$

for

$$Q_{X|W_1}(x|w_1) = \sum_u \bar{P}_U(u) Q_{X|UW_1}(x|u,w_1). \tag{17}$$

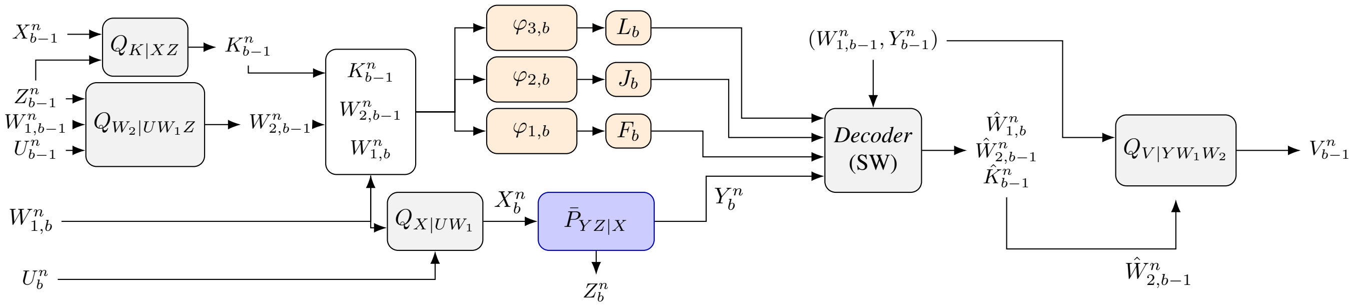

Fig. 2. Random binning scheme for blocks $b \geq 2$.

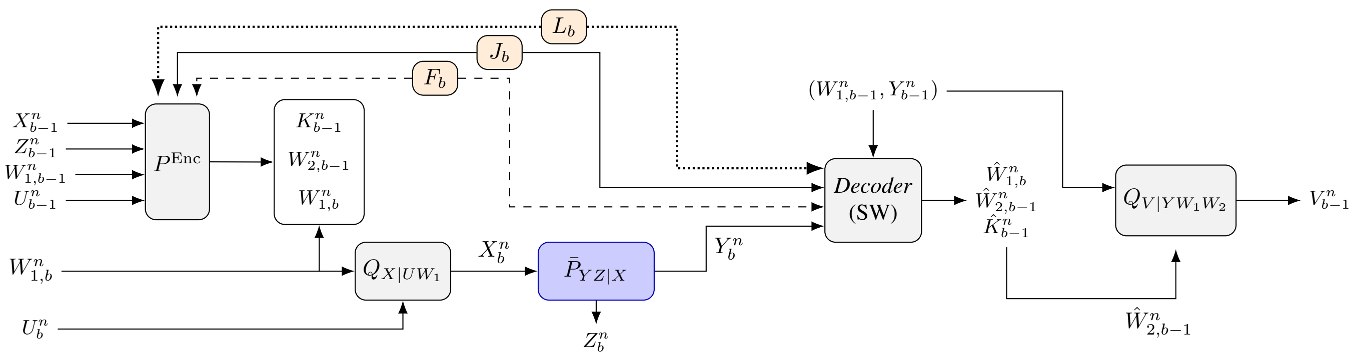

Fig. 3. Random coding scheme for blocks $b \geq 2$.

Also, let $(Y^n_{B+1}, Z^n_{B+1}) \sim \bar{P}^{\otimes n}_{YZ|X}(\cdot,\cdot|X^n_{B+1})$.

In addition, for every $b = 2, \ldots, B+1$, generate three random bin indices [14] of rates $\tilde{R}_F$, $\tilde{R}_J$, and $\tilde{R}_K$:

$$F_b = \varphi_{1,b}(W^n_{1,b}, W^n_{2,b-1}, K^n_{b-1}),$$
$$J_b = \varphi_{2,b}(W^n_{1,b}, W^n_{2,b-1}, K^n_{b-1}),$$
$$L_b = \varphi_{3,b}(W^n_{1,b}, W^n_{2,b-1}, K^n_{b-1}),$$

using independent random binnings $\varphi_{1,b}, \varphi_{2,b}, \varphi_{3,b}$. (For $b = 1$, $W_{2,b-1}$ and K_{b-1} are assumed constant.)

Based on these bin indices and the side information $(Y^n_b, W^n_{1,b-1}, Y^n_{b-1})$, for every $b = 2, \ldots, B+1$, a Slepian-Wolf (SW) decoder [14], [15] is employed to produce the reconstructions $(\hat{W}^n_{1,b}, \hat{W}^n_{2,b-1}, \hat{K}^n_{b-1})$, and finally generate

$$V^n_{b-1} \sim Q^{\otimes n}_{V|YW_1W_2}(\cdot|Y^n_{b-1}, W^n_{1,b-1}, \hat{W}^n_{2,b-1}). \tag{18}$$

For given binnings $\varphi_{1,b}, \varphi_{2,b}, \varphi_{3,b}$, the SW-decoder induces a kernel

$$P^{\mathrm{SW}}_{\hat{W}^n_{1,b}\hat{W}^n_{2,b-1}\hat{K}^n_{b-1}|F_bJ_bL_bW^n_{1,b-1}Y^n_{b-1}Y^n_b}, \tag{19}$$

which will be the decoder in our random coding scheme described in the following subsection. The above construction also induces a kernel

$$P^{\mathrm{RB}}_{W^n_{1,b}W^n_{2,b-1}K^n_{b-1}|U^n_{b-1}W^n_{1,b-1}X^n_{b-1}Z^n_{b-1}F_bJ_bL_b},$$

which will be our encoder. We later refer to the joint distribution induced by the scheme above as P^{RB}.

B. Random Coding Scheme

Split C into equally sized parts $C = (J_1, \ldots, J_{B+1})$ and assume for the moment extra common randomness $(L_1, \ldots, L_{B+1})$ and $(F_1, \ldots, F_{B+1})$ of rates $\tilde{R}_K$ and $\tilde{R}_F$. The randomness L_b will later be replaced by randomness extracted at both the transmitter and the receiver, and the randomness F_b will be eliminated later on.

We start by describing the encoding. In block $b = 1$, the encoder generates $W^n_{1,1} \sim Q^{\otimes n}_{W_1}$ and $X^n_1 \sim Q^{\otimes n}_{X|UW_1}(\cdot|U^n_1, W^n_{1,1})$, and transmits X^n_1 over the channel.

For blocks $b = 2, \ldots, B$ the encoding is described in Figure 3, where we choose

$$P^{\mathrm{Enc}} \triangleq P^{\mathrm{RB}}_{W^n_{1,b}W^n_{2,b-1}K^n_{b-1}|U^n_{b-1}W^n_{1,b-1}X^n_{b-1}Z^n_{b-1}F_bJ_bL_b}. \tag{20}$$

That means, based on the common randomness bits F_b, J_b, and L_b, and the sequences $(U^n_{b-1}W^n_{1,b-1}X^n_{b-1}Z^n_{b-1})$ generated in the previous block, it constructs the sequences $(W^n_{1,b}, W^n_{2,b-1}K^n_{b-1})$, and feeds $W^n_{1,b}$ and U^n_b to the kernel $Q_{X|UW_1}$ to construct $X^n_b \sim Q^{\otimes n}_{X|UW_1}(\cdot|U^n_b, W^n_{1,b})$, which it then sends through the channel. In the termination block $b = B+1$, the same encoder kernel produces $W^n_{1,B+1}$, $W^n_{2,B}$, and K^n_B, and the channel input is generated according to $Q^{\otimes n}_{X|W_1}(\cdot|W^n_{1,B+1})$ as in (17).

Figure 3 also describes the decoding steps at the end of each block $b = 2, \ldots, B+1$. The decoder first applies the kernel $P^{\mathrm{SW}}_{\hat{W}^n_{1,b}\hat{W}^n_{2,b-1}\hat{K}^n_{b-1}|F_bJ_bL_bW^n_{1,b-1}Y^n_{b-1}Y^n_b}$ to the common randomness (F_b, J_b, L_b) and the sequences $(W_{1,b-1}^n, Y_{b-1}^n, Y_b^n)$ to produce $(\hat{W}_{1,b}^n, \hat{W}_{2,b-1}^n, \hat{K}_{b-1}^n)$. It finally, generates

$$V_{b-1}^n \sim Q_{V|YW_1W_2}^{\otimes n}(\cdot|Y_{b-1}^n, W_{1,b-1}^n, \hat{W}_{2,b-1}^n).$$

We later refer to the joint distribution induced by the coding scheme above as P^{RC}.

C. TV Analysis

In a first instance, we assume that $(F_1, \ldots, F_{B+1})$ and $(L_1, \ldots, L_{B+1})$ are additional common randomness and show that when

$$\tilde{R}_F + \tilde{R}_J + \tilde{R}_K > H(W_1|Y) + H(W_2, K|W_1, Y), \tag{21}$$
$$\tilde{R}_F + \tilde{R}_J + \tilde{R}_K < H(W_1) + H(W_2|UZW_1) + H(K|XZ), \tag{22}$$

the pmf P^{RC} satisfies (11) for $H_{1,b}^n = (W_{1,b}^n, X_b^n, Y_b^n, Z_b^n)$ and $H_{2,b}^n = (W_{2,b}^n, K_b^n)$. This is detailed in Appendix A-B, and is done in two steps:

1) Prove that (11) is satisfied by the relevant part of P^{RB} (the distribution in the random binning scheme).
2) Prove that P^{RB} and P^{RC} are TV-vanishingly close.

Then, the triangle inequality will conclude the proof.

Note that analysis is performed on expectation over the random choice of the binning schemes. Ensuring vanishing TV distances on expectation ensures that the TV distances vanish for at least one realization of the binnings.

Moreover, ensuring (11) coordinates the whole sequence of involved random variables, while the effective requirement to coordinate the U and V sequences and the S and $\hat{S}$ sequences is much weaker. In Appendix A-C we show that we can fix an instance of the extra randomness F (as in [3]) and also replace the external randomness L with randomness extracted from previous K-sequences, while still ensuring the desired strong coordination under the condition on the common randomness rate in (5) .

V. Proof of Theorem 2

Consider $\alpha > 0$ and an (N, α, R_C) strong coordination–sensing code with common randomness C, as in Definition 1, that achieves $(Q_{UV}, Q_{S\hat{S}})$ in the sense of Definition 2. Let $N' \leq (1+\alpha)N$ denote the number of channel uses. For ease of notation, let Ξ denote the internal local randomness employed at the stochastic encoder. Then, for each $i \in [1:N]$, define

$$W_{1,i} \triangleq (\Xi, U^{i-1}, Z^{i-1}, Y^{i-1}, C), \qquad W_{2,i} \triangleq Y_{i+1}^{N'}. \tag{23}$$

Let T be uniform over $[1:N]$ and independent of $(U^N, X^{N'}, S^{N'}, Z^{N'}, Y^{N'}, V^N, \hat{S}^N, C)$. Define

$$\begin{aligned}
&W_1 \triangleq (W_{1,T}, T), && W_2 \triangleq (W_{2,T}, T), && U \triangleq U_T, \\
&X \triangleq X_T, && S \triangleq S_T, && Y \triangleq Y_T, \\
&Z \triangleq Z_T, && V \triangleq V_T, && \hat{S} \triangleq \hat{S}_T.
\end{aligned} \tag{24}$$

Then we have:

$$\begin{aligned}
0 &\overset{(a)}{=} \frac{1}{N}\sum_{i=1}^{N} I(U^{i-1}, Z^{i-1}, Y^{i-1}; Y_i | Y_{i+1}^N, C, \Xi, Y_{N+1}^{N'}) \\
&\quad - \frac{1}{N}\sum_{i=1}^{N} I(Y_{i+1}^N; U_i, Z_i, Y_i | U^{i-1}, Z^{i-1}, \\
&\qquad\qquad Y^{i-1}, C, \Xi, Y_{N+1}^{N'}) \qquad (25) \\
&\leq \frac{1}{N}\sum_{i=1}^{N} I(U^{i-1}, Z^{i-1}, Y^{i-1}, C, \Xi; Y_i) + \frac{1}{N} H(Y_{N+1}^{N'}) \\
&\quad - \frac{1}{N}\sum_{i=1}^{N} I(Y_{i+1}^{N'}; U_i, Z_i | U^{i-1}, Z^{i-1}, Y^{i-1}, C, \Xi, Y_i) \qquad (26) \\
&\leq \frac{1}{N}\sum_{i=1}^{N} I(W_{1,i}; Y_i) - \frac{1}{N}\sum_{i=1}^{N} I(W_{2,i}; U_i, Z_i | W_{1,i}, Y_i) \\
&\quad + \frac{N'-N}{N}\log|\mathcal{Y}| \qquad (27) \\
&\leq I(W_1; Y) - I(W_2; U, Z | W_1, Y) + \alpha \log|\mathcal{Y}|. \qquad (28)
\end{aligned}$$

Here, (a) follows from Csiszár's sum identity [16] with $(\Xi, C, Y_{N+1}^{N'})$ included as common side information. Letting $\alpha \to 0$ (achievability needs to hold for all $\alpha > 0$), establishes the information inequality (4) (with K constant) for all blocklengths N.

The proof is then concluded by showing that there exists a subsequence of increasing blocklengths $\{N_i\}$ such that the joint pmf of above random variables $P_{USW_1W_2XYZV\hat{S}}^{(N_i)}$ converges to a limiting distribution of the form

$$\begin{aligned}
\tilde{Q}_{UVSXYZ\hat{S}W_1W_2} =& \bar{P}_U \bar{P}_S Q_{W_1} Q_{X|UW_1} Q_{W_2|UZW_1} \\
&\cdot \bar{P}_{YZ|XS} Q_{V|YW_1W_2} \tilde{Q}_{\hat{S}|XSZYUVW_1W_2}
\end{aligned} \tag{29}$$

with marginals $Q_{S\hat{S}}$ and Q_{UV} and satisfying (4). This holds by standard arguments, like bounding the size of the cardinality of the auxiliary random variables using Carathéory's theorem; invoking the Bolzano-Weierstrass theorem to establish the desired convergent subsequent $\{N_i\}$; exploiting the strong coordination definition; and proving the Markov chains. (Details are provided in Appendix B)

Define now $Q_{\hat{S}|XZS}$ as the conditional marginal of above distribution. We notice that

$$\begin{aligned}
Q_{UVSXYZ\hat{S}W_1W_2} =& \bar{P}_U \bar{P}_S Q_{W_1} Q_{X|UW_1} Q_{W_2|UZW_1} \\
&\cdot \bar{P}_{YZ|XS} Q_{V|YW_1W_2} Q_{\hat{S}|XZS}
\end{aligned} \tag{30}$$

continues to have marginals $Q_{S\hat{S}}$ and Q_{UV} as well as satisfying (4).

The proof is then concluded by noting that $Q_{\hat{S}|XZS} = Q_{\hat{S}|XZ}$ because for each N:

$$P_{S\hat{S}XZ} = P_{\hat{S}|XZ} P_{SXZ} = P_{\hat{S}|XZ} \bar{P}_S \bar{P}_{Z|XS} P_X, \tag{31}$$

where the first equality holds because $P_{S|XZ} = P_{S|XZ\hat{S}}$ by the memoryless law of the channel, which implies that given the input-output pair (X_i, Z_i) the state S_i is conditionally independent of future inputs, outputs, states and $\hat{S}^N$. Further details are given in Appendix B.

Acknowledgment

This work was supported by the ANR grant CLECI.

References

[1] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Transactions on Information Theory*, vol. 56, pp. 4181–4206, Sept. 2010.

[2] M. Le Treust, "Empirical coordination with channel feedback and strictly causal or causal encoding," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, pp. 471–475, June 2015.

[3] F. Haddadpour, M. H. Yassaee, S. Beigi, A. Gohari, and M. R. Aref, "Simulation of a channel with another channel," *IEEE Transactions on Information Theory*, vol. 63, no. 5, pp. 2659–2677, 2016.

[4] G. Cervia, L. Luzzi, M. Le Treust, and M. R. Bloch, "Strong coordination of signals and actions over noisy channels with two-sided state information," *IEEE Transactions on Information Theory*, vol. 66, pp. 4681–4708, Aug. 2020.

[5] S. A. Obead, B. N. Vellambi, and J. Kliewer, "Strong coordination over noisy channels," *IEEE Transactions on Information Theory*, vol. 67, pp. 2716–2738, May 2021.

[6] M. Ahmadipour, M. Kobayashi, M. Wigger, and G. Caire, "An information-theoretic approach to joint sensing and communication," *IEEE Transactions on Information Theory*, vol. 70, pp. 1124–1146, Feb. 2024.

[7] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu, Y. Shen, F. Colone, and K. Chetty, "A survey on fundamental limits of integrated sensing and communication," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 994–1034, 2022.

[8] M. Kobayashi, G. Caire, and G. Kramer, "Joint state sensing and communication: Optimal tradeoff for a memoryless case," in *2018 IEEE International Symposium on Information Theory (ISIT)*, pp. 111–115, 2018.

[9] Y. Cui, F. Liu, C. Masouros, J. Xu, T. X. Han, and Y. C. Eldar, "Integrated sensing and communications: Background and applications," in *Integrated Sensing and Communications*, pp. 3–21, Springer Nature, 2023.

[10] H. Joudeh and F. M. J. Willems, "Joint communication and binary state detection," *IEEE Journal on Selected Areas in Information Theory*, vol. 3, no. 1, pp. 113–124, 2022.

[11] M.-C. Chang, S.-Y. Wang, T. Erdoğan, and M. R. Bloch, "Rate and detection-error exponent tradeoff for joint communication and sensing of fixed channel states," *IEEE Journal on Selected Areas in Information Theory*, vol. 4, pp. 245–259, May 2023.

[12] G. Cervia, L. Luzzi, M. Le Treust, and M. R. Bloch, "Strong coordination over noisy channels with strictly causal encoding," in *Proc. 56th Annual Allerton Conference on Communication, Control, and Computing*, (Monticello, IL, USA), pp. 519–526, Oct. 2018.

[13] P. W. Cuff, *Communication in Networks for Coordinating Behavior*. Ph.d. dissertation, Stanford University, July 2009.

[14] M. H. Yassaee, M. R. Aref, and A. Gohari, "Achievability proof via output statistics of random binning," *IEEE Transactions on Information Theory*, vol. 60, pp. 6760–6786, Nov. 2014.

[15] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, pp. 471–480, July 1973.

[16] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge: Cambridge University Press, second ed., 2011.

[17] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, UK: Cambridge University Press, 2011.

Appendix A
Details for Achievability Proof

A. Proof Step in (12)

Let P_{AB} and Q_{AB} distinct distributions. By the triangle inequality:

$$\begin{aligned}&\|P_{AB} - Q_{AB}\|_{\mathrm{TV}}\\ &\le \|P_{AB} - Q_A P_{B|A}\|_{\mathrm{TV}} + \|Q_A P_{B|A} - Q_{AB}\|_{\mathrm{TV}}.\end{aligned} \tag{32}$$

The first term is upper bounded by

$$\|P_A - Q_A\|_{\mathrm{TV}},$$

since total variation cannot increase under a fixed conditional kernel [13, Lemma 17]. For the second term,

$$\|Q_A P_{B|A} - Q_A Q_{B|A}\|_{\mathrm{TV}} = \mathbb{E}_{Q_A}\Big[\|P_{B|A} - Q_{B|A}\|_{\mathrm{TV}}\Big].$$

Combining the two bounds proves

$$\|P_{AB}-Q_{AB}\|_{\mathrm{TV}} \le \|P_A-Q_A\|_{\mathrm{TV}}+\mathbb{E}_{Q_A}\Big[\|P_{B|A}-Q_{B|A}\|_{\mathrm{TV}}\Big]. \tag{33}$$

B. Detailed Analysis of TV distance

Let $\mathbf{P}^{\mathrm{RC}}$ and $\mathbf{P}^{\mathrm{RB}}$ denote the random pmfs (where the randomness stems from the binning functions $\varphi_{1,b}, \varphi_{2,b}, \varphi_{3,b}$) induced by our binning and coding schemes in Subsections IV-A and IV-B. Recall also the definition of Σ_{b-1} in (15):

$$\Sigma_{b-1}^n = (U_{b-1}^n, W_{1,b-1}^n, X_{b-1}^n, Y_{b-1}^n, Z_{b-1}^n) \tag{34}$$

and define

$$\Lambda_b^n \triangleq (W_{2,b-1}^n, K_{b-1}^n, V_{b-1}^n, U_b^n, W_{1,b}^n, X_b^n, Y_b^n, Z_b^n), \tag{35}$$

$$\widetilde{\Lambda}_b^n \triangleq (\Lambda_b^n, \hat{W}_{1,b}^n, \hat{W}_{2,b-1}^n, \hat{K}_{b-1}^n). \tag{36}$$

1) $\mathbf{P}^{\mathrm{RB}}$ close to i.i.d.: To prove the first point, we define an intermediate distribution $\mathbf{P}^{\mathrm{RB,perf}}$, which is induced by the auxiliary random binning setup under perfect Slepian–Wolf decoding:

$$\begin{aligned}&\mathbf{P}^{\mathrm{RB,perf}}_{F_b,J_b,L_b,\widetilde{\Lambda}_b^n|\Sigma_{b-1}^n}\\ &\quad\triangleq\ \mathbf{P}^{\mathrm{RB}}_{F_b,J_b,L_b|U_{b-1}^n,W_{1,b-1}^n,X_{b-1}^n,Z_{b-1}^n}\\ &\qquad\cdot\mathbf{P}^{\mathrm{RB}}_{W_{1,b}^n,W_{2,b-1}^n,K_{b-1}^n|U_{b-1}^n,W_{1,b-1}^n,X_{b-1}^n,Z_{b-1}^n,F_b,J_b,L_b}\\ &\qquad\cdot\bar{P}_{U_b^n} Q_{X_b^n|U_b^n W_{1,b}^n} \bar{P}_{Y_b^n Z_b^n|X_b^n}\\ &\qquad\cdot\mathbf{1}\big\{(\hat{W}_{1,b}^n, \hat{W}_{2,b-1}^n, \hat{K}_{b-1}^n) = (W_{1,b}^n, W_{2,b-1}^n, K_{b-1}^n)\big\}\\ &\qquad\cdot Q_{V_{b-1}^n|Y_{b-1}^n W_{1,b-1}^n W_{2,b-1}^n}.\end{aligned} \tag{37}$$

Under this distribution, after marginalizing over the bin indices, we have

$$\mathbf{P}^{\mathrm{RB,perf}}_{\Lambda_b^n|\Sigma_{b-1}^n} = Q_{\Lambda|\Sigma}^{\otimes n}, \tag{38}$$

where

$$\begin{aligned}&Q_{\Lambda|\Sigma}\big(w_2, k, v, u', w_1', x', y', z' \mid u, w_1, x, y, z\big)\\ &\quad\triangleq\ Q_{W_2|UZW_1}(w_2|u, z, w_1)\, Q_{K|XZ}(k|x, z)\\ &\qquad\cdot Q_{V|YW_1W_2}(v|y, w_1, w_2)\, \bar{P}_U(u')\, Q_{W_1}(w_1')\\ &\qquad\cdot Q_{X|UW_1}(x'|u', w_1')\, \bar{P}_{YZ|X}(y', z'|x').\end{aligned} \tag{39}$$

Hence, if $\Sigma_{b-1}^n \sim Q_{\Sigma}^{\otimes n}$, where

$$\begin{aligned}Q_\Sigma(u, w_1, x, y, z) \triangleq{}& \bar{P}_U(u)\, Q_{W_1}(w_1)\, Q_{X|UW_1}(x|u, w_1)\\ &\cdot\ \bar{P}_{YZ|X}(y, z|x),\end{aligned} \tag{40}$$

then we have

$$\mathbf{P}^{\mathrm{RB,perf}}_{\Sigma_{b-1}^n\Lambda_b^n} = (Q_\Sigma Q_{\Lambda|\Sigma})^{\otimes n}. \tag{41}$$

Hence, by (41) the first part will be completed if we prove that $\mathbf{P}^{\mathrm{RB,perf}}$ is close to $\mathbf{P}^{\mathrm{RB}}$ in total variation.

This is done by applying [17, Theorem 10.1] for the reconstruction of $(W_{1,b}^n, W_{2,b-1}^n, K_{b-1}^n)$, with side information $(W_{1,b-1}^n, Y_{b-1}^n, Y_b^n)$, and the indices (F_b, J_b, L_b) available to the decoder. In fact, we obtain that if the rates $\tilde{R}_F$, R_C, and $\tilde{R}_K$ of the common randomness F_b, J_b, and L_b satisfy (21), i.e.,

$$\tilde{R}_F + \tilde{R}_J + \tilde{R}_K > H(W_1|Y) + H(W_2, K|W_1, Y), \quad (42)$$

then there exists a vanishing sequence $\delta_n \to 0$ such that

$$\mathbb{E}_{\varphi}\Big[\mathbb{E}_{\mathbf{P}^{\mathrm{RB}}_{\Sigma^n_{b-1}}}\big[\big\|\mathbf{P}^{\mathrm{RB}}_{\tilde{\Lambda}^n_b|\Sigma^n_{b-1}} - \mathbf{P}^{\mathrm{RB,perf}}_{\tilde{\Lambda}^n_b|\Sigma^n_{b-1}}\big\|_{\mathrm{TV}}\big]\Big] \le \delta_n, \quad (43)$$

where here expectation is with respect to the random binning operations $\varphi_{1,b}$, $\varphi_{2,b}$, and $\varphi_{3,b}$, which determine the encoding and decoding rules.

2) $\mathbf{P}^{\mathrm{RC}}$ close to $\mathbf{P}^{\mathrm{RB}}$: Here, we apply [14, Theorem 1] to the bin indices (F_b, J_b, L_b), with side information $(U_{b-1}^n, W_{1,b-1}^n, X_{b-1}^n, Z_{b-1}^n)$. This allows us to conclude that if (22) holds, i.e.,

$$\begin{aligned}\tilde{R}_F + \tilde{R}_J + \tilde{R}_K &< H(W_1) + H(W_2, K|U, W_1, X, Z) \quad (44)\\ &< H(W_1) + H(W_2|UZW_1) + H(K|XZ), \quad (45)\end{aligned}$$

then there exists $\epsilon_n \to 0$ such that

$$\mathbb{E}_{\varphi}\Big[\Big\|\mathbf{P}^{\mathrm{RB}}_{\Sigma^n_{b-1}F_bJ_bL_b} - \mathbf{P}^{\mathrm{RB}}_{\Sigma^n_{b-1}}\, p^U_{F_b} p^U_{J_b} p^U_{L_b}\Big\|_{\mathrm{TV}}\Big] \le \epsilon_n. \quad (46)$$

Since the encoder kernel in $\mathbf{P}^{\mathrm{RC}}$, defined by (20), is equal to the corresponding kernel in $\mathbf{P}^{\mathrm{RB}}$, and also all remaining kernels coincide, [13, Lemma 17] combined with Condition (46) allow to conclude

$$\mathbb{E}_{\varphi}\Big[\mathbb{E}_{\mathbf{P}^{\mathrm{RB}}_{\Sigma^n_{b-1}}}\big[\big\|\mathbf{P}^{\mathrm{RC}}_{\tilde{\Lambda}^n_b|\Sigma^n_{b-1}} - \mathbf{P}^{\mathrm{RB}}_{\tilde{\Lambda}^n_b|\Sigma^n_{b-1}}\big\|_{\mathrm{TV}}\big]\Big] \le \epsilon_n. \quad (47)$$

3) Total Variation of the Block Transitions: Combining (43) and (47), we obtain

$$\begin{aligned}&\mathbb{E}_{\varphi}\Big[\mathbb{E}_{\mathbf{P}^{\mathrm{RB}}_{\Sigma^n_{b-1}}}\big[\big\|\mathbf{P}^{\mathrm{RC}}_{\tilde{\Lambda}^n_b|\Sigma^n_{b-1}} - \mathbf{P}^{\mathrm{RB,perf}}_{\tilde{\Lambda}^n_b|\Sigma^n_{b-1}}\big\|_{\mathrm{TV}}\big]\Big]\\ &\le \epsilon_n + \delta_n. \quad (48)\end{aligned}$$

Marginalization to Λ_b further yields

$$\begin{aligned}&\mathbb{E}_{\varphi}\Big[\mathbb{E}_{\mathbf{P}^{\mathrm{RB}}_{\Sigma^n_{b-1}}}\big[\big\|\mathbf{P}^{\mathrm{RC}}_{\Lambda^n_b|\Sigma^n_{b-1}} - \mathbf{P}^{\mathrm{RB,perf}}_{\Lambda^n_b|\Sigma^n_{b-1}}\big\|_{\mathrm{TV}}\big]\Big]\\ &\le \epsilon_n + \delta_n. \quad (49)\end{aligned}$$

Together with (38), this shows that one block of the random coding scheme simulates the target transition from $(U_{b-1}^n, H_{1,b-1}^n)$ to $(V_{b-1}^n, H_{2,b-1}^n, U_b^n, H_{1,b}^n)$, where

$$H_{1,b}^n = (W_{1,b}^n, X_b^n, Y_b^n, Z_b^n), \qquad H_{2,b}^n = (W_{2,b}^n, K_b^n).$$

Appending S_{b-1}^n and $\hat{S}_{b-1}^n$ through the memoryless kernels induced by $\bar{P}_S\bar{P}_{YZ|XS}$ and $Q_{\hat{S}|XZ}$ then gives (11b).

C. Joint extraction of common randomness and removal of F_b

At the end of block $b = 2, \dots, B+1$, Node 2 reconstructs the sequence K_{b-1}^n previously generated by Node 1. Both terminals extract the block $b+1$ common-randomness L_{b+1} from K_{b-1}^n (or the reconstruction) using an independent random binning

$$\psi_b\colon \mathcal{K}^n \to [1 : 2^{n\tilde{R}_K}],$$

i.e.,

$$L_{b+1} \triangleq \psi_b(K_{b-1}^n).$$

The first two variables L_1 and L_2 cannot be constructed this way, but have to be supplied from the common randomness C. This adds an additional rate of $\frac{2}{B}\tilde{R}_K$ to R_C, which however vanishes as $B \to \infty$. We will simply neglect this edge effect.

Recall $F_b = \varphi_{1,b}(W_{1,b}^n, W_{2,b-1}^n, K_{b-1}^n)$. By [14, Theorem 1] applied to the bin indices F_b and L_{b+1} with side information

$$(U_{b-1}^n, V_{b-1}^n, \Sigma_b^n)$$

we have

$$\mathbb{E}_{\varphi,\psi}\Big[\Big\|\mathbf{P}^{\mathrm{RB,perf}}_{U^n_{b-1}V^n_{b-1}\Sigma^n_bF_bL_{b+1}} - Q^{\otimes n}_{UV}Q^{\otimes n}_{\Sigma}p^U_{F_b}p^U_{L_{b+1}}\Big\|_{\mathrm{TV}}\Big] \to 0, \quad (50)$$

if the following two conditions are satisfied:

$$\tilde{R}_K < H(K|U, V), \quad (51a)$$
$$\tilde{R}_F + \tilde{R}_K < H(W_2, K|U, V). \quad (51b)$$

Similarly, applying [14, Theorem 1] to the same bin-indices and to side-information

$$(S_{b-1}^n, \hat{S}_{b-1}^n, \Sigma_b^n),$$

proves that

$$\mathbb{E}_{\varphi,\psi}\Big[\Big\|\mathbf{P}^{\mathrm{RB,perf}}_{S^n_{b-1}\hat{S}^n_{b-1}\Sigma^n_bF_bL_{b+1}} - Q^{\otimes n}_{S\hat{S}}Q^{\otimes n}_{\Sigma}p^U_{F_b}p^U_{L_{b+1}}\Big\|_{\mathrm{TV}}\Big] \to 0, \quad (52)$$

whenever

$$\tilde{R}_K < H(K|S, \hat{S}), \quad (53a)$$
$$\tilde{R}_F + \tilde{R}_K < H(W_2, K|S, \hat{S}). \quad (53b)$$

Since on expectation the pmfs $\mathbf{P}^{\mathrm{RC}}$ and $\mathbf{P}^{\mathrm{RB,perf}}$ have vanishing TV distance on the marginals $(U_{b-1}^n, V_{b-1}^n, \Sigma_b^n, F_b, L_{b+1})$, and $(S_{b-1}^n, \hat{S}_{b-1}^n, \Sigma_b^n, F_b, L_{b+1})$, the limits in (50) and (52) also hold for $\mathbf{P}^{\mathrm{RC}}$ instead of $\mathbf{P}^{\mathrm{RB,perf}}$. Finally, by applying [14, Lemma 4] to the sum, over all active blocks, of the TV distances in (50), (52), and the corresponding Slepian–Wolf reconstruction-error probabilities, there exists a realization f^B of F^B such that all these summed quantities vanish.

$$\sum_{b=2}^{B+1}\Big\|\mathbf{P}^{\mathrm{RC}}_{U^n_{b-1}V^n_{b-1}\Sigma^n_bL_{b+1}|F^B=f^B} - Q^{\otimes n}_{UV}Q^{\otimes n}_{\Sigma}p^U_{L_{b+1}}\Big\|_{\mathrm{TV}} \to 0, \quad (54)$$

$$\sum_{b=2}^{B+1}\left\|\mathbf{P}^{\mathrm{RC}}_{S^n_{b-1}\hat{S}^n_{b-1}\Sigma^n_b L_{b+1}|F^B=f^B}-Q^{\otimes n}_{S\hat{S}}Q^{\otimes n}_{\Sigma}p^U_{L_{b+1}}\right\|_{\mathrm{TV}}\to 0. \tag{55}$$

Thus, F^B can be fixed. Also, since Node 2 reconstructs K^n_{b-1} with vanishing error, the two terminals agree on the recycled seed $L_{b+1}=\psi_b(K^n_{b-1})$ with vanishing error.

After fixing $F^B=f^B$, the propagation starts from $\Sigma^n_1\sim Q^{\otimes n}_{\Sigma}$ and an independent uniform seed L_2, supplied by the external common randomness. For $b=2,\dots,B$, block b uses the carry (Σ^n_{b-1},L_b), where U^n_{b-1} is contained in Σ^n_{b-1}, to produce V^n_{b-1} for the UV-coordination marginal and $\hat{S}^n_{b-1}$ for the $S\hat{S}$-sensing marginal, and to refresh the carry to (Σ^n_b,L_{b+1}). The bounds (54) and (55) show that this refreshed carry is asymptotically distributed as $Q^{\otimes n}_{\Sigma}p^U_{L_{b+1}}$ and is independent of the corresponding UV- or $S\hat{S}$-block. Hence the same block-propagation and triangle inequality argument as in (11)–(12) applies, now with (Σ^n_b,L_{b+1}) in place of Σ^n_b. The termination block $b=B+1$ is handled in the same way after marginalizing the final carry. Together with the vanishing K-reconstruction error probability, this gives

$$\|P_{U^NV^N}-Q^{\otimes N}_{UV}\|_{\mathrm{TV}}\to 0,\qquad \left\|P_{S^N\hat{S}^N}-Q^{\otimes N}_{S\hat{S}}\right\|_{\mathrm{TV}}\to 0.$$

D. *Proof of the rate condition* (5)

Combining the rate constraints (21), (22), (51), and (53), together with the fact that after fixing F^B and extracting L^B only J^B consumes external common randomness, i.e. $R_C>\tilde{R}_J$, Fourier–Motzkin elimination of $(\tilde{R}_F,\tilde{R}_J,\tilde{R}_K)$ gives (4) and (5).

Appendix B
Details for the Converse Proof

Recall that

$$W_{1,i}=(\Xi,U^{i-1},Z^{i-1},Y^{i-1},C),\qquad W_{2,i}=Y^{N'}_{i+1}.$$

For each $i\in[1:N]$, the following Markov chains hold:

$$W_{1,i}\perp(U_i,S_i), \tag{56}$$
$$X_i-(U_i,W_{1,i})-S_i, \tag{57}$$
$$(Y_i,Z_i)-(X_i,S_i)-(U_i,W_{1,i}), \tag{58}$$
$$V_i-(Y_i,W_{1,i},W_{2,i})-(U_i,Z_i,X_i,S_i), \tag{59}$$
$$W_{2,i}-(U_i,Z_i,W_{1,i})-(X_i,S_i,Y_i). \tag{60}$$

Specifically, (56) holds because (U_i,S_i) is independent of the past variables, the common randomness C, and the encoder private randomness Ξ. The chain (57) holds since X_i is generated from (U^i,Z^{i-1},C,Ξ), which is contained in $(U_i,W_{1,i})$. The chain (58) holds from the memoryless channel law $\bar{P}_{YZ|XS}$. The chain (59) holds because the decoder generates V^N from $(Y^{N'},C)$, which is contained in $(Y_i,W_{1,i},W_{2,i})$.

And (60) holds since

$$\begin{aligned}
&I(W_{2,i};X_i,S_i,Y_i|U_i,Z_i,W_{1,i})\\
&=I(Y^{N'}_{i+1};X_i,S_i,Y_i|U^i,Z^i,Y^{i-1},C,\Xi)\\
&\le I(Y^{N'}_{i+1},Z^{N'}_{i+1};X_i,S_i,Y_i|U^i,Z^i,Y^{i-1},C,\Xi)\\
&=\sum_{t=i+1}^{N'}I(Y_t,Z_t;X_i,S_i,Y_i|U^i,Z^i,Y^{i-1},C,\Xi,Y^{t-1}_{i+1},Z^{t-1}_{i+1})\\
&=0.
\end{aligned} \tag{61}$$

It remains to handle the estimator. The estimator is allowed to use $(U^N,C,X^{N'},Z^{N'})$.

By the support lemma [17, Appendix C], we may restrict the alphabets of W_1 and W_2 to finite sets of bounded size. Hence the time-sharing pmfs are defined over a compact simplex. Moreover, the induced kernels $P^{(N)}_{\hat{S}|XZ}$ defined below are also defined over a compact simplex. Therefore, by the Bolzano–Weierstrass theorem, there exists a subsequence of increasing blocklengths $\{N_i\}$ such that $P^{(N_i)}_{USW_1W_2XYZV\hat{S}}$ converges to a limiting distribution $\tilde{Q}_{USW_1W_2XYZV\hat{S}}$, and the corresponding induced kernels $P^{(N_i)}_{\hat{S}|XZ}$ converge to some kernel $Q_{\hat{S}|XZ}$.

From the Markov chains proved above, the limiting distribution has the form

$$\begin{aligned}
\tilde{Q}_{UVSXYZ\hat{S}W_1W_2}=&\bar{P}_U\bar{P}_SQ_{W_1}Q_{X|UW_1}Q_{W_2|UZW_1}\\
&\cdot\bar{P}_{YZ|XS}Q_{V|YW_1W_2}\tilde{Q}_{\hat{S}|XSZYUVW_1W_2},
\end{aligned} \tag{62}$$

with marginals $Q_{S\hat{S}}$ and Q_{UV}, and satisfying (4) with K constant.

Define now $Q_{\hat{S}|XZS}$ as the conditional marginal of $\hat{S}$ given (X,Z,S) under $\tilde{Q}$. Then

$$\begin{aligned}
Q_{UVSXYZ\hat{S}W_1W_2}=&\bar{P}_U\bar{P}_SQ_{W_1}Q_{X|UW_1}Q_{W_2|UZW_1}\\
&\cdot\bar{P}_{YZ|XS}Q_{V|YW_1W_2}Q_{\hat{S}|XZS}
\end{aligned} \tag{63}$$

continues to have marginals $Q_{S\hat{S}}$ and Q_{UV}, and also satisfies (4).

It remains to note that $Q_{\hat{S}|XZS}=Q_{\hat{S}|XZ}$. For each blocklength N and time $i\in[1:N]$,

$$(U^N,C,X^{N'}_{\sim i},Z^{N'}_{\sim i},\hat{S}_i)-(X_i,Z_i)-S_i,$$

because, given (X_i,Z_i), the state S_i is conditionally independent of the source sequence, the common randomness, and the remaining channel inputs and feedback symbols; moreover, $\hat{S}_i$ is generated from these variables. Hence

$$P^{(N)}_{S_i\hat{S}_iX_iZ_i}=P^{(N)}_{\hat{S}_i|X_iZ_i}P^{(N)}_{S_iX_iZ_i}=P^{(N)}_{\hat{S}_i|X_iZ_i}P^{(N)}_{X_i}\bar{P}_S\bar{P}_{Z|XS}. \tag{64}$$

Averaging over i, and multiplying and dividing by $\frac{1}{N}\sum_{i=1}^{N}P^{(N)}_{X_iZ_i}(x,z)$, define

$$P^{(N)}_{\hat{S}|XZ}(\hat{s}|x,z)\triangleq\frac{\frac{1}{N}\sum_{i=1}^{N}P^{(N)}_{\hat{S}_i|X_iZ_i}(\hat{s}|x,z)P^{(N)}_{X_iZ_i}(x,z)}{\frac{1}{N}\sum_{i=1}^{N}P^{(N)}_{X_iZ_i}(x,z)}, \tag{65}$$

with arbitrary definition when the denominator is zero. Then

$$P^{(N)}_{S_T\hat{S}_TX_TZ_T}=P^{(N)}_X\bar{P}_S\bar{P}_{Z|XS}P^{(N)}_{\hat{S}|XZ}, \tag{66}$$

where $P_X^{(N)}(x) = \frac{1}{N}\sum_{i=1}^N P_{X_i}^{(N)}(x)$. Taking the limit along the subsequence $\{N_i\}$ gives

$$\tilde{Q}_{S\hat{S}XZ} = \tilde{Q}_X \bar{P}_S \bar{P}_{Z|XS} Q_{\hat{S}|XZ} = \tilde{Q}_{SXZ} Q_{\hat{S}|XZ}. \tag{67}$$

Therefore, on the support of $\tilde{Q}_{SXZ}$,

$$Q_{\hat{S}|XZS} = Q_{\hat{S}|XZ},$$

with arbitrary definition on zero-probability points. Substituting this identity into (63) yields a pmf of the form

$$\begin{aligned} Q_{UVSXYZ\hat{S}W_1W_2} =& \bar{P}_U \bar{P}_S Q_{W_1} Q_{X|UW_1} Q_{W_2|UZW_1} \\ & \cdot \bar{P}_{YZ|XS} Q_{V|YW_1W_2} Q_{\hat{S}|XZ}, \end{aligned} \tag{68}$$

whose (U,V)- and $(S,\hat{S})$-marginals are Q_{UV} and $Q_{S\hat{S}}$, respectively.