# State Leakage and Coordination of Actions: Core of the Receiver's Knowledge

Mael Le Treust, Matthieu R Bloch

# State Leakage and Coordination of Actions: Core of the Receiver's Knowledge

Maël Le Treust* and Matthieu Bloch†

* ETIS UMR 8051, Université Paris Seine, Université Cergy-Pontoise, ENSEA, CNRS,
6, avenue du Ponceau, 95014 Cergy-Pontoise CEDEX, FRANCE
Email: mael.le-treust@ensea.fr

† School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, Georgia 30332
Email: matthieu.bloch@ece.gatech.edu



**Abstract**

We revisit the problems of state masking and state amplification through the lens of empirical coordination by considering a state-dependent channel in which the encoder has causal and strictly causal state knowledge. We show that the problem of empirical coordination provides a natural framework in which to jointly study the problems of reliable communication, state masking, and state amplification. We characterize the regions of rate-equivocation-coordination trade-offs for several channel models with causal and strictly causal state knowledge. We introduce the notion of "core of the receiver's knowledge" to capture what the decoder can infer about all the signals involved in the model. We exploit this result to solve a channel state estimation zero-sum game in which the encoder prevents the decoder to estimate the channel state accurately.

**Index Terms**

* Maël Le Treust acknowledges financial support of INS2I CNRS for projects JCJC CoReDe 2015, PEPS StrategicCoo 2016; DIM-RFSI under grant EX032965 and The Paris Seine Initiative 2018. This research has been conducted as part of the Labex MME-DII (ANR11-LBX-0023-01).

† Matthieu Bloch acknowledges financial support of National Science Foundation under award CCF 1320304.

This work was presented in part at the IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, July 2016 [1]. The authors gratefully acknowledge the financial support of SRV ENSEA for visits at Georgia Tech Atlanta in 2014 and at ETIS Cergy in 2017.

Shannon theory, state-dependent channel, state leakage, empirical coordination, state masking, state amplification, causal encoding, two-sided state information, noisy channel feedback.

## I. Introduction

The study of state-dependent channels can be traced back to the early works of Shannon [2], Gelf'and and Pinsker [3], which identified optimal coding strategies to transmit reliably in the presence of a state known at the encoder causally or non-causally, respectively. The insights derived from the models have since proved central to the study of diverse topics including wireless communications [4], information-hiding and watermarking [5], and information transmission in repeated games [6]. The present work relates to the latter application and studies state-dependent channels with causal state knowledge from the perspective of *empirical coordination* [7].

Previous studies that have explored the problem of not only decoding messages at the receiver but also estimating the channel state are particularly relevant to the present work. The *state masking* formulation of the problem [8] aims at characterizing the trade-off between the rate of reliable communication and the minimal leakage about the channel state. The rate-leakage capacity region of state masking has been successfully characterized for both causal and non-causal state knowledge. The *state amplification* formulation [9], in which the state is conveyed to the receiver instead of being masked, aims at characterizing the trade-off between the rate of reliable communication and the reduction of uncertainty about the state. The rate-uncertainty reduction capacity region of state amplification has also been successfully characterized for causal and non-causal state knowledge. The state amplification formulation was subsequently extended in the causal case by replacing the reduction of uncertainty about the state by an average distortion function [10] (this model was dubbed causal *state communication*). The rate-distortion capacity region of state communication has been successfully characterized for causal and strictly causal state knowledge, and has been characterized for noiseless and noisy non-causal state knowledge in the case of Gaussian channels with a quadratic distortion [11], [12]. Both formulations have been combined in [13] to study the trade-off between amplification and leakage rates in a channel with two receivers having opposite objectives. The amplification-leakage capacity region has been investigated for non-causal state knowledge, via generally non-matching inner and outer bounds. As a perhaps more concrete example, [14] has studied the trade-off between amplification and leakage in the context of an energy harvesting scenario.

We revisit in this paper the problems of state masking and state amplification with causal and strictly causal state knowledge through the lens of *empirical coordination* [7], [15]. Empirical coordination refers to the control of the joint histograms of the various sequences such as states, codewords, that

appear in channel models, and is related to the coordination of autonomous decision makers in game theory [6]. Specifically, the study of empirical coordination over state-dependent channels is a proxy for characterizing the utility of autonomous decision makers playing a repeated game in the presence of a environment variable (the state), random [6], [16] or adversarial [17], [18], [19], and of an observation structure (the channel) describing how agents observe each other's actions. The characterization of the empirical coordination capacity requires the design of coding schemes in which the actions of the decision makers are sequences that embed coordination information. The empirical coordination capacity has been studied for state-dependent channels under different constraints including strictly causal and causal encoding [20], for perfect channel [21], for strictly causal and causal decoding [22], with source feedforward [23], for lossless decoding [24], with secrecy constraint [25], with two-sided state information [26] and with channel feedback [27]. Interestingly, empirical coordination is a powerful tool also for controlling the posterior belief of the decoder, e.g. in the problems of "Bayesian persuasion" [28] and "strategic communication" [29], [30].

The main contribution of the present work is to show that empirical coordination provides a natural framework in which to jointly study the problems of reliable communication, state masking, and state amplification. In particular, we obtain the following.

- We introduce and characterize the notion of *core of the receiver's knowledge*, which captures what the decoder can exactly know about the other variables in the system. For instance, this allows us to characterize the rate-leakage-coordination region for the causal state-dependent channel (Theorem II.3). Our definition of leakage refines previous work by exactly characterizing the leakage rate instead of only providing a single-sided bound. When specialized, our result (Theorem II.6) simultaneously recovers the constraints already established both in [8, Section V] and [9, Theorem 2].
- We revisit the problem of causal state communication and characterize the normalized Kullback-Leibler (KL)-divergence between the decoder's posterior beliefs and a target belief induced by coordination (Theorem III.1). This allows us to characterize the rate-distortion trade-off for a zero-sum game, in which the decoder attempts to estimate the state while the encoder tries to mask it (Theorem III.3).
- We extend the results to other models, including two-sided state information (Theorem IV.3), noisy feedback (Theorem IV.5), and strictly causal encoding (Theorem V.2).

The rest of the paper is organized as follows. In Section II, we formally introduce the model, along with necessary definitions and notations, and we state our main results. In Section III, we investigate the

channel state estimation problem by introducing the KL-divergence and the decoder's posterior beliefs. In Section IV and Section V, we present some extensions of our results to different scenarios. The proofs are stated in Appendices A-J.

## II. Problem formulation and main result

### A. Notation

Throughout the paper, capital letters, e.g., $S$, denote random variables while lowercase letters, e.g., $s$. denote their realizations and calligraphic fonts, e.g., $\mathcal{S}$, denote the alphabets in which the realizations take values. All alphabets considered in the paper are assumed finite, i.e., $|\mathcal{S}| < \infty$. Sequences of random variables and realizations are denoted by $S^n = (S_1, \ldots, S_n)$ and $s^n = (s_1, \ldots, s_n)$, respectively. We denote the set of probability distributions over $\mathcal{S}$ by $\Delta(\mathcal{S})$. For a probability distribution $\mathcal{Q}_S \in \Delta(\mathcal{S})$, we drop the subscript and simply write $\mathcal{Q}(s)$ in place of $\mathcal{Q}_S(s)$ for the probability mass assigned to realization $s \in \mathcal{S}$. For two distributions $\mathcal{Q}_X, \mathcal{P}_X \in \Delta(\mathcal{X})$, $||\mathcal{Q}_X - \mathcal{P}_X||_1 = \sum_{x\in\mathcal{X}} |\mathcal{Q}(x) - \mathcal{P}(x)|$ stands for the $\ell_1$-distance between the vectors of probability distributions, see also [31, pp. 370] and in [32, pp. 44]. The notation $Y \multimap X \multimap W$ denotes the Markov chain property corresponding to $\mathcal{P}_{Y|XW} = \mathcal{P}_{Y|X}$. The notation $\mathbb{1}(v = s)$ denotes the indicator function, that is equal to 1 if $v = s$ and 0 otherwise.

For a sequence $s^n \in \mathcal{S}^n$, $\mathsf{N}(s|s^n)$ denotes the occurrence number of symbol $s \in \mathcal{S}$ in the sequence $s^n$. The empirical distribution $Q_S^n \in \Delta(\mathcal{S})$ of sequence $s^n \in \mathcal{S}^n$ is then defined as

$$\forall s \in \mathcal{S} \quad Q^n(s) = \frac{\mathsf{N}(s|s^n)}{n}. \tag{1}$$

Given $\delta > 0$ and a joint distribution $\mathcal{Q}_{SX} \in \Delta(\mathcal{S}\times\mathcal{X})$, $T_\delta(\mathcal{Q}_{SX})$ stands for the set of sequences $(s^n, x^n)$ that are jointly typical with tolerance $\delta > 0$ with respect to the probability distribution $\mathcal{Q}_{SX}$, i.e., such that

$$\Big|\Big|Q_{SX}^n - \mathcal{Q}_{SX}\Big|\Big|_1 = \sum_{s,x} \Big|Q^n(s,x) - \mathcal{Q}(s,x)\Big| \leq \delta. \tag{2}$$

### B. System model

The problem under investigation is illustrated in Figure 1. A uniformly distributed message represented by the random variable $M \in \mathcal{M}$ is to be transmitted over a state dependent memoryless channel characterized by the conditional probability distribution $\mathcal{T}_{Y|XS}$ and a channel state $S^n \in \mathcal{S}^n$ drawn according to the i.i.d. probability distribution $\mathcal{P}_S$. For $n \in \mathbb{N}^*$, the message $M$ and the state sequence $S^n$ are encoded into a codeword $X^n \in \mathcal{X}^n$ using an encoder $\mathcal{E}$, subject to constraints to be precised later. Upon observing the output $Y^n \in \mathcal{Y}^n$ of the noisy channel, the receiver uses a decoder $\mathcal{D}$ to form an estimate $\hat{M} \in \mathcal{M}$ of $M$ and to generate actions $V^n \in \mathcal{V}^n$, whose exact role will be precised shortly.

For now, $V^n$ can be thought of as an estimate of the state sequence $S^n$ but more generally captures the ability of the receiver to coordinate with the transmitter. Both $\mathcal{T}_{Y|XS}$ and $\mathcal{P}_S$ are assumed known to all parties.

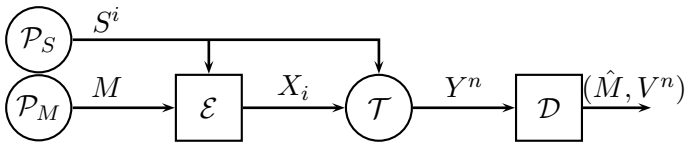


Fig. 1. Memoryless channel $\mathcal{T}_{Y|XS}$ with i.i.d. state drawn according to $\mathcal{P}_S$. The encoding function is causal $f_i : \mathcal{M} \times \mathcal{S}^i \to \mathcal{X}$, for all $i \in \{1, \ldots, n\}$ and the decoding functions $g : \mathcal{Y}^n \to \mathcal{M}$ and $h : \mathcal{Y}^n \to \Delta(\mathcal{V}^n)$ are non-causal.

We are specifically interested in causal encoders formally defined as follows.

**Definition II.1** *A code with causal encoding $c \in \mathcal{C}(n, \mathcal{M})$ is a tuple of functions $c = (\{f_i\}_{i\in\{1,\ldots,n\}}, g, h)$ defined by:*

$$f_i : \mathcal{M} \times \mathcal{S}^i \longrightarrow \Delta(\mathcal{X}), \qquad \forall i \in \{1, \ldots, n\}, \tag{3}$$

$$g : \mathcal{Y}^n \longrightarrow \mathcal{M}, \tag{4}$$

$$h : \mathcal{Y}^n \longrightarrow \Delta(\mathcal{V}^n). \tag{5}$$

*The functions $\{f_i\}_{i\in\{1,\ldots,n\}}$ and $h$ are stochastic and the function $g$ is deterministic.*

The code with causal encoder $c \in \mathcal{C}(n, \mathcal{M})$, the uniform probability distributions of the messages $\mathcal{P}_M$, the source $\mathcal{P}_S$ and the channel $\mathcal{T}_{Y|XS}$ induce a general probability distribution on $(M, S^n, X^n, Y^n, V^n, \hat{M})$:

$$\mathcal{P}_M \prod_{i=1}^{n} \Big[\mathcal{P}_{S_i} f_{X_i|S^iM} \mathcal{T}_{Y_i|X_iS_i}\Big] h_{V^n|Y^n} \mathbb{1}\big(\hat{M} = g(Y^n)\big). \tag{6}$$

Since the sequences $(S^n, X^n, Y^n, V^n)$ are random, the empirical distribution $Q^n_{SXYV}$ is also a random variable.

The performance of codes is measured as follows.

**Definition II.2** *Fix a target rate $\mathsf{R} \geq 0$, a target state leakage $\mathsf{E} \geq 0$ and a target probability distribution $\mathcal{Q}_{SXYV}$. The triple $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable if for all $\varepsilon > 0$, there exists $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$,*

*there exists a code with causal encoding* $c \in \mathcal{C}(n, \mathcal{M})$ *that satisfies:*

$$\begin{aligned}
\frac{\log_2 |\mathcal{M}|}{n} \geq & \mathsf{R} - \varepsilon, \\
\left| \mathcal{L}_{\mathsf{e}}(c) - \mathsf{E} \right| \leq & \varepsilon, \qquad \textit{with} \qquad \mathcal{L}_{\mathsf{e}}(c) = \frac{1}{n} \cdot I(S^n; Y^n), \\
\mathcal{P}_{\mathsf{e}}(c) = & \mathbb{P}\left( M \neq \hat{M} \right) + \mathbb{P}\left( \left| \left| Q^n_{SXYV} - \mathcal{Q}_{SXYV} \right| \right|_1 \geq \varepsilon \right) \leq \varepsilon.
\end{aligned}$$

*We denote by* $\mathcal{A}$ *the set of achievable triples* $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$.

In layman's term, performance is captured along three metrics: i) the rate at which the message $M$ can be reliably transmitted; ii) the information leakage rate about the state sequence $S^n$ at the receiver; and iii) the ability of the encoder to coordinate with the receiver, captured by the empirical coordination with respect to $Q_{SXYV}$.

### *C. Main result*

**Theorem II.3** *Consider a target joint probability distribution* $\mathcal{Q}_{SXYV}$ *that decomposes as* $\mathcal{Q}_{SXYV} = \mathcal{P}_S \mathcal{Q}_{X|S} \mathcal{T}_{Y|XS} \mathcal{Q}_{V|SXY}$. *The triple* $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ *is achievable* if and only if *there exist two auxiliary random variables* $(W_1, W_2)$ *with probability distribution* $\mathcal{Q}_{SW_1W_2XYV} \in \mathbb{Q}_{\mathsf{e}}$ *that satisfy:*

$$I(S; W_1, W_2, Y) \leq \mathsf{E} \leq H(S), \tag{7}$$

$$\mathsf{R} + \mathsf{E} \leq I(W_1, S; Y), \tag{8}$$

*where* $\mathbb{Q}_{\mathsf{e}}$ *is the set of joint distributions* $\mathcal{Q}_{SW_1W_2XYV}$ *with marginal* $\mathcal{Q}_{SXYV}$ *that decompose as*

$$\mathcal{P}_S \mathcal{Q}_{W_1} \mathcal{Q}_{W_2|SW_1} \mathcal{Q}_{X|SW_1} \mathcal{T}_{Y|XS} \mathcal{Q}_{V|YW_1W_2}, \tag{9}$$

*and the supports satisfy* $\max(|\mathcal{W}_1|, |\mathcal{W}_2|) \leq |\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}| + 1$.

Theorem II.3 characterizes the optimal trade-off between reliable transmission, state leakage and empirical coordination. The achievability and converse proofs are provided in Appendices A and B, respectively, while the cardinality bounds are established in Appendix F.

**Remark II.4** *Equation* (8) *and the first inequality of* (7) *imply the information constraints of [10, Theorem 3] for causal state communication and of [20, Theorem 2] for empirical coordination.*

$$\mathsf{R} \leq I(W_1, W_2; Y) - I(W_2; S|W_1). \tag{10}$$

*Indeed, both Markov chains* $X \multimap (S, W_1) \multimap W_2$ *and* $Y \multimap (X, S) \multimap (W_1, W_2)$ *imply* $Y \multimap (W_1, S) \multimap W_2$.

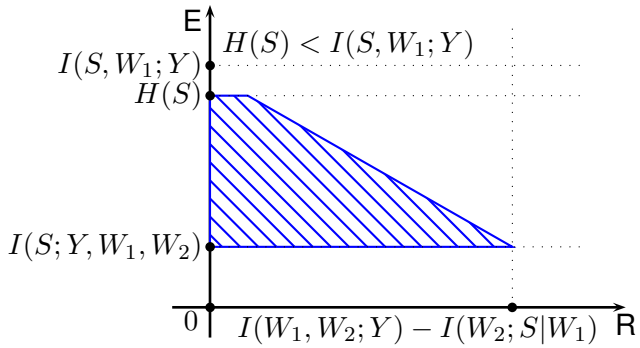


Fig. 2. Region of achievable $(\mathsf{R}, \mathsf{E}) \in \mathcal{A}$ for fixed probability distribution $\mathcal{Q}_{SW_1W_2XYV}$, when $H(S) < I(S, W_1; Y)$.

Theorem II.3 has several important consequences. First, the coordination of both encoder and decoder's actions according to $\mathcal{P}_S \mathcal{Q}_{X|S} \mathcal{T}_{Y|XS} \mathcal{Q}_{V|SXY}$ is compatible with the reliable transmission of additional information at rate $\mathsf{R} \geq 0$. Second, the case of equality in the right-hand-side inequality of (7) corresponds to the full revelation of the channel state $S$ to the decoder. Third, for any achievable pair of rate-distribution $(\mathsf{R}, \mathcal{Q}) \in \mathcal{A}$, the minimal state leakage $\mathsf{E}^\star(\mathsf{R}, \mathcal{Q})$ is given by the first inequality of (7):

$$\mathsf{E}^\star(\mathsf{R}, \mathcal{Q}) = \min_{\substack{\mathcal{Q}_{SW_1W_2XYV} \in \mathbb{Q}_{\mathsf{e}}, \\ \text{s.t. } \mathsf{R} \leq I(W_1, W_2; Y) - I(W_2; S|W_1)}} I(S; W_1, W_2, Y). \tag{11}$$

The reliable transmission of information requires the decoder to know the encoding function, from which it can infer the channel's state parameter $S$. Equation (11) shows that the minimal leakage of state information $I(S^n; Y^n)$ is close to $n \cdot I(S; W_1, W_2, Y)$, as if the sequences $(S^n, Y^n, W_1^n, W_2^n)$ were generated according to the i.i.d. probability distribution $\mathcal{Q}_{SW_1W_2Y}$. In Section III, we investigate the relationship between the state leakage $\mathcal{L}_{\mathsf{e}}(c)$ and the decoder's posterior belief $\mathcal{P}_{S^n|Y^n}$ induced by the coding process.

### *D. Special case without receiver's actions*

We now assume that the decoder does not return an action $V$ coordinated with the other symbols $(S, X, Y)$ to compare our setting with the problems of "state masking" [8, Section V] and "state amplification" [9, Section IV]. Note that these earlier works involve slightly different notions of achievable leakage. In [8], the state leakage is upper bounded by $\mathcal{L}_{\mathsf{e}}(c) = \frac{1}{n} \cdot I(S^n; Y^n) \leq \mathsf{E} + \varepsilon$. In [9], the decoder forms a list $L_n(Y^n) \subseteq \mathcal{S}^n$ with cardinality $\log_2 |L_n(Y^n)| = H(S) - \mathsf{E}$ such that the list decoding error is small $\mathcal{P}(S^n \notin L_n(Y^n)) \leq \varepsilon$, so as to reduce the uncertainty about the state. Here, we require the leakage $\mathcal{L}_{\mathsf{e}}(c) = \frac{1}{n} \cdot I(S^n; Y^n)$ induced by the code to be controlled more precisely as $\left| \mathcal{L}_{\mathsf{e}}(c) - \mathsf{E} \right| \leq \varepsilon$. Nevertheless, we shall see that our definition allows us to obtain the results of [8], [9] as extreme cases.

**Definition II.5** *A code without receiver's actions $c \in \mathcal{C}_{\mathsf{d}}(n, \mathcal{M})$ is a tuple of functions $c = (\{f_i\}_{i\in\{1,\dots,n\}}, g)$ defined by:*

$$f_i : \mathcal{M} \times \mathcal{S}^i \longrightarrow \Delta(\mathcal{X}), \qquad \forall i \in \{1, \dots, n\}, \tag{12}$$

$$g : \mathcal{Y}^n \longrightarrow \mathcal{M}. \tag{13}$$

Compared to Definition II.1, the decoding function $h : \mathcal{Y}^n \longrightarrow \Delta(\mathcal{V}^n)$ in (5) has been removed. The target probability distribution is also restricted to $\mathcal{Q} \in \Delta(\mathcal{S} \times \mathcal{X} \times \mathcal{Y})$, i.e. without receiver's actions. We denote by $\mathcal{A}_{\mathsf{d}}$ the set of achievable triples $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ defined similarly to Definition II.2.

**Theorem II.6** *Consider a target joint probability distribution $\mathcal{Q}_{SXY}$ that decomposes as $\mathcal{Q}_{SXY} = \mathcal{P}_S\mathcal{Q}_{X|S}\mathcal{T}_{Y|XS}$. The triple $(\mathsf{R}, \mathsf{E}, \mathcal{Q}) \in \mathcal{A}_{\mathsf{d}}$ is achievable* if and only if *there exists an auxiliary random variable $W_1$ with probability distribution $\mathcal{Q}_{SW_1XY} \in \mathbb{Q}_{\mathsf{d}}$ that satisfies:*

$$I(S; W_1, Y) \leq \mathsf{E} \leq H(S), \tag{14}$$

$$\mathsf{R} + \mathsf{E} \leq I(W_1, S; Y), \tag{15}$$

*where $\mathbb{Q}_{\mathsf{d}}$ is the set of joint distributions $\mathcal{Q}_{SW_1XY}$ with marginal $\mathcal{Q}_{SXY}$, that decompose as*

$$\mathcal{P}_S\mathcal{Q}_{W_1}\mathcal{Q}_{X|SW_1}\mathcal{T}_{Y|XS}, \tag{16}$$

*and the support satisfies $|\mathcal{W}_1| \leq |\mathcal{S} \times \mathcal{Y}| + 1$.*

The achievability results comes from Theorem II.3 by removing auxiliary random variable $W_2 = \emptyset$ and considering a single block coding instead of block-Markov coding. The converse proof is provided in Appendix G.

**Remark II.7** *The equation* (15)*, the first inequality of* (14) *and the independence between $S$ and $W_1$ imply:*

$$\mathsf{R} \leq I(W_1, S; Y) - \mathsf{E} \leq I(W_1, S; Y) - I(S; W_1, Y) = I(W_1; Y). \tag{17}$$

*Equation* (17) *and the first inequality in* (14) *corresponds to the information constraint stated in [8, pp. 2260], whereas equations* (17)*,* (15) *and second inequality of* (14) *corresponds to the region $\mathcal{R}_0$ stated in [9, Lemma 3]. Formally, the region characterized by Theorem II.6 is the intersection of the regions identified in [8, pp. 2260] and [9, Lemma 3].*

**Remark II.8** *A technical subtlety of the proof is that the requirement of exact leakage $\left|\mathcal{L}_{\mathsf{e}}(c) - \mathsf{E}\right| \leq \varepsilon$ prevents us from replacing the term $I(W_1, S; Y)$ by $I(X, S; Y)$, as in [9, proof of Lemma 2].*

## III. Channel state estimation via distortion function

### A. Decoder's posterior belief

After observing the sequence $Y^n$ of channel outputs, the decoder has posterior belief

$$\mathcal{P}(s^n|y^n) = \frac{\sum_{m,x^n} \mathcal{P}(m, s^n, x^n, y^n)}{\sum_{m,s'^n,x^n} \mathcal{P}(m, s'^n, x^n, y^n)}, \qquad \forall (s^n, y^n). \tag{18}$$

where $\mathcal{P}_{MS^nX^nY^n}$ is the probability distribution induced by the encoding function, defined in (6). The following Theorem upper bounds the KL-divergence between the decoder's posterior belief $\mathcal{P}_{S^n|Y^n}$ and the target conditional distribution $\mathcal{Q}_{S|YW_1W_2}$ for any encoding function, i.e. for any conditional probability distribution $\mathcal{P}_{W_1^nW_2^nX^n|S^n}$.

**Theorem III.1 (Channel state estimation)** *Assume that the probability distribution $\mathcal{Q}_{SW_1W_2XY}$ has full support. For any conditional probability distribution $\mathcal{P}_{W_1^nW_2^nX^n|S^n}$, we have:*

$$\frac{1}{n} \cdot D\bigg(\mathcal{P}_{S^n|Y^n} \bigg|\bigg| \prod_{i=1}^{n} \mathcal{Q}_{S_i|Y_iW_{1,i}W_{2,i}}\bigg) \leq I(S; W_1, W_2, Y) - \mathcal{L}_{\mathsf{e}}(c) + \alpha_1\delta + \alpha_2 \mathbb{P}\Big((S^n, W_1^n, W_2^n, Y^n) \notin T_\delta(\mathcal{Q})\Big). \tag{19}$$

*The parameter $\delta$ is the tolerance of the set of typical sequences $T_\delta(\mathcal{Q})$ and the constants $\alpha_1 = \sum_{\substack{s,w_1,\\ w_2,y}} \log_2 \frac{1}{\mathcal{Q}(s|w_1,w_2,y)}$ and $\alpha_2 = \log_2 \frac{1}{\min_{s,y,w_1,w_2} \mathcal{Q}(s|y,w_1,w_2)}$ are strictly positive.*

The proof of Theorem III.1 is given in Appendix C. Consider a target leakage $\mathsf{E} = I(S; W_1, W_2, Y)$ and a pair $(\mathsf{R}, \mathcal{Q}_{SXYV})$, and assume there exists a probability distribution $\mathcal{Q}_{SW_1W_2XYV} \in \mathbb{Q}_{\mathsf{e}}$ with full support, satisfying (7) and (8). By Theorem II.3, for all $\varepsilon > 0$ and all $\delta > 0$, there exists a $\bar{n}$ such that for all $n \geq \bar{n}$ there exists a code $c \in \mathcal{C}(n, \mathcal{M})$ with two auxiliary sequences $(W_1^n, W_2^n)$, such that:

$$\Big|\mathcal{L}_{\mathsf{e}}(c) - I(S; W_1, W_2, Y)\Big| \leq \varepsilon \quad \text{and} \quad \mathbb{P}\Big((s^n, w_1^n, w_2^n, y^n) \notin T_\delta(\mathcal{Q})\Big) \leq \varepsilon. \tag{20}$$

Hence, by Theorem III.1 we have

$$\frac{1}{n} \cdot D\bigg(\mathcal{P}_{S^n|Y^n} \bigg|\bigg| \prod_{i=1}^{n} \mathcal{Q}_{S_i|Y_iW_{1,i}W_{2,i}}\bigg) \leq \varepsilon + \alpha_1\delta + \alpha_2\varepsilon, \tag{21}$$

where $\epsilon$ and $\delta$ may go to zero when $n$ goes to infinity. The control of the leakage $\mathcal{L}_{\mathsf{e}}(c)$ and the joint typicality of the sequences $(S^n, W_1^n, W_2^n, Y^n) \in T_\delta(\mathcal{Q})$ implies that the decoder's posterior belief $\mathcal{P}_{S^n|Y^n}$ is closely related to the single-letter distribution $\mathcal{Q}_{S|YW_1W_2}$. Based on the triple of symbols $(Y, W_1, W_2)$, the decoder generates action $V$ using the conditional probability distribution $\mathcal{Q}_{V|YW_1W_2}$ and infers the channel state $S$ with the conditional probability distribution $\mathcal{Q}_{S|YW_1W_2}$. We claim that the random variables $(Y, W_1, W_2)$ capture the "*core of the receiver's knowledge*," regarding other random variables $S$ and $V$.

*B. Channel state estimation zero-sum game*

In this section, we introduce the channel state estimation zero-sum game in which the encoder and decoder are opponent players choosing their own encoding and decoding strategies. The encoder seeks to prevent the decoder to return a good estimate $v \in \mathcal{V}$ of the channel state $s \in \mathcal{S}$, evaluated with respect to a distortion function $d(s,v)$. While both players cooperate in transmitting reliably at rate $\mathsf{R}$, the goal of the decoder is to minimize the expected long-run distortion whereas the goal of the encoder is to maximize it.

**Definition III.2** *A target rate $\boldsymbol{R} \geq 0$ and a target distortion $\boldsymbol{D} \geq 0$ are achievable if for all $\varepsilon > 0$, there exists a $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$, there exists a code without receiver's output $c = (\{f_i\}_{i\in\{1,\dots,n\}}, g) \in \mathcal{C}_{\mathsf{d}}(n, \mathcal{M})$ that satisfies:*

$$\frac{\log_2 |\mathcal{M}|}{n} \geq \boldsymbol{R} - \varepsilon, \tag{22}$$

$$\mathcal{P}_{\boldsymbol{e}}(c) = \mathbb{P}\bigg(M \neq \hat{M}\bigg) \leq \varepsilon, \tag{23}$$

$$\bigg| \min_{h_{V^n|Y^n}} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\Big[d(S_i, V_i)\Big] - \boldsymbol{D} \bigg| \leq \varepsilon. \tag{24}$$

*We denote by $\mathcal{A}_{\mathsf{g}}$ the set of achievable pairs $(\boldsymbol{R}, \boldsymbol{D}) \in \mathcal{A}_{\mathsf{g}}$.*

In order to illustration the above definition, we discuss the special case of zero rate $\mathsf{R} = 0$, in which the encoding functions writes $f_{X_i|S^i}$ instead of $f_{X_i|S^iM}$. The channel state estimation zero-sum game reformulates as a maximin

$$\max_{\{f_{X_i|S^i}\}_{i\in\{1,\dots,n\}}} \min_{h_{V^n|Y^n}} \mathbb{E}\bigg[\frac{1}{n}\sum_{i=1}^{n} d(S_i, V_i)\bigg], \tag{25}$$

in which the encoder chooses $\{f_{X_i|S^i}\}_{i\in\{1,\dots,n\}}$ and the decoder choose $h_{V^n|Y^n}$. The following Theorem claims that the single-letter solution is given by:

$$\max_{\mathcal{Q}_{W_1}, \mathcal{Q}_{X|SW_1}} \min_{\mathcal{P}_{V|W_1Y}} \mathbb{E}\Big[d(S,V)\Big]. \tag{26}$$

This Theorem establishes a connexion between the notions of channel state leakage: $\frac{1}{n} I(S^n; Y^n)$, control of the decoder's posterior beliefs: $\frac{1}{n} \cdot D\Big(\mathcal{P}_{S^n|Y^n} \Big|\Big| \prod_{i=1}^{n} \mathcal{Q}_{S_i|Y_iW_{1,i}}\Big)$, and channel state estimation: $\min_{h_{V^n|Y^n}} \mathbb{E}\bigg[\frac{1}{n}\sum_{i=1}^{n} d(S_i, V_i)\bigg]$.

**Theorem III.3 (Zero-sum game)** *A pair of rate and distortion* $(\mathsf{R}, \mathsf{D}) \in \mathcal{A}_{\mathsf{g}}$ *is achievable* if and only if *there exists an auxiliary random variable* $W_1$ *with probability distribution* $\mathcal{Q}_{SW_1XY} \in \mathbb{Q}_{\mathsf{d}}$ *that satisfies:*

$$\mathsf{R} \leq I(W_1;Y), \tag{27}$$

$$\mathsf{D} = \min_{\mathcal{P}_{V|W_1Y}} \mathbb{E}\Big[d(S,V)\Big]. \tag{28}$$

*The set* $\mathbb{Q}_{\mathsf{d}}$ *refers to the set of joint distributions* $\mathcal{P}_S\mathcal{Q}_{W_1}\mathcal{Q}_{X|SW_1}\mathcal{T}_{Y|XS}$ *defined in Theorem II.6.*

The achievability proof of Theorem III.3 is stated in Appendix D and is a consequence of Theorems II.6 and III.1, and of [28, Corollary A.18, Lemma A.22]. The bound on the KL-divergence in equation (19) relates to the notion of "strategic distance" [16, Section 5.2], later used in several articles on repeated game [17], [18], [19], on "Bayesian persuasion" [28] and on "strategic communication" [30]. The converse proof of Theorem is stated in Appendix E.

**Remark III.4 (Maximin-minimax result)** *The optimal distortion-rate function* $\mathsf{D}^\star(\mathsf{R})$ *reformulates as a maximin problem:*

$$\mathsf{D}^\star(\mathsf{R}) = \max_{\substack{\mathcal{Q}_{W_1},\mathcal{Q}_{X|SW_1} \\ \mathsf{R}\leq I(W_1;Y)}} \min_{\mathcal{P}_{V|W_1Y}} \mathbb{E}\Big[d(S,V)\Big] = \min_{\mathcal{P}_{V|W_1Y}} \max_{\substack{\mathcal{Q}_{W_1},\mathcal{Q}_{X|SW_1} \\ \mathsf{R}\leq I(W_1;Y)}} \mathbb{E}\Big[d(S,V)\Big]. \tag{29}$$

*The maximum and the minimum are taken over compact and convex sets and the distortion function is linear. Hence by Sion's Theorem [33] the maximin equal the minimax and the value of this channel state estimation zero-sum game is* $\mathsf{D}^\star(\mathsf{R})$.

**Remark III.5 (One auxiliary random variable)** *The formulation of Theorem III.3 is based on the set of distributions* $\mathbb{Q}_{\mathsf{d}}$ *with only one auxiliary random variable* $W_1$*, instead of the two random variables* $(W_1, W_2)$ *of the set* $\mathbb{Q}_{\mathsf{e}}$*. When the encoder tries to mask the channel state, it does not requires the auxiliary random variable* $W_2$ *anymore.*

$$\mathsf{D}^\circ = \max_{\substack{\mathcal{Q}_{W_1},\mathcal{Q}_{X|SW_1},\mathcal{Q}_{W_2|SW_1} \\ \mathsf{R}\leq I(W_1,W_2;Y)-I(W_2;S|W_1)}} \min_{\mathcal{P}_{V|W_1W_2Y}} \mathbb{E}\Big[d(S,V)\Big] \tag{30}$$

$$\leq \max_{\substack{\mathcal{Q}_{W_1},\mathcal{Q}_{X|SW_1},\mathcal{Q}_{W_2|SW_1} \\ \mathsf{R}\leq I(W_1,W_2;Y)-I(W_2;S|W_1)}} \min_{\mathcal{P}_{V|W_1Y}} \mathbb{E}\Big[d(S,V)\Big] \tag{31}$$

$$\leq \max_{\substack{\mathcal{Q}_{W_1},\mathcal{Q}_{X|SW_1} \\ \mathsf{R}\leq I(W_1;Y)}} \min_{\mathcal{P}_{V|W_1Y}} \mathbb{E}\Big[d(S,V)\Big] = \mathsf{D}^\star, \tag{32}$$

*where* (31) *comes from taking the minimum over* $\mathcal{P}_{V|W_1Y}$ *instead of* $\mathcal{P}_{V|W_1W_2Y}$*;* (32) *comes from the Markov chain* $Y \multimap (S, W_1) \multimap W_2$ *stated in* (9)*, that ensures:* $I(W_1, W_2; Y) - I(W_2; S|W_1) \leq I(W_1; Y)$*. Hence, the information constraint* $\mathsf{R} \leq I(W_1, W_2; Y) - I(W_2; S|W_1)$ *is more restrictive than* $\mathsf{R} \leq I(W_1; Y)$.

**Remark III.6 (Causal state communication)** *In [10], the goal of the encoder is to convey the state information to the receiver so as to minimize the long-run distortion function*

$$\min_{\substack{\{f_{X_i|S^i}\}_{i\in\{1,\ldots,n\}},\\ h_{V^n|Y^n}}} \mathbb{E}\bigg[\frac{1}{n}\sum_{i=1}^{n} d(S_i,V_i)\bigg], \tag{33}$$

*whereas in* (25) *the goal of the encoder is to* mask *the state information to the receiver. The authors of [10] proved the convergence of equation* (33) *to the single-letter optimization problem:*

$$\widehat{D} = \min_{\substack{\mathcal{Q}_{W_1},\mathcal{Q}_{X|SW_1},\mathcal{Q}_{W_2|SW_1},\mathcal{Q}_{V|W_1W_2Y}\\ 0\le I(W_1,W_2;Y)-I(W_2;S|W_1)}} \mathbb{E}\Big[d(S,V)\Big], \tag{34}$$

*where the auxiliary random variable $W_2$ is used to convey a quantized version of the channel state to the decoder.*

## IV. Extensions to more general scenarios

### A. Two-sided state information

The case of two-sided state information is represented by Fig. 3. The distribution $\mathcal{P}_{USZ} \in \Delta(\mathcal{U}\times\mathcal{S}\times\mathcal{Z})$ generates i.i.d. correlated channel state $S$, information source $U$ and state information $Z$ at the decoder.

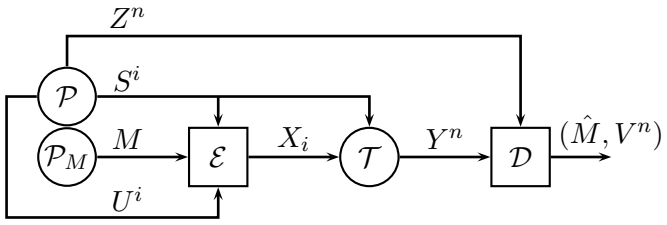


Fig. 3. Causal encoding function $f_i : \mathcal{M}\times\mathcal{U}^i\times\mathcal{S}^i \to \mathcal{X}$, for all $i \in \{1,\ldots,n\}$ and non-causal decoding functions $g : \mathcal{Y}^n\times\mathcal{Z}^n \to \mathcal{M}$ and $h : \mathcal{Y}^n\times\mathcal{Z}^n \to \Delta(\mathcal{V}^n)$.

**Definition IV.1** *A code with two-sided state information $c \in \mathcal{C}_{\mathsf{s}}(n,\mathcal{M})$ is a tuple of functions $c = (\{f_i\}_{i\in\{1,\ldots,n\}}, g, h)$ defined by:*

$$f_i : \mathcal{M}\times\mathcal{U}^i\times\mathcal{S}^i \longrightarrow \Delta(\mathcal{X}), \qquad \forall i\in\{1,\ldots,n\}, \tag{35}$$

$$g : \mathcal{Y}^n\times\mathcal{Z}^n \longrightarrow \mathcal{M}, \tag{36}$$

$$h : \mathcal{Y}^n\times\mathcal{Z}^n \longrightarrow \Delta(\mathcal{V}^n). \tag{37}$$

*The empirical distribution $Q^n_{USZXYV}$ of sequences $(u^n,s^n,z^n,x^n,y^n,v^n)$ is defined by:*

$$Q^n(u,s,z,x,y,v) = \frac{\mathsf{N}(u,s,z,x,y,v|u^n,s^n,z^n,x^n,y^n,v^n)}{n},$$
$$\forall (u,s,z,x,y,v) \in \mathcal{U}\times\mathcal{S}\times\mathcal{Z}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{V}. \tag{38}$$

The code with two-sided state information $c \in \mathcal{C}_{\mathsf{s}}(n, \mathcal{M})$, the uniform probability distributions of the messages $\mathcal{P}_M$, the source $\mathcal{P}_{USZ}$ and the channel $\mathcal{T}_{Y|XS}$ induce a general probability distribution on $(M, U^n, S^n, Z^n, X^n, Y^n, V^n, \hat{M})$:

$$\mathcal{P}_M \prod_{i=1}^{n} \left[ \mathcal{P}_{U_i S_i Z_i} f_{X_i|U^i S^i M} \mathcal{T}_{Y_i|X_i S_i} \right] h_{V^n|Y^n Z^n} \mathbb{1}\left(\hat{M} = g(Y^n, Z^n)\right). \tag{39}$$

**Definition IV.2** *Fix a target rate $\mathsf{R}$, a target state leakage $\mathsf{E}$ and a target probability distribution $\mathcal{Q}_{USZXYV}$. The triple $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable if for all $\varepsilon > 0$, there exists a $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$, there exists a code with two-sided state information $c \in \mathcal{C}_{\mathsf{s}}(n, \mathcal{M})$ that satisfies:*

$$\frac{\log_2 |\mathcal{M}|}{n} \geq \mathsf{R} - \varepsilon, \tag{40}$$

$$\left| \mathcal{L}_{\mathsf{e}}(c) - \mathsf{E} \right| \leq \varepsilon, \qquad \text{with} \qquad \mathcal{L}_{\mathsf{e}}(c) = \frac{1}{n} \cdot I(U^n, S^n; Y^n, Z^n), \tag{41}$$

$$\mathcal{P}_{\mathsf{e}}(c) = \mathbb{P}\left( M \neq \hat{M} \right) + \mathbb{P}\left( \left| \left| Q^n - \mathcal{Q} \right| \right|_1 \geq \varepsilon \right) \leq \varepsilon. \tag{42}$$

**Theorem IV.3 (Two-sided state information)** *We consider a target joint probability distribution $\mathcal{Q}_{USZXYV}$ that decomposes as $\mathcal{Q}_{USZXYV} = \mathcal{P}_{USZ} \mathcal{Q}_{X|US} \mathcal{T}_{Y|XS} \mathcal{Q}_{V|USZXY}$. The triple $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable* if and only if *there exists two auxiliary random variables $(W_1, W_2)$ with probability distribution $\mathcal{Q}_{USZW_1W_2XYV} \in \mathbb{Q}_{\mathsf{s}}$ that satisfy:*

$$I(U, S; W_1, W_2, Y, Z) \leq \mathsf{E} \leq H(U, S), \tag{43}$$

$$\mathsf{R} + \mathsf{E} \leq I(W_1, U, S; Y, Z), \tag{44}$$

*where $\mathbb{Q}_{\mathsf{s}}$ is the set of joint probability distributions $\mathcal{Q}_{USZW_1W_2XYV}$ that decompose as*

$$\mathcal{P}_{USZ} \mathcal{Q}_{W_1} \mathcal{Q}_{W_2|USW_1} \mathcal{Q}_{X|USW_1} \mathcal{T}_{Y|XS} \mathcal{Q}_{V|YZW_1W_2}, \tag{45}$$

*and the support of $(W_1, W_2)$ are bounded by $\max(|\mathcal{W}_1|, |\mathcal{W}_2|) \leq d+1$ with $d = |\mathcal{U} \times \mathcal{S} \times \mathcal{Z} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}|$.*

The achievability proof of Theorem IV.3 follows directly from the proof of Theorem II.3 stated in Section A, by replacing the random variable of the channel state $S$ by the pair $(U, S)$ and the random variable of the channel output $Y$ by the pair $(Y, Z)$. The converse proof of Theorem IV.3 is stated in Appendix H.

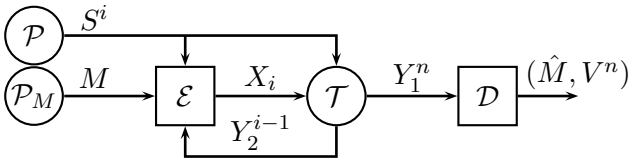


Fig. 4. Noisy feedback $Y_2^{i-1}$ from state-dependent $\mathcal{T}_{Y_1Y_2|XS}$ channel. Encoding writes $f_i : \mathcal{M} \times \mathcal{S}^i \times \mathcal{Y}_2^{i-1} \to \mathcal{X}$, $\forall i \in \{1, \ldots, n\}$.

**Remark IV.4** *The Markov chains* $X \multimap (U, S, W_1) \multimap W_2$ *and* $Y \multimap (X, S) \multimap (U, Z, W_1, W_2)$ *and* $Z \multimap (U, S) \multimap (X, Y, W_1, W_2)$ *imply another Markov chain property:* $(Y, Z) \multimap (W_1, U, S) \multimap W_2$. *Indeed, for all* $(u, s, z, w_1, w_2, x, y)$ *we have:*

$$\begin{aligned}\mathcal{P}(y, z|w_1, w_2, u, s) &= \sum_{x \in \mathcal{X}} \mathcal{P}(x|w_1, w_2, u, s) \cdot \mathcal{P}(y|x, w_1, w_2, u, s) \cdot \mathcal{P}(z|x, w_1, w_2, u, s, y) \\ &= \sum_{x \in \mathcal{X}} \mathcal{Q}(x|u, s, w_1) \cdot \mathcal{T}(y|x, s) \cdot \mathcal{P}_{\mathsf{z|us}}(z|u, s) = \mathcal{P}(y, z|w_1, u, s).\end{aligned}$$

*By combining equations* (43)*,* (44)*, with the Markov chain* $(Y, Z) \multimap (W_1, U, S) \multimap W_2$*, we recover the information constraint of [26, Theorem V.1]:*

$$\mathsf{R} \leq I(W_1, W_2; Y, Z) - I(W_2; U, S|W_1). \tag{46}$$

*B. Noisy channel feedback observed by the encoder*

We characterize the set of achievable triples $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ when the encoder has noisy feedback $Y_2$ from the state-dependent channel $\mathcal{T}_{Y_1Y_2|XS}$ depicted in Fig. 4. The encoding function writes $f_i : \mathcal{M} \times \mathcal{S}^i \times \mathcal{Y}_2^{i-1} \to \mathcal{X}$, $\forall i \in \{1, \ldots, n\}$ whereas the decoding functions and the state leakage remain unchanged.

**Theorem IV.5 (Noisy channel feedback)** *We consider a target joint probability distribution* $\mathcal{Q}_{SXY_1Y_2V}$ *that decomposes as* $\mathcal{Q}_{SXY_1Y_2V} = \mathcal{P}_S \mathcal{Q}_{X|S} \mathcal{T}_{Y_1Y_2|XS} \mathcal{Q}_{V|SXY_1Y_2}$. *The triple* $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ *is achievable* if and only if *there exists two auxiliary random variables* $(W_1, W_2)$ *with probability distribution* $\mathcal{Q}_{SW_1W_2XY_1Y_2V} \in \mathbb{Q}_{\mathsf{f}}$ *that satisfy:*

$$\mathsf{R} \leq I(W_1, W_2; Y_1) - I(W_2; S, Y_2|W_1), \tag{47}$$

$$I(S; W_1, W_2, Y_1) \leq \mathsf{E} \leq H(S), \tag{48}$$

$$\mathsf{R} + \mathsf{E} \leq I(W_1, S; Y_1), \tag{49}$$

*where* $\mathbb{Q}_{\mathsf{f}}$ *is the set of distributions with marginal* $\mathcal{Q}_{SW_1W_2XY_1Y_2V}$ *that decompose as*

$$\mathcal{P}_S \mathcal{Q}_{W_1} \mathcal{Q}_{X|SW_1} \mathcal{T}_{Y_1Y_2|XS} \mathcal{Q}_{W_2|SW_1Y_2} \mathcal{Q}_{V|Y_1W_1W_2},$$

*and the supports of* $(W_1, W_2)$ *are bounded by* $\max(|\mathcal{W}_1|, |\mathcal{W}_2|) \leq d+1$ *with* $d = |\mathcal{S} \times \mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{V}|$.

The achievability proof of Theorem IV.5 follows directly from the proof of Theorem II.3, by replacing the pair $(S^n, W_1^n)$ by the triple $(S^n, W_1^n, Y_2^n)$ when the encoder generates $W_2^n$. The decoding functions and the leakage analysis remain unchanged. The converse proof is stated in Appendix I.

**Remark IV.6 (Noisy feedback improve coordination)** *The channel feedback increases the set of achievable triples* $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$*, since the new conditional distribution* $\mathcal{Q}_{W_2|SW_1Y_2}$ *depends on channel outputs* $Y_2$*. The information constraints of Theorem IV.5 are reduced to that of Theorem II.3 as soon as* $\mathcal{Q}_{W_2|SW_1Y_2} = \mathcal{Q}_{W_2|SW_1} \iff W_2 \multimap (S, W_1) \multimap Y_2 \iff I(W_2; Y_2|S, W_1) = 0$*. The problem of empirical coordination with channel feedback is investigated in [27].*

## V. Strictly causal encoding

**Definition V.1** *A code with strictly causal encoding* $c \in \mathcal{C}_{\mathsf{se}}(n, \mathcal{M})$ *is a tuple of functions* $c = (\{f_i\}_{i\in\{1,\dots,n\}}, g, h)$ *defined by:*

$$f_i : \mathcal{M} \times \mathcal{S}^{i-1} \longrightarrow \Delta(\mathcal{X}), \qquad \forall i \in \{1, \dots, n\}, \tag{50}$$

$$g : \mathcal{Y}^n \longrightarrow \mathcal{M}, \tag{51}$$

$$h : \mathcal{Y}^n \longrightarrow \Delta(\mathcal{V}^n). \tag{52}$$

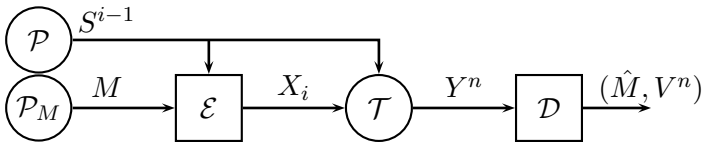


Fig. 5. Strictly causal encoding function $f_i : \mathcal{M} \times \mathcal{S}^{i-1} \to \Delta(\mathcal{X})$, for all $i \in \{1, \dots, n\}$ and non-causal decoding functions $g : \mathcal{Y}^n \to \mathcal{M}$ and $h : \mathcal{Y}^n \to \Delta(\mathcal{V}^n)$.

**Theorem V.2 (Strictly causal encoding)** *We consider a target joint probability distribution* $\mathcal{Q}_{SXYV}$ *that decomposes as* $\mathcal{Q}_{SXYV} = \mathcal{P}_S \mathcal{Q}_X \mathcal{T}_{Y|XS} \mathcal{Q}_{V|SXY}$*. The triple* $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ *is achievable* if and only if *there exists an auxiliary random variable* $W_2$ *with probability distribution* $\mathcal{Q}_{SW_2XYV} \in \mathbb{Q}_{\mathsf{se}}$ *that satisfy:*

$$I(S; X, W_2, Y) \leq \mathsf{E} \leq H(S), \tag{53}$$

$$\mathsf{R}+\mathsf{E}\leq I(X,S;Y), \tag{54}$$

*where $\mathbb{Q}_{\mathsf{se}}$ is the set of joint probability distributions $\mathcal{Q}_{SW_2XYV}$ with marginal $\mathcal{Q}_{SW_2XYV}$ that decompose as*

$$\mathcal{Q}_{SW_2XYV}=\mathcal{P}_S\mathcal{Q}_X\mathcal{Q}_{W_2|SX}\mathcal{T}_{Y|XS}\mathcal{Q}_{V|XYW_2} \tag{55}$$

*and the support of the auxiliary random variable $W_2$ is bounded by $|\mathcal{W}_2|\leq|\mathcal{S}\times\mathcal{X}\times\mathcal{Y}|+1$.*

The achievability proof of Theorem V.2 follows directly from the proof of Theorem II.3 stated in Section A, by replacing the auxiliary random variable $W_1$ by the channel input $X$. The converse proof is stated in Appendix J.

**Remark V.3** *Equation* (54)*, the first inequality of* (53)*, the Markov chain $Y\multimap(X,S)\multimap W_2$ and the independence between $S$ and $X$ imply:*

$$\mathsf{R}\leq I(X,W_2;Y)-I(W_2;S|X). \tag{56}$$

**Corollary V.4 (Strictly causal encoding without receiver's outputs)** *A pair of rate and state leakage $(\mathsf{R},\mathsf{E})$ is achievable* if and only if *there exists a probability distribution $\mathcal{Q}_X$ that satisfies:*

$$I(S;Y|X)\leq\mathsf{E}\leq H(S), \tag{57}$$

$$\mathsf{R}+\mathsf{E}\leq I(X,S;Y). \tag{58}$$

The achievability proof of Corollary V.4 comes from the achievability proof of Theorem V.2. The converse proof is based on standard arguments. Equations (57) and (58) imply $\mathsf{R}\leq I(X;Y)$.

## APPENDIX A
## ACHIEVABILITY PROOF OF THEOREM II.3

### *A. Random coding*

We fix a triple of rate, state leakage and joint probability distribution $(\mathsf{R},\mathsf{E},\mathcal{Q})$ for which there exists a probability distribution $\mathcal{Q}_{SW_1W_2XYV}\in\mathbb{Q}_{\mathsf{e}}$ that satisfy the following equations:

$$I(S;W_1,W_2,Y)\leq\mathsf{E}\leq H(S), \tag{59}$$

$$\mathsf{R}+\mathsf{E}\leq I(W_1,S;Y). \tag{60}$$

We show that $(\mathsf{R},\mathsf{E},\mathcal{Q})$ is achievable by introducing the rate parameters $\mathsf{R}_{\mathsf{L}}$, $\mathsf{R}_{\mathsf{J}}$, $\mathsf{R}_{\mathsf{K}}$ and by considering a block-Markov random code $c\in\mathcal{C}(n\cdot B,\mathcal{M})$ defined over $B\in\mathbb{N}$ blocks of length $n\in\mathbb{N}$. The codebook is defined over one block of length $n\in\mathbb{N}$ and the total length of the code is denoted by $N=n\cdot B\in\mathbb{N}$.

**Random Codebook.**

1) We draw $2^{n(H(S)+\varepsilon)}$ sequences $S^n(l,j)$ according to i.i.d. probability distribution $\mathcal{P}_S$, with indexes $(l,j) \in \mathcal{M}_{\mathsf{L}} \times \mathcal{M}_{\mathsf{J}}$ with cardinalities $|\mathcal{M}_{\mathsf{L}}| = 2^{n\mathsf{R}_{\mathsf{L}}}$ and $|\mathcal{M}_{\mathsf{J}}| = 2^{n\mathsf{R}_{\mathsf{J}}}$.

2) We draw $2^{n(\mathsf{R}+\mathsf{R}_{\mathsf{L}}+\mathsf{R}_{\mathsf{K}})}$ sequences $W_1^n(m,l,k)$ according to the i.i.d. probability distribution $\mathcal{Q}_{W_1}$, with indexes $(m,l,k) \in \mathcal{M} \times \mathcal{M}_{\mathsf{L}} \times \mathcal{M}_{\mathsf{K}}$.

3) For each triple of indexes $(m,l,k) \in \mathcal{M} \times \mathcal{M}_{\mathsf{L}} \times \mathcal{M}_{\mathsf{K}}$, we draw the same number $2^{n(\mathsf{R}+\mathsf{R}_{\mathsf{L}}+\mathsf{R}_{\mathsf{K}})}$ of sequences $W_2^n(m,l,k,\hat{m},\hat{l},\hat{k})$ with indexes $(\hat{m},\hat{l},\hat{k}) \in \mathcal{M} \times \mathcal{M}_{\mathsf{L}} \times \mathcal{M}_{\mathsf{K}}$, according to the i.i.d. conditional probability distribution $\mathcal{Q}(w_2|w_1)$ depending on the sequence $W_1^n(m,l,k)$.

**Encoding function at the beginning of block $b \in \{2,\dots B-1\}$.**

1) The encoder observes the sequence of channel states $S_{b-1}^n$ corresponding to the previous block $b-1$ and finds the indexes $(l_{b-1}, j_{b-1}) \in \mathcal{M}_{\mathsf{L}} \times \mathcal{M}_{\mathsf{J}}$ such that $\big(S^n(l_{b-1}, j_{b-1}), S_{b-1}^n\big) \in T_\delta(\mathcal{Q})$ are jointly typical for the probability distribution $\mathcal{P}_S \mathbb{1}_{\{S=S\}}$.

2) The encoder observes the message $m_b$ and the index $l_{b-1}$ and recalls the sequence $W_1^n(m_{b-1}, l_{b-2}, k_{b-1})$ corresponding to the previous block $b-1$. It finds the index $k_b \in \mathcal{M}_{\mathsf{K}}$ such that the sequences $\big(S_{b-1}^n, W_1^n(m_{b-1}, l_{b-2}, k_{b-1}), W_2^n(m_{b-1}, l_{b-2}, k_{b-1}, m_b, l_{b-1}, k_b)\big) \in T_\delta(\mathcal{Q})$ are jointly typical.

3) The encoder sends the sequence $X_b^n$ drawn from the i.i.d. conditional probability distribution $\mathcal{Q}_{X|SW_1}$ depending on sequences $W_1^n(m_b, l_{b-1}, k_b)$ and $S_b^n$ observed causally on the current block $b \in \{2,\dots B-1\}$.

**Decoding function at the end of block $b \in \{2,\dots B-1\}$.**

1) The receiver recalls the sequence $Y_{b-1}^n$ and the indexes $(m_{b-1}, l_{b-2}, k_{b-1})$ corresponding to the sequence $W_1^n(m_{b-1}, l_{b-2}, k_{b-1})$ decoded at the end of the previous block $b-1$.

2) The receiver observes the sequence $Y_b^n$ and finds the triple of indexes $(m_b, l_{b-1}, k_b)$ such that $\big(Y_b^n, W_1^n(m_b, l_{b-1}, k_b)\big) \in T_\delta(\mathcal{Q})$ and $\big(Y_{b-1}^n, W_1^n(m_{b-1}, l_{b-2}, k_{b-1}), W_2^n(m_{b-1}, l_{b-2}, k_{b-1}, m_b, l_{b-1}, k_b)\big)$ $\in T_\delta(\mathcal{Q})$ are jointly typical.

3) The receiver returns the message $m_b$ corresponding to block $b$.

4) The receiver returns the sequence $V_{b-1}^n$ drawn from the conditional probability distribution $\mathcal{Q}_{V|YW_1W_2}$ depending on sequences $\big(Y_{b-1}^n, W_1^n(m_{b-1}, l_{b-2}, k_{b-1}), W_2^n(m_{b-1}, l_{b-2}, k_{b-1}, m_b, l_{b-1}, k_b)\big)$.

5) The receiver knows that over block $b-1$, the sequences $\big(S_{b-1}^n, W_1^n(m_{b-1}, l_{b-2}, k_{b-1}), W_2^n(m_{b-1}, l_{b-2}, k_{b-1},$ $m_b, l_{b-1}, k_b), X_{b-1}^n, Y_{b-1}^n, V_{b-1}^n\big) \in T_\delta(\mathcal{Q})$ are jointly typical and the sequence of states $S_{b-1}^n$ belongs to the bin with index $l_{b-1} \in \mathcal{M}_{\mathsf{L}}$.

**Initialization of the encoder.** Arbitrary indexes $(m_1, l_0, k_1)$ are given to both encoder and decoder. The encoder sends the sequence $X_{b_1}^n$ drawn according to the conditional probability distribution $\mathcal{Q}_{X|SW_1}$ depending on sequences $\big(S_{b_1}^n, W_1^n(m_1, l_0, k_1)\big)$. At the beginning of the second

block $b_2$, encoder recalls $W_1^n(m_1, l_0, k_1)$, observes message $m_2$, finds the index $l_1$ such that sequences $\left(S_{b_1}^n, S^n(l_1, j_1)\right) \in T_\delta(\mathcal{Q})$ are jointly typical and finds the index $k_2$ such that sequences $\left(S_{b_1}^n, W_1^n(m_1, l_0, k_1), W_2^n(m_1, l_0, k_1, m_2, l_1, k_2)\right) \in T_\delta(\mathcal{Q})$ are jointly typical. The encoder sends the sequence $X_{b_2}^n$ drawn from the conditional probability distribution $\mathcal{Q}_{X|SW_1}$ depending on sequences $\left(S_{b_2}^n, W_1^n(m_2, l_1, k_2)\right)$. The index $m_1$ does not correspond to a message.

**Initialization of the decoder.** At the end of second block $b_2$, the decoder finds the triple of indexes $(m_2, l_1, k_2)$ such that $\left(Y_{b_2}^n, W_1^n(m_2, l_1, k_2)\right) \in T_\delta(\mathcal{Q})$ and $\left(Y_{b_1}^n, W_1^n(m_1, l_0, k_1), W_2^n(m_1, l_0, k_1, m_2, l_1, k_2)\right)$ $\in T_\delta(\mathcal{Q})$ are jointly typical. The decoder returns the message $m_2$ corresponding to the block $b_2$ and the sequence $V_{b_1}^n \in \mathcal{V}^n$ drawn from the conditional probability $\mathcal{Q}_{V|YW_1W_2}$ depending on sequences $\left(Y_{b_1}^n, W_1^n(m_1, l_0, k_1), W_2^n(m_1, l_0, k_1, m_2, l_1, k_2)\right)$. The decoder knows that over the first block $b_1$, the sequence $S_{b_1}^n$ belongs to the bin $l_1$. The sequences $\left(S_{b_1}^n, W_1^n(m_1, l_0, k_1), W_2^n(m_1, l_0, k_1, m_2, l_1, k_2), X_{b_1}^n, Y_{b_1}^n, V_{b_1}^n\right) \in T_\delta(\mathcal{Q})$ and $\left(S_{b_2}^n, W_1^n(m_2, l_1, k_2), W_2^n(m_2, l_1, k_2, m_3, l_2, k_3), X_{b_2}^n, Y_{b_2}^n, V_{b_2}^n\right) \in T_\delta(\mathcal{Q})$ are jointly typical.

**Last block at the encoder.** At the beginning of the last block $B$, the encoder recalls $W_1^n(m_{B-1}$, observes message $m_B$, finds the index $l_{B-1}$ such that sequences $\left(S_{B-1}^n, S^n(l_{B-1}, j_{B-1})\right) \in T_\delta(\mathcal{Q})$ are jointly typical. It finds the index $k_B$ such that sequences $\left(S_{B-1}^n, W_1^n(m_{B-1}, l_{B-2}, k_{B-1}), W_2^n(m_{B-1}, l_{B-2}, k_{B-1}, m_B, l_{B-1}, k_B)\right) \in T_\delta(\mathcal{Q})$ are jointly typical. The encoder sends the sequence $X_B^n$ drawn from the conditional probability distribution $\mathcal{Q}_{X|SW_1}$ depending on the sequences $\left(S_B^n, W_1^n(m_B, l_{B-1}, k_B)\right)$.

**Last block at the decoder.** At the end of the last block $B$, the decoder finds the triple of indexes $(m_B, l_{B-1}, k_B)$ such that $\left(Y_B^n, W_1^n(m_B, l_{B-1}, k_B)\right) \in T_\delta(\mathcal{Q})$ and $\left(Y_{B-1}^n, W_1^n(m_{B-1}, l_{B-2}, k_{B-1}), W_2^n(m_{B-1}, l_{B-2}, k_{B-1}, m_B, l_{B-1}, k_B)\right) \in T_\delta(\mathcal{Q})$ are jointly typical. The decoder returns the message $m_B$ corresponding to the last block $B$ and the sequence $V_{B-1}^n \in \mathcal{V}^n$ drawn from the conditional probability $\mathcal{Q}_{V|YW_1W_2}$ depending on sequences $\left(Y_{B-1}^n, W_1^n(m_{B-1}, l_{B-2}, k_{B-1}), W_2^n(m_{B-1}, l_{B-2}, k_{B-1}, m_B, l_{B-1}, k_B)\right)$. The decoder knows that over the block $B-1$, the sequence $S_{B-1}^n$ belongs to the bin $l_{B-1}$. The sequences $\left(S_{B-1}^n, W_1^n(m_{B-1}, l_{B-2}, k_{B-1}), W_2^n(m_{B-1}, l_{B-2}, k_{B-1}, m_B, l_{B-1}, k_B), X_{B-1}^n, Y_{B-1}^n, V_{B-1}^n\right) \in T_\delta(\mathcal{Q})$ are jointly typical but the sequences $(S_B^n, W_{1,B}^n, W_{2,B}^n, X_B^n, Y_B^n, V_B^n) \notin T_\delta(\mathcal{Q})$ are not jointly typical on the last block $B$. The decoder does not know the index $l_B$ of the bin corresponding to sequence $S_B^n$.

In the following, we introduce the notation $W_{1,b}^n = W_1^n(m_b, l_{b-1}, k_b)$ and $W_{2,b}^n = W_2^n(m_b, l_{b-1}, k_b, m_{b+1}, l_b, k_{b+1})$, with $b \in \{1, \ldots, B-1\}$. If there is no error in the coding scheme, the messages $(m_2, \ldots, m_B)$ are correctly decoded and the decoder knows the bin indexes $(l_1, \ldots, l_{b-1})$ of the sequences $(S_1^n, \ldots, S_{b-1}^n)$. The sequences $(S_b^n, W_{1,b}^n, W_{2,b}^n, X_b^n, Y_b^n, V_b^n) \in T_\delta(\mathcal{Q})$ are jointly typical for each blocks $b \in \{1, \ldots, B-1\}$.

*B. Rate parameters $\mathsf{R}$, $\mathsf{R_L}$ and $\mathsf{R_K}$*

1) At the end of block $b \in \{2, \dots B\}$, the decoder observes $(Y_{b-1}^n, Y_b^n)$ and decodes the sequences $(W_{1,b-1}^n, W_{2,b-1}^n)$ corresponding to the block $b-1$. Intuitively, the observation of the sequences $(W_{1,b-1}^n, W_{2,b-1}^n, Y_{b-1}^n)$ leaks $n \cdot I(S; W_1, W_2, Y) = n \cdot I(S; W_2, Y|W_1)$ bits of information regarding sequence $S_{b-1}^n$. By fixing the rate parameter $\mathsf{R_L} = \mathsf{E} - I(S; W_1, W_2, Y)$, the encoder will transmit $n \cdot \mathsf{R_L} = n \cdot \big(\mathsf{E} - I(S; W_1, W_2, Y)\big)$ additional bits of information corresponding to the state sequence $S_{b-1}^n$. As it will be proven in the Section A-F, the leakage rate $I(S_{b-1}^n; Y_{b-1}^n)$ over block $b-1$ is close to $n \cdot \Big(I(S; W_1, W_2, Y) + \mathsf{E} - I(S; W_1, W_2, Y)\Big) = n \cdot \mathsf{E}$. We fix the rate parameter $\mathsf{R_L}$ equal to:

$$\mathsf{R_L} = \mathsf{E} - I(S; W_1, W_2, Y) - 2\varepsilon \geq 0. \tag{61}$$

The first inequality $I(S; W_1, W_2, Y) \leq \mathsf{E}$ in (59) implies there exists a positive rate parameter $\mathsf{R_L}$. In case of equality $\mathsf{E} = I(S; W_1, W_2, Y)$, then the rate $\mathsf{R_L} = 0$ and no index $l \in \mathcal{M}_\mathsf{L}$ is transmitted to the decoder.

2) The rates parameters $\mathsf{R_L}$, $\mathsf{R_J}$ corresponding to the indexes $(l_{b-1}, j_{b-1})$, guarantee that almost every sequences $S_{b-1}^n$ appear in the codebook.

$$\mathsf{R_L} + \mathsf{R_J} = H(S) + \varepsilon, \tag{62}$$

$$\implies \mathsf{R_J} = H(S) - \mathsf{E} + I(S; W_1, W_2, Y) + 3\varepsilon. \tag{63}$$

The second inequality $\mathsf{E} \leq H(S)$ in (59) implies there exists a positive rate parameter $\mathsf{R_J}$.

3) The rates parameter $\mathsf{R_K}$ corresponding to the index $k_b$, is used by the encoder in order to correlate the sequences $\big(W_1^n(m_{b-1}, l_{b-2}, k_{b-1}), W_2^n(m_{b-1}, l_{b-2}, k_{b-1}, m_b, l_{b-1}, k_b)\big)$ with the sequence of states $S_{b-1}^n$.

$$\mathsf{R_K} = I(W_2; S|W_1) + \varepsilon, \tag{64}$$

Since the random variables $W_1$ and $S$ are independent, we have $I(W_1, W_2; S) = I(W_2; S|W_1)$.

4) The rate parameters $\mathsf{R}$ , $\mathsf{R_L}$ , $\mathsf{R_K}$ are correctly decoded if:

$$\mathsf{R} + \mathsf{R_L} + \mathsf{R_K} \leq I(W_1; Y) + I(W_2; Y|W_1) - \varepsilon, \tag{65}$$

$$\iff \mathsf{R} + \mathsf{E} - I(S; W_1, W_2, Y) - 2\varepsilon + I(W_2; S|W_1) + \varepsilon \leq I(W_1, W_2; Y) - \varepsilon, \tag{66}$$

$$\iff \mathsf{R} + \mathsf{E} \leq I(W_1, W_2, S; Y) \tag{67}$$

$$\iff \mathsf{R} + \mathsf{E} \leq I(W_1, S; Y), \tag{68}$$

where (67) comes from the independence between $X$ and $S$; (68) comes from the Markov chain $Y \multimap (W_1, S) \multimap W_2$ stated in Remark II.4.

Equation (60) implies that for each block $b \in \{2, \dots B\}$, the indexes with rates $\mathsf{R}$ , $\mathsf{R_L}$ , $\mathsf{R_K}$ are recovered by the decoder, with large probability. Hence the rate of the total code of length $N = n \cdot B$ is given by:

$$\frac{1}{n \cdot B} \cdot \sum_{b=2}^{B} \log_2 |\mathcal{M}| = \frac{B-1}{B} \cdot \mathsf{R} = \mathsf{R} - \frac{1}{B} \cdot \mathsf{R} \geq \mathsf{R} - \frac{1}{B} \cdot \log_2 |\mathcal{Y}| \geq \mathsf{R} - \varepsilon. \tag{69}$$

The last equation is satisfied when the number of blocks is sufficiently large: $\frac{1}{B} \cdot \log_2 |\mathcal{Y}| \leq \varepsilon$.

### *C. Case of equality in the information constraints of Theorem II.3*

The choice of rate parameters $(\mathsf{R}, \mathsf{R_L}, \mathsf{R_K})$ is based on the implicit assumption that the information constraint (10) of Theorem II.3 is strictly positive $I(W_1, W_2; Y) - I(W_2; S|W_1) > 0$. Hence, there exists a rate $\mathsf{R_K} > 0$ such that $I(W_1, W_2; Y) - I(W_2; S|W_1) > \mathsf{R_K} > 0$ the above coding scheme works correctly.

Now assume that $I(W_1, W_2; Y) - I(W_2; S|W_1) = 0$, in that case no additional rates as $\mathsf{R} = 0$, $\mathsf{R_L} = 0$ can be transmitted and we have to choose the coordination rate $\mathsf{R_K}$ depending on whether the channel capacity is zero or strictly positive.

1) *First case*, the channel capacity with causal state information is strictly positive: $\max_{\mathcal{P}_{W_1}, \mathcal{P}_{X|W_1S}} I(W_1; Y) > 0$ (stated pp. 176 in [34]), hence the channel is not trivial, and it is possible to send some reliable information $\mathsf{R_K} > 0$. We denote by $\mathcal{P}^\star_{W_1}$, $\mathcal{P}^\star_{X|W_1S}$ the distributions that achieves the maximum. We consider the product of probability distributions $\mathcal{Q}^\star_{SW_1W_2XYV} = \mathcal{P}_S \mathcal{P}^\star_{W_1} \mathcal{P}_{W_2} \mathcal{P}^\star_{X|W_1S} \mathcal{T}_{Y|XS} \mathcal{Q}_V$, with $V$ independent of random variables $(S, W_1, W_2, X, Y)$ and $W_2$ independent of $(S, W_1)$, hence $I(W_2; S|W_1) = 0$. The corresponding information constraint is strictly positive:

$$I(W_1, W_2; Y) - I(W_2; S|W_1) = I(W_1, W_2; Y) \geq I(W_1; Y) > 0. \tag{70}$$

We define the convex combination $\mathcal{Q}^n_{SW_1W_2XYV}$ between the target distribution $\mathcal{Q}_{SW_1W_2XYV}$ and the distribution $\mathcal{Q}^\star_{SW_1W_2XYV}$ with:

$$\begin{aligned} \mathcal{Q}^n(s, w_1, w_2, x, y, v) =& \frac{1}{n} \cdot \Big( (n-1) \cdot \mathcal{Q}(s, w_1, w_2, x, y, v) + \mathcal{Q}^\star(s, w_1, w_2, x, y, v) \Big), \\ & \forall\ (s, w_1, w_2, x, y, v). \end{aligned} \tag{71}$$

The distribution $\mathcal{Q}^n_{SW_1W_2XYV}$ converges to the target $\mathcal{Q}_{SW_1W_2XYV}$ as $n$ goes to infinity. By concavity, the information constraint $I_{\mathcal{Q}^n}(W_1, W_2; Y) - I_{\mathcal{Q}^n}(W_2; S|W_1) > 0$ corresponding to any convex combination $\mathcal{Q}^n_{SW_1W_2XYV}$ is strictly positive and we can construct a coding scheme as described in Section A with

parameter $\mathsf{R_K}$ such that $I(W_1, W_2; Y) - \varepsilon > \mathsf{R_K} > I(W_2; S|W_1) + \varepsilon$.

2) *Second case*, the channel capacity is equal to zero: $\max_{\mathcal{P}_{W_1}, \mathcal{P}_{X|W_1S}} I(W_1; Y) = 0$ and no information $\mathsf{R_K} = 0$ can be transmitted to the decoder. Hence, the following mutual informations are all equal to zero:

$$0 = \max_{\mathcal{P}_{W_1}, \mathcal{P}_{X|W_1S}} I(W_1; Y) \tag{72}$$

$$\geq I(W_1; Y) \tag{73}$$

$$= I(W_2; S|W_1) - I(W_2; Y|W_1) \tag{74}$$

$$= I(W_2; S|W_1, Y) \geq 0, \tag{75}$$

where (72) comes from the hypothesis of zero channel capacity; (73) comes from the fact that the decomposition of the target distribution $\mathcal{Q}_{SW_1W_2XYV} = \mathcal{P}_S \mathcal{Q}_{W_1} \mathcal{Q}_{W_2|SW_1} \mathcal{Q}_{X|SW_1} \mathcal{T}_{Y|XS} \mathcal{Q}_{V|YW_1W_2}$ involves the probability distributions $\mathcal{Q}_{W_1}, \mathcal{Q}_{X|W_1S}$; (74) comes from the hypothesis: the information constraint (10) is equal to zero: $I(W_1; Y) + I(W_2; Y|W_1) - I(W_2; S|W_1) = 0$; (75) comes from the Markov chain $W_2 \multimap (W_1, S) \multimap Y \iff I(W_2; Y|W_1, S) = 0$ that is due to the following equations valid for all $(s, w_1, w_2, x, y)$:

$$\mathcal{P}(y|s, w_1, w_2) = \sum_x \mathcal{P}(x|s, w_1, w_2) \cdot \mathcal{P}(y|x, s, w_1, w_2) \tag{76}$$

$$= \sum_x \mathcal{P}(x|s, w_1) \cdot \mathcal{P}(y|x, s) = \mathcal{P}(y|s, w_1), \tag{77}$$

$$\implies \quad I(W_2; S|W_1) - I(W_2; Y|W_1) = I(W_2; S|W_1) - I(W_2; Y|W_1) + I(W_2; Y|S, W_1) \tag{78}$$

$$= I(W_2; S, Y|W_1) - I(W_2; Y|W_1) = I(W_2; S|W_1, Y). \tag{79}$$

Equation (77) comes from the two Markov chains $W_2 \multimap (W_1, S) \multimap X$ and $Y \multimap (X, S) \multimap (W_1, W_2)$.

Since by hypothesis the channel capacity is zero, we have: $0 = \max_{\mathcal{P}_{W_1}, \mathcal{P}_{X|W_1S}} \geq I(W_2; S|W_1, Y) = 0$. Hence, we have the two following Markov chains:

$$W_2 \multimap (W_1, S) \multimap Y \iff I(W_2; Y|W_1, S) = 0, \tag{80}$$

$$W_2 \multimap (W_1, Y) \multimap S \iff I(W_2; S|W_1, Y) = 0. \tag{81}$$

**Definition A.1 (Connected components of the auxiliary graph)** *We define a graph $G$ associated to the joint probability distribution $\mathcal{P}_{XYZ} \in \Delta(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$.*

*The vertices $v \in V$ are the pairs of symbols $v = (x, y)$ such that $\mathcal{P}(x, y) > 0$.*

*There is an edge $e = (v_1, v_2) \in E$ between two vertices $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2)$ if the first*

*component $x_1 = x_2$ or the second component $y_1 = y_2$ are equal.*
*Two vertices $v_1 = (x_1, y_1)$, $v_2 = (x_2, y_2)$ belong to the same connected component of the graph $G$ if there exists a path from $v_1 = (x_1, y_1)$ to $v_2 = (x_2, y_2)$.*
*We denote by $\mathcal{W}_3$ the set of connected components $w_3 \in \mathcal{W}_3$ of the graph $G$.*

For each symbol $w_1 \in \mathcal{W}_1$, we define a graph $G_{w_1}$ as in Definition A.1, where the vertices $v$ are the pairs of symbols $v = (s, y)$ such that $\mathcal{Q}(s, y|w_1) > 0$. There is an edge $e = (v_1, v_2)$ between two vertices $v_1 = (s_1, y_1)$ and $v_2 = (s_2, y_2)$ if the first component $s_1 = s_2$ or the second component $y_1 = y_2$ are equal. Two vertices $v_1 = (s_1, y_1)$, $v_2 = (s_2, y_2)$ belong to the same connected component $w_3 \in \mathcal{W}_3$ of the graph $G_{w_1}$ if there exists a path from $v_1 = (s_1, y_1)$ to $v_2 = (s_2, y_2)$. Even the channel has zero capacity, the decoder when receiving $Y$, can infer some information about the channel state $S$. The notion of connected component captures the compatibility between the channel states $S$ and the channel outputs $Y$.

We consider any pair of symbols $(s_1, y_1) \in w_3$ and $(s_2, y_2) \in w_3$ that belong to the same connected component $w_3 \in \mathcal{W}_3$ of the associated graph $G_{w_1}$ of definition A.1. By Lemma 1, the two Markov chains (80) and (81) imply that the conditional probability distribution of $w_2 \in \mathcal{W}_2$ are equal:

$$\mathcal{Q}(w_2|w_1, s_1, y_1) = \mathcal{Q}(w_2|w_1, s_2, y_2) = \mathcal{Q}(w_2|w_1, w_3), \qquad \forall\, w_2 \in \mathcal{W}_2. \tag{82}$$

Hence the conditional probability $\mathcal{Q}(w_2|w_1, s, y)$ is constant for any pair of symbols $(s, y) \in w_3$ that belong to the same connected component $w_3 \in \mathcal{W}_3$ of the graph $G_{w_1}$. This defines a conditional probability distribution $\mathcal{Q}(w_2|w_1, w_3)$ depending on the connected component $w_3 \in \mathcal{W}_3$ of the graph $G_{w_1}$ instead of depending on the symbols $(s, y)$. Assume that $\mathcal{Q}(w_1, s, y) > 0$ and $(s, y) \in w_3$, then we have:

$$\mathcal{Q}(w_2|w_1, s, y) = \mathcal{Q}(w_2|w_1, s) = \mathcal{Q}(w_2|w_1, y) = \mathcal{Q}(w_2|w_1, w_3), \qquad \forall\, w_2 \in \mathcal{W}_2. \tag{83}$$

Hence the conditional probability $\mathcal{Q}(w_2|w_1, s)$ of the target probability distribution (84) can be replaced by $\mathcal{Q}(w_2|w_1, w_3)$ where the connected component $w_3$ depends on the pair of symbol $(s, y)$.

$$\mathcal{P}_S \mathcal{Q}_{W_1} \mathcal{Q}_{W_2|SW_1} \mathcal{Q}_{X|SW_1} \mathcal{T}_{Y|XS} \mathcal{Q}_{V|YW_1W_2}, \tag{84}$$

$$\mathcal{P}_S \mathcal{Q}_{W_1} \mathcal{Q}_{X|SW_1} \mathcal{T}_{Y|XS} \mathbb{1}\{W_3 = h(Y)\} \mathcal{Q}_{W_2|W_1W_3} \mathcal{Q}(w_2|w_3, w_1) \mathcal{Q}_{V|YW_1W_2}. \tag{85}$$

The distribution stated in equation (85) is achievable using a trivial coding strategy that does not require an exchange of information $\mathsf{R_K} = 0$:
*Codebook.* We generate a jointly typical sequence $W_1^n \in T_\delta(\tilde{\mathcal{Q}})$ that is known by both the encoder and the decoder.

*Encoder.* At stage $i \in \{1, \dots, n\}$, the encoder observes the symbol $S_i$, recalls the sequence $W_1^n$ and generates a symbol $X_i$ using the conditional distribution $\mathcal{Q}_{X|SW_1}$.
*Decoder.* The decoder observes the sequence of channel output $Y^n$, recalls the pre-defined sequence $W_1^n$ and deduces the sequence of connected components $W_3^n$. Then, it generates the sequence $W_2^n$ using the distribution $\mathcal{Q}_{W_2|W_1W_3}$ and the sequence $V^n$ using the distribution $\mathcal{Q}_{V|W_1W_2Y}$.

Without transmission of information, the target distribution

$$\mathcal{P}_S \mathcal{Q}_{W_1} \mathcal{Q}_{X|SW_1} \mathcal{T}_{Y|XS} \mathbb{1}\{W_3 = h(Y)\} \mathcal{Q}_{W_2|W_1W_3} \mathcal{Q}_{V|YW_1W_2} \tag{86}$$

is achievable.

**Lemma 1 (Two Markov chains)** *We consider a joint probability distribution $\mathcal{P}(x, y, z) \in \Delta(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$. Assume that the two following Markov chains are satisfied:*

$$Z \multimap Y \multimap X, \qquad Z \multimap X \multimap Y. \tag{87}$$

*Then, for any pair of symbols $(x_1, y_1) \in w_3$ and $(x_2, y_2) \in w_3$ that belong to the same connected component $w_3 \in \mathcal{W}_3$ of the graph $G$ of definition A.1, we have equal conditional probability distribution:*

$$\mathcal{P}(z|x_1, y_1) = \mathcal{P}(z|x_2, y_2), \qquad \forall z \in \mathcal{Z}. \tag{88}$$

*In particular, if there exists a unique connected component $w_3 = \mathcal{W}_3$ in the graph $G$, i.e. the graph $G$ is connected, then the random variable $Z$ is independent of $(X, Y)$ and we have $I(Z; X, Y) = 0$.*

*Proof.* [Lemma 1] Consider two pairs of symbols $(x_1, y_1)$ and $(x_2, y_2)$ with strictly positive probability $\mathcal{P}(x_1, y_1) > 0$ and $\mathcal{P}(x_2, y_2) > 0$, that belong to the same connected component $w_3 \in \mathcal{W}_3$ of the graph $G$. For simplicity, we assume that the path connecting $(x_1, y_1)$ and $(x_2, y_2)$ passes through $(x_1, y_2)$. Hence, for any symbol $z \in \mathcal{Z}$, we have the following equations:

$$\mathcal{P}(z|x_1, y_1) = \mathcal{P}(z|x_1) = \mathcal{P}(z|x_1, y_2) \tag{89}$$

$$= \mathcal{P}(z|y_2) = \mathcal{P}(z|x_2, y_2). \tag{90}$$

Equation (89) comes from the Markov chain $Z \multimap X \multimap Y$ and the hypothesis of connected graph that insures $\mathcal{P}(x_1, y_2) > 0$. Otherwise, the pair $(x_1, y_2)$ is not associated with a vertex $v \in V$ of the graph $G$. Equation (90) comes from the Markov chain $Z \multimap Y \multimap X$ and the hypothesis of positive probability $\mathcal{P}(x_2, y_2) > 0$.

If there is a unique connected component in the graph $G$, then for all pair of vertices $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2)$ with positive probability $\mathcal{P}(x_1, y_1) > 0$ and $\mathcal{P}(x_2, y_2) > 0$, the conditional probability distribution $\mathcal{P}(z|x_1, y_1) = \mathcal{P}(z|x_2, y_2)$ is equal. Hence, for any triple of symbols $(x, y, z)$, the conditional

distribution satisfies $\mathcal{P}(z|x,y) = \mathcal{P}(z)$ and $Z$ is independent of the pair $(X,Y)$, $I(Z;X,Y) = 0$. This concludes the proof of Lemma 1. □

### *D. Expected error probability by block*

For each block $b \in \{2,\dots,B\}$, we consider the expected probability of the following error events. The properties of the typical sequences, stated pp. 27, in [34], implies that there exists $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$, the expected probability of the error event is bounded by $\varepsilon$:

$$\mathbb{E}_c\left[\mathbb{P}\left(S_{b-1}^n \notin T_\delta(\mathcal{Q})\right)\right] \leq \varepsilon. \tag{91}$$

From the covering Lemma, stated pp. 208, in [34], equation (62),

$$\mathsf{R}_\mathsf{L} + \mathsf{R}_\mathsf{J} \geq H(S) + \varepsilon,$$

implies that $\exists n_2 \in \mathbb{N}$ such that $\forall n \geq n_2$, the expected probability of the error event is bounded by $\varepsilon > 0$:

$$\mathbb{E}_c\left[\mathbb{P}\left(\forall (L_{b-1}, J_{b-1}) \in \mathcal{M}_\mathsf{L} \times \mathcal{M}_\mathsf{J}, \quad (S^n(L_{b-1}, J_{b-1}), S_{b-1}^n) \notin T_\delta\right)\right] \leq \varepsilon. \tag{92}$$

From the covering Lemma, stated pp. 208, in [34], equation (64),

$$\mathsf{R}_\mathsf{K} \geq I(W_2; S|W_1) + \varepsilon,$$

implies that $\exists n_3 \in \mathbb{N}$ such that $\forall n \geq n_3$, the expected probability of the error event is bounded by $\varepsilon > 0$:

$$\mathbb{E}_c\left[\mathbb{P}\left(\forall K_b \in \mathcal{M}_\mathsf{K},\ (S_{b-1}^n, W_1^n(M_{b-1}, L_{b-2}, K_{b-1}), W_2^n(M_{b-1}, L_{b-2}, K_{b-1}, M_b, L_{b-1}, K_b)) \notin T_\delta(\mathcal{Q})\right)\right] \leq \varepsilon. \tag{93}$$

From the packing Lemma, stated pp. 46, in [34], equation (65),

$$\mathsf{R} + \mathsf{R}_\mathsf{L} + \mathsf{R}_\mathsf{K} \leq I(W_1;Y) + I(W_2;Y|W_1) - \varepsilon,$$

implies that $\exists n_4 \in \mathbb{N}$ such that $\forall n \geq n_4$, the expected probability of the error event is bounded by $\varepsilon > 0$:

$$\mathbb{E}_c\Bigg[\mathbb{P}\Bigg(\exists (M_b, L_{b-1}, K_b) \neq (M_b', L_{b-1}', K_b'),\ \text{s.t.}\ \Big\{(Y_b^n, W_1^n(M_b', L_{b-1}', K_b')) \in T_\delta(\mathcal{Q})\Big\} \cap \tag{94}$$

$$\Big\{(Y_{b-1}^n, W_1^n(M_{b-1}, L_{b-2}, K_{b-1}), W_2^n(M_{b-1}, L_{b-2}, K_{b-1}, M_b', L_{b-1}', K_b')) \in T_\delta(\mathcal{Q})\Big\}\Bigg)\Bigg] \leq \varepsilon. \tag{95}$$

For each block $b \in \{2,\dots,B\}$ and for all $n \geq \bar{n} \geq \max(n_1, n_2, n_3, n_4)$, the expected probability of non-decoding the indexes $(M_b, L_{b-1}, K_b)$ is bounded by:

$$\mathbb{E}_c\left[\mathbb{P}\Big((M_b, L_{b-1}, K_b) \neq (\hat{M}_b, \hat{L}_{b-1}, \hat{M}_b)\Big)\right] \leq 4\varepsilon. \tag{96}$$

### *E. Expected error probability of the block-Markov code*

We evaluate the expected probability of error for the random indexes $(M_b, L_{b-1}, K_b)$, for $b \in \{2, \dots, B\}$ of the block-Markov random code:

$$\mathbb{E}_c\Big[\mathbb{P}\Big(\Big(M_2, L_1, K_2 \dots, M_B, L_{B-1}, K_B\Big) \neq \Big(\hat{M}_2, \hat{L}_1, \hat{K}_2, \dots, \hat{M}_B, \hat{L}_{B-1}, \hat{K}_B\Big)\Big)\Big] \tag{97}$$

$$=1 - \mathbb{E}_c\Big[\mathbb{P}\Big((M_2, L_1, K_2) = (\hat{M}_2, \hat{L}_1, \hat{K}_2)\Big)\Big] \times \dots$$

$$\times\mathbb{E}_c\Big[\mathbb{P}\Big((M_B, L_{B-1}, K_B) = (\hat{M}_B, \hat{L}_{B-1}, \hat{K}_B)\Big|\Big\{(M_2, L_1, K_2) = (\hat{M}_2, \hat{L}_1, \hat{K}_2)\Big\} \tag{98}$$

$$\cap \dots \cap \Big\{(M_{B-1}, L_{B-2}, K_{B-1}) = (\hat{M}_{B-1}, \hat{L}_{B-2}, \hat{K}_{B-1})\Big\}\Big)\Big] \leq 1 - \Big(1 - 4\varepsilon\Big)^{B-1}. \tag{99}$$

We denote by $\widetilde{Q}^N \in \Delta(\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V})$, the empirical distribution of symbols over every blocs $b \in \{1, \dots, B-1\}$ removing the last bloc. We show $\widetilde{Q}^N$ is close to the empirical distribution $Q^N$ over all the $B$ blocks, for a number of blocks $B \in \mathbb{N}$ sufficiently large, *i.e.* for which $\frac{2}{B} \cdot |\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}| \leq \varepsilon$. We denote by $Q_B$, the empirical distribution of symbols over the last bloc.

$$\Big|\Big|Q^N - \widetilde{Q}^N\Big|\Big|_1 = \Big|\Big|\frac{1}{B} \cdot \Big((B-1) \cdot \widetilde{Q}^N + Q_B\Big) - \widetilde{Q}^N\Big|\Big|_1 \tag{100}$$

$$= \frac{1}{B} \cdot \Big|\Big|Q_B - \widetilde{Q}^N\Big|\Big|_1 \leq \frac{2}{B} \cdot \Big|\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}\Big| \leq \varepsilon. \tag{101}$$

Then, the expected probability that the sequences $(S^N, W_1^N, W_2^N, X^N, Y^N, V^N) \notin T_\delta(\mathcal{Q})$ are not jointly typical, is upper bounded by:

$$\mathbb{E}_c\Big[\mathbb{P}\Big(\Big|\Big|Q^N - \mathcal{Q}\Big|\Big|_1 \geq 2\varepsilon\Big)\Big] = \mathbb{E}_c\Big[\mathbb{P}\Big(\Big|\Big|Q^N - \widetilde{Q}^N + \widetilde{Q}^N - \mathcal{Q}\Big|\Big|_1 \geq 2\varepsilon\Big)\Big]$$

$$\leq \mathbb{E}_c\Big[\mathbb{P}\Big(\Big|\Big|Q^N - \widetilde{Q}^N\Big|\Big| + \Big|\Big|\widetilde{Q}^N - \mathcal{Q}\Big|\Big|_1 \geq 2\varepsilon\Big)\Big] \leq \mathbb{E}_c\Big[\mathbb{P}\Big(\Big|\Big|\widetilde{Q}^N - \mathcal{Q}\Big|\Big|_1 \geq 2\varepsilon - \frac{2}{B} \cdot \Big|\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}\Big|\Big)\Big]$$

$$\leq \mathbb{E}_c\Big[\mathbb{P}\Big(\Big|\Big|\widetilde{Q}^N - \mathcal{Q}\Big|\Big|_1 \geq \varepsilon\Big)\Big] \leq 1 - \Big(1 - 4\varepsilon\Big)^{B-1}. \tag{102}$$

Hence, we obtain the following bound on the expected error probability:

$$\mathbb{E}_c\Big[\mathcal{P}_{\mathsf{e}}(c)\Big] = \mathbb{E}_c\Big[\mathbb{P}\Big(M \neq \hat{M}\Big) + \mathbb{P}\Big(\Big|\Big|Q^n - \mathcal{Q}\Big|\Big|_1 \geq \varepsilon\Big)\Big] \leq 2 - 2 \cdot \Big(1 - 4\varepsilon\Big)^{B-1}. \tag{103}$$

This implies the existence of a code $c^\star \in \mathcal{C}(N)$ with an error probability below $2 - 2 \cdot \Big(1 - 4\varepsilon\Big)^{B-1}$ for all $N \geq B \cdot \bar{n}$.

### *F. Expected state leakage rate*

In this section, we provide an upper and a lower bound on the expected state leakage rate, depending on the parameters $\varepsilon_1$ and $\varepsilon_2$ given by equations (144) and (145).

$$\mathsf{E} - \varepsilon_1 - \varepsilon_2 \leq \mathbb{E}_c\Big[\mathcal{L}_{\mathsf{e}}(c)\Big] = \mathbb{E}_c\Big[\frac{1}{n \cdot B} \cdot I(S^{nB}; Y^{nB}|C = c)\Big] = \frac{1}{n \cdot B} \cdot I(S^{nB}; Y^{nB}|C) \leq \mathsf{E} + \varepsilon_1 + \varepsilon_2. \tag{104}$$

Notation $S^{nB}$ denotes the sequence of random variables of channel states of length $N = n \cdot B$, whereas $S_b^n$ denotes the sub-sequence of length $n \in \mathbb{N}$ over the block $b \in \{1, \dots, B\}$.

**Upper bound.** We provide an upper bound on the expected state leakage rate by considering the chain rule, from one block to another.

$$I(S^{nB}; Y^{nB}|C) \leq \sum_{b=b_1}^{B-1} I(S_b^n; Y^{nB}|S_{b+1}^n, \dots, S_B^n, C) + n \cdot \log_2 |\mathcal{S}| \tag{105}$$

$$\leq \sum_{b=b_1}^{B-1} I(S_b^n; Y^{nB}, W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, S_{b+1}^n, \dots, S_B^n|C) + n \cdot \log_2 |\mathcal{S}| \tag{106}$$

$$= \sum_{b=b_1}^{B-1} I(S_b^n; W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b, M_{b+1}|C) + n \cdot \log_2 |\mathcal{S}|$$

$$+ \sum_{b=b_1}^{B-1} I(S_b^n; S_{b+1}^n, \dots, S_B^n|W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b, M_{b+1}, C) \tag{107}$$

$$+ \sum_{b=b_1}^{B-1} I(S_b^n; Y^{nB}|W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b, M_{b+1}, S_{b+1}^n, \dots, S_B^n, C) \tag{108}$$

$$= \sum_{b=b_1}^{B-1} I(S_b^n; W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}|C) + \sum_{b=b_1}^{B-1} I(S_b^n; Y_b^n|W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, C) + n \cdot \log_2 |\mathcal{S}|. \tag{109}$$

The term at line (107) is equal to zero since the strictly causal encoding and the i.i.d. property of the channel state, imply that the random variables $(S_{b+1}^n, \dots, S_B^n)$ are independent of the message $M_{b+1}$ and of the random variables $(S_b^n, W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b)$ of the previous block $b$, and the term at line (108) is equal to zero since in the encoding process, the sequence $S_b^n$ only affects the choice of the bin index $L_b$ and of the sequences $(W_{1,b}^n, W_{2,b}^n, Y_b^n)$ of the current block $b$. This induces the following Markov chain: $S_b^n \multimap (W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b, C) \multimap (Y^{nB}, M_{b+1}, S_{b+1}^n, \dots, S_B^n)$, that is valid for each block $b \in \{1, \dots, B-1\}$. The sequence $S_b^n$ is correlated with the random variables of the other blocks $b' \neq b$ only through $(W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b, C)$.

For each block $b \in \{1, \dots, B-1\}$, the first term in equation (109) satisfies:

$$I(S_b^n; W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}|C) = I(S_b^n; W_{2,b}^n, L_b|W_{1,b}^n, M_{b+1}, C) \tag{110}$$

$$\leq H(W_{2,b}^n, L_b|W_{1,b}^n, M_{b+1}, C) \leq \log_2 |\mathcal{M}_{\mathsf{L}}| + H(W_{2,b}^n|W_{1,b}^n, L_b, M_{b+1}, C) \tag{111}$$

$$\leq \log_2 |\mathcal{M}_{\mathsf{L}}| + \log_2 |\mathcal{M}_{\mathsf{K}}| = n\Big(\mathsf{E} - I(S; W_1, W_2, Y) - 2\varepsilon + I(S; W_2|W_1) + \varepsilon\Big) \tag{112}$$

$$= n\Big(\mathsf{E} - I(S; Y|W_1, W_2) - \varepsilon\Big), \tag{113}$$

where (110) comes from the strictly causal encoding that induces the independence between the sequence auxiliary random variables $W_{1,b}^n$ and the sequence of channel states $S_b^n$, in block $b \in \{1, \dots, B-1\}$.

Hence $S_b^n$ is independent of $(W_{1,b}^n, M_{b+1}, C)$; (111) comes from the cardinality of the set of indexes $\mathcal{M}_{\mathsf{L}}$; (112) comes from the coding scheme described in Section A. By considering a fixed sequence $W_{1,b}^n$ and fixed indexes $(L_b, M_{b+1})$, the encoder chooses an index $K_{b+1} \in \mathcal{M}_{\mathsf{K}}$ corresponding to the sequence $W_{2,b}^n$. Hence, the sequence $W_{2,b}^n$ belongs to the bin of cardinality $|\mathcal{M}_{\mathsf{K}}|$. The rate parameters are given by $\mathsf{R}_{\mathsf{L}} = \mathsf{E} - I(S; W_1, W_2, Y) - 2\varepsilon$ and $\mathsf{R}_{\mathsf{K}} = I(W_2; S|W_1) + \varepsilon$; (113) comes from the independence between $W_1$ and $S$ that induces $I(S; W_2|W_1) = I(S; W_1, W_2)$.

For each block $b \in \{1, \ldots, B-1\}$, we introduce the random event of error $E_b \in \{0, 1\}$ defined with respect to the achievable joint probability distribution $\mathcal{Q}_{SXW_1W_2YV}$, as follows:

$$E_b = \begin{cases} 0 \text{ if } & (S_b^n, X_b^n, W_{1,b}^n, W_{2,b}^n, Y_b^n, V_b^n) \in T_\delta(\mathcal{Q}) \text{ and } (\hat{M}_{b+1}, \hat{L}_b, \hat{K}_{b+1}) = (M_{b+1}, L_b, K_{b+1}), \\ 1 \text{ if } & (S_b^n, X_b^n, W_{1,b}^n, W_{2,b}^n, Y_b^n, V_b^n) \notin T_\delta(\mathcal{Q}) \text{ or } (\hat{M}_{b+1}, \hat{L}_b, \hat{K}_{b+1}) \neq (M_{b+1}, L_b, K_{b+1}). \end{cases} \tag{114}$$

The second term in equation (109) satisfies:

$$\begin{aligned}
&I(S_b^n; Y_b^n | W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, C) \\
=&H(Y_b^n | W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, C) - H(Y_b^n | S_b^n, W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, C) && (115)\\
=&H(Y_b^n | W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, C) - n \cdot H(Y|W_1, W_2, S) && (116)\\
\leq&H(Y_b^n | W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, C, E_b = 0) + 1 + \mathbb{P}(E_b = 1) \cdot n \cdot \log_2 |\mathcal{Y}| - n \cdot H(Y|W_1, W_2, S) && (117)\\
\leq&n\Big(H(Y|W_1, W_2) + \varepsilon\Big) + 1 + \mathbb{P}(E_b = 1) \cdot n \cdot \log_2 |\mathcal{Y}| - n \cdot H(Y|W_1, W_2, S) && (118)\\
=&n\Big(I(S; Y|W_1, W_2) + \varepsilon + \frac{1}{n} + \mathbb{P}(E_b = 1) \cdot \log_2 |\mathcal{Y}|\Big), && (119)
\end{aligned}$$

where (115) comes from the properties of the mutual information; (116) comes from the cascade of memoryless channels $\mathcal{Q}_{X|W_1S}\mathcal{T}_{Y|XS}$ of the coding scheme $C$ coding scheme described in Section A, that implies $H(Y_b^n|S_b^n, W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, C) = n \cdot H(Y|W_1, S) = n \cdot H(Y|W_1, W_2, S)$; (117) is inspired by the proof of Fano's inequality, stated pp. 19, in [34]; (118) comes from the bound on the cardinality of the set of sequences $y^n \in T_\delta(w_1^n, w_2^n)$ that are jointly typical with sequences $(w_1^n, w_2^n)$, as mentioned pp. 26, in [34]. This is possible since no error occurs $E_b = 0$; (119) comes from the properties of the mutual information.

Hence we have the following upper bound on the leakage rate:

$$\begin{aligned}
&n \cdot B \cdot \mathbb{E}_c\Big[\mathcal{L}_{\mathsf{e}}(c)\Big] \leq \sum_{b=b_1}^{B-1} I(S_b^n; W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}|C) + \sum_{b=b_1}^{B-1} I(S_b^n; Y_b^n | W_{1,b}^n, W_{2,b}^n, L_b, M_{b+1}, C) + n \cdot \log_2 |\mathcal{S}| \\
\leq&n \cdot B \cdot \left(\mathsf{E} - I(S; Y|W_1, W_2) - \varepsilon + I(S; Y|W_1, W_2) + \varepsilon + \frac{1}{n} + \mathbb{P}(E_b = 1) \cdot \log_2 |\mathcal{Y}| + \frac{1}{B} \cdot \log_2 |\mathcal{S}|\right)
\end{aligned}$$

$$\leq n \cdot B \cdot \left( \mathsf{E} + \frac{1}{n} + \mathbb{P}(E_b = 1) \cdot \log_2 |\mathcal{Y}| + \frac{1}{B} \cdot \log_2 |\mathcal{S}| \right).$$

**Lower bound.** We provide a lower bound on the expected state leakage rate.

$$n \cdot B \cdot \mathbb{E}_c \left[ \mathcal{L}_{\mathsf{e}}(c) \right] = I(S^{nB}; Y^{nB}|C) = n \cdot B \cdot H(S) - H(S^{nB}|Y^{nB}, C) \tag{120}$$

$$\geq n \cdot B \cdot H(S) - \mathbb{P}(E_b = 1) \cdot n \cdot B \cdot \log_2 |\mathcal{S}| - 1 - H(S^{nB}|Y^{nB}, C, E_b = 0) \tag{121}$$

$$\geq n \cdot B \cdot H(S) - \mathbb{P}(E_b = 1) \cdot n \cdot B \cdot \log_2 |\mathcal{S}| - 1 - \sum_{b=b_1}^{B-1} H(S_b^n|Y^{nB}, S_{b+1}^n, \ldots, S_B^n, C, E_b = 0) - n \cdot \log_2 |\mathcal{S}| \tag{122}$$

$$\begin{aligned} =& n \cdot B \cdot H(S) - \left( \mathbb{P}(E_b = 1) \cdot B + 1 \right) \cdot n \cdot \log_2 |\mathcal{S}| - 1 \\ & - \sum_{b=b_1}^{B-1} H(S_b^n|W_{1,b}^n, W_{2,b}^n, L_b, Y^{nB}, S_{b+1}^n, \ldots, S_B^n, C, E_b = 0) \\ & - \sum_{b=b_1}^{B-1} I(S_b^n; W_{1,b}^n, W_{2,b}^n, L_b|Y^{nB}, S_{b+1}^n, \ldots, S_B^n, C, E_b = 0) \end{aligned} \tag{123}$$

$$= n \cdot B \cdot H(S) - \left( \mathbb{P}(E_b = 1) \cdot B + 1 \right) \cdot n \cdot \log_2 |\mathcal{S}| - 1 - \sum_{b=b_1}^{B-1} H(S_b^n|W_{1,b}^n, W_{2,b}^n, L_b, Y^{nB}, S_{b+1}^n, \ldots, S_B^n, C, E_b = 0) \tag{124}$$

$$\geq n \cdot B \cdot H(S) - \left( \mathbb{P}(E_b = 1) \cdot B + 1 \right) \cdot n \cdot \log_2 |\mathcal{S}| - 1 - \sum_{b=b_1}^{B-1} H(S_b^n|W_{1,b}^n, W_{2,b}^n, L_b, Y_b^n, E_b = 0), \tag{125}$$

where (120) comes from the i.i.d. property of the channel states $S$; (121) is inspired by the proof of Fano's inequality, stated pp. 19, in [34]; (124) comes from the non-error event $E_b = 0$, that implies for all block $b \in \{1, \ldots, B-1\}$, the sequences $(W_{1,b}^n, W_{2,b}^n, L_b)$ are correctly decoded based on the observation of $Y^{nB}$. Hence we have $I(S_b^n; W_{1,b}^n, W_{2,b}^n, L_b|Y^{nB}, S_{b+1}^n, \ldots, S_B^n, C, E_b = 0) = 0$; (125) comes from removing the conditioning over the sequences $(Y_{b_1}^n, \ldots, Y_{b-1}^n, Y_{b+1}^n, \ldots, Y_B^n, S_{b+1}^n, \ldots, S_B^n, C)$ and the random code in the conditional entropy $H(S_b^n|W_{1,b}^n, W_{2,b}^n, L_b, Y_b^n, E_b = 0)$.

In order to provide an upper bound on $H(S_b^n|W_{1,b}^n, W_{2,b}^n, L_b, Y_b^n, E_b = 0)$, we fix an index $l \in \mathcal{M}_{\mathsf{L}}$, some typical sequences $(w_1^n, w_2^n, y^n) \in T_\delta$. We consider the set $\mathcal{S}^\star(w_1^n, w_2^n, y^n, l)$ of sequences $s^n \in \mathcal{S}^n$ that are jointly typical with $(w_1^n, w_2^n, y^n)$ and that belong to the bin with index $l \in \mathcal{M}_{\mathsf{L}}$, denoted by $\mathcal{B}(l)$.

$$\mathcal{S}^\star(w_1^n, w_2^n, y^n, l) = \left\{ s^n \in \mathcal{S}^n, \text{ s.t. } \left\{ (s^n, w_1^n, w_2^n, y^n) \in T_\delta \right\} \cap \left\{ s^n \in \mathcal{B}(l) \right\} \right\} \tag{126}$$

Since the code $C$ is random, the above set $\mathcal{S}^\star(w_1^n, w_2^n, y^n, l)$ is a random set and its expected cardinality satisfies:

$$\mathbb{E}_c\left[\left|\mathcal{S}^\star(W_1^n, W_2^n, Y^n, L)\right|\right] = \mathbb{E}_c\left[\left|\left\{s^n \in \mathcal{S}^n, \text{ s.t. } \left\{(S^n, W_1^n, W_2^n, Y^n) \in T_\delta\right\} \cap \left\{S^n \in \mathcal{B}(L)\right\}\right\}\right|\right] \tag{127}$$

$$= \sum_{S^n \in T_\delta(W_1^n, W_2^n, Y^n)} \mathbb{E}_c\left[\mathbb{1}\left\{S^n \in \mathcal{B}(L)\right\}\right] \tag{128}$$

$$\leq \sum_{S^n \in T_\delta(W_1^n, W_2^n, Y^n)} 2^{-n\cdot \mathsf{R_L}} \leq 2^{n\cdot(H(S|W_1,W_2,Y) - \mathsf{R_L} + \varepsilon)}. \tag{129}$$

Equation (129) comes from the definition of the random code that induces a uniform probability distribution over the bins $\mathcal{B}(l)$ and the properties of the typical sequences stated pp. 27, in [34].

By Markov's inequality, we have the following equation that is valid for all index $l \in \mathcal{M}_L$ and for all typical sequences $(w_1^n, w_2^n, y^n) \in T_\delta$:

$$\mathbb{P}\left[\left|\mathcal{S}^\star(W_1^n, W_2^n, Y^n, L)\right| \geq 2^{n\cdot(H(S|W_1,W_2,Y) - \mathsf{R_L} + 2\varepsilon)}\right] \leq \frac{\mathbb{E}_c\left[\left|\mathcal{S}^\star(W_1^n, W_2^n, Y^n, L)\right|\right]}{2^{n\cdot(H(S|W_1,W_2,Y) - \mathsf{R_L} + 2\varepsilon)}} \tag{130}$$

$$\leq \frac{2^{n\cdot(H(S|W_1,W_2,Y) - \mathsf{R_L} + \varepsilon)}}{2^{n\cdot(H(S|W_1,W_2,Y) - \mathsf{R_L} + 2\varepsilon)}} \leq 2^{-n\varepsilon} \leq \varepsilon. \tag{131}$$

The last equation is valid for $n \geq n_5$. For each block $b \in \{1, \ldots, B-1\}$ we introduce the following random event:

$$F_b = \begin{cases} 1 & \text{if } \left|\mathcal{S}^\star(W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b)\right| \geq 2^{n\cdot(H(S|W_1,W_2,Y) - \mathsf{R_L} + 2\varepsilon)}, \\ 0 & \text{if } \left|\mathcal{S}^\star(W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b)\right| < 2^{n\cdot(H(S|W_1,W_2,Y) - \mathsf{R_L} + 2\varepsilon)}. \end{cases} \tag{132}$$

For each block $b \in \{b_1, \ldots, B-1\}$ we have:

$$H(S_b^n | W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b, E_b = 0)$$

$$\leq 1 + \mathcal{P}(F_b = 1) \cdot n \cdot \log_2 |\mathcal{S}| + H(S_b^n | W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b, E_b = 0, F_b = 0) \tag{133}$$

$$\leq 1 + \varepsilon \cdot n \cdot \log_2 |\mathcal{S}| + H(S_b^n | W_{1,b}^n, W_{2,b}^n, Y_b^n, L_b, E_b = 0, F_b = 0) \tag{134}$$

$$\leq 1 + \varepsilon \cdot n \cdot \log_2 |\mathcal{S}| + \log_2 2^{n\cdot(H(S|W_1,W_2,Y) - \mathsf{R_L} + 2\varepsilon)} \tag{135}$$

$$= n \cdot \left(H(S|W_1, W_2, Y) - \mathsf{R_L} + \frac{1}{n} + \varepsilon \cdot \log_2 |\mathcal{S}| + 2\varepsilon\right) \tag{136}$$

$$= n \cdot \left(H(S|W_1, W_2, Y) - \mathsf{E} + I(S; W_1, W_2, Y) + 2\varepsilon + \frac{1}{n} + \varepsilon \cdot \log_2 |\mathcal{S}| + 2\varepsilon\right) \tag{137}$$

$$= n \cdot \left(H(S) - \mathsf{E} + \frac{1}{n} + \varepsilon \cdot \log_2 |\mathcal{S}| + 4\varepsilon\right), \tag{138}$$

where (133) is inspired by the proof of Fano's inequality, stated pp. 19, in [34]; (134) comes from equation (131), that corresponds to $\mathcal{P}(F_b = 1) \leq \varepsilon$; (135) comes from the definition of event $F_b = 0$, that implies the set $\mathcal{S}^\star(w_1^n, w_2^n, y^n, l)$ has cardinality bounded by $2^{n\cdot(H(S|W_1,W_2,Y)-\mathsf{R_L}+2\varepsilon)}$; (137) comes from the definition of the rate $\mathsf{R_L} = \mathsf{E} - I(S; W_1, W_2, Y) - 2\varepsilon$ stated in equation (61).

This provides the following lower bound:

$$
\begin{aligned}
& n \cdot B \cdot \mathbb{E}_c\Big[\mathcal{L}_{\mathsf{e}}(c)\Big] \\
\geq & n \cdot B \cdot H(S) - \Big(\mathbb{P}(E_b = 1) \cdot B + 1\Big) \cdot n \cdot \log_2 |\mathcal{S}| - 1 - \sum_{b=b_1}^{B-1} H(S_b^n | W_{1,b}^n, W_{2,b}^n, Y_b^n, E_b = 0) && (139) \\
\geq & n \cdot B \cdot H(S) - \Big(\mathbb{P}(E_b = 1) \cdot B + 1\Big) \cdot n \cdot \log_2 |\mathcal{S}| - 1 - (B-1) \cdot n \cdot \left(H(S) - \mathsf{E} + \frac{1}{n} + \varepsilon \cdot \log_2 |\mathcal{S}| + 4\varepsilon\right) && (140) \\
\geq & n \cdot B \cdot H(S) - \Big(\mathbb{P}(E_b = 1) \cdot B + 1\Big) \cdot n \cdot \log_2 |\mathcal{S}| - 1 - B \cdot n \cdot \left(H(S) - \mathsf{E} + \frac{1}{n} + \varepsilon \cdot \log_2 |\mathcal{S}| + 4\varepsilon\right) - n \cdot \log_2 |\mathcal{S}| && (141) \\
\geq & n \cdot B \cdot \left(\mathsf{E} - \left(\mathbb{P}(E_b = 1) + \varepsilon + \frac{2}{B}\right) \cdot \log_2 |\mathcal{S}| - \frac{2}{n \cdot B} - 4\varepsilon\right), && (142)
\end{aligned}
$$

where (139) comes from equation (125); (140) comes from equation (138); (141) comes the lower bound:

$$
n \cdot \left(H(S) - \mathsf{E} + \frac{1}{n} + \varepsilon \cdot \log_2 |\mathcal{S}| + 4\varepsilon\right) \geq -n \cdot \mathsf{E} \geq -n \cdot \log_2 |\mathcal{S}|. \tag{143}
$$

Equation (142) provides the lower bound on the expected state leakage rate.

**Conclusion for Exact state leakage Rate:** We introduce the error parameters $\varepsilon_1$ and $\varepsilon_2$ defined by equations (145) and (144). The upper and lower bound on the expected state leakage rate are given by:

$$
\begin{aligned}
\mathbb{E}_c\Big[\mathcal{L}_{\mathsf{e}}(c)\Big] \leq & \mathsf{E} + \frac{1}{n} + \max_b \mathbb{P}(E_b = 1) \cdot \log_2 |\mathcal{Y}| + \frac{1}{B} \cdot \log_2 |\mathcal{S}| = \mathsf{E} + \varepsilon_2, && (144) \\
\mathbb{E}_c\Big[\mathcal{L}_{\mathsf{e}}(c)\Big] \geq & \mathsf{E} - \left(\max_b \mathbb{P}(E_b = 1) + \varepsilon + \frac{2}{B}\right) \cdot \log_2 |\mathcal{S}| - \frac{2}{n \cdot B} - 4\varepsilon = \mathsf{E} - \varepsilon_1. && (145)
\end{aligned}
$$

Hence, the expected state leakage rate satisfies:

$$
\left|\mathbb{E}_c\Big[\mathcal{L}_{\mathsf{e}}(c)\Big] - \mathsf{E}\right| \leq \varepsilon_1 + \varepsilon_2. \tag{146}
$$

### G. Conclusion of the achievability proof of Theorem II.3

From equations (69), (103) and (146), we obtain the following equations with $\varepsilon_1$ and $\varepsilon_2$ defined by (144) and (145):

$$\frac{1}{n \cdot B} \cdot \sum_{b=2}^{B} \log_2 |\mathcal{M}| \geq \mathsf{R} - \frac{1}{B} \cdot \log_2 |\mathcal{Y}|, \tag{147}$$

$$\mathbb{E}_c\Big[\mathcal{P}_{\mathsf{e}}(c)\Big] \leq 2 - 2 \cdot \left(1 - 4\varepsilon\right)^{B-1}, \tag{148}$$

$$\left|\mathbb{E}_c\Big[\mathcal{L}_{\mathsf{e}}(c)\Big] - \mathsf{E}\right| \leq \varepsilon_1 + \varepsilon_2. \tag{149}$$

By choosing the appropriate number of blocks $B$, error probability $\max_b \mathbb{P}(E_b = 1)$, length of blocks $n$ and parameter for typical sequences $\varepsilon$, we prove that there exists a code $c^\star \in \mathcal{C}(n \cdot B, \mathcal{M})$ such that:

$$\frac{1}{n \cdot B} \cdot \sum_{b=2}^{B} \log_2 |\mathcal{M}| \geq \mathsf{R} - \varepsilon_3, \tag{150}$$

$$\mathcal{P}_{\mathsf{e}}(c^\star) \leq \varepsilon_3, \tag{151}$$

$$\left|\mathcal{L}_{\mathsf{e}}(c^\star) - \mathsf{E}\right| \leq \varepsilon_3. \tag{152}$$

with $\varepsilon_3 = \frac{1}{B} \cdot \log_2 |\mathcal{Y}| + 2 - 2 \cdot \left(1 - 4\varepsilon\right)^{B-1} + \varepsilon_1 + \varepsilon_2$. This concludes the achievability proof of Theorem II.3.

## APPENDIX B
## CONVERSE PROOF OF THEOREM II.3

### A. Information Constraints

Consider that the triple of rate, state leakage and probability distribution $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable with a causal code. We introduce the random event of error $E \in \{0, 1\}$ defined with respect to the achievable joint probability distribution $\mathcal{Q}_{SXYV}$, as follows:

$$E = \begin{cases} 0 \text{ if } & \big|\big|Q^n - \mathcal{Q}\big|\big|_1 \leq \varepsilon \iff (S^n, X^n, Y^n, V^n) \in T_\delta(\mathcal{Q}), \\ 1 \text{ if } & \big|\big|Q^n - \mathcal{Q}\big|\big|_1 > \varepsilon \iff (S^n, X^n, Y^n, V^n) \notin T_\delta(\mathcal{Q}). \end{cases} \tag{153}$$

The event $E = 1$ occurs if the sequences $(S^n, X^n, Y^n, V^n) \notin T_\delta(\mathcal{Q})$ are not jointly typical for the target probability distribution $\mathcal{Q}$. By definition II.1, for all $\varepsilon > 0$, there exists a $\bar{n}$ such that for all $n \geq \bar{n}$, there exists a code $c^\star \in \mathcal{C}(n, \mathcal{M})$ that satisfies the three following equations:

$$\frac{\log_2 |\mathcal{M}|}{n} \geq \mathsf{R} - \varepsilon, \tag{154}$$

$$\left| \mathcal{L}_{\mathsf{e}}(c^\star) - \mathsf{E} \right| = \left| \frac{1}{n} \cdot I(S^n; Y^n) - \mathsf{E} \right| \leq \varepsilon, \tag{155}$$

$$\mathcal{P}_{\mathsf{e}}(c^\star) = \mathbb{P}\left( M \neq \hat{M} \right) + \mathbb{P}\left( \left| \left| Q^n - \mathcal{Q} \right| \right|_1 \geq \varepsilon \right) \leq \varepsilon. \tag{156}$$

We introduce the auxiliary random variables $W_{1,i} = (M, S^{i-1})$ and $W_{2,i} = Y_{i+1}^n$ that satisfy the Markov chains of the set of probability distribution $\mathbb{Q}_{\mathsf{e}}$ for all $i \in \{1, \ldots, n\}$:

$$S_i \text{ independent of } W_{1,i}, \tag{157}$$

$$X_i \multimap (S_i, W_{1,i}) \multimap W_{2,i}, \tag{158}$$

$$Y_i \multimap (X_i, S_i) \multimap (W_{1,i}, W_{2,i}), \tag{159}$$

$$V_i \multimap (Y_i, W_{1,i}, W_{2,i}) \multimap (S_i, X_i), \tag{160}$$

where (157) comes from the i.i.d. property of the source that induces the independence between $S_i$ and $(M, S^{i-1}) = W_{1,i}$; (158) comes from Lemma 2. It is a direct consequence of the causal encoding function and the memoryless property of the channel; (159) comes from the memoryless property of the channel $\mathcal{T}_{Y|XS}$; (160) comes from Lemma 3. It is a direct consequence of the causal encoding function, the non-causal decoding function and the memoryless property of the channel $\mathcal{T}_{Y|XS}$.

We introduce the random variable $T$ that is uniformly distributed over the indices $\{1, \ldots, n\}$ and the corresponding mean random variables $W_{1,T}$, $W_{2,T}$, $S_T$, $X_T$, $Y_T$, $V_T$. The auxiliary random variables $W_1 = (W_{1,T}, T)$ and $W_2 = W_{2,T}$ belong to the set of probability distributions $\mathbb{Q}_{\mathsf{e}}$ and satisfy the three information constraints of Theorem II.3:

$$I(S; W_1, W_2, Y) \leq \mathsf{E} \leq H(S), \tag{161}$$

$$\mathsf{R} + \mathsf{E} \leq I(W_1, S; Y). \tag{162}$$

**First Constraint:**

$$n \cdot \mathsf{E} \geq I(S^n; Y^n) - n \cdot \varepsilon \geq \sum_{i=1}^n I(S_i; Y^n, M | S^{i-1}) - H(M|Y^n) - n \cdot \varepsilon \tag{163}$$

$$\geq \sum_{i=1}^n I(S_i; Y^n, M | S^{i-1}) - n \cdot 2\varepsilon \tag{164}$$

$$= \sum_{i=1}^n I(S_i; Y^n, M, S^{i-1}) - n \cdot 2\varepsilon \geq \sum_{i=1}^n I(S_i; Y_{i+1}^n, M, S^{i-1}, Y_i) - n \cdot 2\varepsilon \tag{165}$$

$$= \sum_{i=1}^{n} I(S_i; W_{1,i}, W_{2,i}, Y_i) - n \cdot 2\varepsilon \tag{166}$$

$$= n \cdot I(S_T; W_{1,T}, W_{2,T}, Y_T|T) - n \cdot 2\varepsilon \tag{167}$$

$$= n \cdot I(S_T; W_{1,T}, W_{2,T}, Y_T, T) - n \cdot 2\varepsilon \tag{168}$$

$$= n \cdot \Big( I(S_T; W_1, W_2, Y_T) - 2\varepsilon \Big) \tag{169}$$

$$\geq n \cdot \Big( I(S_T; W_1, W_2, Y_T|E=0) - 3\varepsilon \Big) \tag{170}$$

$$\geq n \cdot \Big( I(S; W_1, W_2, Y) - 4\varepsilon \Big), \tag{171}$$

where (163) comes from the definition of achievable state leakage rate $\mathsf{E}$, stated in equation (155); (164) comes from equation (156) and Fano's inequality, stated pp. 19, in [34]; (165) comes from the i.i.d. property of the channel states that implies $S_i$ is independent of $S^{i-1}$; (166) comes from the introduction of the auxiliary random variables $W_{1,i} = (M, S^{i-1})$ and $W_{2,i} = Y_{i+1}^n$, for all $i \in \{1, \ldots, n\}$; (167) comes from the introduction of the uniform random variable $T$ over $\{1, \ldots, n\}$ and the corresponding mean random variables $S_T$, $W_{1,T}$, $W_{2,T}$, $Y_T$; (168) comes from the independence between $T$ and $S_T$; (169) comes from identifying the auxiliary random variables$W_1 = (W_{1,T}, T)$ and $W_2 = W_{2,T}$; (170) comes from the empirical coordination requirement as stated in Lemma 6. The sequences of symbols $(S^n, X^n, Y^n, V^n)$ are not jointly typical with small error probability $\mathbb{P}(E = 1)$; (171) comes from Lemma 7 since the probability distribution $\mathcal{P}_{S_T X_T Y_T V_T|E=0}$ is closed to the target probability distribution $\mathcal{Q}_{SXYV}$. The continuity of the entropy function stated pp. 33 in [32] concludes.

**Second Constraint:**

$$n \cdot \mathsf{E} \leq I(S^n; Y^n) + n \cdot \varepsilon \leq H(S^n) + n \cdot \varepsilon \tag{172}$$

$$= n \cdot \Big( H(S) + \varepsilon \Big), \tag{173}$$

where (172) comes from the definition of the achievable state leakage rate $\mathsf{E}$, stated in equation (155); (173) comes from the i.i.d. property of the channel states $S$.

**Third Constraint:**

$$n \cdot \Big( \mathsf{E} + \mathsf{R} \Big) \leq I(S^n; Y^n) + H(M) + n \cdot 2\varepsilon \tag{174}$$

$$= I(S^n; Y^n) + I(M; Y^n) + H(M|Y^n) + n \cdot 2\varepsilon \tag{175}$$

$$\leq I(S^n; Y^n) + I(M; Y^n) + n \cdot 3\varepsilon \tag{176}$$

$$\leq I(S^n; Y^n) + I(M; Y^n|S^n) + n \cdot 3\varepsilon \tag{177}$$

$$\leq I(S^n, M; Y^n) + n \cdot 3\varepsilon \tag{178}$$

$$= \sum_{i=1}^{n} I(S^n, M; Y_i | Y_{i+1}^n) + n \cdot 3\varepsilon \tag{179}$$

$$\leq \sum_{i=1}^{n} I(S^n, M, Y_{i+1}^n; Y_i) + n \cdot 3\varepsilon \tag{180}$$

$$= \sum_{i=1}^{n} I(S_i, M, S^{i-1}; Y_i) + \sum_{i=1}^{n} I(S_{i+1}^n, Y_{i+1}^n; Y_i | S_i, M, S^{i-1}) + n \cdot 3\varepsilon \tag{181}$$

$$= \sum_{i=1}^{n} I(S_i, M, S^{i-1}; Y_i) + n \cdot 3\varepsilon \tag{182}$$

$$= \sum_{i=1}^{n} I(S_i, W_{1,i}; Y_i) + n \cdot 3\varepsilon \tag{183}$$

$$= n \cdot I(S_T, W_{1,T}; Y_T | T) + n \cdot 3\varepsilon \tag{184}$$

$$\leq n \cdot I(S_T, W_{1,T}, T; Y_T) + n \cdot 3\varepsilon \tag{185}$$

$$\leq n \cdot \Big( I(S_T, W_1; Y_T) + 3\varepsilon \Big) \tag{186}$$

$$\leq n \cdot \Big( I(S_T, W_1; Y_T | E = 0) + 4\varepsilon \Big) \tag{187}$$

$$\leq n \cdot \Big( I(S, W_1; Y) + 5\varepsilon \Big), \tag{188}$$

where (174) comes from the definition of achievable rate and information leakage $(\mathsf{R}, \mathsf{E})$, stated in equations (154) and (155); (176) comes from equation (156) and Fano's inequality, stated pp. 19, in [34]; (177) comes from the independence between the message $M$ and the channel states $S^n$, hence $I(M; Y^n) \leq I(M; Y^n, S^n) = I(M; Y^n | S^n)$; (178), (179), (180), (181) comes from the properties of the mutual information; (182) comes from the Markov chain $Y_i \multimap (S_i, M, S^{i-1}) \multimap (S_{i+1}^n, Y_{i+1}^n)$, stated in Lemma 4. It is a direct consequence of the causal encoding function and the memoryless property of the channel; (183) comes from the introduction of the auxiliary random variable $W_{1,i} = (M, S^{i-1})$, for all $i \in \{1, \ldots, n\}$; (184) comes from the introduction of the uniform random variable $T$ over $\{1, \ldots, n\}$ and the corresponding mean random variables $S_T$, $W_{1,T}$, $Y_T$; (186) comes from identifying the auxiliary random variables $W_1 = (W_{1,T}, T)$ and $W_2 = W_{2,T}$; (187) comes from the empirical coordination requirement as stated in Lemma 5. The sequences of symbols $(S^n, X^n, Y^n, V^n)$ are not jointly typical with small error probability $\mathbb{P}(E = 1)$; (188) comes from Lemma 7. The sequences of symbols $(S^n, X^n, Y^n, V^n)$ are jointly typical, hence the distribution of the mean random variables $\mathcal{P}_{S_T X_T Y_T V_T | E=0}$ is closed to the target probability distribution $\mathcal{Q}_{SXYV}$. The continuity of the entropy function stated pp. 33 in [32] concludes.

**Conclusion:** If the triple of rate, state leakage and probability distribution$(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable with

causal encoding, then the following equations are satisfied for all $\varepsilon > 0$:

$$I(S; W_1, W_2, Y) - 4\varepsilon \leq \mathsf{E} \leq H(S) + \varepsilon, \tag{189}$$

$$\mathsf{R} + \mathsf{E} \leq I(S, W_1; Y) + 5\varepsilon. \tag{190}$$

This corresponds to equations (10), (7) and (8) and concludes the converse proof of Theorem II.3.

**Remark B.1** *For the converse proof of Theorem II.3, the causal encoding is not necessarily deterministic. The same optimal performances can be obtained by considering stochastic causal encoding.*

*B. Lemmas for the converse proof with empirical coordination*

**Lemma 2** *The causal encoding function and the memoryless property of the channel induce the Markov chain property corresponding to equation* (158)*:*

$$X_i \multimap (S_i, W_{1,i}) \multimap W_{2,i}. \tag{191}$$

*This Markov chain is satisfied with* $W_{1,i} = (M, S^{i-1})$ *and* $W_{2,i} = Y_{i+1}^n$*, for all* $i \in \{1, \ldots, n\}$*.*

*Proof.* [Lemma 2] The auxiliary random variables $W_{1,i} = (M, S^{i-1})$ and $W_{2,i} = Y_{i+1}^n$ satisfy the following equations for all $(s^n, x^n, w_1^n, w_2^n, y^n, m)$:

$$\begin{aligned}
&\mathcal{P}(w_{2,i}|s_i, w_{1,i}, x_i) \\
=&\mathcal{P}(y_{i+1}^n|s_i, m, s^{i-1}, x_i) && (192)\\
=&\sum_{s_{i+1}^n, x_{i+1}^n} \mathcal{P}(s_{i+1}^n, x_{i+1}^n, y_{i+1}^n|s_i, m, s^{i-1}, x_i) && (193)\\
=&\sum_{s_{i+1}^n, x_{i+1}^n} \mathcal{P}(s_{i+1}^n, x_{i+1}^n|s_i, m, s^{i-1}, x_i) \cdot \mathcal{P}(y_{i+1}^n|s_{i+1}^n, x_{i+1}^n, s_i, m, s^{i-1}, x_i) && (194)\\
=&\sum_{s_{i+1}^n, x_{i+1}^n} \mathcal{P}(s_{i+1}^n, x_{i+1}^n|s_i, m, s^{i-1}) \cdot \mathcal{P}(y_{i+1}^n|s_{i+1}^n, x_{i+1}^n, s_i, m, s^{i-1}, x_i) && (195)\\
=&\sum_{s_{i+1}^n, x_{i+1}^n} \mathcal{P}(s_{i+1}^n, x_{i+1}^n|s_i, m, s^{i-1}) \cdot \mathcal{P}(y_{i+1}^n|s_{i+1}^n, x_{i+1}^n) && (196)\\
=&\sum_{s_{i+1}^n, x_{i+1}^n} \mathcal{P}(s_{i+1}^n, x_{i+1}^n, y_{i+1}^n|s_i, m, s^{i-1}) = \mathcal{P}(y_{i+1}^n|s_i, m, s^{i-1}) = \mathcal{P}(w_{2,i}|s_i, w_{1,i}), && (197)
\end{aligned}$$

where (195) comes from the causal encoding function that induces the Markov chain $X_i \multimap (S_i, M, S^{i-1}) \multimap (S_{i+1}^n, X_{i+1}^n)$; (196) comes from the memoryless property of the channel $Y_{i+1}^n \multimap (S_{i+1}^n, X_{i+1}^n) \multimap (S_i, M, S^{i-1}, X_i)$. This concludes the proof of Lemma 2. □

**Lemma 3** *The causal encoding function, the non-causal decoding function and the memoryless property of the channel induce the Markov chain property corresponding to equation* (160)*:*

$$V_i \multimap (Y_i, W_{1,i}, W_{2,i}) \multimap (S_i, X_i). \tag{198}$$

*This Markov chain is satisfied with* $W_{1,i} = (M, S^{i-1})$ *and* $W_{2,i} = Y_{i+1}^n$*, for all* $i \in \{1, \ldots, n\}$*.*

*Proof.* [Lemma 3] The auxiliary random variables $W_{1,i} = (M, S^{i-1})$ and $W_{2,i} = Y_{i+1}^n$ satisfy the following equations for all $(s^n, x^n, w_1^n, w_2^n, y^n, v^n, m)$:

$$\begin{aligned}
&\mathcal{P}(v_i|y_i, w_{1,i}, w_{2,i}, s_i, x_i) = \mathcal{P}(v_i|y_i, m, s^{i-1}, y_{i+1}^n, s_i, x_i) \\
&= \sum_{x^{i-1}, y^{i-1}} \mathcal{P}(v_i, x^{i-1}, y^{i-1}|y_i, m, s^{i-1}, y_{i+1}^n, s_i, x_i) \qquad (199) \\
&= \sum_{x^{i-1}, y^{i-1}} \mathcal{P}(x^{i-1}|y_i, m, s^{i-1}, y_{i+1}^n, s_i, x_i) \cdot \mathcal{P}(y^{i-1}|y_i, m, s^{i-1}, y_{i+1}^n, s_i, x_i, x^{i-1}) \\
&\quad \cdot \mathcal{P}(v_i|y_i, m, s^{i-1}, y_{i+1}^n, s_i, x_i, x^{i-1}, y^{i-1}). \qquad (200)
\end{aligned}$$

we can remove $(s_i, x_i)$, in the three conditional probability distributions:

$$\mathcal{P}(x^{i-1}|y_i, m, s^{i-1}, y_{i+1}^n, s_i, x_i) = \mathcal{P}(x^{i-1}|m, s^{i-1}), \tag{201}$$

$$\mathcal{P}(y^{i-1}|y_i, m, s^{i-1}, y_{i+1}^n, s_i, x_i, x^{i-1}) = \mathcal{P}(y^{i-1}|s^{i-1}, x^{i-1}), \tag{202}$$

$$\mathcal{P}(v_i|y_i, m, s^{i-1}, y_{i+1}^n, s_i, x_i, x^{i-1}, y^{i-1}) = \mathcal{P}(v_i|y_i, y_{i+1}^n, y^{i-1}), \tag{203}$$

where (201) comes from the causal encoding that induces the following Markov chain $X^{i-1} \multimap (M, S^{i-1}) \multimap (Y_i, Y_{i+1}^n, X_i, S_i)$; (202) comes from the memoryless property of the channel: $Y^{i-1}$ only depends on $(X^{i-1}, S^{i-1})$; (203) comes from the non-causal decoding that induces the following Markov chain $V_i \multimap (Y_i, Y_{i+1}^n, Y^{i-1}) \multimap (M, S^{i-1}, S_i, X_i, X^{i-1})$. Hence we have for all $(s^n, x^n, w_1^n, w_2^n, y^n, v^n, m)$:

$$\begin{aligned}
&\mathcal{P}(v_i|y_i, w_{1,i}, w_{2,i}, s_i, x_i) = \sum_{x^{i-1}, y^{i-1}} \mathcal{P}(v_i, x^{i-1}, y^{i-1}|y_i, m, s^{i-1}, y_{i+1}^n) \qquad (204) \\
&= \mathcal{P}(v_i|y_i, m, s^{i-1}, y_{i+1}^n) = \mathcal{P}(v_i|y_i, w_{1,i}, w_{2,i}). \qquad (205)
\end{aligned}$$

The above equation corresponds to the Markov chain $V_i \multimap (Y_i, W_{1,i}, W_{2,i}) \multimap (S_i, X_i)$ and it concludes the proof of Lemma 3. □

**Lemma 4** *The causal encoding function and the memoryless property of the channel induce the following Markov chain property:*

$$Y_i \multimap (S_i, M, S^{i-1}) \multimap (S_{i+1}^n, Y_{i+1}^n). \tag{206}$$

*This Markov chain is satisfied for all* $i \in \{1, \ldots, n\}$.

*Proof.* [Lemma 4] We have the following equations for all $(s^n, x^n, y^n, m)$:

$$\mathcal{P}(y_i|s_i, m, s^{i-1}, s_{i+1}^n, y_{i+1}^n) = \sum_{x_i} \mathcal{P}(x_i, y_i|s_i, m, s^{i-1}, s_{i+1}^n, y_{i+1}^n) \tag{207}$$

$$= \sum_{x_i} \mathcal{P}(x_i|s_i, m, s^{i-1}, s_{i+1}^n, y_{i+1}^n) \cdot \mathcal{P}(y_i|x_i, s_i, m, s^{i-1}, s_{i+1}^n, y_{i+1}^n) \tag{208}$$

$$= \sum_{x_i} \mathcal{P}(x_i|s_i, m, s^{i-1}) \cdot \mathcal{P}(y_i|x_i, s_i, m, s^{i-1}, s_{i+1}^n, y_{i+1}^n) \tag{209}$$

$$= \sum_{x_i} \mathcal{P}(x_i|s_i, m, s^{i-1}) \cdot \mathcal{P}(y_i|x_i, s_i) = \sum_{x_i} \mathcal{P}(x_i, y_i|s_i, m, s^{i-1}) = \mathcal{P}(y_i|s_i, m, s^{i-1}), \tag{210}$$

where (209) comes from the causal encoding function that induces the Markov chain $X_i \multimap (S_i, M, S^{i-1}) - \!\!\circ\!\!- (S_{i+1}^n, Y_{i+1}^n)$; (210) comes from the memoryless property of the channel that induces the Markov chain: $Y_i \multimap (X_i, S_i) \multimap (M, S^{i-1}, S_{i+1}^n, Y_{i+1}^n)$. This concludes the proof of Lemma 4. □

The random event of error $E \in \{0, 1\}$ of the following Lemmas is defined in equation (153). The event $E = 1$ occurs if the sequences $(S^n, X^n, Y^n, V^n) \notin T_\delta(\mathcal{Q})$ are not jointly typical for the target probability distribution $\mathcal{Q}$.

**Lemma 5** *Fix a probability distribution* $\mathcal{Q}_{S_T W_1 W_2 X_T Y_T V_T} \in \mathbb{Q}_{\mathsf{e}}$ *and suppose that the error probability* $\mathbb{P}(E = 1)$ *is small enough such that* $\mathbb{P}(E = 1) \cdot \log_2 |\mathcal{Y}| + 2 \cdot h_b\Big(\mathbb{P}(E = 1)\Big) \leq \varepsilon$. *Then we have:*

$$\begin{aligned} &I(W_1, W_2; Y_T) - I(S_T; W_2|W_1) \\ \leq &I(W_1, W_2; Y_T|E = 0) - I(S_T; W_2|W_1, E = 0) + \varepsilon. \end{aligned} \tag{211}$$

*Proof.* [Lemma 5]

$$I(W_1, W_2; Y_T) - I(S_T; W_2|W_1) \tag{212}$$

$$= I(W_1, W_2; Y_T|E) - I(S_T; W_2|W_1, E) \tag{213}$$

$$+ I(E; Y_T) - I(E; Y_T|W_1, W_2) + I(S_T; E|W_1, W_2) - I(S_T; E|W_1) \tag{214}$$

$$\leq I(W_1, W_2; Y_T|E) - I(S_T; W_2|W_1, E) + 2 \cdot H(E) \tag{215}$$

$$= \mathbb{P}(E = 0) \cdot \Big(I(W_1, W_2; Y_T|E = 0) - I(S_T; W_2|W_1, E = 0)\Big)$$

$$+ \mathbb{P}(E = 1) \cdot \Big(I(W_1, W_2; Y_T|E = 1) - I(S_T; W_2|W_1, E = 1)\Big) + 2 \cdot H(E) \tag{216}$$

$$= I(W_1, W_2; Y_T|E = 0) - I(S_T; W_2|W_1, E = 0) + \mathbb{P}(E = 1) \cdot \log_2 |\mathcal{Y}| + 2 \cdot h_b\Big(\mathbb{P}(E = 1)\Big) \tag{217}$$

$$\leq I(W_1, W_2; Y_T|E = 0) - I(S_T; W_2|W_1, E = 0) + \varepsilon. \tag{218}$$

The last equation holds since: $\mathbb{P}(E=1)\cdot\log_2|\mathcal{Y}|+2\cdot h_b\Big(\mathbb{P}(E=1)\Big)\leq\varepsilon$. This concludes the proof of Lemma 5. □

**Lemma 6** *Fix a probability distribution $\mathcal{Q}_{S_TW_1W_2X_TY_TV_T}\in\mathbb{Q}_{\mathsf{e}}$ and suppose that the error probability $\mathbb{P}(E=1)$ is small enough such that $\mathbb{P}(E=1)\cdot\log_2|\mathcal{S}|+h_b\Big(\mathbb{P}(E=1)\Big)\leq\varepsilon$. Then we have:*

$$I(S_T;W_1,W_2,Y_T)\geq I(S_T;W_1,W_2,Y_T|E=0)-\varepsilon. \tag{219}$$

*Proof.* [Lemma 6]

$$I(S_T;W_1,W_2,Y_T)=I(S_T;W_1,W_2,Y_T|E)+I(S_T;E)-I(S_T;E|W_1,W_2,Y_T) \tag{220}$$

$$\geq I(S_T;W_1,W_2,Y_T|E)-H(E) \tag{221}$$

$$=\mathbb{P}(E=0)\cdot I(S_T;W_1,W_2,Y_T|E=0)+\mathbb{P}(E=1)\cdot I(S_T;W_1,W_2,Y_T|E=1)-H(E) \tag{222}$$

$$=I(S_T;W_1,W_2,Y_T|E=0)-\mathbb{P}(E=1)\cdot I(S_T;W_1,W_2,Y_T|E=0) \tag{223}$$

$$+\mathbb{P}(E=1)\cdot I(S_T;W_1,W_2,Y_T|E=1)-h_b\Big(\mathbb{P}(E=1)\Big) \tag{224}$$

$$\geq I(S_T;W_1,W_2,Y_T|E=0)-\mathbb{P}(E=1)\cdot\log_2|\mathcal{S}|-h_b\Big(\mathbb{P}(E=1)\Big) \tag{225}$$

$$\geq I(S_T;W_1,W_2,Y_T|E=0)-\varepsilon. \tag{226}$$

The last equation holds since: $\mathbb{P}(E=1)\cdot\log_2|\mathcal{S}|+h_b\Big(\mathbb{P}(E=1)\Big)\leq\varepsilon$. This concludes the proof of Lemma 6. □

We denote by $\mathcal{P}_{S_TX_TY_TV_T|E=0}$ the probability distribution induced by the random variables $(S_T,X_T,Y_T,V_T)$ knowing the event $E=0$ is realized.

**Lemma 7** *Probability distribution defined by $\mathcal{P}_{S_TX_TY_TV_T|E=0}$ is closed to the target probability distribution $\mathcal{Q}_{SXYV}$:*

$$\Big|\mathcal{P}_{S_TX_TY_TV_T|E=0}(s,x,y,v)-\mathcal{Q}(s,x,y,v)\Big|\leq\varepsilon,\qquad\forall(s,x,y,v). \tag{227}$$

*Proof.* [Lemma 7] We evaluate the probability $\mathcal{P}_{S_T|E=0}(s)$ and we show it is closed to the desired

probability $\mathcal{P}(s)$ for all $s \in \mathcal{S}$:

$$\mathcal{P}_{S_T|E=0}(s) = \sum_{s^n \in T_\delta} \sum_{i=1}^{n} \mathbb{P}\big(S^n = s^n, T = i, S_T = s \big| E = 0\big) \tag{228}$$

$$= \sum_{s^n \in T_\delta} \sum_{i=1}^{n} \mathbb{P}\big(S^n = s^n \big| E = 0\big) \cdot \mathbb{P}\big(T = i \big| S^n = s^n, E = 0\big) \cdot \mathbb{P}\big(S_T = s \big| S^n = s^n, T = i, E = 0\big) \tag{229}$$

$$= \sum_{s^n \in T_\delta} \sum_{i=1}^{n} \mathbb{P}\big(S^n = s^n \big| E = 0\big) \cdot \mathbb{P}\big(T = i\big) \cdot \mathbb{P}\big(S_T = s \big| S^n = s^n, T = i, E = 0\big) \tag{230}$$

$$= \sum_{s^n \in T_\delta} \sum_{i=1}^{n} \mathbb{P}\big(S^n = s^n \big| E = 0\big) \cdot \mathbb{P}\big(T = i\big) \cdot \mathbb{1}_{\{s_T = s\}} \tag{231}$$

$$= \sum_{s^n \in T_\delta} \mathbb{P}\big(S^n = s^n \big| E = 0\big) \cdot \sum_{i=1}^{n} \frac{1}{n} \cdot \mathbb{1}_{\{s_T = s\}} = \sum_{s^n \in T_\delta} \mathbb{P}\big(S^n = s^n \big| E = 0\big) \cdot \frac{N(s|s^n)}{n}, \tag{232}$$

where (230) comes from the independence of event $\{T = i\}$ with events $\{S^n = s^n\}$ and $\{E = 0\}$; (232) comes from the definition of the number of occurrence $N(s|s^n) = \sum_{i=1}^{n} \mathbb{1}_{\{s_T = s\}}$.

Since the sequences $s^n \in T_\delta$ are typical, we have for all $s \in \mathcal{S}$:

$$\mathcal{P}(s) - \varepsilon \leq \frac{N(s|s^n)}{n} \leq \mathcal{P}(s) + \varepsilon, \tag{233}$$

which provides an upper bound and a lower bound on $\mathbb{P}(S_T = s|E = 0)$ as

$$\mathcal{P}(s) - \varepsilon = \sum_{s^n \in T_\delta} \mathbb{P}\big(S^n = s^n \big| E = 0\big) \cdot \Big(\mathcal{P}(s) - \varepsilon\Big) \leq \mathbb{P}(S_T = s|E = 0) \tag{234}$$

$$\leq \sum_{s^n \in T_\delta} \mathbb{P}\big(S^n = s^n \big| E = 0\big) \cdot \Big(\mathcal{P}(s) + \varepsilon\Big) = \mathcal{P}(s) + \varepsilon. \tag{235}$$

Using the same arguments, we prove that the probability distribution $\mathcal{P}_{S_T X_T Y_T V_T|E=0}(s, x, y, v)$ induced by the code, is closed to the target probability distribution $\mathcal{Q}(s, x, y, v)$:

$$\Big|\mathcal{P}_{S_T X_T Y_T V_T|E=0}(s, x, y, v) - \mathcal{Q}(s, x, y, v)\Big| \leq \varepsilon, \qquad \forall (s, x, y, v). \tag{236}$$

This concludes the proof of Lemma 7. □

## APPENDIX C
## PROOF OF THEOREM III.1

The conditional probability distribution $\mathcal{P}_{W_1^n W_2^n X^n|S^n}$ combined with $\mathcal{P}_S$ and $\mathcal{T}_{Y|XS}$ define this joint distribution:

$$\mathcal{P}_{S^n W_1^n W_2^n X^n Y^n} = \prod_{i=1}^{n} \mathcal{P}_{S_i} \mathcal{P}_{W_1^n W_2^n X^n|S^n} \prod_{i=1}^{n} \mathcal{T}_{Y_i|X_i S_i} \tag{237}$$

We introduce the random event $F \in \{0,1\}$ depending on whether the random sequences $(S^n, W_1^n, W_2^n, Y^n)$ are jointly typical or not.

$$F = \begin{cases} 0 & \text{if} \quad (S^n, W_1^n, W_2^n, Y^n) \in T_\delta(\mathcal{Q}), \\ 1 & \text{if} \quad (S^n, W_1^n, W_2^n, Y^n) \notin T_\delta(\mathcal{Q}). \end{cases} \tag{238}$$

Since the target probability distribution $\mathcal{Q}_{SW_1W_2XY}$ has full support, the probability distribution $\mathcal{P}_{S^n|Y^n}$ is absolutely continuous with respect to $\prod_{i=1}^n \mathcal{Q}_{S_i|Y_iW_{1,i}W_{2,i}}$, and the conditional KL-divergence is well defined:

$$\begin{aligned}
&\frac{1}{n} \cdot D\left(\mathcal{P}_{S^n|Y^n} \middle|\middle| \prod_{i=1}^n \mathcal{Q}_{S_i|Y_iW_{1,i}W_{2,i}}\right) \\
=&\frac{1}{n} \sum_{w_1^n, w_2^n, y^n} \mathcal{P}(w_1^n, w_2^n, \cdot y^n) \cdot \sum_{s^n} \mathcal{P}(s^n|y^n) \cdot \log_2 \frac{1}{\prod_{i=1}^n \mathcal{Q}(s_i|y_i, w_{1,i}, w_{2,i})} \\
-&\frac{1}{n} \sum_{w_1^n, w_2^n, y^n} \mathcal{P}(w_1^n, w_2^n, y^n) \cdot \sum_{s^n} \mathcal{P}(s^n|y^n) \cdot \log_2 \frac{1}{\mathcal{P}(s^n|y^n)} \qquad (239) \\
=&\frac{1}{n} \sum_{w_1^n, w_2^n, y^n, s^n, F} \mathcal{P}(w_1^n, w_2^n, y^n) \cdot \mathcal{P}(s^n|y^n) \cdot \mathbb{P}(F|s^n, w_1^n, w_2^n, y^n) \cdot \log_2 \frac{1}{\prod_{i=1}^n \mathcal{Q}(s_i|y_i, w_{1,i}, w_{2,i})} - \frac{1}{n} H(S^n|Y^n) \qquad (240) \\
=&\mathbb{P}(F=0) \cdot \frac{1}{n} \sum_{\substack{(w_1^n, w_2^n, y^n, s^n) \\ \in T_\delta(\mathcal{Q})}} \mathbb{P}(s^n, w_1^n, w_2^n, y^n|F=0) \cdot \log_2 \frac{1}{\prod_{i=1}^n \mathcal{Q}(s_i|y_i, w_{1,i}, w_{2,i})} \\
+&\mathbb{P}(F=1) \cdot \frac{1}{n} \sum_{w_1^n, w_2^n, y^n, s^n} \mathbb{P}(s^n, w_1^n, w_2^n, y^n|F=1) \cdot \log_2 \frac{1}{\prod_{i=1}^n \mathcal{Q}(s_i|y_i, w_{1,i}, w_{2,i})} - \frac{1}{n} H(S^n|Y^n) \qquad (241) \\
\leq&\frac{1}{n} \sum_{\substack{(w_1^n, w_2^n, y^n, s^n) \\ \in T_\delta(\mathcal{Q})}} \mathbb{P}(s^n, w_1^n, w_2^n, y^n|F=0) \cdot n \cdot \left(H(S|W_1, W_2, Y) + \delta \cdot \sum_{\substack{s, w_1, \\ w_2, y}} \log_2 \frac{1}{\mathcal{Q}(s|w_1, w_2, y)}\right) \\
+&\mathbb{P}(F=1) \cdot \log_2 \frac{1}{\min_{s,y,w_1,w_2} \mathcal{Q}(s|y, w_1, w_2)} - \frac{1}{n} H(S^n|Y^n) \qquad (242) \\
\leq&H(S|W_1, W_2, Y) - \frac{1}{n} H(S^n|Y^n) + \delta \cdot \sum_{\substack{s, w_1, \\ w_2, y}} \log_2 \frac{1}{\mathcal{Q}(s|w_1, w_2, y)} + \mathbb{P}(F=1) \cdot \log_2 \frac{1}{\min_{s,y,w_1,w_2} \mathcal{Q}(s|y, w_1, w_2)} \qquad (243) \\
=&I(S; W_1, W_2, Y) - \frac{1}{n} I(S^n; Y^n) + \delta \cdot \sum_{\substack{s, w_1, \\ w_2, y}} \log_2 \frac{1}{\mathcal{Q}(s|w_1, w_2, y)} + \mathbb{P}(F=1) \cdot \log_2 \frac{1}{\min_{s,y,w_1,w_2} \mathcal{Q}(s|y, w_1, w_2)} \qquad (244) \\
\leq&I(S; W_1, W_2, Y) - \mathcal{L}_{\mathsf{e}}(c) + \delta \cdot \sum_{\substack{s, w_1, \\ w_2, y}} \log_2 \frac{1}{\mathcal{Q}(s|w_1, w_2, y)} \\
&+ \mathbb{P}\Big((s^n, w_1^n, w_2^n, y^n) \notin T_\delta(\mathcal{Q})\Big) \cdot \log_2 \frac{1}{\min_{s,y,w_1,w_2} \mathcal{Q}(s|y, w_1, w_2)}, \qquad (245)
\end{aligned}$$

where (239) comes from the definition of the KL-divergence; (240) comes from the introduction of the random event $F \in \{0,1\}$; (241) is a reformulation of (240) where $\mathbb{P}(F) \cdot \mathbb{P}(s^n, w_1^n, w_2^n, y^n|F) = \mathcal{P}(w_1^n, w_2^n, y^n) \cdot \mathcal{P}(s^n|y^n) \cdot \mathbb{P}(F|s^n, w_1^n, w_2^n, y^n)$ and where the event $F = 0$ implies that the sequences $(w_1^n, w_2^n, y^n, s^n) \in T_\delta(\mathcal{Q})$ are jointly typical; (242) comes from two properties: 1) the property $2^{-n\cdot\left(H(S|Y,W_1,W_2)+\delta\cdot\sum_{\substack{s,w_1,\\ w_2,y}} \log_2 \frac{1}{\mathcal{Q}(s|w_1,w_2,y)}\right)} \leq \prod_{i=1}^n \mathcal{Q}_{S_i|Y_iW_{1,i}W_{2,i}}$ for typical sequences $(w_1^n, w_2^n, y^n, s^n) \in T_\delta(\mathcal{Q})$, stated in [34, pp. 26]; 2) the hypothesis $\mathcal{Q}_{SW_1W_2XY}$ has full support, which implies that $\min_{s,y,w_1,w_2} \mathcal{Q}(s|y,w_1,w_2) > 0$ is strictly positive; (244) comes from the i.i.d. property of $S$ that implies $H(S) = \frac{1}{n} \cdot H(S^n)$; (245) comes from the definition of the state leakage $\mathcal{L}_{\mathsf{e}}(c) = \frac{1}{n} I(S^n; Y^n)$ and the event $F = 0$.

## APPENDIX D
## ACHIEVABILITY PROOF OF THEOREM III.3

We consider a pair of rate and distortion $(\mathsf{R}, \mathsf{D}) \in \mathcal{A}_{\mathsf{g}}$ that satisfy equations (27) and (28) with probability distribution $\mathcal{P}_S \mathcal{Q}_{W_1} \mathcal{Q}_{X|SW_1} \mathcal{T}_{Y|XS}$. We introduce the target leakage $\mathsf{E} = I(S; W_1, Y)$. Theorem II.6 guarantees that the triple $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable: for all $\varepsilon > 0$, there exists a $\bar{n} \in \mathbb{N}$ such that for all $n \geq \bar{n}$, there exists a code with causal encoding $c \in \mathcal{C}(n, \mathcal{M})$ that satisfies:

$$\frac{\log_2 |\mathcal{M}|}{n} \geq \mathsf{R} - \varepsilon, \tag{246}$$

$$\left| \mathcal{L}_{\mathsf{e}}(c) - \mathsf{E} \right| \leq \varepsilon, \qquad \text{with} \qquad \mathcal{L}_{\mathsf{e}}(c) = \frac{1}{n} \cdot I(S^n; Y^n), \tag{247}$$

$$\mathcal{P}_{\mathsf{e}}(c) = \mathbb{P}\left( M \neq \hat{M} \right) + \mathbb{P}\left( \left|\left| Q^n - \mathcal{Q} \right|\right|_1 \geq \varepsilon \right) \leq \varepsilon. \tag{248}$$

**Remark D.1** *The achievability proof of Theorem II.6 guarantees that the sequences $(S^n, W_1^n, X^n, Y^n) \in A_\delta$ are jointly typical with large probability and that the decoding of $W_1^n$ is correct with large probability: $\mathbb{P}(\hat{W}_1^n \neq W_1^n) \leq \varepsilon$.*

We assume that probability distribution $\mathcal{P}_S \mathcal{Q}_{W_1} \mathcal{Q}_{X|SW_1} \mathcal{T}_{Y|XS}$ has full support. Otherwise, we would consider a sequence of probability distributions of full support, that converges to the target distribution. By replacing the pair $(W_1, W_2)$ by $W_1$ in Theorem III.1, we obtain:

$$\begin{aligned} \frac{1}{n} \cdot D\left( \mathcal{P}_{S^n|Y^n} \middle|\middle| \prod_{i=1}^n \mathcal{Q}_{S_i|Y_iW_{1,i}} \right) \leq & I(S; W_1, Y) - \mathcal{L}_{\mathsf{e}}(c) + \delta \cdot \sum_{s,w_1,y} \log_2 \frac{1}{\mathcal{Q}(s|w_1, y)} \\ & + \mathbb{P}\Big( (s^n, w_1^n, y^n) \notin T_\delta(\mathcal{Q}) \Big) \cdot \log_2 \frac{1}{\min_{s,y,w_1} \mathcal{Q}(s|y, w_1)}. \end{aligned} \tag{249}$$

By combining equations (246) - (248) with (249), we shows that, for any $\varepsilon > 0$, here exists a $\bar{n}$ such that for all $n \geq \bar{n}$ there exists a code with causal encoding $c \in \mathcal{C}(n, \mathcal{M})$ involving an auxiliary sequence $W_1^n$, such that:

$$\frac{\log_2 |\mathcal{M}|}{n} \geq \mathsf{R} - \varepsilon, \tag{250}$$

$$\mathbb{P}\Big(M \neq \hat{M}\Big) \leq \varepsilon, \tag{251}$$

$$\frac{1}{n} \cdot D\Big(\mathcal{P}_{S^n|Y^n} \Big|\Big| \prod_{i=1}^{n} \mathcal{Q}_{S_i|Y_i W_{1,i}}\Big) \leq \varepsilon. \tag{252}$$

Following [28, Notations A.14 and A.16], we define the sets $J_\alpha(w_1^n, y^n)$ and $B_{\alpha,\gamma,\delta}$ depending on small parameters $\alpha > 0$, $\gamma > 0$ and $\delta > 0$:

$$J_\alpha(w_1^n, y^n) = \Big\{ i \in \{1, \dots, n\}, \quad \text{s.t.} \quad D\Big(\mathcal{P}_{S_i|Y^n}(\cdot|y^n) \Big|\Big| \mathcal{Q}_{S_i|Y_i W_{1,i}}(\cdot|y_i, w_{1,i})\Big) \leq \frac{\alpha^2}{2 \ln 2} \Big\}, \tag{253}$$

$$B_{\alpha,\gamma,\delta} = \Big\{ (w_1^n, y^n) \quad \text{s.t.} \quad \frac{|J_\alpha(w_1^n, y^n)|}{n} \geq 1 - \gamma \quad \text{and} \quad (w_1^n, y^n) \in T_\delta(\mathcal{Q}) \Big\}. \tag{254}$$

The notation $B^c_{\alpha,\gamma,\delta}$ stands for the complementary set of $B_{\alpha,\gamma,\delta} \subset \mathcal{W}_1^n \times \mathcal{Y}^n$. The sequences $(w_1^n, y^n)$ belong to the set $B_{\alpha,\gamma,\delta}$ if: 1) they are jointly typical and 2) if the decoder's posterior belief $\mathcal{P}_{S_i|Y^n}(\cdot|y^n)$ is close in K-L divergence to the target belief $\mathcal{Q}_{S_i|Y_i W_{1,i}}(\cdot|y_i, w_{1,i})$, for a large fraction of stages $i \in \{1, \dots, n\}$.

Then [28, Corollary A.18] ensures that for each code with causal encoding $c \in \mathcal{C}(n, \mathcal{M})$, we have:

$$\begin{aligned}
&\Big| \min_{h_{V^n|Y^n}} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\Big[d(S_i, V_i)\Big] - \min_{\mathcal{P}_{V|W_1 Y}} \mathbb{E}\Big[d(S, V)\Big] \Big| \\
&= \Big| \frac{1}{n} \sum_{i=1}^{n} \sum_{y^n} \mathcal{P}(y^n) \min_{h_{V_i|Y^n}} \sum_{s_i} \mathcal{P}(s_i|y^n) \cdot d(s_i, v_i) - \sum_{w_1, y} \mathcal{Q}(w_1, y) \min_{\mathcal{P}_{V|W_1 Y}} \sum_{s} \mathcal{Q}(s|w_1, y) \cdot d(s, v) \Big| \quad (255) \\
&\leq \big(\alpha + 2\gamma + \delta + \mathbb{P}(B^c_{\alpha,\gamma,\delta})\big) \cdot \bar{d}, \quad (256)
\end{aligned}$$

where $\bar{d} = \max_{s,v} d(s, v)$ is the maximal distortion value.

**Lemma 8 (Reformulation of equations (37) - (43) in [28])**

$$\mathbb{P}(B^c_{\alpha,\gamma,\delta}) \leq \frac{2 \ln 2}{\alpha^2 \gamma} \cdot \frac{1}{n} \cdot D\Big(\mathcal{P}_{S^n|Y^n} \Big|\Big| \prod_{i=1}^{n} \mathcal{Q}_{S_i|Y_i W_{1,i}}\Big) + \mathbb{P}((w_1^n, y^n) \notin T_\delta(\mathcal{Q})). \tag{257}$$

We combine equations (252) and (256) with Lemma 8 and we choose the parameters $\alpha > 0$, $\gamma > 0$, $\delta > 0$ small and then $n$ large such as to obtain:

$$\mathbb{P}(B^c_{\alpha,\gamma,\delta}) \leq \varepsilon. \tag{258}$$

This concludes the achievability proof of Theorem III.3.

*A. Proof of Lemma 8*

The union bound implies

$$\mathbb{P}(B^c_{\alpha,\gamma,\delta}) = \mathbb{P}\bigg((W_1^n, Y^n) \quad \text{s.t.} \quad \frac{|J_\alpha(W_1^n, Y^n)|}{n} < 1-\gamma \quad \text{or} \quad (W_1^n, Y^n) \notin T_\delta(\mathcal{Q})\bigg) \tag{259}$$

$$\leq \mathbb{P}\bigg((W_1^n, Y^n) \quad \text{s.t.} \quad \frac{|J_\alpha(W_1^n, Y^n)|}{n} < 1-\gamma\bigg) + \mathbb{P}\Big((W_1^n, Y^n) \notin T_\delta(\mathcal{Q})\Big). \tag{260}$$

Moreover,

$$\mathbb{P}\bigg((W_1^n, Y^n) \quad \text{s.t.} \quad \frac{|J_\alpha(W_1^n, Y^n)|}{n} < 1-\gamma\bigg)$$

$$= \mathbb{P}\bigg(\frac{1}{n}\cdot\bigg|\bigg\{i\in\{1,\dots,n\}, \quad \text{s.t.} \quad D\Big(\mathcal{P}_{S_i|Y^n}(\cdot|Y^n)\Big|\Big|\mathcal{Q}_{S_i|Y_iW_{1,i}}(\cdot|Y_i,W_{1,i})\Big) \leq \frac{\alpha^2}{2\ln 2}\bigg\}\bigg| < 1-\gamma\bigg) \tag{261}$$

$$= \mathbb{P}\bigg(\frac{1}{n}\cdot\bigg|\bigg\{i\in\{1,\dots,n\}, \quad \text{s.t.} \quad D\Big(\mathcal{P}_{S_i|Y^n}(\cdot|Y^n)\Big|\Big|\mathcal{Q}_{S_i|Y_iW_{1,i}}(\cdot|Y_i,W_{1,i})\Big) > \frac{\alpha^2}{2\ln 2}\bigg\}\bigg| \geq \gamma\bigg) \tag{262}$$

$$\leq \frac{2\ln 2}{\alpha^2\gamma}\cdot\mathbb{E}\bigg[\frac{1}{n}\cdot\sum_{i=1}^n D\Big(\mathcal{P}_{S_i|Y^n}(\cdot|Y^n)\Big|\Big|\mathcal{Q}_{S_i|Y_iW_{1,i}}(\cdot|Y_i,W_{1,i})\Big)\bigg] \tag{263}$$

$$\leq \frac{2\ln 2}{\alpha^2\gamma}\cdot\frac{1}{n}\cdot D\Big(\mathcal{P}_{S^n|Y^n}\Big|\Big|\prod_{i=1}^n \mathcal{Q}_{S_i|Y_iW_{1,i}}\Big), \tag{264}$$

where (261)-(262) are reformulations; (263) comes from [28, Lemma A.22]; (264) comes from Lemma 9.

**Lemma 9** *We consider the probability distributions $\mathcal{P}_{A_1A_2B_1B_2}$, $\mathcal{Q}_{B_1|A_1}$ and $\mathcal{Q}_{B_2|A_2}$. We have:*

$$\begin{aligned} &D\Big(\mathcal{P}_{B_1B_2|A_1A_2}\Big|\Big|\mathcal{Q}_{B_1|A_1}\times\mathcal{Q}_{B_2|A_2}\Big) \\ =&D\Big(\mathcal{P}_{B_1|A_1A_2}\Big|\Big|\mathcal{Q}_{B_1|A_1}\Big) + D\Big(\mathcal{P}_{B_2|A_1A_2}\Big|\Big|\mathcal{Q}_{B_2|A_2}\Big) + I(B_1;B_2|A_1,A_2), \end{aligned} \tag{265}$$

*where the mutual information $I(B_1;B_2|A_1,A_2)$ is evaluated with respect to $\mathcal{P}_{A_1A_2B_1B_2}$. In particular, this implies for all $n\geq 1$:*

$$D\bigg(\mathcal{P}_{B^n|A^n}\bigg|\bigg|\prod_{i=1}^n\mathcal{Q}_{B_i|A_i}\bigg) \geq \sum_{i=1}^n D\Big(\mathcal{P}_{B_i|A^n}\Big|\Big|\mathcal{Q}_{B_i|A_i}\Big). \tag{266}$$

*Proof.* [Lemma 9]

$$D\Big(\mathcal{P}_{B_1B_2|A_1A_2}\Big|\Big|\mathcal{Q}_{B_1|A_1}\times\mathcal{Q}_{B_2|A_2}\Big)$$

$$=\sum_{a_1,a_2}\mathcal{P}(a_1,a_2)\sum_{b_1,b_2}\mathcal{P}(b_1,b_2|a_1,a_2)\log_2\frac{\mathcal{P}(b_1,b_2|a_1,a_2)}{\mathcal{Q}(b_1|a_1)\times\mathcal{Q}(b_2|a_2)} \tag{267}$$

$$=\sum_{a_1,a_2}\mathcal{P}(a_1,a_2)\sum_{b_1}\mathcal{P}(b_1|a_1,a_2)\log_2\frac{1}{\mathcal{Q}(b_1|a_1)}+\sum_{a_1,a_2}\mathcal{P}(a_1,a_2)\sum_{b_2}\mathcal{P}(b_2|a_1,a_2)\log_2\frac{1}{\mathcal{Q}(b_2|a_2)}-H(B_1,B_2|A_1,A_2) \tag{268}$$

$$=\sum_{a_1,a_2}\mathcal{P}(a_1,a_2)\sum_{b_1}\mathcal{P}(b_1|a_1,a_2)\log_2\frac{\mathcal{P}(b_1|a_1,a_2)}{\mathcal{Q}(b_1|a_1)}+\sum_{a_1,a_2}\mathcal{P}(a_1,a_2)\sum_{b_2}\mathcal{P}(b_2|a_1,a_2)\log_2\frac{\mathcal{P}(b_2|a_1,a_2)}{\mathcal{Q}(b_2|a_2)} \tag{269}$$

$$+H(B_1|A_1,A_2)+H(B_2|A_1,A_2)-H(B_1,B_2|A_1,A_2) \tag{270}$$

$$=D\Big(\mathcal{P}_{B_1|A_1A_2}\Big|\Big|\mathcal{Q}_{B_1|A_1}\Big)+D\Big(\mathcal{P}_{B_2|A_1A_2}\Big|\Big|\mathcal{Q}_{B_2|A_2}\Big)+I(B_1;B_2|A_1,A_2). \tag{271}$$

□

## APPENDIX E
## CONVERSE PROOF OF THEOREM III.3

We introduce the random event $F\in\{0,1\}$ indicating whether $M$ is correctly decoded or not:

$$F=\begin{cases}0 & \text{if} \quad M=\hat{M},\\ 1 & \text{if} \quad M\neq\hat{M}.\end{cases} \tag{272}$$

We assume that there exists a code with causal encoding $c\in\mathcal{C}(n,\mathcal{M})$ that satisfies:

$$\frac{\log_2|\mathcal{M}|}{n}\geq \mathsf{R}-\varepsilon, \tag{273}$$

$$\mathcal{P}_{\mathsf{e}}(c)=\mathbb{P}\Big(M\neq\hat{M}\Big)\leq\varepsilon, \tag{274}$$

$$\Big|\min_{h_{V^n|Y^n}}\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[d(S_i,V_i)\Big]-\mathsf{D}\Big|\leq\varepsilon. \tag{275}$$

Then, note that

$$\mathsf{R}\leq\frac{\log_2|\mathcal{M}|}{n}+\varepsilon=\frac{1}{n}\cdot H(M)+\varepsilon=\frac{1}{n}\cdot I(M;Y^n)+\frac{1}{n}\cdot H(M|Y^n)+\varepsilon \tag{276}$$

$$=\frac{1}{n}\cdot I(M;Y^n)+2\varepsilon=\frac{1}{n}\cdot\sum_{i=1}^{n}I(M;Y_i|Y^{i-1})+2\varepsilon\leq\frac{1}{n}\cdot\sum_{i=1}^{n}I(M,Y^{i-1},Y_{i+1}^n;Y_i)+2\varepsilon \tag{277}$$

$$=\frac{1}{n}\cdot\sum_{i=1}^{n}I(W_{1,i};Y_i)+2\varepsilon=I(W_{1,T};Y_T|T)+2\varepsilon\leq I(W_{1,T},T;Y_T)+2\varepsilon\leq I(W_1;Y)+2\varepsilon. \tag{278}$$

where (276) comes from assumption (273) and from uniform distribution of the random message $M$; (277) comes from Fano's inequality [31, Theorem 2.10.1] and assumption (274) and adding the mutual informations $I(Y^{i-1};Y_i)$ and $I(Y_{i+1}^n;Y_i|M,Y^{i-1})$; (278) comes from the identification of the auxiliary random variable $W_{1,i}=(M,Y^{i-1},Y_{i+1}^n)$ and the introduction of the uniform random variable $T$ over the indices $\{1,\dots,n\}$ and $(W_{1,T},Y_T)$ and by identifying $W_1=(W_{1,T},T)$ and $Y_T=Y$. This established (27) in Theorem III.3.

The notation $y^{-i}=(y^{i-1},y_{i+1}^n)$ stands for the subsequence of $y^n$ where $y_i$ has been removed. We now assume that the event $F=0$ is realized, hence the average distortion satisfies:

$$\sum_{y^n,m}\mathbb{P}(y^n,m|F=0)\min_{h_{V^n|Y^n}}\sum_{s^n}\mathbb{P}(s^n|y^n,F=0)\cdot\left[\frac{1}{n}\sum_{i=1}^n d(s_i,v_i)\right] \tag{279}$$

$$=\sum_{y^n,m}\mathbb{P}(y^n,m|F=0)\frac{1}{n}\sum_{i=1}^n\min_{h:\mathcal{Y}^n\to\mathcal{V}_i}\sum_{s_i}\mathbb{P}(s_i|y^n,F=0)\cdot d(s_i,v_i) \tag{280}$$

$$=\sum_{y_i,y^{-i},m}\mathbb{P}(y^n,m|F=0)\frac{1}{n}\sum_{i=1}^n\min_{h:\mathcal{Y}^n\times\mathcal{M}\to\mathcal{V}_i}\sum_{s_i}\mathbb{P}(s_i|y_i,y^{-i},m,F=0)\cdot d(s_i,v_i) \tag{281}$$

$$=\sum_{y_i,w_{1,i}}\mathcal{P}(y_i,w_{1,i})\frac{1}{n}\sum_{i=1}^n\min_{h:\mathcal{Y}_i\times\mathcal{W}_{1,i}\to\mathcal{V}_i}\sum_{s_i}\mathcal{P}(s_i|y_i,w_{1,i})\cdot d(s_i,v_i) \tag{282}$$

$$=\sum_{y_T,w_{1,T},T}\mathcal{P}(y_T,w_{1,T},T)\min_{h:\mathcal{Y}\times\mathcal{W}_1\to\mathcal{V}}\sum_{s_T}\mathcal{P}(s_T|y_T,w_{1,T},T)\cdot d(s_T,v_T) \tag{283}$$

$$=\sum_{y,w_1}\mathcal{P}(y,w_1)\min_{h:\mathcal{Y}\times\mathcal{W}_1\to\mathcal{V}}\sum_{s}\mathcal{P}(s|y,w_1)\cdot d(s,v)=\min_{\mathcal{P}_{V|W_1Y}}\mathbb{E}\Big[d(S,V)\Big], \tag{284}$$

where (280) is a reformulation; (281) comes from the hypothesis $F=0$, since the decoder correctly decodes the message $m$ based on the observation of $y^n$; (282) comes from the identification of the auxiliary random variable $w_{1,i}=(m,y^{-i})$; (283) comes from the introduction of the uniform random variable $T$ over the indices $\{1,\dots,n\}$ and the corresponding random variables $(s_T,w_{1,T},y_T,v_T)$; (284) comes from identifying the auxiliary random variables $w_1=(w_{1,T},T)$.

Hence we have:

$$\left|\min_{\mathcal{P}_{V|W_1Y}}\mathbb{E}\Big[d(S,V)\Big]-\mathsf{D}\right| \tag{285}$$

$$\le\left|\min_{\mathcal{P}_{V|W_1Y}}\mathbb{E}\Big[d(S,V)\Big]-\min_{h_{V^n|Y^n}}\frac{1}{n}\sum_{i=1}^n\mathbb{E}\Big[d(S_i,V_i)\Big]\right|+\left|\min_{h_{V^n|Y^n}}\frac{1}{n}\sum_{i=1}^n\mathbb{E}\Big[d(S_i,V_i)\Big]-\mathsf{D}\right| \tag{286}$$

$$\le\left|\min_{\mathcal{P}_{V|W_1Y}}\mathbb{E}\Big[d(S,V)\Big]-\min_{h_{V^n|Y^n}}\frac{1}{n}\sum_{i=1}^n\mathbb{E}\Big[d(S_i,V_i)\Big]\right|+\varepsilon \tag{287}$$

$$=\left|\min_{\mathcal{P}_{V|W_1Y}}\mathbb{E}\Big[d(S,V)\Big]-\sum_{y^n,m,F}\mathbb{P}(y^n,m,F)\times\min_{h_{V^n|Y^n}}\sum_{s^n}\mathbb{P}(s^n|y^n,F)\cdot\left[\frac{1}{n}\sum_{i=1}^n d(s_i,v_i)\right]\right|+\varepsilon \tag{288}$$

$$\leq \left| \min_{\mathcal{P}_{V|W_1Y}} \mathbb{E}\Big[d(S,V)\Big] - \sum_{y^n,m} \mathbb{P}(y^n,m|F=0) \times \min_{h_{V^n|Y^n}} \sum_{s^n} \mathbb{P}(s^n|y^n,F=0) \cdot \left[\frac{1}{n}\sum_{i=1}^{n} d(s_i,v_i)\right] \right|$$

$$+ \mathbb{P}(F=1)\cdot 2\bar{d} + \varepsilon \tag{289}$$

$$= \left| \min_{\mathcal{P}_{V|W_1Y}} \mathbb{E}\Big[d(S,V)\Big] - \min_{\mathcal{P}_{V|W_1Y}} \mathbb{E}\Big[d(S,V)\Big] \right| + \mathcal{P}(F=1)\cdot 2\bar{d} + \varepsilon \tag{290}$$

$$= \mathbb{P}(F=1)\cdot 2\bar{d} + \varepsilon \leq \varepsilon \cdot (2\bar{d}+1), \tag{291}$$

where (286) comes from the triangle inequality; (287) comes from assumption (275); (288) is a reformulation that introduces the random event $F \in \{0,1\}$; (289) comes from removing the event $\mathbb{P}(F=1)\cdot\bar{d}$ from the the triangle inequality; (290) comes from (284); (291) comes from assumption (274). This ensures that (28) of Theorem III.3 holds and concludes the converse proof.

## REFERENCES


[1] M. Le Treust and M. Bloch, "Empirical coordination, state masking and state amplification: Core of the decoder's knowledge," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 895–899.

[2] C. Shannon, "Channels with side information at the transmitter," *IBM Journal of Research and Development*, vol. 2, no. 4, pp. 289–293, Oct 1958.

[3] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[4] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[5] P. Moulin and J. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 563–593, 2003.

[6] O. Gossner, P. Hernández, and A. Neyman, "Optimal use of communication resources," *Econometrica*, vol. 74, no. 6, pp. 1603–1636, 2006.

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

[8] N. Merhav and S. Shamai, "Information rates subject to state masking," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 2254–2261, June 2007.

[9] Y.-H. Kim, A. Sutivong, and T. Cover, "State amplification," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1850–1859, May 2008.

[10] C. Choudhuri, Y.-H. Kim, and U. Mitra, "Causal state communication," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3709–3719, June 2013.

[11] A. Sutivong, M. Chiang, T. M. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1486–1495, Apr. 2005.

[12] C. Tian, B. Bandemer, and S. S. Shitz, "Gaussian state amplification with noisy observations," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4587–4597, Sept 2015.

[13] O. O. Koyluoglu, R. Soundararajan, and S. Vishwanath, "State amplification subject to masking constraints," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6233–6250, Nov. 2016.

[14] K. Tutuncuoglu, O. Ozel, A. Yener, and S. Ulukus, "State amplification and state masking for the binary energy harvesting channel," in *IEEE Information Theory Workshop (ITW)*, Nov 2014, pp. 336–340.

[15] G. Kramer and S. Savari, "Communicating probability distributions," *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 518 – 525, Feb. 2007.

[16] O. Gossner and N. Vieille, "How to play with a biased coin?" *Games and Economic Behavior*, vol. 41, no. 2, pp. 206–226, 2002.

[17] O. Gossner and T. Tomala, "Empirical distributions of beliefs under imperfect observation," *Mathematics of Operation Research*, vol. 31, no. 1, pp. 13–30, 2006.

[18] ——, "Secret correlation in repeated games with imperfect monitoring," *Mathematics of Operation Research*, vol. 32, no. 2, pp. 413–424, 2007.

[19] O. Gossner, R. Laraki, and T. Tomala, "Informationally optimal correlation," *Mathematical Programming*, vol. 116, no. 1-2, pp. 147–172, 2009.

[20] P. Cuff and C. Schieler, "Hybrid codes needed for coordination over the point-to-point channel," in *49th Annual Allerton Conference on Communication, Control, and Computing*, Sept 2011, pp. 235–239.

[21] P. Cuff and L. Zhao, "Coordination using implicit communication," in *IEEE Information Theory Workshop (ITW)*, 2011, pp. 467– 471.

[22] M. Le Treust, "Joint empirical coordination of source and channel," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 5087–5114, Aug 2017.

[23] B. Larrousse, S. Lasaulce, and M. R. Bloch, "Coordination in distributed networks via coded actions with application to power control," *IEEE Transactions on Information Theory*, vol. 64, no. 5, pp. 3633–3654, May 2018.

[24] M. Le Treust, "Correlation between channel state and information source with empirical coordination constraint," in *IEEE Information Theory Workshop (ITW)*, Nov 2014, pp. 272–276.

[25] C. Schieler and P. Cuff, "Rate-distortion theory for secrecy systems," *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7584–7605, Dec 2014.

[26] M. Le Treust, "Empirical coordination with two-sided state information and correlated source and state," in *IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 466–470.

[27] ——, "Empirical coordination with channel feedback and strictly causal or causal encoding," in *IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 471 – 475.

[28] M. Le Treust and T. Tomala, "Persuasion with limited communication capacity," *draft available: https://arxiv.org/abs/1711.04474*, Dec. 2017.

[29] E. Akyol, C. Langbort, and T. Başar, "Information-theoretic approach to strategic communication as a hierarchical game," *Proceedings of the IEEE*, vol. 105, no. 2, pp. 205–218, Feb 2017.

[30] M. Le Treust and T. Tomala, "Information-theoretic limits of strategic communication," *draft available: https://arxiv.org/abs/1807.05147*, July 2018.

[31] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York: 2nd. Ed., Wiley-Interscience, 2006.

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

[33] M. Sion, "On general minimax theorems," *Pacific Journal of mathematics*, vol. 8, no. 1, pp. 171–176, 1958.

[34] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, Dec. 2011.

## APPENDIX F
## CARDINALITY BOUND FOR THEOREM II.3

**Lemma 10 (Cardinality Bound for Theorem II.3)** *We consider the following information constraints with two auxiliary random variables* $(W_1, W_2)$*:*

$$I(S; W_1, W_2, Y) \leq \mathsf{E} \leq H(S), \tag{292}$$

$$\mathsf{R} + \mathsf{E} \leq I(W_1, S; Y). \tag{293}$$

*The cardinality of the supports of the auxiliary random variables* $(W_1, W_2)$ *are bounded by:*

$$\max(|\mathcal{W}_1|, |\mathcal{W}_2|) \leq d + 1, \qquad \textit{with} \qquad d = |\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}|.$$

This result is based on the Lemma of Fenchel-Eggleston-Carathéodory. More details are provided pp. 631 in [34]. Lemma 10 can be adapted for the cardinality bounds of Theorems IV.3 and IV.5.

*Proof.* [Lemma 10] We denote by $d = |\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}|$, the cardinality of the product of the discrete sets. We consider the family of continuous functions $h_i : \Delta(\mathcal{S} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{V}) \mapsto \mathbb{R}$, with $i \in \{1, \ldots, d+1\}$, defined as follows:

$$h_i\Big(\mathcal{P}_{SXYV|W_1W_2}\Big) = \begin{cases} \mathcal{P}_{SXYV|W_1W_2}(i), & \text{for } i \in \big\{1, \ldots, d-1\big\}, \\ H(Y|S, W_1 = w_1), & \text{for } i = d, \\ H(S|Y, W_1 = w_1, W_2 = w_2), & \text{for } i = d+1. \end{cases}$$

Support Lemma stated pp. 631 in [34], implies that there exists a pair of auxiliary random variables $(W_1', W_2') \sim \mathcal{P}_{W_1'W_2'}$ defined on a set $\mathcal{W}_1' \times \mathcal{W}_2'$ with finite cardinality $\max(|\mathcal{W}_1'|, |\mathcal{W}_2'|) \leq d + 1$ such that for all $i \in \{1, \ldots, d+1\}$ we have:

$$\int_{\mathcal{W}_1 \times \mathcal{W}_2} h_i\Big(\mathcal{P}_{SXYV|W_1W_2}\Big) dF(w_1, w_2) = \sum_{(w_1', w_2') \in \mathcal{W}_1' \times \mathcal{W}_2'} h_i\Big(\mathcal{P}_{SXYV|W_1'W_2'}\Big) \mathcal{P}(w_1' w_2').$$

This implies that the probability $\mathcal{P}_{SXYV}$ is preserved and we have:

$$\begin{aligned}
\mathcal{P}_{SXYV}(i) =& \int_{\mathcal{W}_1\times\mathcal{W}_2} \mathcal{P}_{SXYV|W_1W_2}(i)dF(w_1,w_2)\\
=& \sum_{(w_1',w_2')\in\mathcal{W}_1'\times\mathcal{W}_2'} \mathcal{P}_{SXYV|W_1'W_2'}(i)\cdot\mathcal{P}(w_1',w_2'), \quad \text{for } i\in\big\{1,\dots,d-1\big\}\\
H(Y|S,W_1) =& \int_{\mathcal{W}_1} H(Y|S,W_1=w_1)dF(w_1))\\
=& \sum_{(w_1')\in\mathcal{W}_1'} H(Y|S,W_1'=w_1')\cdot\mathcal{P}(w_1') = H(Y|S,W_1'),\\
H(S|Y,W_1,W_2) =& \int_{\mathcal{W}_1\times\mathcal{W}_2} H(S|Y,W_1=w_1,W_2=w_2)dF(w_1,w_2))\\
=& \sum_{(w_1',w_2')\in\mathcal{W}_1'\times\mathcal{W}_2'} H(S|Y,W_1'=w_1',W_2'=w_2')\cdot\mathcal{P}(w_1',w_2') = H(S|Y,W_1',W_2').
\end{aligned}$$

Hence the three information constraints remain equal with $\max(|\mathcal{W}_1'|,|\mathcal{W}_2'|)\leq d+1$.

$$\begin{aligned}
I(S;W_1,W_2,Y) =& H(S)-H(S|W_1',W_2',Y) = I(S;W_1',W_2',Y),\\
I(W_1,S;Y) =& H(Y)-H(Y|W_1',S) = I(W_1',S;Y).
\end{aligned}$$

This concludes the proof of Lemma 10. □

## APPENDIX G
## CONVERSE PROOF OF THEOREM II.6

We consider that the pair of rate and state leakage $(\mathsf{R},\mathsf{E})$ is achievable with a causal code. By definition II.5, for all $\varepsilon>0$, there exists a $\bar{n}$ such that for all $n\geq\bar{n}$, there exists a code $c^\star\in\mathcal{C}(n,\mathcal{M})$ that satisfies the three following equations:

$$\frac{\log_2|\mathcal{M}|}{n} \geq \mathsf{R}-\varepsilon, \tag{294}$$

$$\left|\mathcal{L}_{\mathsf{e}}(c)-\mathsf{E}\right| = \left|\frac{1}{n}\cdot I(S^n;Y^n)-\mathsf{E}\right| \leq \varepsilon, \tag{295}$$

$$\mathcal{P}_{\mathsf{e}}(c) = \mathbb{P}\Big(M\neq\hat{M}\Big)\leq\varepsilon. \tag{296}$$

We introduce the auxiliary random variables $W_{1,i}=(M,S^{i-1})$ that satisfy the Markov chains of the set of probability distribution $\mathbb{Q}_{\mathsf{c}}$ for all $i\in\{1,\dots,n\}$:

$$S_i \text{ independent of } W_{1,i}, \tag{297}$$

$$Y_i \multimap (X_i,S_i) \multimap W_{1,i}, \tag{298}$$

where (297) comes from the i.i.d. property of the source that induces the independence between $S_i$ and $(M,S^{i-1})=W_{1,i}$; (298) comes from the memoryless property of the channel $\mathcal{T}_{Y|XS}$.

We introduce the random variable $T$ that is uniformly distributed over the indices $\{1, \dots, n\}$ and the corresponding mean random variables $(S_T, X_T, W_{1,T}, Y_T)$. The auxiliary random variables $W_1 = (W_{1,T}, T)$ belongs to the set of probability distributions $\mathbb{Q}_{\mathsf{c}}$ and satisfies the four information constraints of Theorem II.6:

$$I(S;Y|W_1) \leq \mathsf{E} \leq H(S), \tag{299}$$

$$\mathsf{R} + \mathsf{E} \leq I(W_1, S; Y). \tag{300}$$

**First Constraint:**

$$\begin{aligned}
n \cdot \mathsf{E} \geq & I(S^n; Y^n) - n \cdot \varepsilon && (301)\\
= & I(S^n; Y^n, M) - I(S^n; M|Y^n) - n \cdot \varepsilon && (302)\\
\geq & \sum_{i=1}^{n} I(S_i; Y^n, M|S^{i-1}) - H(M|Y^n) - n \cdot \varepsilon && (303)\\
\geq & \sum_{i=1}^{n} I(S_i; Y^n, M|S^{i-1}) - n \cdot 2\varepsilon && (304)\\
= & \sum_{i=1}^{n} I(S_i; Y^n|M, S^{i-1}) - n \cdot 2\varepsilon && (305)\\
\geq & \sum_{i=1}^{n} I(S_i; Y_i|M, S^{i-1}) - n \cdot 2\varepsilon && (306)\\
= & \sum_{i=1}^{n} I(S_i; Y_i|W_{1,i}) - n \cdot 2\varepsilon && (307)\\
= & n \cdot I(S_T; Y_T|W_{1,T}, T) - n \cdot 2\varepsilon && (308)\\
= & n \cdot \Big( I(S;Y|W_1) - 2\varepsilon \Big), && (309)
\end{aligned}$$

where (301) comes from the definition of achievable state leakage rate $\mathsf{E}$, stated in equation (295); (302) and (303) come from the properties of the mutual information; (304) comes from equation (296) and Fano's inequality, stated pp. 19, in [34]; (305) comes from the independence between the message $M$ and the channel states $(S^{i-1}, S_i)$; (306) comes from the properties of the mutual information; (307) comes from the introduction of the auxiliary random variable $W_{1,i} = (M, S^{i-1})$, for all $i \in \{1, \dots, n\}$; (308) comes from the introduction of the uniform random variable $T$ over $\{1, \dots, n\}$ and the corresponding mean random variables $S_T$, $W_{1,T}$, $Y_T$; (309) comes from identifying $W_1 = (W_{1,T}, T)$ and $S = S_T$, $Y = Y_T$.

**Second Constraint:**

$$n \cdot \mathsf{E} \leq I(S^n; Y^n) + n \cdot \varepsilon \tag{310}$$

$$\leq H(S^n) + n \cdot \varepsilon \tag{311}$$

$$= n \cdot \Big( H(S) + \varepsilon \Big), \tag{312}$$

where (310) comes from the definition of the achievable state leakage rate $\mathsf{E}$, stated in equation (295); (311) comes from the properties of the mutual information; (312) comes from the i.i.d. property of the channel states $S$.

**Third Constraint:**

$$n \cdot \Big( \mathsf{E} + \mathsf{R} \Big) \leq I(S^n; Y^n) + H(M) + n \cdot 2\varepsilon \tag{313}$$

$$= I(S^n; Y^n) + I(M; Y^n) + H(M|Y^n) + n \cdot 2\varepsilon \tag{314}$$

$$\leq I(S^n; Y^n) + I(M; Y^n) + n \cdot 3\varepsilon \tag{315}$$

$$\leq I(S^n; Y^n) + I(M; Y^n|S^n) + n \cdot 3\varepsilon \tag{316}$$

$$\leq I(S^n, M; Y^n) + n \cdot 3\varepsilon \tag{317}$$

$$= \sum_{i=1}^{n} I(S^n, M; Y_i|Y_{i+1}^n) + n \cdot 3\varepsilon \tag{318}$$

$$\leq \sum_{i=1}^{n} I(S^n, M, Y_{i+1}^n; Y_i) + n \cdot 3\varepsilon \tag{319}$$

$$= \sum_{i=1}^{n} I(S_i, S^{i-1}, M; Y_i) + \sum_{i=1}^{n} I(S_{i+1}^n, Y_{i+1}^n; Y_i|S_i, S^{i-1}, M) + n \cdot 3\varepsilon \tag{320}$$

$$= \sum_{i=1}^{n} I(S_i, S^{i-1}, M; Y_i) + n \cdot 3\varepsilon \tag{321}$$

$$= \sum_{i=1}^{n} I(S_i, W_{1,i}; Y_i) + n \cdot 3\varepsilon \tag{322}$$

$$= n \cdot I(S_T, W_{1,T}; Y_T|T) + n \cdot 3\varepsilon \tag{323}$$

$$\leq n \cdot I(S_T, W_{1,T}, T; Y_T) + n \cdot 3\varepsilon \tag{324}$$

$$\leq n \cdot \Big( I(S, W_1; Y) + 3\varepsilon \Big), \tag{325}$$

where (313) comes from the definition of achievable rate and information leakage $(\mathsf{R}, \mathsf{E})$, stated in equations (294) and (295); (314) comes from the definitions of the mutual information; (315) comes from equation (296) and Fano's inequality, stated pp. 19, in [34]; (316) comes from the independence

between the message $M$ and the channel states $S^n$, hence $I(M;Y^n) \leq I(M;Y^n,S^n) = I(M;Y^n|S^n)$; (317), (318), (319), (320) comes from the properties of the mutual information; (321) comes from the Markov chain $Y_i \multimap (S_i, M, S^{i-1}) \multimap (S_{i+1}^n, Y_{i+1}^n)$, stated in Lemma 4. It is a direct consequence of the causal encoding function and the memoryless property of the channel; (322) comes from the introduction of the auxiliary random variable $W_{1,i} = (M, S^{i-1})$, for all $i \in \{1,\dots,n\}$; (323) comes from the introduction of the uniform random variable $T$ over $\{1,\dots,n\}$ and the corresponding mean random variables $S_T$, $W_{1,T}$, $Y_T$; (324) comes from the properties of the mutual information; (325) comes from identifying $W_1 = (W_{1,T}, T)$ and $S = S_T$, $Y = Y_T$.

**Conclusion:** If the pair of rate and state leakage $(\mathsf{R}, \mathsf{E})$ is achievable with causal encoding, then the following equations are satisfied for all $\varepsilon > 0$:

$$I(S;Y|W_1) - 2\varepsilon \leq \mathsf{E} \leq H(S) + \varepsilon, \tag{326}$$

$$\mathsf{R} + \mathsf{E} \leq I(S, W_1; Y) + 3\varepsilon. \tag{327}$$

This corresponds to equations (17), (14) and (15) and concludes the converse proof of Theorem II.6.

**Remark G.1** *For the converse proof of Theorem II.6, the causal encoding is not necessarily deterministic. The same optimal performances can be obtained by considering stochastic causal encoding.*

## APPENDIX H
## CONVERSE PROOF OF THEOREM IV.3

Consider that the triple of rate, state leakage and probability distribution $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable with a causal code. We introduce the random event of error $E \in \{0,1\}$ defined with respect to the achievable joint probability distribution $\mathcal{Q}_{USZXYV}$, as follows:

$$E = \begin{cases} 0 \text{ if } & \left|\left|Q^n - \mathcal{Q}\right|\right|_1 \leq \varepsilon \iff (U^n, S^n, Z^n, X^n, Y^n, V^n) \in T_\delta(\mathcal{Q}), \\ 1 \text{ if } & \left|\left|Q^n - \mathcal{Q}\right|\right|_1 > \varepsilon \iff (U^n, S^n, Z^n, X^n, Y^n, V^n) \notin T_\delta(\mathcal{Q}). \end{cases} \tag{328}$$

The event $E = 1$ occurs if the sequences $(U^n, S^n, Z^n, X^n, Y^n, V^n) \notin T_\delta(\mathcal{Q})$ are not jointly typical for the target probability distribution $\mathcal{Q}$. By definition IV.1, for all $\varepsilon > 0$, there exists a $\bar{n}$ such that for all $n \geq \bar{n}$, there exists a code $c^\star \in \mathcal{C}(n, \mathcal{M})$ that satisfies the three following equations:

$$\frac{\log_2 |\mathcal{M}|}{n} \geq \mathsf{R} - \varepsilon, \tag{329}$$

$$\left|\mathcal{L}_{\mathsf{e}}(c^\star) - \mathsf{E}\right| = \left|\frac{1}{n} \cdot I(U^n, S^n; Y^n, Z^n) - \mathsf{E}\right| \leq \varepsilon, \tag{330}$$

$$\mathcal{P}_{\mathsf{e}}(c^\star) = \mathbb{P}\left(M \neq \hat{M}\right) + \mathbb{P}\left(\left|\left|Q^n - \mathcal{Q}\right|\right|_1 \geq \varepsilon\right) \leq \varepsilon. \tag{331}$$

We introduce the auxiliary random variables $W_{1,i} = (M, U^{i-1}, S^{i-1})$ and $W_{2,i} = (Y_{i+1}^n, Z_{i+1}^n)$, that satisfy the Markov chain of the set of probability distribution $\mathbb{Q}_2$ for all $i \in \{1, \dots, n\}$:

$$(U_i, S_i) \text{ independent of } W_{1,i}, \tag{332}$$

$$X_i \multimap (U_i, S_i, W_{1,i}) \multimap W_{2,i}, \tag{333}$$

$$Y_i \multimap (X_i, S_i) \multimap (U_i, Z_i, W_{1,i}, W_{2,i}), \tag{334}$$

$$Z_i \multimap (U_i, S_i) \multimap (X_i, Y_i, W_{1,i}, W_{2,i}), \tag{335}$$

$$V_i \multimap (Y_i, Z_i, W_{1,i}, W_{2,i}) \multimap (U_i, S_i, X_i), \tag{336}$$

where (332) comes from the i.i.d. property of the source that induces the independence between $(U_i, S_i)$ and $(M, U^{i-1}, S^{i-1}) = W_{1,i}$; (333) comes from Lemma 11. It is a direct consequence of the causal encoding function, the memoryless property of the channel and the i.i.d. property of the source; (334) comes from the memoryless property of the channel $\mathcal{T}_{Y|XS}$; (335) comes from the i.i.d. property of the source $\mathcal{P}_{USZ}$; (336) comes from Lemma 12. It is a direct consequence of the causal encoding function, the non-causal decoding function, the memoryless property of the channel and the i.i.d. property of the source.

We introduce the random variable $T$ that is uniformly distributed over the indices $\{1, \dots, n\}$ and the corresponding mean random variables $W_{1,T}$, $W_{2,T}$, $U_T$, $S_T$, $Z_T$, $X_T$, $Y_T$, $V_T$. The auxiliary random variables $W_1 = (W_{1,T}, T)$ and $W_2 = W_{2,T}$ belong to the set of probability distributions $\mathbb{Q}_2$ and satisfy the four information constraints of Theorem IV.3:

$$I(U, S; W_1, W_2, Y, Z) \leq \mathsf{E} \leq H(U, S), \tag{337}$$

$$\mathsf{R} + \mathsf{E} \leq I(W_1, U, S; Y, Z). \tag{338}$$

**First Constraint:**

$$\begin{aligned} n \cdot \mathsf{E} \geq & I(U^n, S^n; Y^n, Z^n) - n \cdot \varepsilon & (339)\\ = & I(U^n, S^n; Y^n, Z^n, M) - I(U^n, S^n; M | Y^n, Z^n) - n \cdot \varepsilon & (340)\\ \geq & \sum_{i=1}^{n} I(U_i, S_i; Y^n, Z^n, M | U^{i-1}, S^{i-1}) - H(M | Y^n, Z^n) - n \cdot \varepsilon & (341)\\ \geq & \sum_{i=1}^{n} I(U_i, S_i; Y^n, Z^n, M | U^{i-1}, S^{i-1}) - n \cdot 2\varepsilon & (342)\\ = & \sum_{i=1}^{n} I(U_i, S_i; Y^n, Z^n, M, U^{i-1}, S^{i-1}) - n \cdot 2\varepsilon & (343) \end{aligned}$$

$$\geq \sum_{i=1}^{n} I(U_i, S_i; Y_{i+1}^n, Z_{i+1}^n, M, U^{i-1}, S^{i-1}, Y_i, Z_i) - n \cdot 2\varepsilon \tag{344}$$

$$= \sum_{i=1}^{n} I(U_i, S_i; W_{1,i}, W_{2,i}, Y_i, Z_i) - n \cdot 2\varepsilon \tag{345}$$

$$= n \cdot I(U_T, S_T; W_{1,T}, W_{2,T}, Y_T, Z_T|T) - n \cdot 2\varepsilon \tag{346}$$

$$= n \cdot I(U_T, S_T; W_{1,T}, W_{2,T}, Y_T, Z_T, T) - n \cdot 2\varepsilon \tag{347}$$

$$= n \cdot \Big( I(U_T, S_T; W_1, W_2, Y_T, Z_T) - 2\varepsilon \Big) \tag{348}$$

$$\geq n \cdot \Big( I(U_T, S_T; W_1, W_2, Y_T, Z_T|E = 0) - 3\varepsilon \Big) \tag{349}$$

$$\geq n \cdot \Big( I(U, S; W_1, W_2, Y, Z) - 4\varepsilon \Big), \tag{350}$$

where (339) comes from the definition of achievable state leakage rate $\mathsf{E}$, stated in equation (330); (340) and (341) come from the properties of the mutual information; (342) comes from equation (331) and Fano's inequality, stated pp. 19, in [34]; (343) comes from the i.i.d. property of the channel states that implies $(U_i, S_i)$ is independent of $(U^{i-1}, S^{i-1})$; (344) comes from the properties of the mutual information; (345) comes from the introduction of the auxiliary random variables $W_{1,i} = (M, U^{i-1}, S^{i-1})$ and $W_{2,i} = (Y_{i+1}^n, Z_{i+1}^n)$, for all $i \in \{1, \ldots, n\}$; (346) comes from the introduction of the uniform random variable $T$ over $\{1, \ldots, n\}$ and the corresponding mean random variables $U_T$, $S_T$, $W_{1,T}$, $W_{2,T}$, $Y_T$ , $Z_T$; (347) comes from the independence between $T$ and $(U_T, S_T)$; (348) comes from identifying the auxiliary random variables$W_1 = (W_{1,T}, T)$ and $W_2 = W_{2,T}$; (349) comes from the empirical coordination requirement as stated in Lemma 6. The sequences of symbols $(U^n, S^n, Z^n, X^n, Y^n, V^n)$ are not jointly typical with small error probability $\mathbb{P}(E = 1)$; (350) comes from Lemma 7. The sequences of symbols $(U^n, S^n, Z^n, X^n, Y^n, V^n)$ are jointly typical, hence the distribution of the mean random variables $\mathcal{P}_{U_T S_T Z_T X_T Y_T V_T|E=0}$ is closed to the target probability distribution $\mathcal{Q}_{USZXYV}$. The continuity of the entropy function stated pp. 33 in [32] concludes.

**Second Constraint:**

$$n \cdot \mathsf{E} \leq I(U^n, S^n; Y^n, Z^n) + n \cdot \varepsilon \tag{351}$$

$$\leq H(U^n, S^n) + n \cdot \varepsilon \tag{352}$$

$$= n \cdot \Big( H(U, S) + \varepsilon \Big), \tag{353}$$

where (351) comes from the definition of the achievable state leakage rate $\mathsf{E}$, stated in equation (330); (352) comes from the properties of the mutual information; (353) comes from the i.i.d. property of the channel states $(U, S)$.

**Third Constraint:**

$$
\begin{aligned}
n\cdot\Big(\mathsf{E}+\mathsf{R}\Big) \leq & I(U^n,S^n;Y^n,Z^n)+H(M)+n\cdot 2\varepsilon && (354)\\
= & I(U^n,S^n;Y^n,Z^n)+I(M;Y^n,Z^n)+H(M|Y^n,Z^n)+n\cdot 2\varepsilon && (355)\\
\leq & I(U^n,S^n;Y^n,Z^n)+I(M;Y^n,Z^n)+n\cdot 3\varepsilon && (356)\\
\leq & I(U^n,S^n;Y^n,Z^n)+I(M;Y^n,Z^n|U^n,S^n)+n\cdot 3\varepsilon && (357)\\
\leq & I(U^n,S^n,M;Y^n,Z^n)+n\cdot 3\varepsilon && (358)\\
= & \sum_{i=1}^{n} I(U^n,S^n,M;Y_i,Z_i|Y_{i+1}^n,Z_{i+1}^n)+n\cdot 3\varepsilon && (359)\\
\leq & \sum_{i=1}^{n} I(U^n,S^n,M,Y_{i+1}^n,Z_{i+1}^n;Y_i,Z_i)+n\cdot 3\varepsilon && (360)\\
= & \sum_{i=1}^{n} I(U_i,S_i,M,U^{i-1},S^{i-1};Y_i,Z_i)\\
& +\sum_{i=1}^{n} I(U_{i+1}^n,S_{i+1}^n,Y_{i+1}^n,Z_{i+1}^n;Y_i|U_i,S_i,M,U^{i-1},S^{i-1})\\
& +\sum_{i=1}^{n} I(U_{i+1}^n,S_{i+1}^n,Y_{i+1}^n,Z_{i+1}^n;Z_i|U_i,S_i,M,U^{i-1},S^{i-1},Y_i)+n\cdot 3\varepsilon && (361)\\
= & \sum_{i=1}^{n} I(U_i,S_i,M,U^{i-1},S^{i-1};Y_i,Z_i)\\
& +\sum_{i=1}^{n} I(U_{i+1}^n,S_{i+1}^n,Y_{i+1}^n,Z_{i+1}^n;Z_i|U_i,S_i,M,U^{i-1},S^{i-1},Y_i)+n\cdot 3\varepsilon && (362)\\
= & \sum_{i=1}^{n} I(U_i,S_i,M,U^{i-1},S^{i-1};Y_i,Z_i)+n\cdot 3\varepsilon && (363)\\
= & \sum_{i=1}^{n} I(U_i,S_i,W_{1,i};Y_i,Z_i)+n\cdot 3\varepsilon && (364)\\
= & n\cdot I(U_T,S_T,W_{1,T};Y_T,Z_T|T)+n\cdot 3\varepsilon && (365)\\
\leq & n\cdot I(U_T,S_T,W_{1,T},T;Y_T,Z_T)+n\cdot 3\varepsilon && (366)\\
\leq & n\cdot\Big(I(U_T,S_T,W_1;Y_T,Z_T)+3\varepsilon\Big) && (367)\\
\leq & n\cdot\Big(I(U_T,S_T,W_1;Y_T,Z_T|E=0)+4\varepsilon\Big) && (368)\\
\leq & n\cdot\Big(I(U,S,W_1;Y,Z)+5\varepsilon\Big), && (369)
\end{aligned}
$$

where (354) comes from the definition of achievable rate and information leakage $(\mathsf{R},\mathsf{E})$, stated

in equations (329) and (330); (356) comes from equation (331) and Fano's inequality, stated pp. 19, in [34]; (357) comes from the independence between the message $M$ and the channel states $(U^n, S^n)$, hence $I(M;Y^n,Z^n) \leq I(M;Y^n,Z^n,U^n,S^n) = I(M;Y^n,Z^n|U^n,S^n)$; (362) comes from the Markov chain $Y_i \multimap (U_i,S_i,M,U^{i-1},S^{i-1}) \multimap (U_{i+1}^n,S_{i+1}^n,Y_{i+1}^n,Z_{i+1}^n)$ stated in Lemma 13. It is a direct consequence of the causal encoding function and the memoryless property of the channel; (363) comes from the i.i.d. property of the source that induces the Markov chain $Z_i \multimap (U_i,S_i) \multimap (U_{i+1}^n,S_{i+1}^n,Y_{i+1}^n,Z_{i+1}^n,M,U^{i-1},S^{i-1},Y_i)$, hence we have: $I(U_{i+1}^n,S_{i+1}^n,Y_{i+1}^n,Z_{i+1}^n;Z_i|U_i,S_i,M,U^{i-1},S^{i-1},Y_i) = 0$; (364) comes from the introduction of the auxiliary random variable $W_{1,i} = (M,U^{i-1},S^{i-1})$, for all $i \in \{1,\dots,n\}$; (365) comes from the introduction of the uniform random variable $T$ over $\{1,\dots,n\}$ and the corresponding mean random variables $U_T, S_T$, $W_{1,T}$, $Y_T, Z_T$; (367) comes from identifying the auxiliary random variable $W_1 = (W_{1,T},T)$; (368) comes from the empirical coordination requirement as stated in Lemma 5. The sequences of symbols $(U^n,S^n,Z^n,X^n,Y^n,V^n)$ are not jointly typical with small error probability $\mathbb{P}(E=1)$; (369) comes from Lemma 7. The sequences of symbols $(U^n,S^n,Z^n,X^n,Y^n,V^n)$ are jointly typical, hence the distribution of the mean random variables $\mathcal{P}_{U_TS_TZ_TX_TY_TV_T|E=0}$ is closed to the target probability distribution $\mathcal{Q}_{USZXYV}$. The continuity of the entropy function stated pp. 33 in [32] concludes.

**Conclusion:** If the triple of rate, state leakage and probability distribution$(\mathsf{R},\mathsf{E},\mathcal{Q})$ is achievable with causal encoding, then the following equations are satisfied for all $\varepsilon > 0$:

$$I(U,S;W_1,W_2,Y,Z) - 4\varepsilon \leq \mathsf{E} \leq H(S) + \varepsilon, \tag{370}$$

$$\mathsf{R} + \mathsf{E} \leq I(U,S,W_1;Y,Z) + 5\varepsilon. \tag{371}$$

This corresponds to equations (43) and (44) and concludes the converse proof of Theorem IV.3.

**Remark H.1** *For the converse proof of Theorem IV.3, the causal encoding is not necessarily deterministic. The same optimal performances can be obtained by considering stochastic causal encoding.*

**Lemma 11** *The causal encoding function, the memoryless property of the channel and the i.i.d. property of the source induce the Markov chain property corresponding to equation* (333)*:*

$$X_i \multimap (U_i,S_i,W_{1,i}) \multimap W_{2,i}. \tag{372}$$

*This Markov chain is satisfied with* $W_{1,i} = (M,U^{i-1},S^{i-1})$ *and* $W_{2,i} = (Y_{i+1}^n,Z_{i+1}^n)$*, for all* $i \in \{1,\dots,n\}$.

*Proof.* [Lemma 11] The auxiliary random variables $W_{1,i} = (M, U^{i-1}, S^{i-1})$ and $W_{2,i} = (Y_{i+1}^n, Z_{i+1}^n)$ satisfy the following equations for all $(u^n, s^n, z^n, x^n, y^n, v^n, m)$:

$$\mathcal{P}(w_{2,i}|u_i, s_i, w_{1,i}, x_i) = \mathcal{P}(y_{i+1}^n, z_{i+1}^n|u_i, s_i, m, u^{i-1}, s^{i-1}, x_i)$$

$$= \sum_{u_{i+1}^n, s_{i+1}^n, x_{i+1}^n} \mathcal{P}(u_{i+1}^n, s_{i+1}^n, x_{i+1}^n, y_{i+1}^n, z_{i+1}^n|u_i, s_i, m, u^{i-1}, s^{i-1}, x_i) \tag{373}$$

$$= \sum_{u_{i+1}^n, s_{i+1}^n, x_{i+1}^n} \mathcal{P}(u_{i+1}^n, s_{i+1}^n, x_{i+1}^n|u_i, s_i, m, u^{i-1}, s^{i-1}, x_i)$$

$$\cdot \mathcal{P}(y_{i+1}^n, z_{i+1}^n|u_{i+1}^n, s_{i+1}^n, x_{i+1}^n, u_i, s_i, m, u^{i-1}, s^{i-1}, x_i) \tag{374}$$

$$= \sum_{u_{i+1}^n, s_{i+1}^n, x_{i+1}^n} \mathcal{P}(u_{i+1}^n, s_{i+1}^n, x_{i+1}^n|u_i, s_i, m, u^{i-1}, s^{i-1})$$

$$\cdot \mathcal{P}(y_{i+1}^n, z_{i+1}^n|u_{i+1}^n, s_{i+1}^n, x_{i+1}^n, u_i, s_i, m, u^{i-1}, s^{i-1}, x_i) \tag{375}$$

$$= \sum_{u_{i+1}^n, s_{i+1}^n, x_{i+1}^n} \mathcal{P}(u_{i+1}^n, s_{i+1}^n, x_{i+1}^n|u_i, s_i, m, u^{i-1}, s^{i-1}) \cdot \mathcal{P}(y_{i+1}^n|s_{i+1}^n, x_{i+1}^n) \cdot \mathcal{P}(z_{i+1}^n|u_{i+1}^n, s_{i+1}^n) \tag{376}$$

$$= \sum_{u_{i+1}^n, s_{i+1}^n, x_{i+1}^n} \mathcal{P}(u_{i+1}^n, s_{i+1}^n, x_{i+1}^n, y_{i+1}^n, z_{i+1}^n|u_i, s_i, m, u^{i-1}, s^{i-1}) \tag{377}$$

$$= \mathcal{P}(y_{i+1}^n, z_{i+1}^n|u_i, s_i, m, u^{i-1}, s^{i-1}) = \mathcal{P}(w_{2,i}|u_i, s_i, w_{1,i}), \tag{378}$$

where (375) comes from the causal encoding function that induces the Markov chain $X_i \multimap (U_i, S_i, M, U^{i-1}, S^{i-1}) \multimap (U_{i+1}^n, S_{i+1}^n, X_{i+1}^n)$; (376) comes from the memoryless property of the channel $Y_{i+1}^n \multimap (S_{i+1}^n, X_{i+1}^n) \multimap (U_{i+1}^n, U_i, S_i, M, U^{i-1}, S^{i-1}, X_i)$ and the i.i.d. property of the source $Z_{i+1}^n \multimap (S_{i+1}^n, U_{i+1}^n) \multimap (U_{i+1}^n, U_i, S_i, M, U^{i-1}, S^{i-1}, X_i, Y_{i+1}^n)$. This concludes the proof of Lemma 11. □

**Lemma 12** *The causal encoding function, the non-causal decoding function, the memoryless property of the channel and the i.i.d. property of the source induce the Markov chain property corresponding to equation* (336)*:*

$$V_i \multimap (Y_i, Z_i, W_{1,i}, W_{2,i}) \multimap (U_i, S_i, X_i). \tag{379}$$

*This Markov chain is satisfied with* $W_{1,i} = (M, U^{i-1}, S^{i-1})$ *and* $W_{2,i} = (Y_{i+1}^n, Z_{i+1}^n)$*, for all* $i \in \{1, \ldots, n\}$.

*Proof.* [Lemma 12] The auxiliary random variables $W_{1,i} = (M, U^{i-1}, S^{i-1})$ and $W_{2,i} = (Y_{i+1}^n, Z_{i+1}^n)$

satisfy the following equations for all $(u^n, s^n, z^n, w_1^n, w_2^n, x^n, y^n, v^n, m)$:

$$\begin{aligned}
&\mathcal{P}(v_i|y_i, z_i, w_{1,i}, w_{2,i}, u_i, s_i, x_i)\\
=&\mathcal{P}(v_i|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i)\\
=&\sum_{x^{i-1}, y^{i-1}, z^{i-1}} \mathcal{P}(v_i, x^{i-1}, y^{i-1}, z^{i-1}|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i) \qquad (380)\\
=&\sum_{x^{i-1}, y^{i-1}, z^{i-1}} \mathcal{P}(z^{i-1}|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i)\\
&\cdot \mathcal{P}(x^{i-1}|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i, z^{i-1})\\
&\cdot \mathcal{P}(y^{i-1}|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i, z^{i-1}, x^{i-1})\\
&\cdot \mathcal{P}(v_i|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i, z^{i-1}, x^{i-1}, y^{i-1}). \qquad (381)
\end{aligned}$$

we can remove $(u_i, s_i, x_i)$, in the four conditional probability distributions:

$$\mathcal{P}(z^{i-1}|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i) =\mathcal{P}(z^{i-1}|u^{i-1}, s^{i-1}), \qquad (382)$$

$$\mathcal{P}(x^{i-1}|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i, z^{i-1}) =\mathcal{P}(x^{i-1}|m, u^{i-1}, s^{i-1}), \qquad (383)$$

$$\mathcal{P}(y^{i-1}|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i, z^{i-1}, x^{i-1}) =\mathcal{P}(y^{i-1}|s^{i-1}, x^{i-1}), \qquad (384)$$

$$\mathcal{P}(v_i|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n, u_i, s_i, x_i, z^{i-1}, x^{i-1}, y^{i-1}) =\mathcal{P}(v_i|y_i, z_i, y_{i+1}^n, z_{i+1}^n, z^{i-1}, y^{i-1}), \qquad (385)$$

where (382) comes from the i.i.d. property of the information source: $Z^{i-1}$ only depends on $(U^{i-1}, S^{i-1})$; (383) comes from the causal encoding that induces the Markov chain $X^{i-1} \multimap (M, U^{i-1}, S^{i-1}) \multimap (Y_i, Z_i, Y_{i+1}^n, Z_{i+1}^n, X_i, Z^{i-1}, U_i, S_i)$; (384) comes from the memoryless property of the channel: $Y^{i-1}$ only depends on $(X^{i-1}, S^{i-1})$; (385) comes from the non-causal decoding that induces the Markov chain $V_i \multimap (Y_i, Z_i, Y_{i+1}^n, Z_{i+1}^n, Z^{i-1}, Y^{i-1}) \multimap (M, U^{i-1}, S^{i-1}, U_i, S_i, X_i, X^{i-1})$. Hence we have for all $(u^n, s^n, z^n, x^n, y^n, v^n, m)$:

$$\begin{aligned}
&\mathcal{P}(v_i|y_i, z_i, w_{1,i}, w_{2,i}, u_i, s_i, x_i)\\
=&\sum_{x^{i-1}, y^{i-1}, z^{i-1}} \mathcal{P}(v_i, x^{i-1}, y^{i-1}, z^{i-1}|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n) \qquad (386)\\
=&\mathcal{P}(v_i|y_i, z_i, m, u^{i-1}, s^{i-1}, y_{i+1}^n, z_{i+1}^n) \qquad (387)\\
=&\mathcal{P}(v_i|y_i, z_i, w_{1,i}, w_{2,i}). \qquad (388)
\end{aligned}$$

The above equation corresponds to the Markov chain $V_i \multimap (Y_i, Z_i, W_{1,i}, W_{2,i}) \multimap (U_i, S_i, X_i)$ and it concludes the proof of Lemma 12. □

**Lemma 13** *The causal encoding function and the memoryless property of the channel induce the following Markov chain property:*

$$Y_i \multimap (U_i, S_i, M, U^{i-1}, S^{i-1}) \multimap (U_{i+1}^n, S_{i+1}^n, Y_{i+1}^n, Z_{i+1}^n). \tag{389}$$

*This Markov chain is satisfied for all* $i \in \{1, \ldots, n\}$.

*Proof.* [Lemma 13] We have the following equations for all $(u^n, s^n, z^n, x^n, y^n, v^n, m)$:

$$\begin{aligned}
&\mathcal{P}(y_i|u_i, s_i, m, u^{i-1}, s^{i-1}, u_{i+1}^n, s_{i+1}^n, y_{i+1}^n, z_{i+1}^n) \\
&= \sum_{x_i} \mathcal{P}(x_i, y_i|u_i, s_i, m, u^{i-1}, s^{i-1}, u_{i+1}^n, s_{i+1}^n, y_{i+1}^n, z_{i+1}^n) && (390)\\
&= \sum_{x_i} \mathcal{P}(x_i|u_i, s_i, m, u^{i-1}, s^{i-1}, u_{i+1}^n, s_{i+1}^n, y_{i+1}^n, z_{i+1}^n) \\
&\quad \cdot \mathcal{P}(y_i|x_i, u_i, s_i, m, u^{i-1}, s^{i-1}, u_{i+1}^n, s_{i+1}^n, y_{i+1}^n, z_{i+1}^n) && (391)\\
&= \sum_{x_i} \mathcal{P}(x_i|u_i, s_i, m, u^{i-1}, s^{i-1}) \cdot \mathcal{P}(y_i|x_i, u_i, s_i, m, u^{i-1}, s^{i-1}, u_{i+1}^n, s_{i+1}^n, y_{i+1}^n, z_{i+1}^n) && (392)\\
&= \sum_{x_i} \mathcal{P}(x_i|u_i, s_i, m, u^{i-1}, s^{i-1}) \cdot \mathcal{P}(y_i|x_i, s_i) && (393)\\
&= \sum_{x_i} \mathcal{P}(x_i, y_i|u_i, s_i, m, u^{i-1}, s^{i-1}) = \mathcal{P}(y_i|u_i, s_i, m, u^{i-1}, s^{i-1}), && (394)
\end{aligned}$$

where (392) comes from the causal encoding function that induces the Markov chain $X_i \multimap (U_i, S_i, M, U^{i-1}, S^{i-1}) \multimap (U_{i+1}^n, S_{i+1}^n, Y_{i+1}^n, Z_{i+1}^n)$; (393) comes from the memoryless property of the channel that induces the Markov chain: $Y_i \multimap (X_i, S_i) \multimap (U_i, M, U^{i-1}, S^{i-1}, U_{i+1}^n, S_{i+1}^n, Y_{i+1}^n, Z_{i+1}^n)$. This concludes the proof of Lemma 13. □

## APPENDIX I
## CONVERSE PROOF OF THEOREM IV.5

The converse proof for information constraints (48) and (49) are the same as for Theorem II.3, in Appendix B. We prove the converse result for the information constraint (47). We consider the joint probability distribution $\mathcal{Q}_{SXY_1Y_2V}$ and we introduce the random event of error $E \in \{0, 1\}$ defined as follows:

$$E = \begin{cases} 0 \text{ if } & \big|\big|Q^n - \mathcal{Q}\big|\big|_1 \leq \varepsilon \iff (S^n, X^n, Y_1^n, Y_2^n, V^n) \in T_\delta(\mathcal{Q}), \\ 1 \text{ if } & \big|\big|Q^n - \mathcal{Q}\big|\big|_1 > \varepsilon \iff (S^n, X^n, Y_1^n, Y_2^n, V^n) \notin T_\delta(\mathcal{Q}). \end{cases} \tag{395}$$

Consider a sequence of code $c(n) \in \mathcal{C}$ that achieves the probability distribution $\mathcal{Q}_{SXY_1Y_2V}$, *i.e.* for which the probability of error $\mathcal{P}_{\mathsf{e}}(c) = \mathbb{P}(E = 1)$ goes to zero. We have:

$$
\begin{aligned}
n \cdot \mathsf{R} \leq & \log_2 |\mathcal{M}| + n \cdot \varepsilon && (396)\\
= & H(M) + n \cdot \varepsilon && (397)\\
= & I(M; Y_1^n) + H(M|Y_1^n) + n \cdot \varepsilon && (398)\\
\leq & I(M; Y_1^n) + n \cdot 2\varepsilon && (399)\\
= & \sum_{i=1}^{n} I(M; Y_{1,i}|Y_{1,i+1}^n) + n \cdot \varepsilon && (400)\\
= & \sum_{i=1}^{n} I(M, S^{i-1}, Y_2^{i-1}; Y_{1,i}|Y_{1,i+1}^n) - \sum_{i=1}^{n} I(S^{i-1}, Y_2^{i-1}; Y_{1,i}|Y_{1,i+1}^n, M) + n \cdot 2\varepsilon && (401)\\
= & \sum_{i=1}^{n} I(M, S^{i-1}, Y_2^{i-1}; Y_{1,i}|Y_{1,i+1}^n) - \sum_{i=1}^{n} I(Y_{1,i+1}^n; S_i, Y_{2,i}|S^{i-1}, Y_2^{i-1}, M) + n \cdot 2\varepsilon && (402)\\
\leq & \sum_{i=1}^{n} I(M, S^{i-1}, Y_2^{i-1}, Y_{1,i+1}^n; Y_{1,i}) - \sum_{i=1}^{n} I(Y_{1,i+1}^n; S_i, Y_{2,i}|S^{i-1}, Y_2^{i-1}, M) + n \cdot 2\varepsilon && (403)\\
= & \sum_{i=1}^{n} I(W_{1,i}, W_{2,i}; Y_{1,i}) - \sum_{i=1}^{n} I(W_{2,i}; S_i, Y_{2,i}|W_{1,i}) + n \cdot 2\varepsilon, && (404)
\end{aligned}
$$

where (396) comes from the definition of achievable rate $\mathsf{R}$, stated in equation (154); (397) comes from the uniform distribution of the random message $M$; (398) comes from the definition of the mutual information; (399) comes from equation (156) and Fano's inequality, stated pp. 19, in [34]; (402) comes from Csiszár Sum Identity stated pp. 25 in [34]; (404) comes from the introduction of the auxiliary random variables $W_{1,i} = (M, S^{i-1}, Y_2^{i-1})$ and $W_{2,i} = Y_{1,i+1}^n$, that satisfy the properties corresponding to the set of probability distributions $\mathbb{Q}_{\mathsf{f}}$, as proved in Lemma 14.

**Lemma 14** *For all $i \in \{1, \ldots, n\}$, the auxiliary random variables $W_{1,i} = (M, S^{i-1}, Y_2^{i-1})$ and $W_{2,i} = Y_{1,i+1}^n$ satisfy following the properties corresponding to the set of probability distributions $\mathbb{Q}_{\mathsf{f}}$:*

$$
\begin{aligned}
& (S_i) \text{ are independent of } W_{1,i}, && (405)\\
& (Y_{1,i}, Y_{2,i}) \multimap (X_i, S_i) \multimap W_{1,i}, && (406)\\
& W_{2,i} \multimap (S_i, Y_{2,i}, W_{1,i}) \multimap (X_i, Y_{1,i}), && (407)\\
& V_i \multimap (Y_{1,i}, W_{1,i}, W_{2,i}) \multimap (X_i, S_i, Y_{2,i}). && (408)
\end{aligned}
$$

Equation (404) gives:

$$\begin{aligned}
n \cdot \mathsf{R} \leq & \sum_{i=1}^{n} I(W_{1,i}, W_{2,i}; Y_{1,i}) - \sum_{i=1}^{n} I(W_{2,i}; S_i, Y_{2,i} | W_{1,i}) + n \cdot 2\varepsilon \\
= & n \cdot \Big( I(W_{1,T}, W_{2,T}; Y_{1,T} | T) - I(W_{2,T}; S_T, Y_{2,T} | W_{1,T}, T) + 2\varepsilon \Big) && (409) \\
\leq & n \cdot \Big( I(T, W_{1,T}, W_{2,T}; Y_{1,T}) - I(W_{2,T}; S_T, Y_{2,T} | W_{1,T}, T) + 2\varepsilon \Big) && (410) \\
\leq & n \cdot \max_{\mathcal{Q} \in \mathbb{Q}_{\mathsf{f}}} \Big( I(W_1, W_2; Y_{1,T}) - I(W_2; S_T, Y_{2,T} | W_1) + 2\varepsilon \Big) && (411) \\
\leq & n \cdot \max_{\mathcal{Q} \in \mathbb{Q}_{\mathsf{f}}} \Big( I(W_1, W_2; Y_{1,T} | E = 0) - I(W_2; S_T, Y_{2,T} | W_1, E = 0) + 3\varepsilon \Big) && (412) \\
\leq & n \cdot \max_{\mathcal{Q} \in \mathbb{Q}_{\mathsf{f}}} \Big( I(W_1, W_2; Y_1) - I(W_2; S, Y_2 | W_1) + 4\varepsilon \Big), && (413)
\end{aligned}$$

where (409) comes from the introduction of the uniform random variable $T$ over $\{1, \ldots, n\}$ and the introduction of the corresponding mean random variables $S_T$, $W_{1,T}$, $W_{2,T}$, $X_T$, $Y_{1,T}$, $Y_{2,T}$, $V_T$; (410) and comes from the properties of the mutual information; (411) comes from the identifying $W_1$ with $(W_{1,T}, T)$, $W_2$ with $W_{2,T}$ and taking the maximum over the probability distributions that belong to the set $\mathbb{Q}_{\mathsf{f}}$; (412) comes from the empirical coordination requirement as stated in Lemma 5 in Section B-B, since the sequences are not jointly typical with small error probability $\mathbb{P}(E = 1)$; (413) comes from Lemma 7 in Appendix B-B, that states that the probability distribution induced by the coding scheme $\mathcal{P}_{S_T X_T Y_{1,T} Y_{2,T} V_T | E=0}$ is closed to the target probability distribution $\mathcal{Q}(s, x, y_1, y_2, v)$. The continuity of the entropy function stated pp. 33 in [32] concludes. This concludes the converse proof of Theorem IV.5.

*Proof.* [Lemma 14] Equation (405) comes from the i.i.d. property of the source $S$, the independence of $S$ with respect to the message $M$ and the causal encoding function, hence $S_i$ is independent of the past channel inputs $X^{i-1}$, hence $S_i$ is independent of $Y_2^{i-1}$; (406) comes from the memoryless property of the channel, hence $(Y_{1,i}, Y_{2,i})$ is drawn with $(X_i, S_i)$; (407) comes from the following equations:

$$\begin{aligned}
& \mathcal{P}(x_i, y_{1,i} | s_i, y_{2,i}, w_{1,i}, w_{2,i}) && (414) \\
= & \mathcal{P}(x_i | s_i, y_{2,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n) \cdot \mathcal{P}(y_{1,i} | x_i, s_i, y_{2,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n) && (415) \\
= & \mathcal{P}(x_i | s_i, y_{2,i}, m, s^{i-1}, y_2^{i-1}) \cdot \mathcal{P}(y_{1,i} | x_i, s_i, y_{2,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n) && (416) \\
= & \mathcal{P}(x_i | s_i, y_{2,i}, m, s^{i-1}, y_2^{i-1}) \cdot \mathcal{P}(y_{1,i} | x_i, s_i, y_{2,i}, m, s^{i-1}, y_2^{i-1}) && (417) \\
= & \mathcal{P}(x_i, y_{1,i} | s_i, y_{2,i}, w_{1,i}) \qquad \forall (s^n, x^n, y_1^n, y_2^n, w_1^n, w_2^n). && (418)
\end{aligned}$$

Equation (415) comes from the choice of the auxiliary random variables $W_{1,i} = (M, S^{i-1}, Y_2^{i-1})$ and $W_{2,i} = Y_{1,i+1}^n$; (416) comes from the causal encoding that implies $X_i$ is a function of $(S_i, M, S^{i-1}, Y_2^{i-1})$

but not of $Y_{1,i+1}^n$; (417) comes from the memoryless property of the channel that implies $Y_{1,i}$ is drawn depending on $(X_i, S_i, Y_{2,i})$ and not on $Y_{1,i+1}^n$; (418) concludes that the Markov chain (407) holds.

Equation (408) comes from the following equations for all $(u^n, s^n, z^n, x^n, y^n, v^n, m)$:

$$\mathcal{P}(v_i|y_{1,i}, w_{1,i}, w_{2,i}, x_i, s_i, y_{2,i}) \tag{419}$$

$$= \sum_{x^{i-1}, y_1^{i-1}} \mathcal{P}(v_i, x^{i-1}, y_1^{i-1}|y_{1,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n, x_i, s_i, y_{2,i}) \tag{420}$$

$$= \sum_{x^{i-1}, y_1^{i-1}} \mathcal{P}(x^{i-1}|y_{1,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n, x_i, s_i, y_{2,i})$$

$$\cdot \mathcal{P}(y_1^{i-1}|x^{i-1}, y_{1,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n, x_i, s_i, y_{2,i})$$

$$\cdot \mathcal{P}(v_i|x^{i-1}, y_1^{i-1}, y_{1,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n, x_i, s_i, y_{2,i}) \tag{421}$$

$$= \sum_{x^{i-1}, y_1^{i-1}} \mathcal{P}(x^{i-1}|y_{1,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n) \tag{422}$$

$$\cdot \mathcal{P}(y_1^{i-1}|x^{i-1}, y_{1,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n) \tag{423}$$

$$\cdot \mathcal{P}(v_i|x^{i-1}, y_1^{i-1}, y_{1,i}, m, s^{i-1}, y_2^{i-1}, y_{1,i+1}^n) \tag{424}$$

$$= \mathcal{P}(v_i|y_{1,i}, w_{1,i}, w_{2,i}), \tag{425}$$

where (420) comes from the choice of the auxiliary random variables $W_{1,i} = (M, S^{i-1}, Y_2^{i-1})$ and $W_{2,i} = Y_{1,i+1}^n$; (421) comes from the decomposition of the probability; (422) comes from the causal encoding that implies $X^{i-1}$ is a function of $(M, S^{i-1}, Y_2^{i-2})$ but not of $(X_i, S_i, Y_{2,i})$; (423) comes from the memoryless property of the channel that implies $Y_1^{i-1}$ depends only on $(X^{i-1}, S^{i-1}, Y_2^{i-1})$ and not on $(X_i, S_i, Y_{2,i})$; (424) comes from the non-causal decoding that implies $V_i$ is a function of $(Y_1^{i-1}, Y_{1,i}, Y_{1,i+1}^n)$ but not of $(X_i, S_i, Y_{2,i})$; (425) concludes that the Markov chain (408) holds. This concludes the proof of Lemma 14. □

## APPENDIX J
## CONVERSE PROOF OF THEOREM V.2

We consider that the triple of rate, information leakage and probability distribution $(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable with a strictly causal code. We introduce the random event of error $E \in \{0, 1\}$ defined with respect to the achievable joint probability distribution $\mathcal{Q}_{SXYV}$, as follows:

$$E = \begin{cases} 0 & \text{if} \quad \left|\left|Q^n - \mathcal{Q}\right|\right|_{\mathsf{tv}} \leq \varepsilon \quad \Longleftrightarrow \quad (S^n, X^n, Y^n, V^n) \in T_\delta(\mathcal{Q}), \\ 1 & \text{if} \quad \left|\left|Q^n - \mathcal{Q}\right|\right|_{\mathsf{tv}} > \varepsilon \quad \Longleftrightarrow \quad (S^n, X^n, Y^n, V^n) \notin T_\delta(\mathcal{Q}). \end{cases} \tag{426}$$

The event $E = 1$ occurs if the sequences $(S^n, X^n, Y^n, V^n) \notin T_\delta(\mathcal{Q})$ are not jointly typical for the target probability distribution $\mathcal{Q}$. By definition V.1, for all $\varepsilon > 0$, there exists a $\bar{n}$ such that for all $n \geq \bar{n}$, there exists a code $c^\star \in \mathcal{C}(n, \mathcal{M})$ that satisfies the three following equations:

$$\begin{aligned} \frac{\log_2 |\mathcal{M}|}{n} &\geq \mathsf{R} - \varepsilon, && (427) \\ \left| \mathcal{L}_{\mathsf{e}}(c^\star) - \mathsf{E} \right| &= \left| \frac{1}{n} \cdot I(S^n; Y^n) - \mathsf{E} \right| \leq \varepsilon, && (428) \\ \mathcal{P}_{\mathsf{e}}(c^\star) &= \mathbb{P}\left( M \neq \hat{M} \right) + \mathbb{P}\left( \left|\left| Q^n - \mathcal{Q} \right|\right|_{\mathsf{tv}} \geq \varepsilon \right) \leq \varepsilon. && (429) \end{aligned}$$

We introduce the auxiliary random variables $W_{2,i} = (M, S^{i-1}, Y_{i+1}^n)$ that satisfy the Markov chains of the set of probability distribution $\mathbb{Q}_{\mathsf{se}}$ for all $i \in \{1, \ldots, n\}$:

$$S_i \text{ independent of } X_i, \tag{430}$$

$$Y_i \multimap (X_i, S_i) \multimap W_{2,i}, \tag{431}$$

$$V_i \multimap (Y_i, X_i, W_{2,i}) \multimap S_i, \tag{432}$$

where (430) comes from the strictly causal encoding function; (431) comes from the memoryless property of the channel $\mathcal{T}_{Y|XS}$; (432) comes from slight modification of Lemma 3 where only the random variable $S_i$ is removed from the equations (201), (202), (203). It is a direct consequence of the strictly causal encoding function, the non-causal decoding function and the memoryless property of the channel $\mathcal{T}_{Y|XS}$.

We introduce the random variable $T$ that is uniformly distributed over the indices $\{1, \ldots, n\}$ and the corresponding mean random variables $W_{2,T}$, $S_T$, $X_T$, $Y_T$, $V_T$. The auxiliary random variable $W_2 = (W_{2,T}, T)$ belongs to the set of probability distributions $\mathbb{Q}_{\mathsf{se}}$ and satisfies the information constraints of Theorem V.2:

$$\begin{aligned} I(S; X, W_2, Y) \leq \quad \mathsf{E} \quad &\leq H(S), && (433) \\ \mathsf{R} + \mathsf{E} \quad &\leq I(X, S; Y). && (434) \end{aligned}$$

**First Constraint:**

$$\begin{aligned} n \cdot \mathsf{E} \geq & I(S^n; Y^n) - n \cdot \varepsilon && (435) \\ = & I(S^n; Y^n, M) - I(S^n; M | Y^n) - n \cdot \varepsilon && (436) \\ \geq & n \cdot H(S) - H(S^n | Y^n, M) - H(M | Y^n) - n \cdot \varepsilon && (437) \\ \geq & n \cdot H(S) - H(S^n | Y^n, M) - n \cdot 2\varepsilon && (438) \\ = & n \cdot H(S) - \sum_{i=1}^{n} H(S_i | Y^n, M, S^{i-1}) - n \cdot 2\varepsilon && (439) \end{aligned}$$

$$\geq n \cdot H(S) - \sum_{i=1}^{n} H(S_i|Y_{i+1}^n, Y_i, M, S^{i-1}) - n \cdot 2\varepsilon \tag{440}$$

$$= n \cdot H(S) - \sum_{i=1}^{n} H(S_i|Y_{i+1}^n, Y_i, M, S^{i-1}, X_i) - n \cdot 2\varepsilon \tag{441}$$

$$= n \cdot H(S) - \sum_{i=1}^{n} H(S_i|W_{2,i}, X_i, Y_i) - n \cdot 2\varepsilon \tag{442}$$

$$= n \cdot H(S) - n \cdot H(S_T|W_{2,T}, X_T, Y_T, T) - n \cdot 2\varepsilon \tag{443}$$

$$= n \cdot H(S) - n \cdot H(S_T|W_2, X_T, Y_T) - n \cdot 2\varepsilon \tag{444}$$

$$\geq n \cdot H(S) - n \cdot H(S_T|W_2, X_T, Y_T, E = 0) - n \cdot 3\varepsilon \tag{445}$$

$$\geq n \cdot H(S) - n \cdot H(S|W_2, X, Y) - n \cdot 4\varepsilon \tag{446}$$

$$= n \cdot \Big( I(S; W, X, Y) - 4\varepsilon \Big), \tag{447}$$

where (435) comes from the definition of achievable information leakage rate $\mathsf{E}$, stated in equation (428); (438) comes from equation (429) and Fano's inequality, stated pp. 19, in [34]; (441) comes from the strictly causal encoding $X_i = f_i(M, S^{i-1})$ that implies $I(S_i; X_i|Y_{i+1}^n, Y_i, M, S^{i-1}) = 0$, for all $i \in \{1, \dots, n\}$; (442) comes from the introduction of the auxiliary random variable $W_{2,i} = (M, S^{i-1}, Y_{i+1}^n)$, for all $i \in \{1, \dots, n\}$; (443) comes from the introduction of the uniform random variable $T$ over $\{1, \dots, n\}$ and the introduction of the corresponding mean random variables $S_T$, $W_{2,T}$, $X_T$, $Y_T$; (444) comes from identifying $W_2 = (W_{2,T}, T)$; (445) comes from the empirical coordination requirement as stated in Lemma 6. The sequences of symbols $(S^n, X^n, Y^n, V^n)$ are not jointly typical with small error probability $\mathbb{P}(E = 1)$; (446) comes from Lemma 7. The sequences of symbols $(S^n, X^n, Y^n, V^n)$ are jointly typical, hence the distribution of the mean random variables $\mathcal{P}_{S_T X_T Y_T V_T|E=0}$ is closed to the target probability distribution $\mathcal{Q}_{SXYV}$. The continuity of the entropy function stated pp. 33 in [32] concludes.

**Second Constraint:**

$$n \cdot \mathsf{E} \leq I(S^n; Y^n) + n \cdot \varepsilon \leq H(S^n) + n \cdot \varepsilon = n \cdot \Big( H(S) + \varepsilon \Big), \tag{448}$$

where (448) comes from the definition of the leakage rate $\mathsf{E}$, stated in equation (428) and the i.i.d. property of the channel states $S$.

**Third Constraint:**

$$
\begin{aligned}
n \cdot \Big( \mathsf{E} + \mathsf{R} \Big) &\leq I(S^n; Y^n) + H(M) + n \cdot 2\varepsilon && (449)\\
&= I(S^n; Y^n) + I(M; Y^n) + H(M|Y^n) + n \cdot 2\varepsilon && (450)\\
&\leq I(S^n; Y^n) + I(M; Y^n) + n \cdot 3\varepsilon && (451)\\
&\leq I(S^n; Y^n) + I(M; Y^n|S^n) + n \cdot 3\varepsilon && (452)\\
&\leq I(S^n, M; Y^n) + n \cdot 3\varepsilon && (453)\\
&\leq I(S^n, X^n; Y^n) + n \cdot 3\varepsilon && (454)\\
&\leq n \cdot \Big( I(S, X; Y) + 3\varepsilon \Big), && (455)
\end{aligned}
$$

where (449) comes from the definition of achievable rate and information leakage $(\mathsf{R}, \mathsf{E})$, stated in equations (427) and (428); (451) comes from equation (429) and Fano's inequality, stated pp. 19, in [34]; (452) comes from the independence between the message $M$ and the channel states $S^n$, hence $I(M; Y^n) \leq I(M; Y^n, S^n) = I(M; Y^n|S^n)$; (454) comes from the Markov chain $Y^n \multimap (X^n, S^n) \multimap M$, induced by the channel; (455) comes from the memoryless property of the channel that implies: $H(Y^n|S^n, X^n) = n \cdot H(Y|X, S)$.

**Conclusion:** If the triple of rate, information leakage and probability distribution$(\mathsf{R}, \mathsf{E}, \mathcal{Q})$ is achievable with strictly causal encoding, then the following equations are satisfied for all $\varepsilon > 0$:

$$
\begin{aligned}
I(S; X, Y, W_2) - 4\varepsilon \leq \quad \mathsf{E} \quad &\leq H(S) + \varepsilon, && (456)\\
\mathsf{R} + \mathsf{E} \quad &\leq I(X, S; Y) + 3\varepsilon. && (457)
\end{aligned}
$$

This corresponds to equations (53) and (54) and concludes the converse proof of Theorem V.2.