IEEE Signal Processing Society | Transactions on Information Forensics and Security

Full Secrecy for the Wiretap Channel with State Information Available at Both Encoder and Decoder

Journal:	*Transactions on Information Forensics & Security*
Manuscript ID:	T-IFS-02794-2012
Manuscript Type:	Regular Paper
Date Submitted by the Author:	31-Jul-2012
Complete List of Authors:	Le Treust, Mael; Université Paris-Est Marne la Vallée, LIGM Zaidi, Abdellatif; Université Paris-Est Marne la Vallée, LIGM
EDICS:	INF-SECC-Security over channels (wire-tap, broadcast, multiple access, interference, multi-user, etc.) < INF-INFORMATION THEORETIC SECURITY

Full Secrecy for the Wiretap Channel with State Information Available at Both Encoder and Decoder

Maël Le Treust and Abdellatif Zaidi
Laboratoire d'Informatique Gaspard Monge
Université Paris-Est Marne La Vallée, 77454, Marne La Vallée Cedex 2, France
Email: {mael.letreust,abdellatif.zaidi}@univ-mlv.fr

***Abstract*—We investigate the problem of secure communication over a wiretap channel with state information available at both encoder and decoder. In this framework, the secrecy capacity is strongly related to the knowledge, by both encoder and decoder, of the past, the current and the future channel states. The secrecy capacity has not been characterized yet, for neither causal, nor non-causal state information. Our results provide a better understanding of the fundamental problems that arise in secure communication with state information and establish a connection with the wiretap channel with shared key. We introduce the concept of anti-causal state information i.e. the length of the sequence of states available at both encoder and decoder is arbitrarily larger than the length of the transmission block. The analysis reveals that the state information can be utilized as a secret key that is shared among the encoder and the decoder. We establish the secrecy capacity for two interesting cases. The first is the discrete channel with anti-causal state information and the second is the Gaussian channel with non-causal state information. For both cases, our results show that the encoder and the decoder can use the channel state information in order to secure all the information that can be transmitted reliably.**

***Index Terms*—Physical layer security, Shannon theory, wiretap channel, state information, secrecy capacity.**

I. Introduction

The problem of secure communication is modeled by the wiretap channel in which an encoder sends information to a decoder, while keeping it secret from the eavesdropper. In [1], Shannon shows that secure communication is possible when a secret key is available at both encoder and decoder, but not at the eavesdropper. Wyner [2] investigates the problem of degraded wiretap channel and exploits the statistics of the channel in order to secure information. Csiszar and Körner [3] extended the result of Wyner [2] to the case of general wiretap channel where the legitimate decoder and the eavesdropper have also to decode a common message.

We investigate the problem of secure communication over a wiretap channel with state information available at both encoder and decoder [4] but not at the eavesdropper. The communication possibilities for a channel with states are deeply related to the knowledge of the state parameter by both encoder and decoder. This knowledge is "strictly causal" (resp. "causal") when at a given instant $t \in \{1, \dots, n\}$ of the transmission block, the encoder only observes the past (resp. the past and the current) states of the channel. This knowledge is "non-causal" when both encoder and decoder observe the sequence of states that corresponds to the transmission block. Figure 1 represents the channel with non-causal state information. This knowledge is "anti-causal" (resp. "infinite") when both encoder and decoder observe a sequence of states that is arbitrarily larger than the transmission block (resp. that is infinite). Figure 2 represents the channel with anti-causal state information. In [5], Chen and Vinck establish a lower bound on the secrecy capacity for the wiretap channel with state information non-causally available at the encoder. This result is obtained using a combination of the Gel'fand and Pinsker coding [4] and the Wyner coding [2] and it extends easily to the broadcast wiretap channel with state information [6]. The authors of [7] strengthen the result of Chen and Vinck [5] and provide a lower bound on the secrecy capacity when distinct channel state information are non-causally available at encoder and decoder.

Subsequently, the article of Chia and El Gamal [8] introduces a promising perspective for the study of the wiretap channel with state information at both encoder and decoder. Considering the channel state information as a secret key shared by the encoder and decoder, enhances secure communication significantly. The two main results presented in [8] are bounds on the secrecy capacity. The first result is a lower bound for the case of causal state information which is strictly larger than the lower bound for the non-causal case stated in [5]. The second result is an upper bound for the case of non-causal state information. An important contribution of [8] is to connect the literature devoted to the wiretap channel with state information and the literature devoted to the wiretap channel with secret shared key. In fact, the secrecy capacity when considering a secret shared key is characterized in [9] for the less noisy wiretap channel and in [10] for the general wiretap channel.

The secrecy capacity for the general wiretap channel with state information available at both encoder and decoder is not available yet in the literature for neither the strictly causal, nor the causal, nor the non-causal case. In this article, we introduce the concept of anti-causal state information and we characterize the secrecy capacity for two interesting cases. The first one, stated in Theorem 11, considers the discrete channel with anti-causal state information, and the second one, stated in Theorem 16, considers the Gaussian channel with non-causal state information. Remarkably, in both situations the secrecy capacity is shown to be equal to the capacity of the same channel without eavesdropper. We show that the knowledge of the sequence of states allow the encoder and the decoder

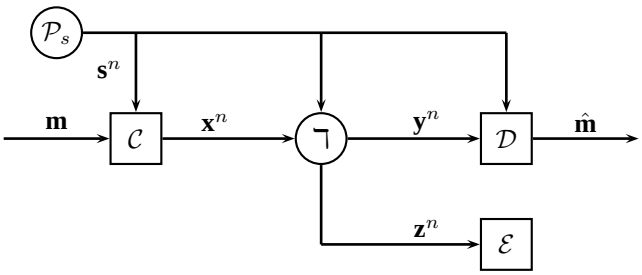

Fig. 1. Wiretap channel $\urcorner(\mathbf{y}, \mathbf{z}|\mathbf{x}, \mathbf{s})$ with **non-causal** state information $\mathbf{s}^n \in \mathcal{S}^n$ drawn according to the i.i.d. probability distribution $\mathcal{P}_s \in \Delta(\mathcal{S})$ and available at both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ but not at the eavesdropper. Encoder $\mathcal{C}$ represents the message $\mathbf{m} \in \mathcal{M}$ by a sequence of inputs $\mathbf{x}^n \in \mathcal{X}^n$. Decoder $\mathcal{D}$ observes a sequence of outputs $\mathbf{y}^n \in \mathcal{Y}^n$ and reconstructs the message $\hat{\mathbf{m}} \in \mathcal{M}$. The eavesdropper $\mathcal{E}$ observes a sequence of outputs $\mathbf{z}^n \in \mathcal{Z}^n$ and tries to decode the message $\mathbf{m} \in \mathcal{M}$ sent by encoder $\mathcal{C}$.

to secure all the information that can be transmitted reliably. Ultimately, we prove that the secrecy capacity with anti-causal state information is larger (and sometimes strictly larger) than the secrecy capacity with non-causal state information.

Section II presents the channel model under investigation and the concepts of strictly causal, causal, non-causal, anti-causal and infinite knowledge of the state information. The results for the discrete case are stated in Section III where the first contribution (Theorem 11) is devoted to the case of anti-causal state information. The results for the Gaussian case are stated in Section IV where the second contribution (Theorem 16) is devoted to the case of non-causal state information. Section V contains the proof of the Theorem 11 and Section VI concludes the article.

II. Preliminaries

The problems under investigation are depicted in Figures 1 and 2. We denote by $\mathbf{m}$, $\mathbf{s}^n$, $\mathbf{x}^n$, $\mathbf{y}^n$, $\mathbf{z}^n$, and $\hat{\mathbf{m}}$ the random variables of the message $m \in \mathcal{M}$, the sequences of states $s^n = (s_1, \ldots, s_n) \in \mathcal{S}^n$, of channel inputs $x^n \in \mathcal{X}^n$, of legitimate channel outputs $y^n \in \mathcal{Y}^n$, of eavesdropper channel outputs $z^n \in \mathcal{Z}^n$ and of the message $\hat{m} \in \mathcal{M}$ obtained as output by the decoder. The notation $\Delta(\mathcal{X})$ stands for the set of the probability distributions $\mathcal{P}(\mathbf{x})$ over the set $\mathcal{X}$ and $\mathcal{S}^\infty$ stands for the set of infinite sequences $s^\infty \in \mathcal{S}^\infty$ of channel states. A wiretap channel with state is defined by a probability distribution $\mathcal{P}_s \in \Delta(\mathcal{S})$ over the channel states and a transition probability $\urcorner : \mathcal{X} \times \mathcal{S} \longrightarrow \Delta(\mathcal{Y} \times \mathcal{Z})$. The statistics of the transition probability $\urcorner$ and the probability distribution $\mathcal{P}_s$ are known by both encoder $\mathcal{C}$ and decoder $\mathcal{D}$.

The information message $\mathbf{m} \in \mathcal{M}$ is supposed to be drawn according to the uniform probability distribution over $\mathcal{M}$. Encoder $\mathcal{C}$ aims at transmitting the realization of the message $m \in \mathcal{M}$ to decoder $\mathcal{D}$ using a transmission block of $n \in \mathbb{N}$ input symbols $x^n \in \mathcal{X}^n$. The state $s \in \mathcal{S}$ of the channel is drawn according to the i.i.d. probability distribution $\mathcal{P}_s$ and it's observation by both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ should be defined carefully. Indeed, the capacity of the wiretap channel with state information is strongly related to the length of the sequence of channel states observed by the encoder $\mathcal{C}$ and the decoder $\mathcal{D}$.

- **Strictly causal case.** At instant $t \in \{1, \ldots, n\}$ of the transmission block, encoder $\mathcal{C}$ observes the sequence of past states $s^{t-1} = (s_1, \ldots, s_{t-1}) \in \mathcal{S}^{t-1}$, whereas decoder $\mathcal{D}$ observes the sequence of the states $s^n = (s_1, \ldots, s_n) \in \mathcal{S}^n$ that corresponds to the transmission block.
- **Causal case.** At instant $t \in \{1, \ldots, n\}$ of the transmission block, encoder $\mathcal{C}$ observes the sequence of past and current states $s^t = (s_1, \ldots, s_t) \in \mathcal{S}^t$, whereas decoder $\mathcal{D}$ observes the sequence of the states $s^n = (s_1, \ldots, s_n) \in \mathcal{S}^n$ that corresponds to the transmission block.
- **Non-causal case.** At the first instant of the transmission block, both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ observe the sequence of states $s^n = (s_1, \ldots, s_n) \in \mathcal{S}^n$ that corresponds to the transmission block (see Figure 1 and Definition 3).

The secrecy capacity is not available in the literature for neither the strictly causal, nor the causal, nor the non-causal case. In this article, we introduce the concept of anti-causal and infinite state information and we characterize the secrecy capacity for two interesting cases (Theorems 11 and 16). The results presented in this article provide a better understanding of the fundamental problems that arise in secure communication with channel state information.

- **Anti-causal case.** At the first instant of the transmission block, both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ observe the sequence of states $s^{n+k} = (s_1, \ldots, s_n, \ldots, s_{n+k}) \in \mathcal{S}^{n+k}$ whose length $n + k \in \mathbb{N}$ is arbitrarily larger than the length $n \in \mathbb{N}$ of the transmission block (see Figure 2 and Definition 1).
- **Infinite case.** At the first instant of the transmission block, both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ observe the infinite sequence of states $s^\infty = (s_1, \ldots, s_n, \ldots, s_{n+k}, \ldots) \in \mathcal{S}^\infty$.

Encoder $\mathcal{C}$ uses the sequence of input symbols $x^n \in \mathcal{X}^n$ and the knowledge of the appropriate sequence of states $(s^{t-1}, s^t, s^n, s^{n+k}, s^\infty)$ in order to transmit the messages $m \in \mathcal{M}$ to decoder $\mathcal{D}$. The pair of sequences of channel outputs $(y^n, z^n) \in \mathcal{Y}^n \times \mathcal{Z}^n$ is drawn according to the discrete and memoryless transition probability defined by equation (1).

$$\urcorner^{\otimes n}(y^n, z^n|x^n, s^n) = \prod_{t=1}^{n} \urcorner(y_t, z_t|x_t, s_t). \tag{1}$$

Decoder $\mathcal{D}$ observes the sequence of channel outputs $y^n \in \mathcal{Y}^n$ and reconstruct the message $\hat{m} \in \mathcal{M}$. The eavesdropper $\mathcal{E}$ observes the sequence of outputs $z^n \in \mathcal{Z}^n$ and tries to decode the message $m \in \mathcal{M}$ sent by encoder $\mathcal{C}$.

Definition 1 *An anti-causal code $c \in \mathcal{AC}(n, k, M)$ is a pair of functions (f, g) defined by equations (2) and (3).*

$$f: \quad \mathcal{M} \times \mathcal{S}^{n+k} \longrightarrow \mathcal{X}^n, \tag{2}$$

$$g: \quad \mathcal{Y}^n \times \mathcal{S}^{n+k} \longrightarrow \mathcal{M}. \tag{3}$$

We denote by $\mathcal{AC}(n, k, M)$ the set of anti-causal codes for which the parameters $n \in \mathbb{N}$, $k \in \mathbb{N}$ and the cardinality $M = |\mathcal{M}| \in \mathbb{N}$ are fixed.

Remark 2 Inspired from control theory, the notion of anti-causal state information corresponds to the situation described by Figure 2, where the length $n+k \in \mathbb{N}$ of the sequence of state information is arbitrarily larger than the length $n \in \mathbb{N}$ of the transmission block.

Non-causal codes are defined similarly to anti-causal code, where the set of sequences $\mathcal{S}^{n+k}$ is replaced by the set of sequences $\mathcal{S}^n$.

Definition 3 *A non-causal code $c \in \mathcal{NC}(n, M)$ is a pair of functions (f, g) defined by equations (4) and (5).*

$$f: \quad \mathcal{M} \times \mathcal{S}^n \longrightarrow \mathcal{X}^n, \tag{4}$$

$$g: \quad \mathcal{Y}^n \times \mathcal{S}^n \longrightarrow \mathcal{M}. \tag{5}$$

We denote by $\mathcal{NC}(n, M)$ the set of non-causal codes for which the parameter $n \in \mathbb{N}$ and the cardinality $M = |\mathcal{M}| \in \mathbb{N}$ are fixed.

Each anti-causal or non-causal code induces an error probability and a leakage rate stated formally in Definition 4. We denote by $\hat{\mathbf{m}} = g(\mathbf{y}^n, \mathbf{s}^n)$ the random variable of the message obtained as output of the decoding function.

Definition 4 *For each anti-causal code $c \in \mathcal{AC}(n, k, M)$ (resp. non-causal code $c \in \mathcal{NC}(n, M)$), the error probability $\mathcal{P}_e(c)$ and the information leakage rate $\mathcal{L}_e(c)$ are defined by equations (6) and (7).*

$$\mathcal{P}_e(c) = \quad \mathcal{P}(\boldsymbol{m} \neq \hat{\boldsymbol{m}} |\, c\,), \tag{6}$$

$$\mathcal{L}_e(c) = \quad \frac{I(\boldsymbol{m}; \boldsymbol{z}^n |\, c\,)}{n}. \tag{7}$$

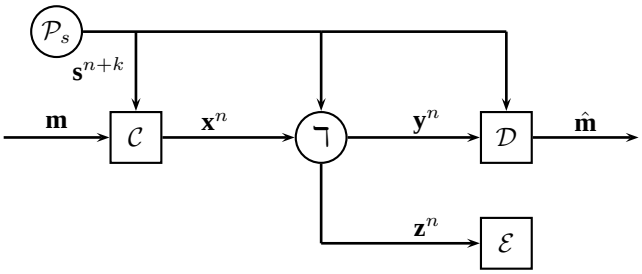

Fig. 2. Wiretap channel $\urcorner(\mathbf{y}, \mathbf{z}|\mathbf{x}, \mathbf{s})$ with **anti-causal** state information $\mathbf{s}^{n+k} \in \mathcal{S}^{n+k}$ available at both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ but not at the eavesdropper. The length $n+k \in \mathbb{N}$ of the sequence of states is arbitrarily larger than the length $n \in \mathbb{N}$ of the sequence of states observed in the non-causal case.

To make the proof of Theorem 11 easier to understand, we formally define the notion of random code for the case of anti-causal state information.

Definition 5 *A random code $r \in \Delta\Big(\mathcal{AC}(n, k, M)\Big)$ is a probability distribution over the set of anti-causal codes $\mathcal{AC}(n, k, M)$. The expected error probability $\mathbb{E}_c\Big[\mathcal{P}_e(\boldsymbol{c})\Big]$ and the expected information leakage rate $\mathbb{E}_c\Big[\mathcal{L}_e(\boldsymbol{c})\Big]$ are defined by equations (8) and (9) where $\boldsymbol{c} \in \mathcal{AC}(n, k, M)$ denote the random variable of the code.*

$$\mathbb{E}_c\Big[\mathcal{P}_e(\boldsymbol{c})\Big] = \quad \sum_{c \in \mathcal{AC}(n,k,M)} r(\boldsymbol{c} = c)\mathcal{P}(\boldsymbol{m} \neq \hat{\boldsymbol{m}} | \boldsymbol{c} = c), \tag{8}$$

$$\mathbb{E}_c\Big[\mathcal{L}_e(\boldsymbol{c})\Big] = \quad \sum_{c \in \mathcal{AC}(n,k,M)} r(\boldsymbol{c} = c)\frac{I(\boldsymbol{m}; \boldsymbol{z}^n | \boldsymbol{c} = c)}{n}. \tag{9}$$

The aim of our study is to prove the existence of an anti-causal code $c \in \mathcal{AC}(n, k, M)$ (resp. a non-causal code $c \in \mathcal{NC}(n, M)$) with the maximal information rate $\frac{\log_2 M}{n}$ under the constraint that the error probability $\mathcal{P}(\hat{\mathbf{m}} \neq \mathbf{m} |\, c\,)$ and the information leakage rate $\frac{I(\mathbf{m}; \mathbf{z}^n |\, c\,)}{n}$ are arbitrarily small.

Definition 6 *A rate R is anti-causally achievable (resp. non-causally achievable), if for all $\varepsilon > 0$, there exists some parameters $n \in \mathbb{N}$, $k \in \mathbb{N}$, $M \in \mathbb{N}$ and a code $c \in \mathcal{AC}(n, k, M)$ (resp. there exists some parameters $n \in \mathbb{N}$, $M \in \mathbb{N}$ and a code $c \in \mathcal{NC}(n, M)$) such that:*

$$\frac{\log_2 M}{n} \quad \geq \quad R - \varepsilon, \tag{10}$$

$$\mathcal{P}_e(c) \quad \leq \quad \varepsilon, \tag{11}$$

$$\mathcal{L}_e(c) \quad \leq \quad \varepsilon. \tag{12}$$

The anti-causal secrecy capacity C_{ac} (resp. non-causal secrecy capacity C_{nc}) of the wiretap channel with state information available at both encoder and decoder is defined as the supremum of the anti-causally (resp. non-causally) achievable rates R.

Remark 7 For the case of anti-causal state information, both the length $n \in \mathbb{N}$ of the transmission block and the length $k \in \mathbb{N}$ of the additional sequence of states can be made arbitrarily large.

The anti-causal secrecy capacity C_{ac} (resp. non-causal secrecy capacity C_{nc}) characterizes the maximal amount of information that can be transmitted reliably using anti-causal codes (resp. non-causal codes) while keeping it secret from the eavesdropper.

III. Discrete Channel

In this section, we present two interesting results regarding the wiretap channel with state information available at both encoder and decoder. Theorem 8 presents the lower bound on the secrecy capacity C_{nc} for the non-causal case stated by Chen and Vinck in [5]. The first contribution, stated in Theorem 11, is the characterization of the secrecy capacity C_{ac} for the case of anti-causal state information.

Theorem 8 (Chen and Vinck [5]) *The non-causal secrecy capacity C_{nc} of the wiretap channel with state information available at both encoder and decoder satisfies equation (13).*

$$\mathsf{C}_{\mathsf{nc}} \geq \max_{\mathcal{P}(\boldsymbol{u}, \boldsymbol{x} | \boldsymbol{s})} \Big[I(\boldsymbol{u}; \boldsymbol{y}, \boldsymbol{s}) - \max \Big[I(\boldsymbol{u}; \boldsymbol{s}), I(\boldsymbol{u}; \boldsymbol{z}) \Big] \Big]. \tag{13}$$

Remark 9 The authors of [5] investigate a slightly different model where the state information is non-causally available at the encoder but not at the decoder. However, their result can be adapted to the model depicted in Figure 1 by considering the channel output of the legitimate receiver $(\mathbf{y}, \mathbf{s})$ instead of $\mathbf{y}$. Theorem 8 is also a consequence of the result of [7] when considering equal state information at both encoder and decoder. A detailed proof of Theorem 8 can be found in [6] when considering a channel with only one legitimate decoder.

Remark 10 The achievability bound stated in [8] for the causal state information is larger than the acheivability bound stated in equation (13) of Theorem 8. However, we show in Section IV that the rate provided by Theorem 8 achieves the secrecy capacity in the Gaussian case. A converse result for the non-causal secrecy capacity is also stated in [8].

Theorem 11 characterizes the secrecy capacity of the wiretap channel with anti-causal state information depicted in Figure 2.

Theorem 11 (First Contribution) *Suppose that $H(\boldsymbol{s}) > 0$, the anti-causal secrecy capacity C_{ac} of the wiretap channel with state information available at both encoder and decoder is given by equation (14).*

$$\mathsf{C}_{ac} = \max_{\mathcal{P}(\boldsymbol{x}|\boldsymbol{s})} I(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s}). \tag{14}$$

Remark 12 The anti-causal secrecy capacity is equal to the capacity of the channel without eavesdropper. The additional sequence of states $s^k \in \mathcal{S}^k$ can be interpreted as a secret key shared by the encoder and the decoder. If the length $k \in \mathbb{N}$ is sufficiently high, the amount of randomness provided by the secret key $s^k \in \mathcal{S}^k$ is large enough so that the encoder and the decoder can secure all the transmitted information. The capacity result we derive in Theorem 11 is still valid when considering the case of infinite state information $s^\infty = (s_1, \ldots, s_n, \ldots, s_{n+k}, \ldots) \in \mathcal{S}^\infty$ because the amount of randomness is infinite.

Remark 13 Comparing the result of Theorem 11 with the outer bound provided in [8], we conclude that the anti-causal secrecy capacity is larger $\mathsf{C}_{\mathsf{ac}} \geq \mathsf{C}_{\mathsf{nc}}$ than the non-causal secrecy capacity and in some cases the inequality is strict.

Remark 14 The first difference with the problem investigated in [10] is that the rate $k \cdot H(\mathbf{s})$ of the shared key $s^k \in \mathcal{S}^k$ can be made arbitrarily large by choosing a large parameter $k \in \mathbb{N}$. Thus the encoder and the decoder are able to secure all the information that can be transmitted reliably. The second difference with the model presented in [10] is that the secret shared key does not influence the statistics of the channel.

Remark 15 If $H(\mathbf{s}) = 0$, the anti-causal secrecy capacity reduces to the one characterized in [3] for the wiretap channel without states. Hence, the anti-causal secrecy capacity is discontinuous between $H(\mathbf{s}) = 0$ and $H(\mathbf{s}) > 0$.

The proof of Theorem 11 is provided in section V. The achievability part relies on a one-time pad argument where the sequence of future states $s^k \in \mathcal{S}^k$ is treated as a secret key shared by both encoder and decoder. The converse follows by considering the state information $\mathbf{s}$ as a part of the channel output.

IV. Gaussian Channel

In this section we present the second contribution of our work which is based on Theorem 8. Theorem 16 states that the presence of the eavesdropper does not affect the secrecy capacity of the Gaussian wiretap channel with non-causal state information. Even if the eavesdropper has a better observation than the legitimate receiver, the encoder and the decoder are able to secure all the information that can be transmitted reliably.

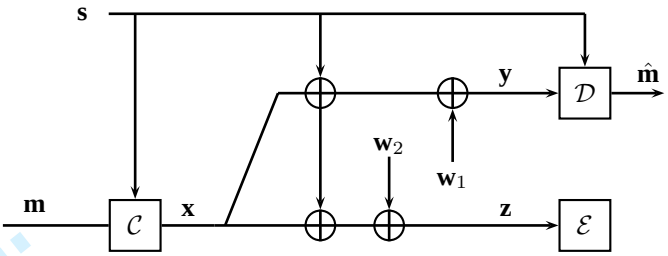

Fig. 3. Gaussian wiretap channel with state information $\mathbf{s}$ available non-causally at both encoder and decoder but not at the eavesdropper. The random variables of the noise $\mathbf{w}_1$ and $\mathbf{w}_2$ and the state $\mathbf{s}$ are Gaussian random variable with zero mean and variance N_1, N_2 and Q.

Let us consider the Gaussian wiretap channel with additive state and noise depicted in Figure 3. The signals received by the legitimate receiver and by the eavesdropper are defined by the equations (15) and (16)

$$\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{w}_1, \tag{15}$$
$$\mathbf{z} = \mathbf{x} + \mathbf{s} + \mathbf{w}_2, \tag{16}$$

where the random variables $\mathbf{w}_1$, $\mathbf{w}_1$ and $\mathbf{s}$ are Gaussian with mean 0 and variance equal to N_1, N_2 and Q.

$$\mathbf{w}_1 \sim \mathcal{N}(0, N_1), \tag{17}$$
$$\mathbf{w}_2 \sim \mathcal{N}(0, N_2), \tag{18}$$
$$\mathbf{s} \sim \mathcal{N}(0, Q). \tag{19}$$

The channel input $\mathbf{x}$ must satisfy the power constraint (20).

$$\mathbb{E}\left[\mathbf{x}^2\right] \leq P. \tag{20}$$

Based on the result stated in Theorem 8 we characterize the secrecy capacity of the Gaussian wiretap channel with non-causal state information available at both encoder and decoder.

Theorem 16 (Second Contribution) *Suppose that $H(\boldsymbol{s}) > 0$, the non-causal secrecy capacity C_{nc} of the Gaussian wiretap*

channel with state information available at both encoder and decoder is given by equation (21).

$$C_{nc} \quad = \quad \max_{\mathcal{P}(\boldsymbol{x}|\boldsymbol{s})} I(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s}) = \frac{1}{2}\log\left(1+\frac{P}{N_1}\right). \qquad (21)$$

Remark 17 The anti-causal secrecy capacity and the non-causal secrecy capacity are equal for the Gaussian channel. Even if the variance $Q > 0$ of the random variable $\mathbf{s}$ is small, the parameter $\alpha \in \mathbb{R}$ can be chosen appropriately such that the term $I(\mathbf{u};\mathbf{y},\mathbf{s}) - I(\mathbf{u};\mathbf{z})$ in equation (13) of Theorem 8 is arbitrarily large and thus C_{nc} satisfies equation (21).

The proof of the Theorem 16 is presented in Section IV-B and consists in evaluating the equation (13) of Theorem 8 with the auxiliary random variable

$$\mathbf{u} = \mathbf{x} + \alpha \cdot \mathbf{s} \quad \sim \quad \mathcal{N}(0, P + \alpha^2 Q), \qquad (22)$$

and $\alpha \in]-\infty, +\infty[$. The knowledge of the sequence $s^n \in \mathcal{S}^n$ of state information allows the encoder to transmit the maximal rate of information to the legitimate decoder while keeping it secret from the decoder.

A. Numerical result

We provide a numerical illustration of the result stated in Theorem 16. Figure 4 represents four information rates in terms of parameter $\alpha \in]-15, 15[$. We can see that the rate $I(\mathbf{u};\mathbf{y},\mathbf{s}) - I(\mathbf{u};\mathbf{s})$ provided by Gel'fand Pinsker's coding [4] is constant whereas the rate $I(\mathbf{u};\mathbf{y},\mathbf{s}) - I(\mathbf{u};\mathbf{z})$ provided by Wyner's coding [2] corresponds to the opposite curve of the one described by Costa in [11]. The parameters α_1 and α_2 given by equations (34), (35) and correspond to the values for which the mutual information $I(\mathbf{u};\mathbf{s})$ and $I(\mathbf{u};\mathbf{z})$ are equal. For the whole range of parameters $\alpha \in]-\infty, \alpha_1] \cup [\alpha_2, +\infty[$, the secure rate provided by Theorem 8 is equal to the channel capacity $C_{nc} = \max_{\mathcal{P}(\mathbf{x}|\mathbf{s})} I(\mathbf{x};\mathbf{y}|\mathbf{s})$.

B. Proof of Theorem 16

In order to prove Theorem 16, we evaluate the achievable rate stated in Theorem 8 for the Gaussian case with the random variable $\mathbf{u}$ defined by equation (22). This leads to the following

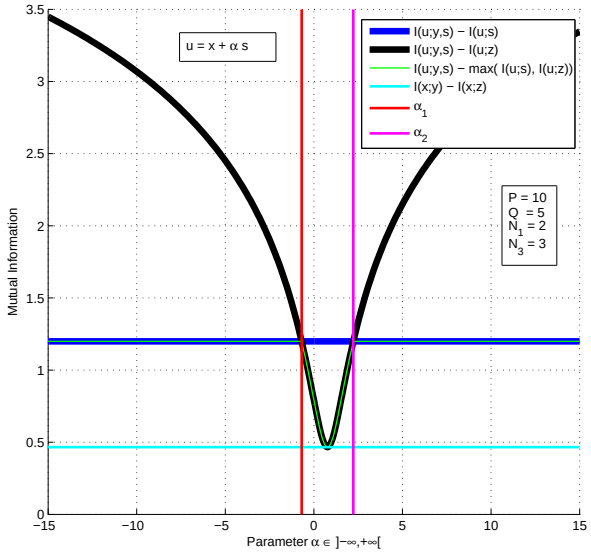

Fig. 4. The achievable rate characterized by Chen and Vinck [5] equals the capacity $C_{nc} = \max_{\mathcal{P}(\mathbf{x}|\mathbf{s})} I(\mathbf{x};\mathbf{y}|\mathbf{s})$ of the channel without eavesdropper for all $\alpha \in]-\infty, \alpha_1] \cup [\alpha_2, +\infty[$.

equations.

$$C_{nc} \geq \max_{\mathcal{P}(\mathbf{u},\mathbf{x}|\mathbf{s})} \left[I(\mathbf{u};\mathbf{y},\mathbf{s}) - \max\left(I(\mathbf{u};\mathbf{s}), I(\mathbf{u};\mathbf{z}) \right) \right] \qquad (23)$$

$$= \max_{\mathcal{P}(\mathbf{u},\mathbf{x}|\mathbf{s})} \left[I(\mathbf{u};\mathbf{y}|\mathbf{s}) - \max\left(0, I(\mathbf{u};\mathbf{z}) - I(\mathbf{u};\mathbf{s}) \right) \right] \qquad (24)$$

$$= \max_{\mathcal{P}(\mathbf{u},\mathbf{x}|\mathbf{s})} \Big[h(\mathbf{y},\mathbf{s}) - h(\mathbf{y},\mathbf{u},\mathbf{s}) - h(\mathbf{s}) + h(\mathbf{u},\mathbf{s}) \qquad (25)$$

$$- \max\left(0, h(\mathbf{z}) - h(\mathbf{u},\mathbf{z}) - h(\mathbf{s}) + h(\mathbf{u},\mathbf{s}) \right) \Big] \qquad (26)$$

$$= \frac{1}{2} \max_{\alpha \in \mathbb{R}} \left[\log\left(\frac{(P+N_1)QPQ}{PQN_1Q} \right) - \max\left(0, \log\left(\frac{(P+Q+N_2)PQ}{(PQ(1-\alpha)^2 + N_2(P+\alpha^2 Q))Q} \right) \right) \right] \qquad (27)$$

$$= \frac{1}{2}\log\left(1+\frac{P}{N_1}\right) - \frac{1}{2} \max_{\alpha \in \mathbb{R}} \left[0, \log\left(\frac{(P+Q+N_2)P}{PQ(1-\alpha)^2 + N_2(P+\alpha^2 Q)} \right) \right]. \qquad (28)$$

By hypothesis, we have that $H(\mathbf{s}) > 0$ which is equivalent to the fact that the variance $Q > 0$. The sign of the term

$I(\mathbf{u};\mathbf{z}) - I(\mathbf{u};\mathbf{s})$ is characterized by the following equations.

$$\begin{aligned} & 0 \geq I(\mathbf{u};\mathbf{z}) - I(\mathbf{u};\mathbf{s}) && (29)\\ \iff \quad & 0 \geq \log\left(\frac{(P+Q+N_2)P}{PQ(1-\alpha)^2 + N_2(P+\alpha^2 Q)}\right) && (30)\\ \iff \quad & 1 \geq \frac{(P+Q+N_2)P}{PQ(1-\alpha)^2 + N_2(P+\alpha^2 Q)} && (31)\\ \iff \quad & 0 \leq \alpha^2 Q(P+N_2) - 2\alpha PQ - P^2 && (32)\\ \iff \quad & \alpha \in]-\infty, \alpha_1] \cup [\alpha_2, +\infty[. && (33)\end{aligned}$$

where α_1 and α_2 are defined by equations (34) and (35).

$$\alpha_1 = \frac{P(Q - \sqrt{Q(P+Q+N_2)})}{Q(P+N_2)}, \quad (34)$$

$$\alpha_2 = \frac{P(Q + \sqrt{Q(P+Q+N_2)})}{Q(P+N_2)}. \quad (35)$$

Any choice of the parameter α such that $\alpha \in]-\infty, \alpha_1] \cup [\alpha_2, +\infty[$ makes the term $I(\mathbf{u};\mathbf{z}) - I(\mathbf{u};\mathbf{s})$ non-positive and so the rate given by equation (21) is equal to the channel capacity $\mathsf{C}_{\mathsf{nc}} = \max_{\mathcal{P}(\mathbf{x}|\mathbf{s})} I(\mathbf{x};\mathbf{y}|\mathbf{s})$. This conclude the proof of Theorem 16.

V. Proof of Theorem 11

A. Achievability

For each $\varepsilon > 0$, the notion of random code (see Definition 5) allow us to prove the existence of parameters $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $M \in \mathbb{N}$ and a code $c \in \mathcal{AC}(n,k,M)$ which satisfy the following conditions:

$$\frac{\log_2 M}{n} = \mathsf{C}_{\mathsf{ac}} - 5\varepsilon, \qquad \mathcal{P}_e(c) \leq 7\varepsilon, \qquad \mathcal{L}_e(c) \leq 7\varepsilon. \quad (36)$$

Coding parameters. Let $\mathcal{P}^\star(\mathbf{x}|\mathbf{s}) : \mathcal{S} \longrightarrow \Delta(\mathcal{X})$ the transition probability which attains the maximum in equation (14). Fix $\varepsilon > 0$ and choose two auxiliary parameters $\tilde{\varepsilon} > 0$ and $\gamma \in \mathbb{Q}$ positive, which satisfy the following equation:

$$\left| I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 6\varepsilon - \gamma H(\mathbf{s}) \right| < \varepsilon - \gamma\tilde{\varepsilon}. \quad (37)$$

Remark 18 The hypothesis $H(\mathbf{s}) > 0$ insures the existence of the parameters $\gamma \in \mathbb{Q}$ and $\tilde{\varepsilon} > 0$.

Parameters $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $M \in \mathbb{N}$ are defined in the following manner.

- $n \in \mathbb{N}$ is larger than $n_1 \in \mathbb{N}$ et $n_2 \in \mathbb{N}$ defined by the equations (43) and (44) in terms of the parameter $\varepsilon > 0$.
- $k \in \mathbb{N}$ is larger than $k_1 \in \mathbb{N}$, $k_2 \in \mathbb{N}$ and $k_3 \in \mathbb{N}$ defined by the equations (43), (67) and (90) in term of the parameter $\tilde{\varepsilon} > 0$.
- $n \in \mathbb{N}$ and $k \in \mathbb{N}$ satisfy the equation $\gamma = \frac{k}{n}$.
- $M \in \mathbb{N}$ is equal to $2^{n(I(\mathbf{x};\mathbf{y}|\mathbf{s})-5\varepsilon)}$ which is supposed to be an integer.

Remark 19 Substituting $\gamma = \frac{k}{n}$, the equation (37) is equivalent to the pair of equations (38) and (39).

$$n\Big(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 7\varepsilon\Big) < k\Big(H(\mathbf{s}) - \tilde{\varepsilon}\Big), \quad (38)$$

$$k\Big(H(\mathbf{s}) + \tilde{\varepsilon}\Big) < n\Big(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon\Big). \quad (39)$$

Random coding scheme. We define the random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ using a random codebook, an encoding and a decoding function.

- *Random codebook* $(\mathbf{x}^n(m))_{m\in\mathcal{M}}$: The sequence of states $s^{n+k} \in \mathcal{S}^{n+k}$ is divided into two parts $s^n \in \mathcal{S}^n$ and $s^k \in \mathcal{S}^k$. For each sequence of states $s^n \in \mathcal{S}^n$, we generate a random codebook consisting of $M = 2^{n(I(\mathbf{x};\mathbf{y}|\mathbf{s})-5\varepsilon)}$ sequences $\mathbf{x}^n(m) \in \mathcal{X}^n$ indexed by $m \in \mathcal{M}$ and drawn according to the product probability distribution $\mathcal{P}^\star(\mathbf{x}^n|\mathbf{s}^n) = \prod_{i=1}^n \mathcal{P}^\star(\mathbf{x}_i|\mathbf{s}_i)$ set forth above. Then, each typical sequence of states $s^k \in A_{\tilde{\varepsilon}}^{\star k}(S)$ for the parameter $\tilde{\varepsilon} > 0$, is indexed by $i \in \{1,\dots,I\} = \mathcal{I}$ with $I = |A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S})|$. For each codebook realization, we define the encoding and decoding functions as follows.

- *Encoder* observes the realizations of the message $m \in \mathcal{M}$ and the sequence of states $s^{n+k} \in \mathcal{S}^{n+k}$. It generates the auxiliary message $m' = m \oplus i \in \mathcal{M}$ by performing the addition modulo-M between index $i \in \mathcal{I}$ of the sequence of states $s^k \in \mathcal{S}^k$ and index $m \in \mathcal{M}$ of the message. The encoder sends through the channel, the inputs sequence $x^n(m') \in \mathcal{X}^n$ that corresponds to the message $m' \in \mathcal{M}$ and the sequence $s^n \in \mathcal{S}^n$.

- *Decoder* observes the realizations of the sequence of outputs $y^n \in \mathcal{Y}^n$ and of the sequence of states $s^{n+k} \in \mathcal{S}^{n+k}$. Decoder finds the auxiliary message $\hat{m}' \in \mathcal{M}$ for which the sequence of inputs $x^n(\hat{m}') \in A_\varepsilon^{\star n}(s^n, y^n)$ is jointly typical for the parameter $\varepsilon > 0$ with the sequences of state and output (s^n, y^n). The decoder obtains the original message $\hat{m} \in \mathcal{M}$ by performing the substraction $\hat{m} = \hat{m}' \ominus i = \hat{m}' \oplus (M - i)$ modulo-M between the index of the auxiliary message $\hat{m}' \in \mathcal{M}$ and the index $i \in \mathcal{I}$ of the sequence of states $s^k(i) \in \mathcal{S}^k$.

- *An error* is declared if $s^k \notin A_{\tilde{\varepsilon}}^{\star n}(\mathcal{S})$ or if $(x^n(m'), s^n, y^n) \notin A_\varepsilon^{\star n}(\mathcal{X}\times\mathcal{S}\times\mathcal{Y})$.

Remark 20 The sequence of states $s^k \in \mathcal{S}^k$ is considered, here above, as a secret key shared by the encoder and the decoder. Equations (38) and (39) guarantee that the sets $\mathcal{I}$ and $\mathcal{M}$ have almost the same cardinality. Hence the sequence of states $s^k \in \mathcal{S}^k$ induces a random index $\mathbf{i}$ distributed almost uniformly over $\mathcal{M}$. As a consequence, the transmitted message $\mathbf{m}' = \mathbf{m} \oplus \mathbf{i}$ is almost statistically independent of the original message $\mathbf{m}$.

Expected error probability. *First,* we provide an upper bound over the expected error probability for the random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ knowing that the realization of the message

is $m = 1$.

$$\begin{aligned} & \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = 1)\Big] \\ = & \sum_{c \in \mathcal{AC}(n,k,M)} r(\mathbf{c} = c)\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = 1, \mathbf{c} = c). \end{aligned} \quad (40)$$

Let us denote by $m' = 1 \oplus i$ and define the following error events:

- $E_1 = \Big\{ s^n \notin A_\varepsilon^{\star n}(\mathcal{S}) \Big\}$ the sequence of states $s^n \in \mathcal{S}^n$ is not ε-typical.
- $E_2 = \Big\{ s^k \notin A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S}) \Big\}$ the sequence of states $s^k \in \mathcal{S}^k$ is not $\tilde{\varepsilon}$-typical.
- $E_3 = \Big\{ x^n(m') \notin A_\varepsilon^{\star n}(\mathcal{X}|s^n) \Big\}$ the sequence of input symbols $x^n(m') \in \mathcal{X}^n$ is not jointly ε-typical with the sequence of states $s^n \in \mathcal{S}^n$.
- $E_4 = \Big\{ y^n \notin A_\varepsilon^{\star n}(\mathcal{Y}|x^n(m'), s^n) \Big\}$ the sequence of output symbols $y^n \in \mathcal{Y}^n$ is not jointly ε-typical with the sequences of input symbols and states $(x^n(m'), s^n) \in \mathcal{X}^n \times \mathcal{S}^n$.
- $E_5 = \Big\{ \exists \tilde{m} \neq m' \in \mathcal{M},\ x^n(\tilde{m}) \in A_\varepsilon^{\star n}(\mathcal{X}|y^n, s^n) \Big\}$ there exists a sequence of symbols $x^n(\tilde{m}) \in \mathcal{X}^n$ corresponding to a message $\tilde{m} \in \mathcal{M}$ which is different from the message sent $m' \in \mathcal{M}$ and the sequence $x^n(\tilde{m}) \in \mathcal{X}^n$ is jointly ε-typical with the sequences of output symbols and states $(y^n, s^n) \in \mathcal{Y}^n \times \mathcal{S}^n$.

Let us denote by $\mathbf{E}_i$ the random variable of the event E_i with $i \in \{1, 2, 3, 4, 5\}$. The expected error probability is upper bounded using Boole's inequality:

$$\begin{aligned} & \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = 1)\Big] \\ \leq & \ \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_1 \cup \mathbf{E}_2 \cup \mathbf{E}_3 \cup \mathbf{E}_4 \cup \mathbf{E}_5)\Big] \quad (41) \\ \leq & \ \mathcal{P}(\mathbf{E}_1) + \mathcal{P}(\mathbf{E}_2) + \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_3)\Big] \\ + & \ \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_4)\Big] + \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_5)\Big]. \quad (42) \end{aligned}$$

Random events $\mathbf{E}_1$ and $\mathbf{E}_2$ are independents of the realization of the random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$. The properties of ε-typical sequences (see [12], property (33) page 26 and the "Conditional Typicality Lemma" page 27) imply that there exists $n_1 \in \mathbb{N}$ and $k_1 \in \mathbb{N}$ such that for all $n \geq n_1$ and $k \geq k_1$ we have:

$$\max\Big(\mathcal{P}(\mathbf{E}_1), \mathcal{P}(\mathbf{E}_2), \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_3)\Big], \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_4)\Big]\Big) \leq \varepsilon. \quad (43)$$

The random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ allows to bound the expected error probability $\mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_5)\Big]$. From the "jointly typical Lemma" (see [12], page 29), there exists a $n_2 \in \mathbb{N}$ such that for all $n \geq n_2$, we have the following property:

$$\frac{\log_2 M}{n} = I(\mathbf{x}; \mathbf{y}|\mathbf{s}) - 5\varepsilon \Longrightarrow \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_5)\Big] \leq \varepsilon. \quad (44)$$

The expected error probability of the random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ is upper bounded by:

$$\mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = 1)\Big] \quad \leq \quad 5\varepsilon. \quad (45)$$

Second, we provide an upper bound over the expected error probability. Since the codebook is drawn randomly, the expected error probability does not depend on the realized message $m \in \mathcal{M}$. We obtain the following equation:

$$\begin{aligned} & \sum_{c \in \mathcal{AC}(n,k,M)} r(\mathbf{c} = c)\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = 1, \mathbf{c} = c) \\ = & \sum_{c \in \mathcal{AC}(n,k,M)} r(\mathbf{c} = c)\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = m, \mathbf{c} = c), \quad \forall m \in \mathcal{M}, \end{aligned} \quad (46)$$

$$\Longleftrightarrow \quad \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = 1)\Big] = \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = m)\Big], \quad \forall m \in \mathcal{M}. \quad (47)$$

Consequently, the expected error probability can be written:

$$\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c})\Big] \quad (48)$$

$$= \sum_{\substack{c \in \mathcal{AC}(n,k,M),\\ m \in \mathcal{M}}} r(\mathbf{c} = c)\mathcal{P}(\mathbf{m} = m)\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = m, \mathbf{c} = c) \quad (49)$$

$$= \sum_{m \in \mathcal{M}} \mathcal{P}(\mathbf{m} = m)\mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = m)\Big] \quad (50)$$

$$= \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = 1)\Big] \sum_{m \in \mathcal{M}} \mathcal{P}(\mathbf{m} = m) \quad (51)$$

$$= \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m} = 1)\Big] \quad (52)$$

$$\leq 5\varepsilon. \quad (53)$$

We prove the expected error probability of random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ is lower than 5ε.

$$\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c})\Big] \quad \leq 5\varepsilon. \quad (54)$$

Information rate. For all codes $c \in \mathcal{AC}(n,k,M)$ belonging to the support of random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$, the information rate is given by the equation (55).

$$\frac{\log M}{n} \quad = \quad \mathsf{C}_{\mathsf{ac}} - 5\varepsilon. \quad (55)$$

Expected information leakage rate. Equations (38), (39) and Lemma 21 allow us to obtain a bound on the expected rate of

equivocation $\mathbb{E}_c\Big[\mathcal{L}_e(\mathbf{c})\Big] \leq 2\varepsilon$.

$$\begin{aligned}
\mathbb{E}_c\Big[I(\mathbf{m};\mathbf{z}^n|\ \mathbf{c}\ )\Big] &\leq \mathbb{E}_c\Big[I(\mathbf{m};\mathbf{m}'|\ \mathbf{c}\ )\Big] && (56)\\
&= I(\mathbf{m};\mathbf{m}') && (57)\\
&= H(\mathbf{m}') - H(\mathbf{m}'|\mathbf{m}) && (58)\\
&\leq \log M - H(\mathbf{m}\oplus\mathbf{i}|\mathbf{m}) && (59)\\
&= n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - H(\mathbf{i}|\mathbf{m})\\
&- H(\mathbf{m}\oplus\mathbf{i}|\mathbf{i},\mathbf{m}) + H(\mathbf{i}|\mathbf{m},\mathbf{m}\oplus\mathbf{i}) && (60)\\
&= n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - H(\mathbf{i}|\mathbf{m})\\
&+ H(\mathbf{i}|\mathbf{m},\mathbf{m}\oplus\mathbf{i}) && (61)\\
&= n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - H(\mathbf{i}|\mathbf{m}) && (62)\\
&= n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - H(\mathbf{i}) && (63)\\
&\leq n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - k(H(\mathbf{s}) - \tilde{\varepsilon}) && (64)\\
&< n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon)\\
&- n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 7\varepsilon) && (65)\\
&\leq n2\varepsilon. && (66)
\end{aligned}$$

• Inequality (56) follows from the fact that for each code $c \in \mathcal{AC}(n,k,M)$, the Markov chain $\mathbf{m} - \mathbf{m}' - \mathbf{z}^n$ is satisfied. Indeed, the sequence of channel outputs $\mathbf{z}^n$ depends on $\mathbf{m}$ only through $\mathbf{m}'$. The conditional probability can be written $\mathcal{P}(z^n|m,m') = \mathcal{P}(z^n|m')$ for all $m \in \mathcal{M}$, $m' \in \mathcal{M}$, $z^n \in \mathcal{Z}^n$ and this for any code $c \in \mathcal{AC}(n,k,M)$. From the data processing inequality (see [12], page 24), the inequality $I(\mathbf{m};\mathbf{z}^n|\ \mathbf{c} = c\ ) \leq I(\mathbf{m};\mathbf{m}'|\ \mathbf{c} = c\ )$ is valid for all codes $c \in \mathcal{AC}(n,k,M)$ and this proves the inequality (56).
• Inequality (57) is due to the fact that the probability distribution $\mathcal{P}(\mathbf{m},\mathbf{m}') \in \mathcal{M}\times\mathcal{M}$ of the messages $\mathbf{m}$ et $\mathbf{m}'$ is independent of the code $\mathbf{c} \in \mathcal{AC}(n,k,M)$ choused at random.
• Inequality (59) is due to the fact that the addition $\oplus$ is performed modulo M and then the message $\mathbf{m}' = \mathbf{m}\oplus\mathbf{i}$ belongs to the set $\mathcal{M}$ of cardinality M.
• Equality (61) is due to the fact that $\mathbf{m}\oplus\mathbf{i}$ is a deterministic function of $\mathbf{i}$ and $\mathbf{m}$, then $H(\mathbf{m}\oplus\mathbf{i}|\mathbf{i},\mathbf{m}) = 0$.
• Equality (62) is due to the condition (39) and to the properties of the set of $\tilde{\varepsilon}$-typical sequences presented in [12] by the property 2 page 26. Indeed, there exists $k_2 \in \mathbb{N}$ such that for all $k \geq k_2$:

$$|A^{\star k}_{\tilde{\varepsilon}}(\mathcal{S})| \leq 2^{k(H(\mathbf{s})+\tilde{\varepsilon})}. \quad (67)$$

Equations (39) and (67) allow us to obtain the following equation:

$$I = |A^{\star k}_{\tilde{\varepsilon}}(\mathcal{S})| \leq 2^{k(H(\mathbf{s})+\tilde{\varepsilon})} < 2^{n(I(\mathbf{x};\mathbf{y}|\mathbf{s})-5\varepsilon)} = M. \quad (68)$$

Because $I < M$, the realizations $m \in \mathcal{M}$ and $m\oplus i \in \mathcal{M}$ allow us to characterize a unique $i \in \mathcal{I} \subset \mathcal{M}$. As a consequence $H(\mathbf{i}|\mathbf{m},\mathbf{m}\oplus\mathbf{i}) = 0$.
• Equality (63) is due to the fact that the index $\mathbf{i}$ and the message $\mathbf{m}$ are drawn independently.
• Inequality (64) is due to the Lemma 21.
• Inequality (65) is due to the condition (38).

Lemme 21 *For all $\tilde{\varepsilon}$, there exists $k_3 \in \mathbb{N}$ such that for all $k \geq k_3$:*

$$H(\boldsymbol{i}) \geq k(H(\boldsymbol{s}) - \tilde{\varepsilon}). \quad (69)$$

Remark 22 The Lemma 21 guarantee that the secret key $\mathbf{i} \in \mathcal{I}$ shared by the encoder and the decoder has a rate closed to $k\cdot H(\mathbf{s})$ and hence closed to the anti-causal secrecy capacity C_{ac}.

The proof of Theorem 11 is outlined in Section V-C.

Existence of a code. The random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ we defined here above, satisfies the following three conditions:

1) Every code $c \in \mathcal{AC}(n,k,M)$ belonging to the support of $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ have a rate equal to

$$\frac{\log M}{n} = \mathsf{C}_{\mathsf{ac}} - 5\varepsilon. \quad (70)$$

2) The expected error probability is bounded by

$$\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c})\Big] \leq 5\varepsilon. \quad (71)$$

2) The expected information leakage rate is bounded by

$$\mathbb{E}_c\Big[\mathcal{L}_e(\mathbf{c})\Big] \leq 2\varepsilon. \quad (72)$$

The bounds on the expected error probability and the expected information leakage rate imply:

$$\begin{aligned}
&\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c})\Big] + \mathbb{E}_c\Big[\mathcal{L}_e(\mathbf{c})\Big] && \leq 7\varepsilon && (73)\\
\iff &\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c}) + \mathcal{L}_e(\mathbf{c})\Big] && \leq 7\varepsilon && (74)\\
\iff &\textstyle\sum_{c\in\mathcal{AC}(n,k,M)} r(\mathbf{c}=c)\Big[\mathcal{P}_e(c) + \mathcal{L}_e(c)\Big] && \leq 7\varepsilon && (75)\\
\implies &\min_{c\in\mathcal{AC}(n,k,M)}\Big[\mathcal{P}_e(c) + \mathcal{L}_e(c)\Big] && \leq 7\varepsilon && (76)\\
\iff &\exists c^\star \in \mathcal{AC}(n,k,M), \quad \mathcal{P}_e(c^\star) + \mathcal{L}_e(c^\star) && \leq 7\varepsilon. && (77)
\end{aligned}$$

This demonstrates that there exists a code $c^\star \in \mathcal{AC}(n,k,M)$ in the support of $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ such that the error probability and the information leakage rate are bounded below 7ε.

$$\begin{aligned}
\mathcal{P}_e(c^\star) &\leq 7\varepsilon, && (78)\\
\mathcal{L}_e(c^\star) &\leq 7\varepsilon. && (79)
\end{aligned}$$

To conclude, we showed the existence of a code $c^\star \in \mathcal{AC}(n,k,M)$ whose rate is equal to $\frac{\log M}{n} = \mathsf{C}_{\mathsf{ac}} - 5\varepsilon$, the probability of error is bounded by $\mathcal{P}_e(c^\star) \leq 7\varepsilon$ and the information leakage rate is bounded by $\mathcal{L}_e(c^\star) \leq 7\varepsilon$.

B. Converse

The converse of Theorem 11 is obtained from the converse result of the point to point channel coding result [13], [12] considering the pair $(\mathbf{y}^n, \mathbf{s}^n)$ as the channel output instead of $\mathbf{y}^n$.

C. Proof of Lemma 21

The random variable $\mathbf{E}$ is defined by equation (80).

$$\mathbf{E} = \begin{cases} 0 \text{ si } s^k \in A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S}) \\ 1 \text{ si } s^k \notin A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S}). \end{cases} \tag{80}$$

The following equalities are due to the definition of entropy:

$$\begin{aligned} & H(\mathbf{s}^k, \mathbf{E}) = H(\mathbf{E}) + H(\mathbf{s}^k|\mathbf{E}) \\ & = H(\mathbf{s}^k) + H(\mathbf{E}|\mathbf{s}^k) = H(\mathbf{s}^k) && (81) \\ \Longrightarrow \quad & H(\mathbf{s}^k|\mathbf{E}=0)\mathcal{P}(\mathbf{E}=0) \\ & = H(\mathbf{s}^k) - H(\mathbf{E}) - H(\mathbf{s}^k|\mathbf{E}=1)\mathcal{P}(\mathbf{E}=1). && (82) \end{aligned}$$

Let us denote by $H_b(\delta)$ the entropy of the binary random variable $\{0,1\}$ drawn according to the probabilities $(\delta, 1-\delta)$ with parameter $\delta \in [0,1]$. The random variable $\mathbf{i} \in \mathcal{I}$ is defined as the index of the sequence of states $\mathbf{s}^k(\mathbf{i}) \in A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S})$ that belong to the set of $\tilde{\varepsilon}$-typical sequences Its probability distribution is given by the following equation:

$$\begin{aligned} \mathcal{P}(\mathbf{i}=i) &= \mathcal{P}(\mathbf{s}^k = s^k(i)|\mathbf{E}=0) && (83) \\ &= \frac{\mathcal{P}(\mathbf{s}^k = s^k(i), \mathbf{E}=0)}{\mathcal{P}(\mathbf{E}=0)}. && (84) \end{aligned}$$

The entropy of this random variable satisfies the following equations:

$$\begin{aligned} H(\mathbf{i}) &= H(\mathbf{s}^k|\mathbf{E}=0) && (85) \\ &\geq H(\mathbf{s}^k|\mathbf{E}=0)\mathcal{P}(\mathbf{E}=0) && (86) \\ &= H(\mathbf{s}^k) - H(\mathbf{E}) - H(\mathbf{s}^k|\mathbf{E}=1)\mathcal{P}(\mathbf{E}=1) && (87) \\ &\geq kH(\mathbf{s}) - H_b(\mathcal{P}(\mathbf{E}=1)) - \mathcal{P}(\mathbf{E}=1)k\log|\mathcal{S}| && (88) \\ &\geq k\left(H(\mathbf{s}) - \left(\frac{H_b(\mathcal{P}(\mathbf{E}=1))}{k} + \mathcal{P}(\mathbf{E}=1)\log|\mathcal{S}|\right)\right). && (89) \end{aligned}$$

From the properties of ε-typical sequences (see [12], property 3 page 26), for all $\tilde{\varepsilon}$, there exists a $k_3 \in \mathbb{N}$ such that for all $k \geq k_3$ we have:

$$\frac{H_b(\mathcal{P}(\mathbf{E}=1))}{k} + \mathcal{P}(\mathbf{E}=1)\log|\mathcal{S}| \quad \leq \quad \tilde{\varepsilon}. \tag{90}$$

We showed that for all $\tilde{\varepsilon}$, there exists a $k_3 \in \mathbb{N}$ such that for all $k \geq k_3$ we have:

$$H(\mathbf{i}) \geq k(H(\mathbf{s}) - \tilde{\varepsilon}). \tag{91}$$

This concludes the proof of Lemma 21.

VI. Conclusion

This article is devoted to the problem of secure communication over a wiretap channel with state information available at both encoder and decoder but not at the eavesdropper. The secrecy capacity for such a channel is not available yet in the literature for neither the causal, nor the non-causal case. We introduce the concept of anti-causal state information i.e. the length of the sequence of states available at both encoder and decoder is arbitrarily larger than the length of the transmission block. We characterize the secrecy capacity for the discrete channel with anti-causal state information and for the Gaussian channel with non-causal state information. These two cases are of particular interest because we show that the encoder and the decoder can use the state information in order to secure all the information that can be transmitted reliably.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[2] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

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

[5] Y. Chen and H. Vinck, "Wiretap channel with side information," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 395–402, 2008.

[6] M. LeTreust, A. Zaidi, and S. Lasaulce, "An achievable rate region for the broadcast wiretap channel with asymmetric side information," *IEEE Proc. of the 49th Allerton conference, Monticello, Illinois*, 2011.

[7] W. Liu and B. Chen, "Wiretap channel with two-sided channel state information," *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, pp. 893–897, 2007.

[8] Y. K. Chia and A. E. Gamal, "Wiretap channel with causal state information," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2838–2849, 2012.

[9] H. Yamamoto, "Rate-distortion theory for the shannon cipher system," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 827–835, 1997.

[10] W. Kang and N. Liu, "Wiretap channel with shared key," in *Proc. IEEE Int. Symp. Information Theory (ISIT'10)*, Sep. 2010.

[11] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439–441, 1983.

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

[13] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

Full Secrecy for the Wiretap Channel with State Information Available at Both Encoder and Decoder

Maël Le Treust and Abdellatif Zaidi
Laboratoire d'Informatique Gaspard Monge
Université Paris-Est Marne La Vallée, 77454, Marne La Vallée Cedex 2, France
Email: {mael.letreust,abdellatif.zaidi}@univ-mlv.fr

Abstract

We investigate the problem of secure communication over a wiretap channel with state information available at both encoder and decoder. In this framework, the secrecy capacity is strongly related to the knowledge, by both encoder and decoder, of the past, the current and the future channel states. The secrecy capacity has not been characterized yet, for neither causal, nor non-causal state information. Our results provide a better understanding of the fundamental problems that arise in secure communication with state information and establish a connection with the wiretap channel with shared key. We introduce the concept of anti-causal state information i.e. the length of the sequence of states available at both encoder and decoder is arbitrarily larger than the length of the transmission block. The analysis reveals that the state information can be utilized as a secret key that is shared among the encoder and the decoder. We establish the secrecy capacity for two interesting cases. The first is the discrete channel with anti-causal state information and the second is the Gaussian channel with non-causal state information. For both cases, our results show that the encoder and the decoder can use the channel state information in order to secure all the information that can be transmitted reliably.

Index Terms

Physical layer security, Shannon theory, wiretap channel, state information, secrecy capacity.

I. Introduction

The problem of secure communication is modeled by the wiretap channel in which an encoder sends information to a decoder, while keeping it secret from the eavesdropper. In [1], Shannon shows that

secure communication is possible when a secret key is available at both encoder and decoder, but not at the eavesdropper. Wyner [2] investigates the problem of degraded wiretap channel and exploits the statistics of the channel in order to secure information. Csiszar and Körner [3] extended the result of Wyner [2] to the case of general wiretap channel where the legitimate decoder and the eavesdropper have also to decode a common message.

We investigate the problem of secure communication over a wiretap channel with state information available at both encoder and decoder [4] but not at the eavesdropper. The communication possibilities for a channel with states are deeply related to the knowledge of the state parameter by both encoder and decoder. This knowledge is "strictly causal" (resp. "causal") when at a given instant $t \in \{1, \ldots, n\}$ of the transmission block, the encoder only observes the past (resp. the past and the current) states of the channel. This knowledge is "non-causal" when both encoder and decoder observe the sequence of states that corresponds to the transmission block. Figure 1 represents the channel with non-causal state information. This knowledge is "anti-causal" (resp. "infinite") when both encoder and decoder observe a sequence of states that is arbitrarily larger than the transmission block (resp. that is infinite). Figure 2 represents the channel with anti-causal state information. In [5], Chen and Vinck establish a lower bound on the secrecy capacity for the wiretap channel with state information non-causally available at the encoder. This result is obtained using a combination of the Gel'fand and Pinsker coding [4] and the Wyner coding [2] and it extends easily to the broadcast wiretap channel with state information [6]. The authors of [7] strengthen the result of Chen and Vinck [5] and provide a lower bound on the secrecy capacity when distinct channel state information are non-causally available at encoder and decoder.

Subsequently, the article of Chia and El Gamal [8] introduces a promising perspective for the study of the wiretap channel with state information at both encoder and decoder. Considering the channel state information as a secret key shared by the encoder and decoder, enhances secure communication significantly. The two main results presented in [8] are bounds on the secrecy capacity. The first result is a lower bound for the case of causal state information which is strictly larger than the lower bound for the non-causal case stated in [5]. The second result is an upper bound for the case of non-causal state information. An important contribution of [8] is to connect the literature devoted to the wiretap channel with state information and the literature devoted to the wiretap channel with secret shared key. In fact, the secrecy capacity when considering a secret shared key is characterized in [9] for the less noisy wiretap channel and in [10] for the general wiretap channel.

The secrecy capacity for the general wiretap channel with state information available at both encoder and decoder is not available yet in the literature for neither the strictly causal, nor the causal, nor the non-

causal case. In this article, we introduce the concept of anti-causal state information and we characterize the secrecy capacity for two interesting cases. The first one, stated in Theorem 11, considers the discrete channel with anti-causal state information, and the second one, stated in Theorem 16, considers the Gaussian channel with non-causal state information. Remarkably, in both situations the secrecy capacity is shown to be equal to the capacity of the same channel without eavesdropper. We show that the knowledge

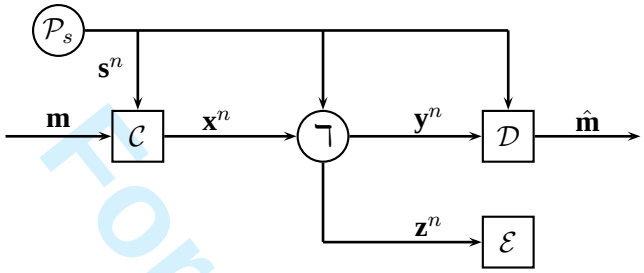

Fig. 1. Wiretap channel $\urcorner(\mathbf{y}, \mathbf{z}|\mathbf{x}, \mathbf{s})$ with **non-causal** state information $\mathbf{s}^n \in \mathcal{S}^n$ drawn according to the i.i.d. probability distribution $\mathcal{P}_s \in \Delta(\mathcal{S})$ and available at both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ but not at the eavesdropper. Encoder $\mathcal{C}$ represents the message $\mathbf{m} \in \mathcal{M}$ by a sequence of inputs $\mathbf{x}^n \in \mathcal{X}^n$. Decoder $\mathcal{D}$ observes a sequence of outputs $\mathbf{y}^n \in \mathcal{Y}^n$ and reconstructs the message $\hat{\mathbf{m}} \in \mathcal{M}$. The eavesdropper $\mathcal{E}$ observes a sequence of outputs $\mathbf{z}^n \in \mathcal{Z}^n$ and tries to decode the message $\mathbf{m} \in \mathcal{M}$ sent by encoder $\mathcal{C}$.

of the sequence of states allow the encoder and the decoder to secure all the information that can be transmitted reliably. Ultimately, we prove that the secrecy capacity with anti-causal state information is larger (and sometimes strictly larger) than the secrecy capacity with non-causal state information.

Section II presents the channel model under investigation and the concepts of strictly causal, causal, non-causal, anti-causal and infinite knowledge of the state information. The results for the discrete case are stated in Section III where the first contribution (Theorem 11) is devoted to the case of anti-causal state information. The results for the Gaussian case are stated in Section IV where the second contribution (Theorem 16) is devoted to the case of non-causal state information. Section V contains the proof of the Theorem 11 and Section VI concludes the article.

II. Preliminaries

The problems under investigation are depicted in Figures 1 and 2. We denote by $\mathbf{m}$, $\mathbf{s}^n$, $\mathbf{x}^n$, $\mathbf{y}^n$, $\mathbf{z}^n$, and $\hat{\mathbf{m}}$ the random variables of the message $m \in \mathcal{M}$, the sequences of states $s^n = (s_1, \ldots, s_n) \in \mathcal{S}^n$, of channel

inputs $x^n \in \mathcal{X}^n$, of legitimate channel outputs $y^n \in \mathcal{Y}^n$, of eavesdropper channel outputs $z^n \in \mathcal{Z}^n$ and of the message $\hat{m} \in \mathcal{M}$ obtained as output by the decoder. The notation $\Delta(\mathcal{X})$ stands for the set of the probability distributions $\mathcal{P}(\mathbf{x})$ over the set $\mathcal{X}$ and $\mathcal{S}^\infty$ stands for the set of infinite sequences $s^\infty \in \mathcal{S}^\infty$ of channel states. A wiretap channel with state is defined by a probability distribution $\mathcal{P}_s \in \Delta(\mathcal{S})$ over the channel states and a transition probability $\urcorner : \mathcal{X} \times \mathcal{S} \longrightarrow \Delta(\mathcal{Y} \times \mathcal{Z})$. The statistics of the transition probability $\urcorner$ and the probability distribution $\mathcal{P}_s$ are known by both encoder $\mathcal{C}$ and decoder $\mathcal{D}$.

The information message $\mathbf{m} \in \mathcal{M}$ is supposed to be drawn according to the uniform probability distribution over $\mathcal{M}$. Encoder $\mathcal{C}$ aims at transmitting the realization of the message $m \in \mathcal{M}$ to decoder $\mathcal{D}$ using a transmission block of $n \in \mathbb{N}$ input symbols $x^n \in \mathcal{X}^n$. The state $s \in \mathcal{S}$ of the channel is drawn according to the i.i.d. probability distribution $\mathcal{P}_s$ and it's observation by both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ should be defined carefully. Indeed, the capacity of the wiretap channel with state information is strongly related to the length of the sequence of channel states observed by the encoder $\mathcal{C}$ and the decoder $\mathcal{D}$.

- **Strictly causal case.** At instant $t \in \{1, \ldots, n\}$ of the transmission block, encoder $\mathcal{C}$ observes the sequence of past states $s^{t-1} = (s_1, \ldots, s_{t-1}) \in \mathcal{S}^{t-1}$, whereas decoder $\mathcal{D}$ observes the sequence of the states $s^n = (s_1, \ldots, s_n) \in \mathcal{S}^n$ that corresponds to the transmission block.
- **Causal case.** At instant $t \in \{1, \ldots, n\}$ of the transmission block, encoder $\mathcal{C}$ observes the sequence of past and current states $s^t = (s_1, \ldots, s_t) \in \mathcal{S}^t$, whereas decoder $\mathcal{D}$ observes the sequence of the states $s^n = (s_1, \ldots, s_n) \in \mathcal{S}^n$ that corresponds to the transmission block.
- **Non-causal case.** At the first instant of the transmission block, both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ observe the sequence of states $s^n = (s_1, \ldots, s_n) \in \mathcal{S}^n$ that corresponds to the transmission block (see Figure 1 and Definition 3).

The secrecy capacity is not available in the literature for neither the strictly causal, nor the causal, nor the non-causal case. In this article, we introduce the concept of anti-causal and infinite state information and we characterize the secrecy capacity for two interesting cases (Theorems 11 and 16). The results presented in this article provide a better understanding of the fundamental problems that arise in secure communication with channel state information.

- **Anti-causal case.** At the first instant of the transmission block, both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ observe the sequence of states $s^{n+k} = (s_1, \ldots, s_n, \ldots, s_{n+k}) \in \mathcal{S}^{n+k}$ whose length $n + k \in \mathbb{N}$ is arbitrarily larger than the length $n \in \mathbb{N}$ of the transmission block (see Figure 2 and Definition 1).
- **Infinite case.** At the first instant of the transmission block, both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ observe the infinite sequence of states $s^\infty = (s_1, \ldots, s_n, \ldots, s_{n+k}, \ldots) \in \mathcal{S}^\infty$.

Encoder $\mathcal{C}$ uses the sequence of input symbols $x^n \in \mathcal{X}^n$ and the knowledge of the appropriate sequence of states $(s^{t-1}, s^t, s^n, s^{n+k}, s^\infty)$ in order to transmit the messages $m \in \mathcal{M}$ to decoder $\mathcal{D}$. The pair of sequences of channel outputs $(y^n, z^n) \in \mathcal{Y}^n \times \mathcal{Z}^n$ is drawn according to the discrete and memoryless transition probability defined by equation (1).

$$\urcorner^{\otimes n}(y^n, z^n | x^n, s^n) = \prod_{t=1}^{n} \urcorner(y_t, z_t | x_t, s_t). \tag{1}$$

Decoder $\mathcal{D}$ observes the sequence of channel outputs $y^n \in \mathcal{Y}^n$ and reconstruct the message $\hat{m} \in \mathcal{M}$. The eavesdropper $\mathcal{E}$ observes the sequence of outputs $z^n \in \mathcal{Z}^n$ and tries to decode the message $m \in \mathcal{M}$ sent by encoder $\mathcal{C}$.

Definition 1 *An anti-causal code $c \in \mathcal{AC}(n, k, M)$ is a pair of functions (f, g) defined by equations (2) and (3).*

$$f: \ \mathcal{M} \times \mathcal{S}^{n+k} \longrightarrow \mathcal{X}^n, \tag{2}$$

$$g: \ \mathcal{Y}^n \times \mathcal{S}^{n+k} \longrightarrow \mathcal{M}. \tag{3}$$

We denote by $\mathcal{AC}(n, k, M)$ the set of anti-causal codes for which the parameters $n \in \mathbb{N}$, $k \in \mathbb{N}$ and the cardinality $M = |\mathcal{M}| \in \mathbb{N}$ are fixed.

Remark 2 Inspired from control theory, the notion of anti-causal state information corresponds to the situation described by Figure 2, where the length $n + k \in \mathbb{N}$ of the sequence of state information is arbitrarily larger than the length $n \in \mathbb{N}$ of the transmission block.

Non-causal codes are defined similarly to anti-causal code, where the set of sequences $\mathcal{S}^{n+k}$ is replaced by the set of sequences $\mathcal{S}^n$.

Definition 3 *A non-causal code $c \in \mathcal{NC}(n, M)$ is a pair of functions (f, g) defined by equations (4) and (5).*

$$f: \ \mathcal{M} \times \mathcal{S}^{n} \longrightarrow \mathcal{X}^n, \tag{4}$$

$$g: \ \mathcal{Y}^n \times \mathcal{S}^{n} \longrightarrow \mathcal{M}. \tag{5}$$

We denote by $\mathcal{NC}(n, M)$ the set of non-causal codes for which the parameter $n \in \mathbb{N}$ and the cardinality $M = |\mathcal{M}| \in \mathbb{N}$ are fixed.

Each anti-causal or non-causal code induces an error probability and a leakage rate stated formally in Definition 4. We denote by $\hat{\mathbf{m}} = g(\mathbf{y}^n, \mathbf{s}^n)$ the random variable of the message obtained as output of the decoding function.

Definition 4 *For each anti-causal code $c \in \mathcal{AC}(n,k,M)$ (resp. non-causal code $c \in \mathcal{NC}(n,M)$), the error probability $\mathcal{P}_e(c)$ and the information leakage rate $\mathcal{L}_e(c)$ are defined by equations (6) and (7).*

$$\mathcal{P}_e(c) = \mathcal{P}(\boldsymbol{m} \neq \hat{\boldsymbol{m}} |\, c\,), \tag{6}$$

$$\mathcal{L}_e(c) = \frac{I(\boldsymbol{m}; \boldsymbol{z}^n |\, c\,)}{n}. \tag{7}$$

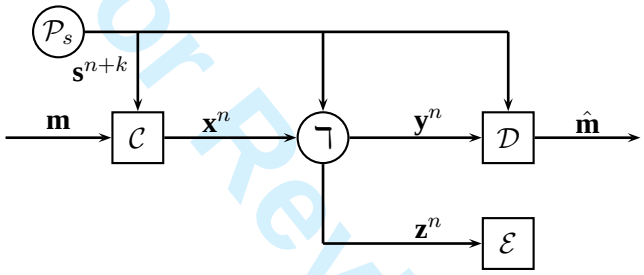

Fig. 2. Wiretap channel $\urcorner(\mathbf{y}, \mathbf{z}|\mathbf{x}, \mathbf{s})$ with **anti-causal** state information $\mathbf{s}^{n+k} \in \mathcal{S}^{n+k}$ available at both encoder $\mathcal{C}$ and decoder $\mathcal{D}$ but not at the eavesdropper. The length $n+k \in \mathbb{N}$ of the sequence of states is arbitrarily larger than the length $n \in \mathbb{N}$ of the sequence of states observed in the non-causal case.

To make the proof of Theorem 11 easier to understand, we formally define the notion of random code for the case of anti-causal state information.

Definition 5 *A random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ is a probability distribution over the set of anti-causal codes $\mathcal{AC}(n,k,M)$. The expected error probability $\mathbb{E}_c\Big[\mathcal{P}_e(\boldsymbol{c})\Big]$ and the expected information leakage rate $\mathbb{E}_c\Big[\mathcal{L}_e(\boldsymbol{c})\Big]$ are defined by equations (8) and (9) where $\boldsymbol{c} \in \mathcal{AC}(n,k,M)$ denote the random variable of the code.*

$$\mathbb{E}_c\Big[\mathcal{P}_e(\boldsymbol{c})\Big] = \sum_{c \in \mathcal{AC}(n,k,M)} r(\boldsymbol{c} = c)\mathcal{P}(\boldsymbol{m} \neq \hat{\boldsymbol{m}} | \boldsymbol{c} = c), \tag{8}$$

$$\mathbb{E}_c\Big[\mathcal{L}_e(\boldsymbol{c})\Big] = \sum_{c \in \mathcal{AC}(n,k,M)} r(\boldsymbol{c} = c)\frac{I(\boldsymbol{m}; \boldsymbol{z}^n | \boldsymbol{c} = c)}{n}. \tag{9}$$

The aim of our study is to prove the existence of an anti-causal code $c \in \mathcal{AC}(n,k,M)$ (resp. a non-causal code $c \in \mathcal{NC}(n,M)$) with the maximal information rate $\frac{\log_2 M}{n}$ under the constraint that the error probability $\mathcal{P}(\hat{\mathbf{m}} \neq \mathbf{m}|\ c\ )$ and the information leakage rate $\frac{I(\mathbf{m};\mathbf{z}^n|\ c\ )}{n}$ are arbitrarily small.

Definition 6 *A rate R is anti-causally achievable (resp. non-causally achievable), if for all $\varepsilon > 0$, there exists some parameters $n \in \mathbb{N}$, $k \in \mathbb{N}$, $M \in \mathbb{N}$ and a code $c \in \mathcal{AC}(n,k,M)$ (resp. there exists some parameters $n \in \mathbb{N}$, $M \in \mathbb{N}$ and a code $c \in \mathcal{NC}(n,M)$) such that:*

$$\frac{\log_2 M}{n} \geq R - \varepsilon, \tag{10}$$

$$\mathcal{P}_e(c) \leq \varepsilon, \tag{11}$$

$$\mathcal{L}_e(c) \leq \varepsilon. \tag{12}$$

The anti-causal secrecy capacity C_{ac} (resp. non-causal secrecy capacity C_{nc}) of the wiretap channel with state information available at both encoder and decoder is defined as the supremum of the anti-causally (resp. non-causally) achievable rates R.

Remark 7 For the case of anti-causal state information, both the length $n \in \mathbb{N}$ of the transmission block and the length $k \in \mathbb{N}$ of the additional sequence of states can be made arbitrarily large.

The anti-causal secrecy capacity C_{ac} (resp. non-causal secrecy capacity C_{nc}) characterizes the maximal amount of information that can be transmitted reliably using anti-causal codes (resp. non-causal codes) while keeping it secret from the eavesdropper.

III. Discrete Channel

In this section, we present two interesting results regarding the wiretap channel with state information available at both encoder and decoder. Theorem 8 presents the lower bound on the secrecy capacity C_{nc} for the non-causal case stated by Chen and Vinck in [5]. The first contribution, stated in Theorem 11, is the characterization of the secrecy capacity C_{ac} for the case of anti-causal state information.

Theorem 8 (Chen and Vinck [5]) *The non-causal secrecy capacity C_{nc} of the wiretap channel with state information available at both encoder and decoder satisfies equation (13).*

$$C_{nc} \geq \max_{\mathcal{P}(\boldsymbol{u},\boldsymbol{x}|\boldsymbol{s})} \Big[I(\boldsymbol{u};\boldsymbol{y},\boldsymbol{s}) - \max \Big[I(\boldsymbol{u};\boldsymbol{s}), I(\boldsymbol{u};\boldsymbol{z}) \Big] \Big]. \tag{13}$$

Remark 9 The authors of [5] investigate a slightly different model where the state information is non-causally available at the encoder but not at the decoder. However, their result can be adapted to the

model depicted in Figure 1 by considering the channel output of the legitimate receiver $(\mathbf{y}, \mathbf{s})$ instead of $\mathbf{y}$. Theorem 8 is also a consequence of the result of [7] when considering equal state information at both encoder and decoder. A detailed proof of Theorem 8 can be found in [6] when considering a channel with only one legitimate decoder.

Remark 10 The achievability bound stated in [8] for the causal state information is larger than the acheivability bound stated in equation (13) of Theorem 8. However, we show in Section IV that the rate provided by Theorem 8 achieves the secrecy capacity in the Gaussian case. A converse result for the non-causal secrecy capacity is also stated in [8].

Theorem 11 characterizes the secrecy capacity of the wiretap channel with anti-causal state information depicted in Figure 2.

Theorem 11 (First Contribution) *Suppose that $H(\boldsymbol{s}) > 0$, the anti-causal secrecy capacity C_{ac} of the wiretap channel with state information available at both encoder and decoder is given by equation (14).*

$$\mathsf{C}_{\mathsf{ac}} = \max_{\mathcal{P}(\boldsymbol{x}|\boldsymbol{s})} I(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s}). \tag{14}$$

Remark 12 The anti-causal secrecy capacity is equal to the capacity of the channel without eavesdropper. The additional sequence of states $s^k \in \mathcal{S}^k$ can be interpreted as a secret key shared by the encoder and the decoder. If the length $k \in \mathbb{N}$ is sufficiently high, the amount of randomness provided by the secret key $s^k \in \mathcal{S}^k$ is large enough so that the encoder and the decoder can secure all the transmitted information. The capacity result we derive in Theorem 11 is still valid when considering the case of infinite state information $s^\infty = (s_1, \ldots, s_n, \ldots, s_{n+k}, \ldots) \in \mathcal{S}^\infty$ because the amount of randomness is infinite.

Remark 13 Comparing the result of Theorem 11 with the outer bound provided in [8], we conclude that the anti-causal secrecy capacity is larger $\mathsf{C}_{\mathsf{ac}} \geq \mathsf{C}_{\mathsf{nc}}$ than the non-causal secrecy capacity and in some cases the inequality is strict.

Remark 14 The first difference with the problem investigated in [10] is that the rate $k \cdot H(\mathbf{s})$ of the shared key $s^k \in \mathcal{S}^k$ can be made arbitrarily large by choosing a large parameter $k \in \mathbb{N}$. Thus the encoder and the decoder are able to secure all the information that can be transmitted reliably. The second difference with the model presented in [10] is that the secret shared key does not influence the statistics of the channel.

Remark 15 If $H(\mathbf{s}) = 0$, the anti-causal secrecy capacity reduces to the one characterized in [3] for the wiretap channel without states. Hence, the anti-causal secrecy capacity is discontinuous between $H(\mathbf{s}) = 0$ and $H(\mathbf{s}) > 0$.

The proof of Theorem 11 is provided in section V. The achievability part relies on a one-time pad argument where the sequence of future states $s^k \in \mathcal{S}^k$ is treated as a secret key shared by both encoder and decoder. The converse follows by considering the state information $\mathbf{s}$ as a part of the channel output.

IV. Gaussian Channel

In this section we present the second contribution of our work which is based on Theorem 8. Theorem 16 states that the presence of the eavesdropper does not affect the secrecy capacity of the Gaussian wiretap channel with non-causal state information. Even if the eavesdropper has a better observation than the legitimate receiver, the encoder and the decoder are able to secure all the information that can be transmitted reliably.

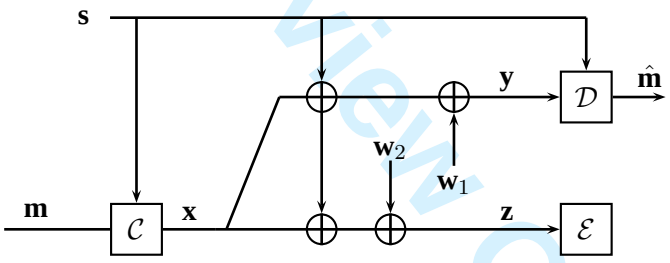

Fig. 3. Gaussian wiretap channel with state information $\mathbf{s}$ available non-causally at both encoder and decoder but not at the eavesdropper. The random variables of the noise $\mathbf{w}_1$ and $\mathbf{w}_2$ and the state $\mathbf{s}$ are Gaussian random variable with zero mean and variance N_1, N_2 and Q.

Let us consider the Gaussian wiretap channel with additive state and noise depicted in Figure 3. The signals received by the legitimate receiver and by the eavesdropper are defined by the equations (15) and (16)

$$\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{w}_1, \tag{15}$$

$$\mathbf{z} = \mathbf{x} + \mathbf{s} + \mathbf{w}_2, \tag{16}$$

where the random variables $\mathbf{w}_1$, $\mathbf{w}_1$ and $\mathbf{s}$ are Gaussian with mean 0 and variance equal to N_1, N_2 and Q.

$$\mathbf{w}_1 \sim \mathcal{N}(0, N_1), \tag{17}$$

$$\mathbf{w}_2 \sim \mathcal{N}(0, N_2), \tag{18}$$

$$\mathbf{s} \sim \mathcal{N}(0, Q). \tag{19}$$

The channel input $\mathbf{x}$ must satisfy the power constraint (20).

$$\mathbb{E}\left[\mathbf{x}^2\right] \leq P. \tag{20}$$

Based on the result stated in Theorem 8 we characterize the secrecy capacity of the Gaussian wiretap channel with non-causal state information available at both encoder and decoder.

Theorem 16 (Second Contribution) *Suppose that $H(\boldsymbol{s}) > 0$, the non-causal secrecy capacity C_{nc} of the Gaussian wiretap channel with state information available at both encoder and decoder is given by equation (21).*

$$\mathsf{C}_{nc} = \max_{\mathcal{P}(\boldsymbol{x}|\boldsymbol{s})} I(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{s}) = \frac{1}{2}\log\left(1+\frac{P}{N_1}\right). \tag{21}$$

Remark 17 The anti-causal secrecy capacity and the non-causal secrecy capacity are equal for the Gaussian channel. Even if the variance $Q > 0$ of the random variable $\mathbf{s}$ is small, the parameter $\alpha \in \mathbb{R}$ can be chosen appropriately such that the term $I(\mathbf{u};\mathbf{y},\mathbf{s}) - I(\mathbf{u};\mathbf{z})$ in equation (13) of Theorem 8 is arbitrarily large and thus C_{nc} satisfies equation (21).

The proof of the Theorem 16 is presented in Section IV-B and consists in evaluating the equation (13) of Theorem 8 with the auxiliary random variable

$$\mathbf{u} = \mathbf{x} + \alpha \cdot \mathbf{s} \sim \mathcal{N}(0, P + \alpha^2 Q), \tag{22}$$

and $\alpha \in]-\infty, +\infty[$. The knowledge of the sequence $s^n \in \mathcal{S}^n$ of state information allows the encoder to transmit the maximal rate of information to the legitimate decoder while keeping it secret from the decoder.

A. Numerical result

We provide a numerical illustration of the result stated in Theorem 16. Figure 4 represents four information rates in terms of parameter $\alpha \in]-15, 15[$. We can see that the rate $I(\mathbf{u};\mathbf{y},\mathbf{s}) - I(\mathbf{u};\mathbf{s})$

provided by Gel'fand Pinsker's coding [4] is constant whereas the rate $I(\mathbf{u};\mathbf{y},\mathbf{s}) - I(\mathbf{u};\mathbf{z})$ provided by Wyner's coding [2] corresponds to the opposite curve of the one described by Costa in [11]. The parameters α_1 and α_2 given by equations (34), (35) and correspond to the values for which the mutual information $I(\mathbf{u};\mathbf{s})$ and $I(\mathbf{u};\mathbf{z})$ are equal. For the whole range of parameters $\alpha \in]-\infty,\alpha_1] \cup [\alpha_2,+\infty[$, the secure rate provided by Theorem 8 is equal to the channel capacity $\mathsf{C}_{\mathsf{nc}} = \max_{\mathcal{P}(\mathbf{x}|\mathbf{s})} I(\mathbf{x};\mathbf{y}|\mathbf{s})$.

B. Proof of Theorem 16

In order to prove Theorem 16, we evaluate the achievable rate stated in Theorem 8 for the Gaussian case with the random variable $\mathbf{u}$ defined by equation (22). This leads to the following equations.

$$\mathsf{C}_{\mathsf{nc}} \geq \max_{\mathcal{P}(\mathbf{u},\mathbf{x}|\mathbf{s})} \left[I(\mathbf{u};\mathbf{y},\mathbf{s}) - \max\left(I(\mathbf{u};\mathbf{s}), I(\mathbf{u};\mathbf{z}) \right) \right] \tag{23}$$

$$= \max_{\mathcal{P}(\mathbf{u},\mathbf{x}|\mathbf{s})} \left[I(\mathbf{u};\mathbf{y}|\mathbf{s}) - \max\left(0, I(\mathbf{u};\mathbf{z}) - I(\mathbf{u};\mathbf{s}) \right) \right] \tag{24}$$

$$= \max_{\mathcal{P}(\mathbf{u},\mathbf{x}|\mathbf{s})} \Big[h(\mathbf{y},\mathbf{s}) - h(\mathbf{y},\mathbf{u},\mathbf{s}) - h(\mathbf{s}) + h(\mathbf{u},\mathbf{s}) \tag{25}$$

$$- \max\left(0, h(\mathbf{z}) - h(\mathbf{u},\mathbf{z}) - h(\mathbf{s}) + h(\mathbf{u},\mathbf{s}) \right) \Big] \tag{26}$$

$$= \frac{1}{2} \max_{\alpha \in \mathbb{R}} \left[\log\left(\frac{(P+N_1)QPQ}{PQN_1Q} \right) - \max\left(0, \log\left(\frac{(P+Q+N_2)PQ}{(PQ(1-\alpha)^2 + N_2(P+\alpha^2 Q))Q} \right) \right) \right] \tag{27}$$

$$= \frac{1}{2} \log\left(1 + \frac{P}{N_1} \right) - \frac{1}{2} \max_{\alpha \in \mathbb{R}} \left[0, \log\left(\frac{(P+Q+N_2)P}{PQ(1-\alpha)^2 + N_2(P+\alpha^2 Q)} \right) \right]. \tag{28}$$

By hypothesis, we have that $H(\mathbf{s}) > 0$ which is equivalent to the fact that the variance $Q > 0$. The sign

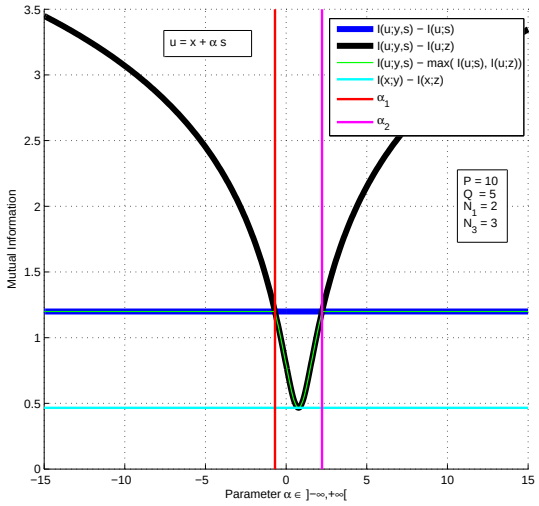

Fig. 4. The achievable rate characterized by Chen and Vinck [5] equals the capacity $\mathsf{C}_{\mathsf{nc}} = \max_{\mathcal{P}(\mathbf{x}|\mathbf{s})} I(\mathbf{x};\mathbf{y}|\mathbf{s})$ of the channel without eavesdropper for all $\alpha \in]-\infty, \alpha_1] \cup [\alpha_2, +\infty[$.

of the term $I(\mathbf{u};\mathbf{z}) - I(\mathbf{u};\mathbf{s})$ is characterized by the following equations.

$$
\begin{aligned}
0 &\geq I(\mathbf{u};\mathbf{z}) - I(\mathbf{u};\mathbf{s}) && (29)\\
\iff \quad 0 &\geq \log\left(\frac{(P+Q+N_2)P}{PQ(1-\alpha)^2 + N_2(P+\alpha^2 Q)}\right) && (30)\\
\iff \quad 1 &\geq \frac{(P+Q+N_2)P}{PQ(1-\alpha)^2 + N_2(P+\alpha^2 Q)} && (31)\\
\iff \quad 0 &\leq \alpha^2 Q(P+N_2) - 2\alpha PQ - P^2 && (32)\\
\iff \quad \alpha &\in]-\infty, \alpha_1] \cup [\alpha_2, +\infty[. && (33)
\end{aligned}
$$

where α_1 and α_2 are defined by equations (34) and (35).

$$
\begin{aligned}
\alpha_1 &= \frac{P(Q - \sqrt{Q(P+Q+N_2)})}{Q(P+N_2)}, && (34)\\
\alpha_2 &= \frac{P(Q + \sqrt{Q(P+Q+N_2)})}{Q(P+N_2)}. && (35)
\end{aligned}
$$

Any choice of the parameter α such that $\alpha \in]-\infty, \alpha_1] \cup [\alpha_2, +\infty[$ makes the term $I(\mathbf{u};\mathbf{z}) - I(\mathbf{u};\mathbf{s})$ non-positive and so the rate given by equation (21) is equal to the channel capacity $\mathsf{C}_{\mathsf{nc}} = \max_{\mathcal{P}(\mathbf{x}|\mathbf{s})} I(\mathbf{x};\mathbf{y}|\mathbf{s})$. This conclude the proof of Theorem 16.

V. Proof of Theorem 11

A. Achievability

For each $\varepsilon > 0$, the notion of random code (see Definition 5) allow us to prove the existence of parameters $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $M \in \mathbb{N}$ and a code $c \in \mathcal{AC}(n,k,M)$ which satisfy the following conditions:

$$\frac{\log_2 M}{n} = \mathsf{C}_{\mathsf{ac}} - 5\varepsilon, \qquad \mathcal{P}_e(c) \leq 7\varepsilon, \qquad \mathcal{L}_e(c) \leq 7\varepsilon. \tag{36}$$

Coding parameters. Let $\mathcal{P}^\star(\mathbf{x}|\mathbf{s}) : \mathcal{S} \longrightarrow \Delta(\mathcal{X})$ the transition probability which attains the maximum in equation (14). Fix $\varepsilon > 0$ and choose two auxiliary parameters $\tilde{\varepsilon} > 0$ and $\gamma \in \mathbb{Q}$ positive, which satisfy the following equation:

$$\Big| I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 6\varepsilon - \gamma H(\mathbf{s}) \Big| < \varepsilon - \gamma\tilde{\varepsilon}. \tag{37}$$

Remark 18 The hypothesis $H(\mathbf{s}) > 0$ insures the existence of the parameters $\gamma \in \mathbb{Q}$ and $\tilde{\varepsilon} > 0$.

Parameters $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $M \in \mathbb{N}$ are defined in the following manner.

- $n \in \mathbb{N}$ is larger than $n_1 \in \mathbb{N}$ et $n_2 \in \mathbb{N}$ defined by the equations (43) and (44) in terms of the parameter $\varepsilon > 0$.
- $k \in \mathbb{N}$ is larger than $k_1 \in \mathbb{N}$, $k_2 \in \mathbb{N}$ and $k_3 \in \mathbb{N}$ defined by the equations (43), (67) and (90) in term of the parameter $\tilde{\varepsilon} > 0$.
- $n \in \mathbb{N}$ and $k \in \mathbb{N}$ satisfy the equation $\gamma = \frac{k}{n}$.
- $M \in \mathbb{N}$ is equal to $2^{n(I(\mathbf{x};\mathbf{y}|\mathbf{s})-5\varepsilon)}$ which is supposed to be an integer.

Remark 19 Substituting $\gamma = \frac{k}{n}$, the equation (37) is equivalent to the pair of equations (38) and (39).

$$n\Big(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 7\varepsilon\Big) < k\Big(H(\mathbf{s}) - \tilde{\varepsilon}\Big), \tag{38}$$

$$k\Big(H(\mathbf{s}) + \tilde{\varepsilon}\Big) < n\Big(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon\Big). \tag{39}$$

Random coding scheme. We define the random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ using a random codebook, an encoding and a decoding function.

- *Random codebook* $(\mathbf{x}^n(m))_{m\in\mathcal{M}}$: The sequence of states $s^{n+k} \in \mathcal{S}^{n+k}$ is divided into two parts $s^n \in \mathcal{S}^n$ and $s^k \in \mathcal{S}^k$. For each sequence of states $s^n \in \mathcal{S}^n$, we generate a random codebook

consisting of $M = 2^{n(I(\mathbf{x};\mathbf{y}|\mathbf{s})-5\varepsilon)}$ sequences $\mathbf{x}^n(m) \in \mathcal{X}^n$ indexed by $m \in \mathcal{M}$ and drawn according to the product probability distribution $\mathcal{P}^\star(\mathbf{x}^n|\mathbf{s}^n) = \prod_{i=1}^n \mathcal{P}^\star(\mathbf{x}_i|\mathbf{s}_i)$ set forth above. Then, each typical sequence of states $s^k \in A_{\tilde{\varepsilon}}^{\star k}(S)$ for the parameter $\tilde{\varepsilon} > 0$, is indexed by $i \in \{1, \ldots, I\} = \mathcal{I}$ with $I = |A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S})|$. For each codebook realization, we define the encoding and decoding functions as follows.

- *Encoder* observes the realizations of the message $m \in \mathcal{M}$ and the sequence of states $s^{n+k} \in \mathcal{S}^{n+k}$. It generates the auxiliary message $m' = m \oplus i \in \mathcal{M}$ by performing the addition modulo-M between index $i \in \mathcal{I}$ of the sequence of states $s^k \in \mathcal{S}^k$ and index $m \in \mathcal{M}$ of the message. The encoder sends through the channel, the inputs sequence $x^n(m') \in \mathcal{X}^n$ that corresponds to the message $m' \in \mathcal{M}$ and the sequence $s^n \in \mathcal{S}^n$.

- *Decoder* observes the realizations of the sequence of outputs $y^n \in \mathcal{Y}^n$ and of the sequence of states $s^{n+k} \in \mathcal{S}^{n+k}$. Decoder finds the auxiliary message $\hat{m}' \in \mathcal{M}$ for which the sequence of inputs $x^n(\hat{m}') \in A_\varepsilon^{\star n}(s^n, y^n)$ is jointly typical for the parameter $\varepsilon > 0$ with the sequences of state and output (s^n, y^n). The decoder obtains the original message $\hat{m} \in \mathcal{M}$ by performing the substraction $\hat{m} = \hat{m}' \ominus i = \hat{m}' \oplus (M - i)$ modulo-M between the index of the auxiliary message $\hat{m}' \in \mathcal{M}$ and the index $i \in \mathcal{I}$ of the sequence of states $s^k(i) \in \mathcal{S}^k$.

- *An error* is declared if $s^k \notin A_{\tilde{\varepsilon}}^{\star n}(\mathcal{S})$ or if $(x^n(m'), s^n, y^n) \notin A_\varepsilon^{\star n}(\mathcal{X} \times \mathcal{S} \times \mathcal{Y})$.

Remark 20 The sequence of states $s^k \in \mathcal{S}^k$ is considered, here above, as a secret key shared by the encoder and the decoder. Equations (38) and (39) guarantee that the sets $\mathcal{I}$ and $\mathcal{M}$ have almost the same cardinality. Hence the sequence of states $s^k \in \mathcal{S}^k$ induces a random index $\mathbf{i}$ distributed almost uniformly over $\mathcal{M}$. As a consequence, the transmitted message $\mathbf{m}' = \mathbf{m} \oplus \mathbf{i}$ is almost statistically independent of the original message $\mathbf{m}$.

Expected error probability. *First,* we provide an upper bound over the expected error probability for the random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ knowing that the realization of the message is $m = 1$.

$$\begin{aligned} & \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}} | \mathbf{m} = 1)\Big] \\ = & \sum_{c \in \mathcal{AC}(n,k,M)} r(\mathbf{c} = c)\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}} | \mathbf{m} = 1, \mathbf{c} = c). \end{aligned} \tag{40}$$

Let us denote by $m' = 1 \oplus i$ and define the following error events:

- $E_1 = \Big\{ s^n \notin A_\varepsilon^{\star n}(\mathcal{S}) \Big\}$ the sequence of states $s^n \in \mathcal{S}^n$ is not ε-typical.
- $E_2 = \Big\{ s^k \notin A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S}) \Big\}$ the sequence of states $s^k \in \mathcal{S}^k$ is not $\tilde{\varepsilon}$-typical.
- $E_3 = \Big\{ x^n(m') \notin A_\varepsilon^{\star n}(\mathcal{X}|s^n) \Big\}$ the sequence of input symbols $x^n(m') \in \mathcal{X}^n$ is not jointly ε-typical with the sequence of states $s^n \in \mathcal{S}^n$.
- $E_4 = \Big\{ y^n \notin A_\varepsilon^{\star n}(\mathcal{Y}|x^n(m'), s^n) \Big\}$ the sequence of output symbols $y^n \in \mathcal{Y}^n$ is not jointly ε-typical with the sequences of input symbols and states $(x^n(m'), s^n) \in \mathcal{X}^n \times \mathcal{S}^n$.
- $E_5 = \Big\{ \exists \tilde{m} \neq m' \in \mathcal{M},\ x^n(\tilde{m}) \in A_\varepsilon^{\star n}(\mathcal{X}|y^n, s^n) \Big\}$ there exists a sequence of symbols $x^n(\tilde{m}) \in \mathcal{X}^n$ corresponding to a message $\tilde{m} \in \mathcal{M}$ which is different from the message sent $m' \in \mathcal{M}$ and the sequence $x^n(\tilde{m}) \in \mathcal{X}^n$ is jointly ε-typical with the sequences of output symbols and states $(y^n, s^n) \in \mathcal{Y}^n \times \mathcal{S}^n$.

Let us denote by $\mathbf{E}_i$ the random variable of the event E_i with $i \in \{1,2,3,4,5\}$. The expected error probability is upper bounded using Boole's inequality:

$$\begin{aligned} & \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}} | \mathbf{m} = 1)\Big] \\ \leq & \ \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_1 \cup \mathbf{E}_2 \cup \mathbf{E}_3 \cup \mathbf{E}_4 \cup \mathbf{E}_5)\Big] \qquad (41) \\ \leq & \ \mathcal{P}(\mathbf{E}_1) + \mathcal{P}(\mathbf{E}_2) + \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_3)\Big] \\ + & \ \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_4)\Big] + \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_5)\Big]. \qquad (42) \end{aligned}$$

Random events $\mathbf{E}_1$ and $\mathbf{E}_2$ are independents of the realization of the random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$. The properties of ε-typical sequences (see [12], property (33) page 26 and the "Conditional Typicality Lemma" page 27) imply that there exists $n_1 \in \mathbb{N}$ and $k_1 \in \mathbb{N}$ such that for all $n \geq n_1$ and $k \geq k_1$ we

have:

$$\max\left(\mathcal{P}(\mathbf{E}_1), \mathcal{P}(\mathbf{E}_2), \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_3)\Big], \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_4)\Big]\right) \leq \varepsilon. \tag{43}$$

The random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ allows to bound the expected error probability $\mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_5)\Big]$. From the "jointly typical Lemma" (see [12], page 29), there exists a $n_2 \in \mathbb{N}$ such that for all $n \geq n_2$, we have the following property:

$$\frac{\log_2 M}{n} = I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon \Longrightarrow \mathbb{E}_c\Big[\mathcal{P}(\mathbf{E}_5)\Big] \leq \varepsilon. \tag{44}$$

The expected error probability of the random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ is upper bounded by:

$$\mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=1)\Big] \quad \leq \quad 5\varepsilon. \tag{45}$$

Second, we provide an upper bound over the expected error probability. Since the codebook is drawn randomly, the expected error probability does not depend on the realized message $m \in \mathcal{M}$. We obtain the following equation:

$$\begin{aligned} & \sum_{c \in \mathcal{AC}(n,k,M)} r(\mathbf{c}=c)\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=1, \mathbf{c}=c) \\ = & \sum_{c \in \mathcal{AC}(n,k,M)} r(\mathbf{c}=c)\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=m, \mathbf{c}=c), \quad \forall m \in \mathcal{M}, \end{aligned} \tag{46}$$

$$\begin{aligned} \Longleftrightarrow \quad & \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=1)\Big] \\ = \quad & \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=m)\Big], \quad \forall m \in \mathcal{M}. \end{aligned} \tag{47}$$

Consequently, the expected error probability can be written:

$$\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c})\Big] \tag{48}$$

$$= \sum_{\substack{c \in \mathcal{AC}(n,k,M), \\ m \in \mathcal{M}}} r(\mathbf{c}=c)\mathcal{P}(\mathbf{m}=m)\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=m, \mathbf{c}=c) \tag{49}$$

$$= \sum_{m \in \mathcal{M}} \mathcal{P}(\mathbf{m}=m)\mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=m)\Big] \tag{50}$$

$$= \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=1)\Big] \sum_{m \in \mathcal{M}} \mathcal{P}(\mathbf{m}=m) \tag{51}$$

$$= \mathbb{E}_c\Big[\mathcal{P}(\mathbf{m} \neq \hat{\mathbf{m}}|\mathbf{m}=1)\Big] \tag{52}$$

$$\leq 5\varepsilon. \tag{53}$$

We prove the expected error probability of random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ is lower than 5ε.

$$\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c})\Big] \quad \leq 5\varepsilon. \tag{54}$$

Information rate. For all codes $c \in \mathcal{AC}(n,k,M)$ belonging to the support of random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$, the information rate is given by the equation (55).

$$\frac{\log M}{n} \quad = \quad \mathsf{C_{ac}} - 5\varepsilon. \tag{55}$$

Expected information leakage rate. Equations (38), (39) and Lemma 21 allow us to obtain a bound on the expected rate of equivocation $\mathbb{E}_c\Big[\mathcal{L}_e(\mathbf{c})\Big] \leq 2\varepsilon$.

$$\begin{aligned}
\mathbb{E}_c\Big[I(\mathbf{m};\mathbf{z}^n|\ \mathbf{c}\ )\Big] &\leq \mathbb{E}_c\Big[I(\mathbf{m};\mathbf{m}'|\ \mathbf{c}\ )\Big] && (56)\\
&= I(\mathbf{m};\mathbf{m}') && (57)\\
&= H(\mathbf{m}') - H(\mathbf{m}'|\mathbf{m}) && (58)\\
&\leq \log M - H(\mathbf{m}\oplus\mathbf{i}|\mathbf{m}) && (59)\\
&= n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - H(\mathbf{i}|\mathbf{m})\\
&- H(\mathbf{m}\oplus\mathbf{i}|\mathbf{i},\mathbf{m}) + H(\mathbf{i}|\mathbf{m},\mathbf{m}\oplus\mathbf{i}) && (60)\\
&= n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - H(\mathbf{i}|\mathbf{m})\\
&+ H(\mathbf{i}|\mathbf{m},\mathbf{m}\oplus\mathbf{i}) && (61)\\
&= n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - H(\mathbf{i}|\mathbf{m}) && (62)\\
&= n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - H(\mathbf{i}) && (63)\\
&\leq n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon) - k(H(\mathbf{s}) - \tilde{\varepsilon}) && (64)\\
&< n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 5\varepsilon)\\
&- n(I(\mathbf{x};\mathbf{y}|\mathbf{s}) - 7\varepsilon) && (65)\\
&\leq n2\varepsilon. && (66)
\end{aligned}$$

• Inequality (56) follows from the fact that for each code $c \in \mathcal{AC}(n,k,M)$, the Markov chain $\mathbf{m}-\mathbf{m}'-\mathbf{z}^n$ is satisfied. Indeed, the sequence of channel outputs $\mathbf{z}^n$ depends on $\mathbf{m}$ only through $\mathbf{m}'$. The conditional probability can be written $\mathcal{P}(z^n|m,m') = \mathcal{P}(z^n|m')$ for all $m \in \mathcal{M}$, $m' \in \mathcal{M}$, $z^n \in \mathcal{Z}^n$ and this for any code $c \in \mathcal{AC}(n,k,M)$. From the data processing inequality (see [12], page 24), the inequality $I(\mathbf{m};\mathbf{z}^n|\ \mathbf{c}=c\ ) \leq I(\mathbf{m};\mathbf{m}'|\ \mathbf{c}=c\ )$ is valid for all codes $c \in \mathcal{AC}(n,k,M)$ and this proves the inequality (56).

• Inequality (57) is due to the fact that the probability distribution $\mathcal{P}(\mathbf{m},\mathbf{m}') \in \mathcal{M}\times\mathcal{M}$ of the messages $\mathbf{m}$ et $\mathbf{m}'$ is independent of the code $\mathbf{c} \in \mathcal{AC}(n,k,M)$ choused at random.

• Inequality (59) is due to the fact that the addition $\oplus$ is performed modulo M and then the message $\mathbf{m}' = \mathbf{m}\oplus\mathbf{i}$ belongs to the set $\mathcal{M}$ of cardinality M.

• Equality (61) is due to the fact that $\mathbf{m}\oplus\mathbf{i}$ is a deterministic function of $\mathbf{i}$ and $\mathbf{m}$, then $H(\mathbf{m}\oplus\mathbf{i}|\mathbf{i},\mathbf{m}) = 0$.

• Equality (62) is due to the condition (39) and to the properties of the set of $\tilde{\varepsilon}$-typical sequences presented in [12] by the property 2 page 26. Indeed, there exists $k_2 \in \mathbb{N}$ such that for all $k \geq k_2$:

$$|A^{\star k}_{\tilde{\varepsilon}}(\mathcal{S})| \leq 2^{k(H(\mathbf{s})+\tilde{\varepsilon})}. \tag{67}$$

Equations (39) and (67) allow us to obtain the following equation:

$$I = |A^{\star k}_{\tilde{\varepsilon}}(\mathcal{S})| \leq 2^{k(H(\mathbf{s})+\tilde{\varepsilon})} < 2^{n(I(\mathbf{x};\mathbf{y}|\mathbf{s})-5\varepsilon)} = M. \tag{68}$$

Because $I < M$, the realizations $m \in \mathcal{M}$ and $m\oplus i \in \mathcal{M}$ allow us to characterize a unique $i \in \mathcal{I} \subset \mathcal{M}$. As a consequence $H(\mathbf{i}|\mathbf{m},\mathbf{m}\oplus\mathbf{i}) = 0$.

• Equality (63) is due to the fact that the index $\mathbf{i}$ and the message $\mathbf{m}$ are drawn independently.

• Inequality (64) is due to the Lemma 21.

• Inequality (65) is due to the condition (38).

Lemme 21 *For all $\tilde{\varepsilon}$, there exists $k_3 \in \mathbb{N}$ such that for all $k \geq k_3$:*

$$H(\boldsymbol{i}) \geq k(H(\boldsymbol{s}) - \tilde{\varepsilon}). \tag{69}$$

Remark 22 The Lemma 21 guarantee that the secret key $\mathbf{i} \in \mathcal{I}$ shared by the encoder and the decoder has a rate closed to $k \cdot H(\mathbf{s})$ and hence closed to the anti-causal secrecy capacity C_{ac}.

The proof of Theorem 11 is outlined in Section V-C.

Existence of a code. The random code $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ we defined here above, satisfies the following three conditions:

1) Every code $c \in \mathcal{AC}(n,k,M)$ belonging to the support of $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ have a rate equal to

$$\frac{\log M}{n} = \mathsf{C}_{\mathsf{ac}} - 5\varepsilon. \tag{70}$$

2) The expected error probability is bounded by

$$\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c})\Big] \leq 5\varepsilon. \tag{71}$$

2) The expected information leakage rate is bounded by

$$\mathbb{E}_c\Big[\mathcal{L}_e(\mathbf{c})\Big] \leq 2\varepsilon. \tag{72}$$

The bounds on the expected error probability and the expected information leakage rate imply:

$$\mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c})\Big] + \mathbb{E}_c\Big[\mathcal{L}_e(\mathbf{c})\Big] \leq 7\varepsilon \tag{73}$$

$$\iff \quad \mathbb{E}_c\Big[\mathcal{P}_e(\mathbf{c}) + \mathcal{L}_e(\mathbf{c})\Big] \leq 7\varepsilon \tag{74}$$

$$\iff \quad \textstyle\sum_{c\in\mathcal{AC}(n,k,M)} r(\mathbf{c}=c)\Big[\mathcal{P}_e(c) + \mathcal{L}_e(c)\Big] \leq 7\varepsilon \tag{75}$$

$$\implies \quad \min_{c\in\mathcal{AC}(n,k,M)}\Big[\mathcal{P}_e(c) + \mathcal{L}_e(c)\Big] \leq 7\varepsilon \tag{76}$$

$$\iff \quad \exists c^\star \in \mathcal{AC}(n,k,M), \quad \mathcal{P}_e(c^\star) + \mathcal{L}_e(c^\star) \leq 7\varepsilon. \tag{77}$$

This demonstrates that there exists a code $c^\star \in \mathcal{AC}(n,k,M)$ in the support of $r \in \Delta\Big(\mathcal{AC}(n,k,M)\Big)$ such that the error probability and the information leakage rate are bounded below 7ε.

$$\mathcal{P}_e(c^\star) \leq 7\varepsilon, \tag{78}$$

$$\mathcal{L}_e(c^\star) \leq 7\varepsilon. \tag{79}$$

To conclude, we showed the existence of a code $c^\star \in \mathcal{AC}(n,k,M)$ whose rate is equal to $\frac{\log M}{n} = \mathsf{C}_{\mathsf{ac}} - 5\varepsilon$, the probability of error is bounded by $\mathcal{P}_e(c^\star) \leq 7\varepsilon$ and the information leakage rate is bounded by $\mathcal{L}_e(c^\star) \leq 7\varepsilon$.

B. Converse

The converse of Theorem 11 is obtained from the converse result of the point to point channel coding result [13], [12] considering the pair $(\mathbf{y}^n, \mathbf{s}^n)$ as the channel output instead of $\mathbf{y}^n$.

C. Proof of Lemma 21

The random variable $\mathbf{E}$ is defined by equation (80).

$$\mathbf{E} = \begin{cases} 0 \text{ si } s^k \in A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S}) \\ 1 \text{ si } s^k \notin A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S}). \end{cases} \tag{80}$$

The following equalities are due to the definition of entropy:

$$\begin{aligned} & H(\mathbf{s}^k, \mathbf{E}) = H(\mathbf{E}) + H(\mathbf{s}^k|\mathbf{E}) \\ & = H(\mathbf{s}^k) + H(\mathbf{E}|\mathbf{s}^k) = H(\mathbf{s}^k) \qquad (81) \\ \Longrightarrow \quad & H(\mathbf{s}^k|\mathbf{E}=0)\mathcal{P}(\mathbf{E}=0) \\ & = H(\mathbf{s}^k) - H(\mathbf{E}) - H(\mathbf{s}^k|\mathbf{E}=1)\mathcal{P}(\mathbf{E}=1). \qquad (82) \end{aligned}$$

Let us denote by $H_b(\delta)$ the entropy of the binary random variable $\{0,1\}$ drawn according to the probabilities $(\delta, 1-\delta)$ with parameter $\delta \in [0,1]$. The random variable $\mathbf{i} \in \mathcal{I}$ is defined as the index of the sequence of states $\mathbf{s}^k(\mathbf{i}) \in A_{\tilde{\varepsilon}}^{\star k}(\mathcal{S})$ that belong to the set of $\tilde{\varepsilon}$-typical sequences Its probability distribution is given by the following equation:

$$\begin{aligned} \mathcal{P}(\mathbf{i}=i) &= \mathcal{P}(\mathbf{s}^k = s^k(i)|\mathbf{E}=0) \qquad (83) \\ &= \frac{\mathcal{P}(\mathbf{s}^k = s^k(i), \mathbf{E}=0)}{\mathcal{P}(\mathbf{E}=0)}. \qquad (84) \end{aligned}$$

The entropy of this random variable satisfies the following equations:

$$\begin{aligned} H(\mathbf{i}) &= H(\mathbf{s}^k|\mathbf{E}=0) \qquad (85) \\ &\geq H(\mathbf{s}^k|\mathbf{E}=0)\mathcal{P}(\mathbf{E}=0) \qquad (86) \\ &= H(\mathbf{s}^k) - H(\mathbf{E}) - H(\mathbf{s}^k|\mathbf{E}=1)\mathcal{P}(\mathbf{E}=1) \qquad (87) \\ &\geq kH(\mathbf{s}) - H_b(\mathcal{P}(\mathbf{E}=1)) - \mathcal{P}(\mathbf{E}=1)k\log|\mathcal{S}| \qquad (88) \\ &\geq k\left(H(\mathbf{s}) - \left(\frac{H_b(\mathcal{P}(\mathbf{E}=1))}{k} + \mathcal{P}(\mathbf{E}=1)\log|\mathcal{S}|\right)\right). \\ & \qquad (89) \end{aligned}$$

From the properties of ε-typical sequences (see [12], property 3 page 26), for all $\tilde{\varepsilon}$, there exists a $k_3 \in \mathbb{N}$ such that for all $k \geq k_3$ we have:

$$\frac{H_b(\mathcal{P}(\mathbf{E}=1))}{k} + \mathcal{P}(\mathbf{E}=1)\log|\mathcal{S}| \quad \leq \quad \tilde{\varepsilon}. \tag{90}$$

We showed that for all $\tilde{\varepsilon}$, there exists a $k_3 \in \mathbb{N}$ such that for all $k \geq k_3$ we have:

$$H(\mathbf{i}) \geq k(H(\mathbf{s}) - \tilde{\varepsilon}). \tag{91}$$

This concludes the proof of Lemma 21.

VI. Conclusion

This article is devoted to the problem of secure communication over a wiretap channel with state information available at both encoder and decoder but not at the eavesdropper. The secrecy capacity for such a channel is not available yet in the literature for neither the causal, nor the non-causal case. We introduce the concept of anti-causal state information i.e. the length of the sequence of states available at both encoder and decoder is arbitrarily larger than the length of the transmission block. We characterize the secrecy capacity for the discrete channel with anti-causal state information and for the Gaussian channel with non-causal state information. These two cases are of particular interest because we show that the encoder and the decoder can use the state information in order to secure all the information that can be transmitted reliably.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[2] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

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

[5] Y. Chen and H. Vinck, "Wiretap channel with side information," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 395–402, 2008.

[6] M. LeTreust, A. Zaidi, and S. Lasaulce, "An achievable rate region for the broadcast wiretap channel with asymmetric side information," *IEEE Proc. of the 49th Allerton conference, Monticello, Illinois*, 2011.

[7] W. Liu and B. Chen, "Wiretap channel with two-sided channel state information," *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, pp. 893–897, 2007.

[8] Y. K. Chia and A. E. Gamal, "Wiretap channel with causal state information," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2838–2849, 2012.

[9] H. Yamamoto, "Rate-distortion theory for the shannon cipher system," *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 827–835, 1997.

[10] W. Kang and N. Liu, "Wiretap channel with shared key," in *Proc. IEEE Int. Symp. Information Theory (ISIT'10)*, Sep. 2010.

[11] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439–441, 1983.

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

[13] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.