On the Secrecy Capacity of Arbitrary Wiretap Channels

Matthieu Bloch and J. Nicholas Laneman

***Abstract*— We investigate the fundamental secrecy limits of arbitrary wiretap channels using the information-spectrum approach and we provide a random coding theorem for the secrecy capacity under various secrecy metrics. We show how our result specializes to several recent results, e.g., compound channels, parallel channels, and fading channels. As a side benefit, our analysis shows that earlier results hold under more stringent secrecy metrics than previously established.**

I. Introduction

A. Motivation

Although cryptography is traditionally handled at the application layer without regard to the imperfections of the lower layers, many contributions support the idea that the inherent noise of communication channels can be exploited for security. The idea of designing *physical-layer* security schemes first appeared in the seminal works of Wyner [1] and Csiszár and Körner [2], who investigated a channel model coined the *wiretap channel*. In this model, a transceiver attempts to communicate reliably and securely with a legitimate receiver over a noisy channel, while its messages are being eavesdropped by a passive adversary through another noisy channel. In sharp contrast with Shannon's disappointing result [3] on the impracticality of information-theoretic security, [1] and [2] proved the existence of coding schemes achieving information-theoretically secure communications over certain wiretap channels.

With the growth of wireless communications, which are extremely susceptible to eavesdropping by nature and whose ubiquitous deployment makes security a crucial issue, there has recently been a renewed interest for the analysis of wiretap channel models. The fundamental secrecy limits of various fading wiretap channels have been characterized in [4], [5], [6], and results on extensions of wiretap channels to multi-user scenarios have been provided in [7], [8], [9], [10]. Despite the surge of recent results, the mathematical tools used to analyze wiretap channels are still largely based on [1], [2], which suggests the existence of a more fundamental connection between many models. Hayashi [11] recently analyzed the connection between coding for the wiretap channel and resolvability using the information-spectrum approach [12], and derived the secrecy capacity of arbitrary wiretap channels for the case in which the output alphabet at the eavesdropper is finite. With the notable exception of [13], [14], Hayashi's results have caught little attention in spite of their generality.

M. Bloch and J. N. Laneman are with the Department of Electrical Engineering, University of Notre Dame, Fitzpatrick Hall, Notre Dame, IN, 46556, USA. {mbloch1,jnl}@nd.edu

Motivated by the perspective of a general result encompassing many wiretap channel models, we also investigate the information-spectrum approach to arbitrary wiretap channels. Section I-B reviews classic definitions of information spectra and sets the notation used throughout the paper. Section II establishes our main result, a closed-form expression of the secrecy capacity of arbitrary wiretap channels under various secrecy criteria. The usefulness of our approach is illustrated in Section III, where we show that earlier contributions can be obtained by reducing the general result to special cases.

B. Information-Spectrum information theory

Consider two random variables X and Y taking values in alphabets $\mathcal{X}$ and $\mathcal{Y}$. Sample values of X and Y are denoted by x and y, respectively; the joint probability law is denoted by $p_{\mathsf{XY}}(\mathsf{x},\mathsf{y})$, and the marginal probabilities are denoted by $p_{\mathsf{X}}(\mathsf{x})$ and $p_{\mathsf{Y}}(\mathsf{y})$, respectively. Unless mentioned otherwise, alphabets are assumed to be abstract alphabets, including countably infinite or continuous alphabets. The *mutual information* between X and Y is the random variable [15], [16], [17], [18]

$$\mathrm{I}(\mathsf{X};\mathsf{Y}) := \log \frac{p_{\mathsf{XY}}(\mathsf{X},\mathsf{Y})}{p_{\mathsf{X}}(\mathsf{X})\,p_{\mathsf{Y}}(\mathsf{Y})}.$$

We refer to the average of the mutual information random variable as the *average mutual information*, and denote it by

$$\mathbb{I}(\mathsf{X};\mathsf{Y}) := \mathbb{E}[\mathrm{I}(\mathsf{X};\mathsf{Y})] = \sum_{\mathsf{x}\in\mathcal{X}}\sum_{\mathsf{y}\in\mathcal{Y}} p_{\mathsf{XY}}(\mathsf{x},\mathsf{y}) \log \frac{p_{\mathsf{XY}}(\mathsf{x},\mathsf{y})}{p_{\mathsf{X}}(\mathsf{x})\,p_{\mathsf{Y}}(\mathsf{y})}.$$

The *variational distance* between two random variables $\mathsf{X}\in\mathcal{X}$ and $\mathsf{X}'\in\mathcal{X}$ is defined as

$$d\left(p_{\mathsf{X}},p_{\mathsf{X}'}\right) := \sum_{\mathsf{x}\in\mathcal{X}} |p_{\mathsf{X}}(\mathsf{x}) - p_{\mathsf{X}'}(\mathsf{x})|.$$

Given two sequence of random variables $\{\mathsf{X}^n\}_{n=1}^{\infty}$ and $\{\mathsf{Y}^n\}_{n=1}^{\infty}$, where X^n and Y^n are arbitrary random variables in $\mathcal{X}^n$ and $\mathcal{Y}^n$, respectively, the *mutual information spectrum* and *rate-mutual information spectrum* are defined as the probability distribution of $\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n)$ and $\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n)$, respectively. In addition, the *spectral sup-mutual information rate* and the *spectral inf-mutual information rate* are defined as [12]

$$\begin{aligned}\operatorname*{p-limsup}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n)&\\ := \inf\left\{\alpha : \lim_{n\to\infty}\mathbb{P}\left[\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n)>\alpha\right]=0\right\},&\end{aligned}$$

and

$$\underset{n\to\infty}{\text{p-liminf}}\ \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n) := \sup\left\{\beta : \lim_{n\to\infty}\mathbb{P}\left[\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n)<\beta\right]=0\right\},$$

respectively. We will see that these two quantities play a central role in establishing the secrecy capacity of an arbitrary wiretap channel.

II. Secrecy capacity of arbitrary wiretap channels

Definition 1: A general wiretap channel denoted by $\left(\mathcal{X},\mathcal{Y},\mathcal{Z},\left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$ consists of an arbitrary input alphabet $\mathcal{X}$, two arbitrary output alphabets $\mathcal{Y}$ and $\mathcal{Z}$, and a sequence of transition probabilities $\left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\right\}_{n=0}^{\infty}$ such that

$$\forall n\in\mathbb{N}^*\ \forall \mathsf{x}^n\in\mathcal{X}^n \sum_{\mathsf{y}^n\in\mathcal{Y}^n,\mathsf{z}^n\in\mathcal{Z}^n} p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)=1.$$

A communication scheme over a wiretap channel is illustrated in Figure 1.

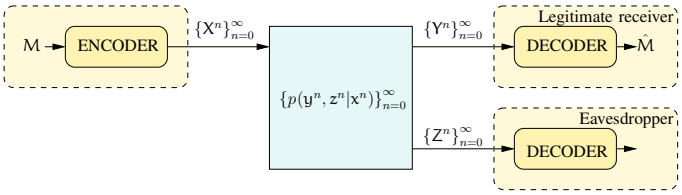

Fig. 1. Communication scheme for an arbitrary wiretap channel.

The marginal probabilities of the transition probabilities $p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)$ with respect to Y^n or Z^n define two channels $\left(\mathcal{X},\mathcal{Y},\left\{p_{\mathsf{Y}^n|\mathsf{X}^n}(\mathsf{y}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$ and $\left(\mathcal{X},\mathcal{Z},\left\{p_{\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$, called the *main channel* and the *eavesdropper's channel*, respectively. Accordingly, the observer of channel output $\mathcal{Y}$ is named the *legitimate receiver* at the observer of channel output $\mathcal{Z}$ is named the *eavesdropper*.

Definition 2: An $(n,M_n,\epsilon_n,\delta_n)$ wiretap code for a wiretap channel $\left(\mathcal{X},\mathcal{Y},\mathcal{Z},\left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$ consists of the following:

- a message set $\mathcal{M}_n=\{1,\dots,M_n\}$;
- a stochastic encoding function $\varphi_n:\mathcal{M}_n\to\mathcal{X}^n$ characterized by a transition probability $p_{\mathsf{X}^n|\mathsf{M}}(\mathsf{x}^n|\mathsf{m})$;
- a decoding function $\phi_n:\mathcal{Y}^n\to\mathcal{M}_n$.

The rate of a wiretap code is defined as $r_n=\frac{1}{n}\log M_n$. Let M be the random variable denoting the uniform choice of message i in $\mathcal{M}_n$, and Y^n (Z^n) be the random variable representing the legitimate receiver (eavesdropper) output corresponding to $\varphi(\mathsf{M})$. The average error probability is defined as

$$\epsilon_n=\mathbb{P}[\phi(\mathsf{Y}^n)\neq\mathsf{M}],$$

and the secrecy at the eavesdropper is measured by one the six following metrics $\delta_n^{(i)}$ ($i\in\{1,\dots,6\}$), which characterize in various ways the dependencies between M and Z^n.

$$\delta_n^{(1)}=\mathbb{I}(\mathsf{M};\mathsf{Z}^n), \tag{1}$$
$$\delta_n^{(2)}=d\left(p_{\mathsf{M}\mathsf{Z}^n},p_{\mathsf{M}}p_{\mathsf{Z}^n}\right) \tag{2}$$
$$\delta_n^{(3)}=\mathrm{I}(\mathsf{M};\mathsf{Z}^n), \tag{3}$$
$$\delta_n^{(4)}=\frac{\mathbb{I}(\mathsf{M};\mathsf{Z}^n)}{n}, \tag{4}$$
$$\delta_n^{(5)}=\frac{d\left(p_{\mathsf{M}\mathsf{Z}^n},p_{\mathsf{M}}p_{\mathsf{Z}^n}\right)}{n}, \tag{5}$$
$$\delta_n^{(6)}=\frac{\mathrm{I}(\mathsf{M};\mathsf{Z}^n)}{n}, \tag{6}$$

Definition 3: A rate R is achievable over a wiretap channel $\left(\mathcal{X},\mathcal{Y},\mathcal{Z},\left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$ under secrecy criterion (i) if there exists a sequence of $\left(n,M_n,\epsilon_n,\delta_n^{(i)}\right)$ codes such that

$$\liminf_{n\to\infty} r_n\geq R,\quad \limsup_{n\to\infty}\epsilon_n=0,\ \text{and}\ \underset{n\to\infty}{\text{p-lim}}\ \delta_n^{(i)}=0,$$

where p-lim denotes convergence in probability.

Definition 4: The *secrecy capacity* of a general wiretap channel is defined as

$$C_s^{(i)}=\sup\{R: R\text{ is achievable under secrecy criterion }(i)\}$$

Regardless of which metric (i) is used, the requirements p-lim $\delta_n^{(i)}=0$ express the idea that the random variable M and Z^n should become independent as the block length n gets large. Achievability under secrecy criteria (1) and (4) corresponds to the usual notions of *strong secrecy* [19] and *weak secrecy*, respectively. Although the former is obviously stronger than the latter, most of the recent literature on wiretap channels assesses security based on (4). Achievability under (2) and (5) settles for the asymptotic independence of M and Z^n in terms of the variational distance and normalized variational distance, respectively. Finally, achievability under (3) and (6) requires the convergence in probability to zero of the mutual information spectrum and of the rate-mutual information spectrum, respectively.

It is of both theoretical and practical importance to have means of comparing the various secrecy metrics. To that end, we define an ordering of the secrecy requirements as follows. We say that secrecy requirement (j) is *stronger* than secrecy requirement (i) (or, equivalently, (i) is *weaker* than (j)), and we write $(j)\succeq(i)$, if

$$\left(\underset{n\to\infty}{\text{p-lim}}\ \delta_n^{(j)}=0\right)\Rightarrow\left(\underset{n\to\infty}{\text{p-lim}}\ \delta_n^{(i)}=0\right).$$

If $(i)\succeq(j)$ and $(j)\succeq(i)$, we simply write $(i)\equiv(j)$. By definition, it is clear that (1) is stronger than (4), (2) is stronger than (5), and (3) is stronger than (6); however, the following lemma establishes a more precise ordering.

Lemma 1 (Ordering of secrecy metrics): The secrecy criteria (1)-(6) can be ordered as follows.

$$(1)\succeq(2)\succeq(3)\succeq(4)\succeq(5)\succeq(6).$$

Proof: See Appendix I. ∎

We note that, in [11], Hayashi analyzes the security of wiretap channels with the following metric.

$$\frac{1}{M_n}\frac{1}{M_n-1}\sum_{m=1}^{M_n}\sum_{m'=1:m\neq n}^{M_n} d\left(p_{Z^n|m}, p_{Z^n|m'}\right) \qquad (7)$$

It can be easily verified that $(7) \equiv (2)$; however, achievability under secrecy criterion (2) will be simpler to analyze.

The remainder of the section establishes our main result, which is a single expression for the secrecy capacity of arbitrary wiretap channels under criteria (2)-(6). Lemma 2 and Lemma 3 characterize a set of achievable rates under secrecy criterion (2), and Lemma 4 provides a converse result under secrecy criterion (6). According to the ordering established in Lemma 1, the upper bound of the capacity under secrecy criterion (6) established by Lemma 4 is also a valid upper bound of the secrecy capacity under all other criteria. We will see that the maximum achievable rate identified by Lemma 3 matches the outer bound in Lemma 4, allowing us to conclude that the secrecy capacities under criteria (2)-(6) are identical.

Lemma 2 (Achievability): For a wiretap channel $\left(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$, all rates R_s satisfying

$$R_s < \max_{\{\mathsf{X}^n\}_{n=1}^{\infty}} \Bigg(\text{p-liminf}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n)\Bigg)$$

are achievable under secrecy criterion (2).

Proof: We briefly highlight the key ideas of the proof, and we refer the reader to Appendix II for a more rigorous treatment. The result is proven using a random coding argument, and by constructing codes having the structure illustrated in Figure 2. These code contain M_nL_n codewords

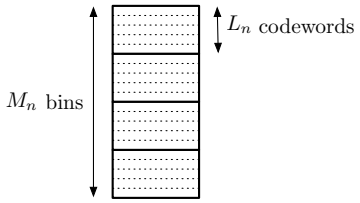

Fig. 2. Random code structure

binned in M_n subcodebooks of size L_n. To transmit a message $i \in \{1, \ldots, M_n\}$, the transmitter selects a codeword at random in bin i. All M_nL_n codewords are meant to be decoded at the output of the main channel, which can be guaranteed if the following upper bound on the total size of the code is satisfied.

$$M_nL_n \leq \text{p-liminf}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n).$$

By noticing that criterion (2) is similar to a *channel resolvability* criterion [12] for each subcodebook, the convergence to zero can be obtained by selecting the subcodebook size such that

$$L_n \geq \text{p-limsup}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n),$$

and the lemma follows by combining the two constraints above. ∎

Interestingly, this proof formalizes the intuition that a wiretap code should exploit some randomness to confuse the eavesdropper by connecting explicitly the secrecy criterion to the problem of channel resolvability. An important consequence is that a code of rate

$$\text{p-liminf}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n) - \beta$$

for some $\beta > 0$, can be generated by selecting

$$M_n = \text{p-liminf}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n) - \beta$$

and

$$L_n = \text{p-limsup}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n) + \beta.$$

In general, this flexibility with subcodebook size cannot be obtained directly from the result of Wyner [1] or Csiszár and Körner [2].

Lemma 2 extends and strengthens [11, Lemma 4], which states a similar result for situations where the output alphabet $\mathcal{Z}$ is finite under a weak secrecy criterion; however, our result does not supersede [11, Theorem 3] where the existence of wiretap codes is characterized in a non-asymptotic setting. We do not have any result providing achievable rates under secrecy criterion (1), although we conjecture that the same set of rates is achievable for the class of stationary memoryless channels.

Lemma 3 (Achievability): For a wiretap channel $\left(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$, all rates R_s satisfying

$$R_s < \max_{\{\mathsf{V}^n,\mathsf{X}^n\}_{n=1}^{\infty}} \Bigg(\text{p-liminf}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n)\Bigg),$$

where the process $\{\mathsf{V}^n, \mathsf{X}^n\}_{n=1}^{\infty}$ is subject to

$$\mathsf{V}^n \to \mathsf{X}^n \to \mathsf{Z}^n\mathsf{Y}^n \quad \forall n \in \mathbb{N}^*,$$

are achievable under secrecy criterion (2).

Proof: The result follows by introducing an arbitrary prefix channel characterized by the sequence of transition probabilities $\left\{p_{\mathsf{X}^n|\mathsf{V}^n}(\mathsf{x}^n|\mathsf{v}^n)\right\}_{n=0}^{\infty}$ and applying Lemma 2 to the concatenated channel characterized by $\left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n\mathsf{z}^n|\mathsf{x}^n)\, p_{\mathsf{X}^n|\mathsf{V}^n}(\mathsf{x}^n|\mathsf{v}^n)\right\}_{n=0}^{\infty}$. ∎

Lemma 4 (Converse): For a wiretap channel $\left(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$, the achievable

rates under secrecy criterion (6) are upper bounded as

$$R_s \leq \max_{\{\mathsf{V}^n,\mathsf{X}^n\}_{n=1}^{\infty}} \left(\text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n) \right),$$

where the process $\{\mathsf{V}^n, \mathsf{X}^n\}_{n=1}^{\infty}$ is subject to

$$\mathsf{V}^n \to \mathsf{X}^n \to \mathsf{Z}^n\mathsf{Y}^n \quad \forall n \in \mathbb{N}^*.$$

Proof: Consider a sequence of $(n, M_n, \epsilon_n, \delta_n)$ wiretap codes achieving a rate R_s under secrecy criterion (6). Let V^n denote the random variable representing the uniform choice of a message in $\mathcal{M}_n$. To model stochastic encoding, we introduce a prefix channel with transition probability $p_{\mathsf{X}^n|\mathsf{V}^n}(\mathsf{x}^n|\mathsf{v}^n)$ that accounts for the randomness used in the encoder. By assumption,

$$\forall \epsilon > 0 \ \lim_{n\to\infty} \mathbb{P}\left[\left| \frac{\mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n)}{n} \right| > \epsilon \right] = 0,$$

therefore

$$\text{p-limsup}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n) = 0.$$

The Verdú-Han Lemma [12, Lemma 3.2.2] ensures that

$$R_s \leq \text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n),$$

therefore, we obtain

$$\begin{aligned} R_s &\leq \text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) \\ &= \text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) \\ &\leq \max_{\{\mathsf{V}^n,\mathsf{X}^n\}_{n=1}^{\infty}} \left(\text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n) \right). \end{aligned}$$

■

Clearly, the upper bound of the set of rates achievable with a wiretap code under criterion (6) is also an upper bound of the set of rates achievable under (2); therefore, combining Lemma 1, Lemma 3, and Lemma 4, we obtain the following theorem.

Theorem 1: The secrecy capacity of a wiretap channel $\left(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$ under secrecy criterion (2)-(6) is

$$C_s = \max_{\{\mathsf{V}^n,\mathsf{X}^n\}_{n=1}^{\infty}} \left(\text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n) \right),$$

where the process $\{\mathsf{V}^n, \mathsf{X}^n\}_{n=1}^{\infty}$ is subject to

$$\mathsf{V}^n \to \mathsf{X}^n \to \mathsf{Z}^n\mathsf{Y}^n \quad \forall n \in \mathbb{N}^*.$$

Theorem 1 is easily generalized to include cost constraints. For $n \in \mathbb{N}$, let $c_n : \mathcal{X}^n \to \mathbb{R}$ be an arbitrary mapping. A sequence of $(n, M_n, \epsilon_n, \delta_n)$ wiretap codes is said to satisfy cost constraint Γ if it satisfies

$$\forall n \in \mathbb{N}^* \, \forall i \in \mathcal{M}_n \quad \frac{1}{n} c_n\left(\varphi(i)\right) \leq \Gamma,$$

and we call it a sequence of $(n, M_n, \epsilon_n, \delta_n, \Gamma)$ codes. The definitions of achievable rates and secrecy capacity with cost constraint Γ are obtained by replacing $(n, M_n, \epsilon_n, \delta_n)$ codes by $(n, M_n, \epsilon_n, \delta_n, \Gamma)$ codes in Definitions 3 and 4, and we have the following result.

Theorem 2: The secrecy capacity of a wiretap channel $\left(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \left\{p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n|\mathsf{x}^n)\right\}_{n=1}^{\infty}\right)$ with cost constraint Γ under secrecy criterion (2)-(6) is

$$C_s = \max_{\{\mathsf{V}^n,\mathsf{X}^n\}_{n=1}^{\infty}} \left(\text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \text{p-limsup}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n) \right),$$

where the processes $\{\mathsf{V}^n, \mathsf{X}^n\}_{n=1}^{\infty}$ is subject to

$$\mathsf{V}^n \to \mathsf{X}^n \to \mathsf{Z}^n\mathsf{Y}^n \quad \forall n \in \mathbb{N}^*,$$

and

$$\mathbb{P}\left[\mathsf{X}^n \in \left\{\mathsf{x}^n \in \mathcal{X}^n : \frac{1}{n} c_n(\mathsf{x}^n) \leq \Gamma\right\}\right] = 1.$$

III. Applications

Specializing the general result obtained in Theorem 1 to simpler channel models is in general non-trivial. The main difficulty stems from the the nature of the spectral sup and spectral inf mutual information rates, which is fundamentally different from the mutual information usually used to characterize communications rates. In this final section, we justify the usefulness of Theorem 1 by showing how some of the results that have appeared in the recent literature are indeed corollaries of Theorem 1. As a side benefit, the results re-derived by reduction of Theorem 1 hold under secrecy criterion (2) and are stronger than their original counterparts, which were only established under weak secrecy.

A. Stationary Memoryless Channels

Corollary 1 (Stationary memoryless channels): Consider a stationary memoryless wiretap channel $\left(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, p_{\mathsf{YZ}|\mathsf{X}}(\mathsf{y}, \mathsf{z}|\mathsf{x})\right)$ with arbitrary input and output alphabets. The following rates are achievable secrecy rates under secrecy criteria (2)-(6).

$$R_s < \sup_{\mathsf{V}\to\mathsf{X}\to\mathsf{YZ}} \left[\mathbb{I}(\mathsf{V};\mathsf{Y}) - \mathbb{I}(\mathsf{V};\mathsf{Z})\right].$$

Moreover, if alphabets $\mathcal{Z}$ is a finite set, the secrecy capacity is given by

$$C_s = \sup_{\mathsf{V}\to\mathsf{X}\to\mathsf{YZ}} \left[\mathbb{I}(\mathsf{V};\mathsf{Y}) - \mathbb{I}(\mathsf{V};\mathsf{Z})\right].$$

Proof: Let V and X be the random variables attaining $\sup_{\mathsf{V}\to\mathsf{X}\to\mathsf{YZ}} [\mathbb{I}(\mathsf{V};\mathsf{Y}) - \mathbb{I}(\mathsf{V};\mathsf{Z})]$, and denote by Y and Z the

corresponding outputs to the channel. Define V^n, X^n, Y^n, and Z^n such that

$$p_{\mathsf{Y}^n\mathsf{Z}^n\mathsf{X}^n\mathsf{V}^n}(\mathsf{y}^n,\mathsf{z}^n,\mathsf{x}^n,\mathsf{v}^n) = \prod_{i=1}^{n} p_{\mathsf{YZ}|\mathsf{X}}(\mathsf{y}_i,\mathsf{z}_i|\mathsf{x}_i)\, p_{\mathsf{X}|\mathsf{V}}(\mathsf{x}_i|\mathsf{v}_i)\, p_{\mathsf{V}}(\mathsf{v}_i)\,,$$

and consider the random processes $\{\mathsf{V}^n\}_{n=1}^{\infty}$, $\{\mathsf{X}^n\}_{n=1}^{\infty}$, $\{\mathsf{Y}^n\}_{n=1}^{\infty}$, and $\{\mathsf{Z}^n\}_{n=1}^{\infty}$. By the law of large number, we know that

$$\frac{\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n)}{n} = \frac{1}{n}\sum_{i=1}^{n}\mathrm{I}(\mathsf{V};\mathsf{Y}) \to \mathbb{I}(\mathsf{V};\mathsf{Y}) \text{ a.s. as } n\to\infty;$$

therefore,

$$\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) = \mathbb{I}(\mathsf{V};\mathsf{Y}). \tag{8}$$

Following the same reasoning, we can also prove that

$$\operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n) = \mathbb{I}(\mathsf{V};\mathsf{Z}). \tag{9}$$

Substituting (8) and (9) in Lemma 3 yields the first part of the corollary.

The second part of the corollary follows by recalling that

$$\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) \le \liminf_{n\to\infty}\frac{1}{n}\mathbb{I}(\mathsf{V}^n;\mathsf{Y}^n), \tag{10}$$

$$\operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n) \ge \limsup_{n\to\infty}\frac{1}{n}\mathbb{I}(\mathsf{V}^n;\mathsf{Z}^n), \tag{11}$$

when $|\mathcal{Z}|$ is finite [12, Theorem 3.5.2]; therefore,

$$\begin{aligned}&\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n)\\ &\le \liminf_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) + \liminf_{n\to\infty}\frac{-1}{n}\mathbb{I}(\mathsf{V}^n;\mathsf{Z}^n)\\ &= \frac{1}{n}\liminf_{n\to\infty}\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n).\end{aligned}$$

Let Q be a random variable independent of all others and uniformly distributed over $\{1,\dots,n\}$. Expanding the above expression in a way similar to that of [2, Section V] and setting,

$$\begin{aligned}&\mathsf{U}' := \mathsf{Q}\mathsf{Y}^{\mathsf{Q}-1}\tilde{\mathsf{Z}}^{\mathsf{Q}+1} \quad \mathsf{V}' := \mathsf{U}_{\mathsf{Q}}\mathsf{V}^n \quad \mathsf{X}' := \mathsf{X}_{\mathsf{Q}}\\ &\mathsf{Y}' := \mathsf{Y}_{\mathsf{Q}} \quad \mathsf{Z}' := \mathsf{Z}_{\mathsf{Q}},\end{aligned}$$

we obtain

$$\begin{aligned}&\mathrm{I}(\mathsf{V}^n;\mathsf{Y}^n) - \mathrm{I}(\mathsf{V}^n;\mathsf{Z}^n)\\ &\le \sum_{i=1}^{n}\frac{\mathbb{I}\Big(\mathsf{V}^n;\mathsf{Y}_i|\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\Big) - \mathbb{I}\Big(\mathsf{V}^n;\mathsf{Z}_i|\mathsf{Y}^{i-1}\tilde{\mathsf{Z}}^{i+1}\Big)}{n}\\ &= \mathbb{I}(\mathsf{V}';\mathsf{Y}'|\mathsf{U}') - \mathbb{I}(\mathsf{V}';\mathsf{Z}'|\mathsf{U}')\\ &\le \sup_{\mathsf{U}\to\mathsf{V}\to\mathsf{X}\to\mathsf{YZ}}[\mathbb{I}(\mathsf{V};\mathsf{Y}|\mathsf{U}) - \mathbb{I}(\mathsf{V};\mathsf{Z}|\mathsf{U})]\\ &= \sup_{\mathsf{V}\to\mathsf{X}\to\mathsf{YZ}}[\mathbb{I}(\mathsf{V};\mathsf{Y}) - \mathbb{I}(\mathsf{V};\mathsf{Z})]\,,\end{aligned}$$

which concludes the proof. ∎

The achievability of rates below the secrecy capacity of a Gaussian wiretap channel [20] follows in a similar way from the reduction of Theorem 2 and [12, Theorem 3.6.2].

For discrete stationary memoryless channels, we note that Maurer and Wolf showed that the rates in Corollary 1 are also achievable under secrecy criterion (1) [19].

B. Mixed Channels

Definition 5: Consider two wiretap channels $\Big(\mathcal{X},\mathcal{Y},\mathcal{Z},\big\{p_{\mathsf{Y}_1^n\mathsf{Z}_1^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\big\}_{n=1}^{\infty}\Big)$ and $\Big(\mathcal{X},\mathcal{Y},\mathcal{Z},\big\{p_{\mathsf{Y}_2^n\mathsf{Z}_2^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\big\}_{n=1}^{\infty}\Big)$, and let $\alpha_1,\alpha_2 > 0$ be constants such that $\alpha_1+\alpha_2 = 1$. We define the *mixed wiretap channel* as the channel $\big(\mathcal{X}, p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n),\mathcal{Y},\mathcal{Z}\big)$, where

$$\begin{aligned}p_{\mathsf{Y}^n\mathsf{Z}^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n) &= \alpha_1 p_{\mathsf{Y}_1^n\mathsf{Z}_1^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\\ &\quad + \alpha_2 p_{\mathsf{Y}_2^n\mathsf{Z}_2^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\,.\end{aligned}$$

Lemma 5: The spectral mutual information rates of a mixed wiretap channel satisfy

$$\begin{aligned}&\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n) =\\ &\quad\min\left(\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}_1^n),\ \operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Y}_2^n)\right),\\ &\operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n) =\\ &\quad\max\left(\operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}_1^n),\ \operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}_2^n)\right),\end{aligned}$$

Proof: The results follow from steps similar to those of [12, Lemma 1.4.2]. ∎

In view of the previous lemma, we immediately obtain the following result.

Theorem 3 (Mixed channels): The secrecy capacity of a mixed wiretap channel under criteria (2)-(6) is given by

$$\max_{\{\mathsf{V}^n,\mathsf{X}^n\}_{n=1}^{\infty}}\left(\min_{i\in\{1,2\}}\operatorname*{p-liminf}_{n\to\infty}\frac{1}{n}\mathrm{I}(\mathsf{V}^n;\mathsf{Y}_i^n) - \max_{j\in\{1,2\}}\operatorname*{p-limsup}_{n\to\infty}\frac{1}{n}\mathrm{I}\big(\mathsf{V}^n;\mathsf{Z}_j^n\big)\right),$$

where the process $\{\mathsf{V}^n,\mathsf{X}^n\}_{n=1}^{\infty}$ is subject to

$$\mathsf{V}^n \to \mathsf{X}^n \to \mathsf{Z}_i^n\mathsf{Y}_i^n \quad \forall n\in\mathbb{N}^*\,\forall i\in\{1,2\}.$$

Corollary 2 (Stationary memoryless mixed channels): For a stationary memoryless mixed wiretap channel, all rates R_s such that

$$R_s < \sup_{\mathsf{V}\to\mathsf{X}\to\mathsf{Y}_i\mathsf{Z}_j}\left(\min_{i\in\{1,2\}}\mathbb{I}(\mathsf{V};\mathsf{Y}_i) - \max_{j\in\{1,2\}}\mathbb{I}(\mathsf{V};\mathsf{Z}_j)\right),$$

are achievable under criteria (2)-(6).

Proof: Let V, X be the random variables attaining $\sup_{\mathsf{V}\to\mathsf{X}\to\mathsf{Y}_i\mathsf{Z}_j}\left(\min_{i\in\{1,2\}}\mathbb{I}(\mathsf{V};\mathsf{Y}_i) - \max_{j\in\{1,2\}}\mathbb{I}(\mathsf{V};\mathsf{Z}_j)\right)$. and denote by Y_i, Z_i ($i\in\{1,2\}$) the corresponding outputs of wiretap channel $\big(\mathcal{X},\mathcal{Y},\mathcal{Z},p_{\mathsf{Y}_i^n\mathsf{Z}_i^n|\mathsf{X}^n}(\mathsf{y}^n,\mathsf{z}^n|\mathsf{x}^n)\big)$. Define

V^n,X^n, Y_i^n, and Z_i^n ($i \in \{1,2\}$) such that

$$\forall i \in \{1,2\} \quad p_{\mathsf{Y}_i^n \mathsf{Z}_i^n \mathsf{X}^n \mathsf{V}^n}(\mathsf{y}^n, \mathsf{z}^n, \mathsf{x}^n, \mathsf{v}^n) = \prod_{j=1}^{n} p_{\mathsf{Y}_i \mathsf{Z}_i | \mathsf{X}}(\mathsf{y}_j, \mathsf{z}_j | \mathsf{x}_j)\, p_{\mathsf{X}|\mathsf{V}}(\mathsf{x}_j|\mathsf{v}_j)\, p_{\mathsf{V}}(\mathsf{v}_j)\,,$$

and consider the random processes $\{\mathsf{V}^n\}_{n=1}^{\infty}$, $\{\mathsf{X}^n\}_{n=1}^{\infty}$, $\{\mathsf{Y}_i^n\}_{n=1}^{\infty}$, and $\{\mathsf{Z}_i^n\}_{n=1}^{\infty}$. By the law of large numbers, we have that

$$\forall i \in \{1,2\} \quad \text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n; \mathsf{Y}_i^n) = \mathbb{I}(\mathsf{V}; \mathsf{Y}_i) \quad \text{and} \quad \text{p-limsup}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n; \mathsf{Z}_i^n) = \mathbb{I}(\mathsf{V}; \mathsf{Z}_i),$$

which yields the desired result. ∎

The above result can be easily generalized to a mixture of $j > 2$ channels, and to an infinite mixture of channels.

Definition 6: Given infinitely many wiretap channels $\big(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, p_{\mathsf{Y}_i^n \mathsf{Z}_i^n | \mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n | \mathsf{x}^n)\big)$ ($i \in \mathbb{N}^*$) and infinitely many constants $\alpha_i > 0$ such that $\sum_{i=1}^{\infty} \alpha_i = 1$, we define the *infinitely mixed channel* as the wiretap channel $\big(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, p_{\mathsf{Y}^n \mathsf{Z}^n | \mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n | \mathsf{x}^n)\big)$ characterized by

$$p_{\mathsf{Y}^n \mathsf{Z}^n | \mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n | \mathsf{x}^n) = \sum_{i=1}^{\infty} \alpha_i p_{\mathsf{Y}_i^n \mathsf{Z}_i^n | \mathsf{X}^n}(\mathsf{y}^n, \mathsf{z}^n | \mathsf{x}^n)$$

Corollary 3 (Infinitely mixed channels): The secrecy capacity of an infinitely mixed wiretap channel is given by

$$C_s = \max_{\{\mathsf{V}^n, \mathsf{X}^n\}_{n=1}^{\infty}} \left(\inf_{i\alpha_i > 0} \text{p-liminf}_{n\to\infty} \frac{1}{n} \mathrm{I}(\mathsf{V}^n; \mathsf{Y}_i^n) - \sup_{j:\alpha_j>0} \text{p-limsup}_{n\to\infty} \frac{1}{n} \mathrm{I}\big(\mathsf{V}^n; \mathsf{Z}_j^n\big) \big) \right).$$

where the process $\{\mathsf{V}^n, \mathsf{X}^n\}_{n=1}^{\infty}$ is subject to

$$\mathsf{V}^n \to \mathsf{X}^n \to \mathsf{Z}_i^n \mathsf{Y}_i^n \quad \forall n \in \mathbb{N}^* \, \forall i \in \mathbb{N}^*.$$

Proof: The proof follows from steps similar to those used in [12, Theorem 3.3.3]. ∎

It is easy to show that the secrecy capacity of a compound wiretap channel is identical to that of a mixed channel. Hence, the achievable rates obtained for memoryless compounds wiretap channels [7] and memoryless wiretap channels with multiple eavesdroppers [8] follow from the general results.

We conclude this section with remarks regarding future extensions of this work. We emphasize that specializing the general secrecy capacity result of Theorem 1 to obtain achievable secure rates is in general much easier than specializing the converse and therefore obtaining the exact secrecy capacity. Nevertheless, our result can probably be specialized to obtain general achievable rates for large classes of channels, such as fading channels with non i.i.d. Gaussian noise, etc. These promising aspects will be investigated in future work.

IV. Acknowledgment

This research has been supported in part by the NSF CAREER grant CCF05-46618.

Appendix I
Proof of Lemma 1

The implications (1) $\succeq$ (2) $\succeq$ (3) and (4) $\succeq$ (5) $\succeq$ (6) follow directly from [18, Corollary p.16] and [18, Corollary p.18], and we only need to prove that (3) $\succeq$ (4).

Let $\epsilon > 0$. The average mutual information rate can be written as

$$\begin{aligned}\frac{\mathbb{I}(\mathsf{M}; \mathsf{Z}^n)}{n} &= \mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n}\right], \\ &= \mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \mathbf{1}\left\{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n) \le -\epsilon\right\}\right] \\ &\quad + \mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \mathbf{1}\left\{-\epsilon < \mathrm{I}(\mathsf{M}; \mathsf{Z}^n) \le \epsilon\right\}\right] \\ &\quad + \mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \mathbf{1}\left\{\epsilon < \mathrm{I}(\mathsf{M}; \mathsf{Z}^n) \le \log M_n\right\}\right] \\ &\quad + \mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \mathbf{1}\left\{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n) > \log M_n\right\}\right].\end{aligned}$$

Clearly, we have

$$\mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \mathbf{1}\left\{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n) \le -\epsilon\right\}\right] < 0,$$

$$\begin{aligned}&\mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \mathbf{1}\left\{-\epsilon < \mathrm{I}(\mathsf{M}; \mathsf{Z}^n) \le \epsilon\right\}\right] \\ &\qquad \le \mathbb{E}\left[\frac{|\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)|}{n} \mathbf{1}\left\{|\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)| \le \epsilon\right\}\right] \le \frac{\epsilon}{n},\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \mathbf{1}\left\{\epsilon < \mathrm{I}(\mathsf{M}; \mathsf{Z}^n) \le \log M_n\right\}\right] \\ &\qquad \le \frac{\log M_n}{n} \mathbb{P}[\epsilon < \mathrm{I}(\mathsf{M}; \mathsf{Z}^n)] = R\, \mathbb{P}[\epsilon < \mathrm{I}(\mathsf{M}; \mathsf{Z}^n)].\end{aligned}$$

To bound the fourth term, recall that, by definition, M is uniformly distributed over the set $\{1, \ldots, M_n\}$, therefore,

$$\mathrm{I}(\mathsf{M}; \mathsf{Z}^n) = \log \frac{\mathbb{P}_{\mathsf{M}|\mathsf{Z}^n}[\mathsf{M}|\mathsf{Z}^n]}{\mathbb{P}_{\mathsf{M}}[\mathsf{M}]} \le \log \frac{1}{\mathbb{P}_{\mathsf{M}}[\mathsf{M}]} = \log M_n,$$

and consequently,

$$\mathbf{1}\left\{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n) > \log M_n\right\} = 0$$

$$\text{and} \quad \mathbb{E}\left[\frac{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \mathbf{1}\left\{\mathrm{I}(\mathsf{M}; \mathsf{Z}^n) > \log M_n\right\}\right] = 0.$$

All in all, we obtain

$$\begin{aligned}\forall \epsilon > 0 \quad 0 &\le \lim_{n\to\infty} \frac{\mathbb{I}(\mathsf{M}; \mathsf{Z}^n)}{n} \\ &\le \lim_{n\to\infty} \left[\frac{\epsilon}{n} + R\, \mathbb{P}[\epsilon < \mathrm{I}(\mathsf{M}; \mathsf{Z}^n)]\right] = 0,\end{aligned}$$

which concludes the proof.

APPENDIX II
PROOF OF LEMMA 2

We prove Lemma 2 using a random coding argument that shows that the average error probability and average variational distance over an ensemble of random codes vanish when the block length gets large. The essence of the proof is similar to that of [11], although we target directly an asymptotic result.

A. Random code generation

Let M_n and L_n be two integers, and fix a distribution $p_{\mathsf{X}^n}(\mathsf{x}^n)$ on $\mathcal{X}^n$. Generate M_nL_n sequences in $\mathcal{X}^n$ according to $p_{\mathsf{X}^n}(\mathsf{x}^n)$, and label them

$$\mathsf{u}^n(i,j) \quad \text{with} \quad (i,j) \in \{1,\dots,M_n\} \times \{1,\dots,L_n\}.$$

The random variable representing the randomly generated sequence $\mathsf{u}(i,j)$ is denoted by $\mathsf{X}^n_{i,j}$.

B. Encoding and decoding procedures

Let $\mathcal{B}^n_i = \{\mathsf{u}^n(i,j) : j \in \{1,\dots,L_n\}\}$. To send a message index $i \in \{1,\dots,M_n\}$, the emitter chooses a codeword $\mathsf{u}^n(i,j)$ uniformly at random in $\mathcal{B}^n_i$. Consequently, we have

$$\forall i \in \{1,\dots,M_n\}\, \forall j \in \{1,\dots,L_n\} \qquad p_{\mathsf{u}^n(i,1),\dots\mathsf{u}^n(i,L_n)}(\mathsf{u}^n(i,j)) = \frac{1}{L_n}.$$

For any $\gamma > 0$, we define the decoder of the legitimate receiver as follows. Define the set $T_n \in \mathcal{X}^n \times \mathcal{Y}^n$ as

$$T_n = \left\{ \mathsf{x}^n, \mathsf{y}^n : \frac{1}{n} \log \mathrm{I}(\mathsf{x}^n; \mathsf{y}^n) \geq \frac{1}{n} \log M_n L_n + \gamma \right\}.$$

Upon observing a channel output y^n, the receiver declares $\mathsf{u}^n(i,j)$ as the decoded message if $\mathsf{u}^n(i,j)$ is the unique codeword such that $(\mathsf{u}^n(i,j), \mathsf{y}^n) \in T_n$. Otherwise, it declares an error.

C. Analysis of probability of error

Define the event

$$E(i,j) = \{(\mathsf{u}^n(i,j), \mathsf{y}^n) \in T_n \text{ given } \mathsf{u}^n(1,1) \text{ sent}\},$$

and denote its complimentary event by $E^c(i,j)$. By symmetry of the random code construction, the average error probability $\bar{\epsilon}_n = \mathbb{E}_{\mathsf{X}^n_{1,1},\dots,\mathsf{X}^n_{M_n,L_n}}[\epsilon_n]$ is equal to the error probability given the transmission of any specific codeword. Without loss of generality, we assume that the codeword $\mathsf{u}^n(1,1)$ was sent; therefore, we have

$$\begin{aligned}
\bar{\epsilon}_n &= \mathbb{E}_{\mathsf{X}^n_{1,1},\dots,\mathsf{X}^n_{M_n,L_n}}[\epsilon_n] \\
&= \mathbb{P}_{\mathsf{Y}^n\mathsf{X}^n}\left[E^c(1,1) \cup \bigcup_{(i,j)\neq(1,1)} E(i,j) \right], \\
&\leq \mathbb{P}_{\mathsf{Y}^n\mathsf{X}^n}[E^c(1,1)] + \sum_{(i,j)\neq(1,1)} \mathbb{P}_{\mathsf{Y}^n\mathsf{X}^n}[E(i,j)].
\end{aligned}$$

Noticing that for all $(i,j) \neq (1,1)$

$$\begin{aligned}
\mathbb{P}_{\mathsf{Y}^n\mathsf{X}^n}[E(i,j)] &= \sum_{(\mathsf{x}^n,\mathsf{y}^n)\in T_n} p_{\mathsf{X}^n}(\mathsf{x}^n)\, p_{\mathsf{Y}^n}(\mathsf{y}^n), \\
&\leq \sum_{(\mathsf{x}^n,\mathsf{y}^n)\in T_n} p_{\mathsf{X}^n}(\mathsf{x}^n)\, p_{\mathsf{Y}^n|\mathsf{X}^n}(\mathsf{y}^n|\mathsf{x}^n)\, \frac{2^{-n\gamma}}{M_nL_n}, \\
&\leq \frac{2^{-\gamma}}{M_nL_n},
\end{aligned}$$

we obtain

$$\bar{\epsilon}_n \leq \mathbb{P}_{\mathsf{Y}^n\mathsf{X}^n}[E^c(1,1)] + 2^{-n\gamma}.$$

By definition of T_n, if we choose

$$\frac{1}{n} \log M_nL_n \leq \text{p-liminf}\, \frac{1}{n} \mathrm{I}(\mathsf{X}^n; \mathsf{Y}^n) - 2\gamma, \tag{12}$$

then $\lim_{n\to\infty} \mathbb{P}_{\mathsf{Y}^n\mathsf{X}^n}[E^c(1,1)] = 0$ and $\lim_{n\to\infty} \bar{\epsilon}_n = 0$.

D. Analysis of leaked information

Let $\tilde{\mathsf{Z}}^n$ denote the random variable obtained at the output of the channel if X^n, distributed according to $p_{\mathsf{X}^n}(\mathsf{x}^n)$ independently of $\mathsf{X}^n_{i,j}$ for all i,j, is present at the input. We can rewrite the variational distance in (2) as

$$\begin{aligned}
\delta_n &= d\left(p_{\mathsf{M}\mathsf{Z}^n}, p_{\mathsf{M}} p_{\mathsf{Z}^n}\right) \\
&= \sum_{i=1}^{M_n} \sum_{\mathsf{z}^n} \left| p_{\mathsf{M}\mathsf{Z}^n}(i,\mathsf{z}^n) - p_{\mathsf{M}}(i)\, p_{\mathsf{Z}^n}(\mathsf{z}^n) \right| \\
&= \sum_{i=1}^{M_n} \sum_{\mathsf{z}^n} \frac{1}{M_n} \left| p_{\mathsf{Z}^n|\mathsf{M}}(\mathsf{z}^n|i) - p_{\mathsf{Z}^n}(\mathsf{z}^n) \right| \\
&\leq \sum_{i=1}^{M_n} \frac{1}{M_n} \sum_{\mathsf{z}^n} \left| p_{\mathsf{Z}^n|\mathsf{M}}(\mathsf{z}^n|i) - p_{\tilde{\mathsf{Z}}^n}(\mathsf{z}^n) \right| \\
&\qquad\qquad + \left| p_{\tilde{\mathsf{Z}}^n}(\mathsf{z}^n) - p_{\mathsf{Z}^n}(\mathsf{z}^n) \right| \\
&\leq 2 \sum_{i=1}^{M_n} \frac{1}{M_n} \sum_{\mathsf{z}^n} \left| p_{\mathsf{Z}^n|\mathsf{M}}(\mathsf{z}^n|i) - p_{\tilde{\mathsf{Z}}^n}(\mathsf{z}^n) \right| \\
&= 2 \sum_{i=1}^{M_n} \frac{1}{M_n} d\left(p_{\mathsf{Z}^n|i}, p_{\tilde{\mathsf{Z}}^n}\right).
\end{aligned}$$

Therefore, by symmetry of the random argument, we have

$$\begin{aligned}
\bar{\delta}_n &= \mathbb{E}_{\mathsf{X}^n_{1,1},\dots,\mathsf{X}^n_{M_n,L_n}}[\delta_n] \\
&\leq 2 \sum_{i=1}^{M_n} \frac{1}{M_n} \mathbb{E}_{\mathsf{X}^n_{1,1},\dots,\mathsf{X}^n_{M_n,L_n}}\left[d\left(p_{\mathsf{Z}^n|i}, p_{\tilde{\mathsf{Z}}^n}\right) \right] \\
&= 2\mathbb{E}_{\mathsf{X}^n_{1,1},\dots,\mathsf{X}^n_{M_n,L_n}}\left[d\left(p_{\mathsf{Z}^n|1}, p_{\tilde{\mathsf{Z}}^n}\right) \right]
\end{aligned} \tag{13}$$

Now, recall that according to the encoding procedure defined earlier, we have

$$p_{\mathsf{Z}^n|1}(\mathsf{z}^n|1) = \frac{1}{L_n} \sum_{j=1}^{L_n} p_{\mathsf{Z}^n|\mathsf{X}^n}\left(\mathsf{z}^n|\mathsf{X}^n_{1,j}\right). \tag{14}$$

From (13) and (14), we see that the secrecy criterion can be satisfied if all subcodes $\mathcal{B}^n_i$ are *resolvability* codes [12, Chapter 6]. To simplify notation, we denote the random variable $Z^n|1$ simply by $\hat{Z}^n$. The following results hold.

Lemma 6 ([12], Lemma 6.3.1): For all $\tau > 0$, it holds that

$$d\left(p_{\hat{Z}^n}, p_{\tilde{Z}^n}\right) \leq 2\tau + 2\mathbb{P}_{\hat{Z}^n}\left[\log \frac{p_{\hat{Z}^n}\left(\hat{Z}^n\right)}{p_{\tilde{Z}^n}\left(\hat{Z}^n\right)} > \tau\right].$$

Lemma 7 ([12], adapted from proof of Theorem 6.3.1): For all $\gamma > 0$, it holds that

$$\begin{aligned}&\mathbb{E}\left[\mathbb{P}_{\hat{Z}^n}\left[\log \frac{p_{\hat{Z}^n}\left(\hat{Z}^n\right)}{p_{\tilde{Z}^n}\left(\hat{Z}^n\right)} > \tau\right]\right] \\ &\leq \mathbb{P}\left[\frac{1}{n}\log \frac{p_{\tilde{Z}^n|\mathsf{X}^n}\left(\tilde{Z}^n|\mathsf{X}^n\right)}{p_{\tilde{Z}^n}\left(\tilde{Z}^n\right)} > \frac{\log L_n}{n} - \gamma\right] + e^{-n\gamma}.\end{aligned}$$

Clearly, if we choose

$$\frac{1}{n}\log L_n \geq \operatorname*{p\text{-}limsup}_{n\to\infty} \frac{1}{n}\mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n) + 2\gamma, \tag{15}$$

then

$$\lim_{n\to\infty} \bar{\delta}_n = 0.$$

E. Expurgation of ensemble

For any $n \in \mathbb{N}^*$, the Markov inequality ensures

$$\begin{aligned}\mathbb{P}\left[\epsilon_n(\mathsf{X}^n_{1,1},\dots\mathsf{X}^n_{M_n,L_n}) \geq 3\bar{\epsilon}_n\right] &\leq \frac{1}{3},\\ \mathbb{P}\left[\delta_n(\mathsf{X}^n_{1,1},\dots\mathsf{X}^n_{M_n,L_n}) \geq 3\bar{\delta}_n\right] &\leq \frac{1}{3};\end{aligned}$$

moreover, if

$$\begin{aligned}\frac{1}{n}\log M_n \leq \operatorname*{p\text{-}liminf}_{n\to\infty} &\frac{1}{n}\log \mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n)\\ &- \operatorname*{p\text{-}limsup}_{n\to\infty} \frac{1}{n}\log \mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n) - 4\gamma,\end{aligned}$$

we can find L_n such that Equations (12) and (15) are satisfied; therefore, for any $\gamma > 0$ there exists a sequence of codes $(n, M_n, \epsilon_n, \delta_n)$ such that

$$\begin{aligned}\liminf_{n\to\infty} r_n = \operatorname*{p\text{-}liminf}_{n\to\infty} &\frac{1}{n}\log \mathrm{I}(\mathsf{X}^n;\mathsf{Y}^n)\\ &- \operatorname*{p\text{-}limsup}_{n\to\infty} \frac{1}{n}\log \mathrm{I}(\mathsf{X}^n;\mathsf{Z}^n) - 4\gamma,\\ \lim_{n\to\infty} \epsilon_n \leq \lim_{n\to\infty} \bar{\epsilon}_n = 0,&\\ \lim_{n\to\infty} \delta_n \leq \lim_{n\to\infty} \bar{\delta}_n = 0,&\end{aligned}$$

which concludes the proof.

Although the random coding scheme exploits a binning structure similar to that of [1], [2], we highlight that the proof differs fundamentally in the secrecy criterion (variational distance) and analysis technique (channel resolvability). In [1], [2], the random coding argument is used to prove the existence of a code with a certain structure, and the equivocation is calculated on the specific code; here, the random coding argument shows *directly* the existence of a code satisfying both reliably and security.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1367, October 1975.

[2] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[3] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, pp. 656–715, 1948.

[4] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *Proc. IEEE International Symposium on Information Theory*, Seattle, USA, 2006, pp. 356–360.

[5] P. K. Gopala, L. Lai, and H. El-Gamal, "On the secrecy capacity of fading channels," in *Proc. of IEEE International Symposium on Information Theory*, Nice, France, June 2007, see also cs.IT/0610103.

[6] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[7] Y. Liang, G. Kramer, H. V. Poor, and S. S. (Shitz), "Compound wiretap channels," in *Proc. of the Allerton Conference on Communications, Control, and Computing*, Allerton, IL, September 2007, pp. 136–143.

[8] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, 2008.

[9] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, March 2008.

[10] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, 2008.

[11] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, April 2006.

[12] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer, 2002.

[13] D. Kobayashi, H. Yamamoto, and T. Ogawa, "How to attain the ordinary channel capacity securely in wiretap channels," in *Proc. IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security*, 2005, pp. 13–18.

[14] H. Koga and N. Sato, "On an upper bound of the secrecy capacity for a general wiretap channel," in *Proc. International Symposium on Information Theory ISIT 2005*, Adelaide, Australia, September 2005, pp. 1641–1645.

[15] C. E. Shannon, "Certain results in coding theory for noisy channels," *Inform. Contr.*, vol. 1, pp. 6–25, September 1957.

[16] R. M. Fano, *Transmission of Information: A Statistical Theory of Communication*. New-York: John Wiley & Sons, Inc., 1961.

[17] R. G. Gallager, *Information Theory and Reliable Communication*. Wyley, 1968.

[18] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. Holden Day, 1964.

[19] U. M. Maurer and S. Wolf, "Information-Theoretic Key Agreement: From Weak to Strong Secrecy for Free," in *Advances in Cryptology - Eurocrypt 2000*, Lecture Notes in Computer Science. B. Preneel, 2000, p. 351.

[20] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian Wire-Tap Channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, July 1978.