Reliability and Security Functions of the Wiretap Channel under Cost Constraint

Te Sun HAN,[†] Hiroyuki ENDO,[‡] Masahide SASAKI[§]

September 18, 2013

[†]T. S. Han is with the Quantum ICT Laboratory, National Institute of Information and Communications Technology (NICT), Nukui-kitamachi 4-2-1, Koganei, Tokyo,184-8795, Japan (email: han@is.uec.ac.jp, han@nict.go.jp)

[‡]H. Endo is with the Department of Applied Physics, Waseda University, Okubo 3-4-1, Shinjuku, Tokyo, Japan, and is also a collaborating research fellow of the Quantum ICT Laboratory, NICT (email: h-endo-1212@ruri.waseda.jp, h-endo@nict.go.jp)

[§]M. Sasaki is with the Quantum ICT Laboratory, NICT, Nukui-kitamachi 4-2-1, Koganei, Tokyo,184-8795, Japan (email: psasaki@nict.go.jp)

Abstract: The wiretap channel has been devised and studied first by Wyner, and subsequently extended to the case with non-degraded general wiretap channels by Csiszár and Körner. Focusing mainly on the Poisson wiretap channel with cost constraint, we newly introduce the notion of reliability and security functions as a fundamental tool to analyze and/or design the performance of an efficient wiretap channel system. Compact formulae for those functions are explicitly given for stationary memoryless wiretap channels. It is also demonstrated that, based on such a *pair* of reliability and security functions, we can control the tradeoff between reliability and security (usually conflicting), both with exponentially decreasing rates as block length n becomes large. Two ways to do so are given on the basis of concatenation and rate exchange. In this framework, the notion of the δ secrecy capacity is defined and shown to attain the strongest security standard among others. The maximized vs. averaged security measures is also discussed.

Index terms: reliability function, security function, security measures, Poisson wiretapp channel, cost constraint, Gaussian wiretap channel, binary symmetric wiretap channel, tradeoff between reliability and security, concatenation, rate exchange

1 Introduction

Contemporary cryptographic protocols are based on algorithmic security, assuming that certain mathematical problems are practically impossible to solve with current computer resources and well-known attacks. Since they can be implemented by computer algorithms as software, they can be used ubiquitously in network systems, not restricted by phycisal devices and communications channels. In fact, all cryptographic protocols widely used in the Internet such as RSA and AES are implemmented in the network layer and the transport layer, which correspond to the third and forth layer, respectively, according to the OSI network model. This kind of algorithmic techniques are, however, not provably secure, that is, in-principle, vulnerable to off-line attacks that can occur long after a secret message has been sent and which exploit unforeseen mathematical insights.

A well known scheme of provably secure protocol is Vernam's one-time pad (OTP) [1] whose perfect secrecy was proven by Shannon [2]. In this protocol, a truly random key whose length is the same as that of a plain text should be shared between the legitimate sender and the legitimate receiver for encryption and decryption simply by exclusive OR with the plain text, and is used only once. Quantum key distribution (QKD) [3] can provide a means to share such a truly random key on demand through a physical channel of photon transmission. Its security can be proven against the eavesdropper with unbounded abilities. Combination of QKD and OTP constitutes quantum cryptography. It is an unconditionally secure cryptography implemented in the physical layer, which corresponds to the first layer of network.

Actually the provable security is more or less based upon the analysis of the physical properties of the communications channels. In other words, given a channel model, security proofs are made in an information theoretic manner by showing the existence of channel codings or error correcting codes that can effectively establish the statistical independence between the legitimate users and the eavesdropper. The theoretical basis was laid by Wyner [4] and later by Csiszár and Körner [4, 5]. These kinds of schemes are therefore referred to as the physical layer cryptography, and its provable security is often called the information-theoretic security.

Quantum cryptography is an ultimate manifestation of the physical layer cryptography with the strongest security. It has become in practical use. However, its speed and distance are still limited, hindering wider adoption. Actually, the extreme assumption on an eavesdropper who can have unbounded ability allowed by physics, enables one to prove the unconditional security in a clear manner. It instead imposes the stringent conditions on the implementation, not only limiting the transmission rate and distance but also narrowing an operation margin. Quantum repeater is known to be a means to extend the distance, however, it further narrows the operation margin. While quantum cryptography in the present form can be an option for high-end applications, another kind of scheme not only with provable security but also with improved usability should be pursued, even by restricting the range of security assumptions into a compromised one.

From this point of view, the traditional apporach by Wyner [4] and Csiszár and Körner [5], based on the wiretap channel, provides a strong impetus to find a new scheme of the physical layer cryptography in a good balance of usability and security. In the scheme of wiretap channel, an assumption is made on a physical channel that channel characteristics for an eavesdropper (Eve) is worse than that for the legitimate receiver (Bob), and the existence of channel coding is proven, which ensures a transmission with asymptotically vanishing error to Bob at a finite rate, while making the leaked information to Eve infinitesimally small. In [4], Wyner formulated the tradeoff between the transmission rate for Bob and the leaked information to Eve. The maximum attainable rate for Bob is referred to as the secrecy capacity. Since then there have been extensive studies on various kinds of wiretap channels, which are nicely summarized in Laourine and Wagner [6] along with the secrecy capacity formula for the Poisson wiretap channel without cost constraint.

However, interesting questions are left open, including precise formulation of asymptotic characteristics on how fast the decoding error for Bob and the leaked information to Eve decreases as exponential functions of the code length when the cost (available transmission energy, bandwidth, and so on) constraint is imposed. It is actually a more involved task than the derivation of the reliability function only in the absence of Eve. Security measures should also be carefully defined for required tasks and their levels of security. For example, usual averaged quantities against Eve, such as the mutual information between the legitimate sender Alice and the eavesdropper Eve, are sometimes insufficient for ensuring a required level of security, because such measures cannot exclude less likely but serious security incidents. Hayashi [7], [10] have derived several error exponents for averaged quantities in the case where no cost constraint is considered.

In this paper we establish formulas to characterize the reliability function for Bob and the exponents of various kinds of security measures against Eve *with* *cost constraints*, by introducing the notion of the security function. We then fully characterize the tradeoff between the transmission rate-reliability and the security for several important wiretap channel models.

The paper is organized as follows. In Section 2, the definitions of the wiretap channel and related notions (including the information-spectrum ones) are introduced along with various kinds of security measures. In Section 3, we review the known results on the exponents of decreasing rates of the decoding error of Bob and of the divergence distance for Eve, providing more compact proofs, in the case where there is nothing to do with cost constraint. In particular, we introduce the notion of the security function for the exponent of the divergence distance. The bases are thus prepared for more quantitative analysis/design of the reliability-security tradeoff. Two ways for tradeoff are demonstrated: one is by concatenation and the other by rate exchange. Also, superiority of the maximum security criterion to the average security criterion is discussed. In Section 4, we extend the results presented in the preceding section to the case where cost constraint is imposed on Alice's input into the channel. The δ-secrecy capacity formula (with the strongest secrecy) is given under cost constraint. In Section 5, we give a fundamental formula to simultaneously evaluate a *pair* of reliability and security functions under cost constraint for a general channel, which is then particularized to stationary and memoryless wiretap channels. This is one of the main results in this paper. We also present their numerical examples to see how the reliability and security performances vary depending on the channel and cost parameters. In Section 6, the formula of the δ-secrecy capacity is applied to the Poisson wiretap channel with cost constraint, which is a practical model for free-space Laser communication with a photon counter. In Section 7, the reliability function and the security function are derived for the Poisson wiretap channel with cost constraint. The tradeoff between reliability and security is discussed with numerical results. In Section 8, the formula of the δ-secrecy capacity is applied to the Gaussian wiretap channel with cost constraint. This channel model covers a wide variety of wireless communications including RF and Laser signals. In Section 9, the reliability function and the security function are derived for the Gaussian wiretap channel with cost constraint. Two kinds of reliability and security functions are introduced and compared with each other. In Section 10, we investigate the effects of channel concatenation with an auxiliary channel for the Poisson wiretap channel. In Section 11, we conclude the paper.

2 Preliminaries and basic concepts

In this section we give the definition of the wiretap channel. There are several levels and ways to specify the superiority of the legitimate users, Alice and Bob, to the eavesdropper, Eve, such as physically degraded for Eve, less noisy for Bob, (statistically) degraded for Eve, and more capable for Bob. In this paper, we are interested mainly in the last criterion because the other ones imply the last one (cf. Csiszár and Körner [14]).

We introduce here necessary notions and notations to quantify the reliability and the security of this kind of wiretap channel model. In particular, we define several kinds of security metrics, including the strongest criterion based on the divergence distance with reference to a target output distribution, while the notion of concatenation of channels is also introduced to construct a possible way to control tradeoff between the reliability and the security.

A. Wiretap channel

Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be arbitrary alphabets (not necessarily finite), where $\mathcal{X}$ is called an *input alphabet*, and $\mathcal{Y}, \mathcal{Z}$ are called *output alphabets.* A general *wiretap channel* consists of two general channels, i.e., $W_B^n : \mathcal{X}^n \to \mathcal{Y}^n$ (from Alice for Bob) and $W_E^n : \mathcal{X}^n \to \mathcal{Z}^n$ (from Alice against Eve), where $W_B^n(\mathbf{y}|\mathbf{x})$, $W_E^n(\mathbf{z}|\mathbf{x})$ are the conditional probabilities of $\mathbf{y} \in \mathcal{Y}^n, \mathbf{z} \in \mathcal{Z}^n$ given $\mathbf{x} \in \mathcal{X}^n$ (of block length n), respectively. Alice wants to communicate with Bob as reliably as possible but as secretly as possible against Eve. We let (W_B^n, W_E^n) indicate such a wiretap channel.

B. Coding reliability and security measure

Let $\mathcal{M}_n \equiv \{1, 2, \cdots, M_n\}$ be a message set, and let $\varphi_n : \mathcal{M}_n \to \mathcal{X}^n$ be a *stochastic* encoder for Alice; and $\psi_n^B : \mathcal{Y}^n \to \mathcal{M}_n$ a decoder for Bob.

The error probability ϵ_n^B (*measure of reliability*) via channel W_B^n for Bob is defined to be

$$\epsilon_n^B \equiv \frac{1}{M_n} \sum_{i \in \mathcal{M}_n} \Pr\left\{\psi_n^B(\varphi_n(i)) \neq i\right\}, \tag{2.1}$$

whereas the divergence distance (*measure 1 of security*) δ_n^E and the variational distance (*measure 2 of security*) ∂_n^E via channel W_E^n against Eve are defined to be

$$\delta_n^E \equiv \frac{1}{M_n} \sum_{i \in \mathcal{M}_n} D(P_n^{(i)} || \pi_n), \tag{2.2}$$

$$\partial_n^E \equiv \frac{1}{M_n} \sum_{i \in \mathcal{M}_n} d(P_n^{(i)}, \pi_n) \tag{2.3}$$

with

$$D(P_1||P_2) = \sum_{u \in \mathcal{U}} P_1(u) \log \frac{P_1(u)}{P_2(u)},$$

$$d(P_1, P_2) = \sum_{u \in \mathcal{U}} |P_1(u) - P_2(u)|;$$

where $P_n^{(i)}$ denotes the output probability distribution on $\mathcal{Z}^n$ via channel W_E^n due to the input $\varphi_n(i)$, and π_n is the (*target*) output probability distribution on $\mathcal{Z}^n$ via channel W_E^n due to an arbitrarily prescribed input distribution on $\mathcal{X}^n$. In this paper the logarithm is taken to the natural base e.

With these two typical measures of security, we can define two kinds of criteria for *achievability*:

$$\epsilon_n^B \to 0,\ \delta_n^E \to 0 \quad \text{as } n \to \infty, \tag{2.4}$$

$$\epsilon_n^B \to 0,\ \partial_n^E \to 0 \quad \text{as } n \to \infty. \tag{2.5}$$

We say that a rate R is δ-*achievable* if there exists a pair (φ_n, ψ_n^B) of encoder and decoder satisfying criterion (2.4) and

$$\liminf_{n \to \infty} \frac{1}{n} \log M_n \geq R. \tag{2.6}$$

Similarly, we say that a rate R is ∂-*achievable* if there exists a pair (φ_n, ψ_n^B) of encoder and decoder satisfying criterion (2.5) and (2.6). It should be noted here that criterion (2.4) implies criterion (2.5), owing to Pinsker inequality [15]:

$$\left(\partial_n^E\right)^2 \leq 2\delta_n^E,$$

which means that criterion (2.4) is stronger than criterion (2.5).

Which criterion of (2.4) and (2.5) is preferred depends on the context of the problem in consideration. On the other hand, some people (e.g., Csiszár [12], Hayashi [7]) have used, instead of measure (2.2), the *mutual information*:

$$I_n^E \equiv \frac{1}{M_n} \sum_{i \in \mathcal{M}_n} D(P_n^{(i)}||P_n), \quad P_n = \frac{1}{M_n} \sum_{i \in \mathcal{M}_n} P_n^{(i)}. \tag{2.7}$$

With this measure (*measure 3 of security*), we may consider one more criterion for achievability (called i-achievability):

$$\epsilon_n^B \to 0,\ I_n^E \to 0 \quad \text{as } n \to \infty. \tag{2.8}$$

However, it is shown that the inequality $I_n^E \le \delta_n^E$ holds for any π_n (cf. [7]). Thus, δ_n^E is a stronger measure than I_n^E. Moreover, since

$$d_n^E \equiv \frac{1}{M_n}\sum_{i\in\mathcal{M}_n} d(P_n^{(i)}, P_n) \le \frac{2}{M_n}\sum_{i\in\mathcal{M}_n} d(P_n^{(i)}, \pi_n) = 2\partial_n^E$$

always holds by virtue of the triangle axiom of the variational distance, ∂_n^E is stronger than d_n^E (*measure 4 of security*: cf. [12]), so that criterion (2.5) is stronger than the d-achievability:

$$\epsilon_n^B \to 0,\ d_n^E \to 0 \quad \text{as } n \to \infty. \tag{2.9}$$

Furthermore, one may sometimes prefer to consider the following achievability (called w-achievability):

$$\epsilon_n^B \to 0,\ \frac{1}{n}I_n^E \to 0 \quad \text{as } n \to \infty, \tag{2.10}$$

which is nothing but the so-called weak secrecy (*measure 5 of security*). Indeed, this is the *weakest* criterion among others; its illustrating example will appear in Remark 6.1, while criterion (2.4) is the *strongest* one. Fig.1 shows the implication scheme among these five measures of security.

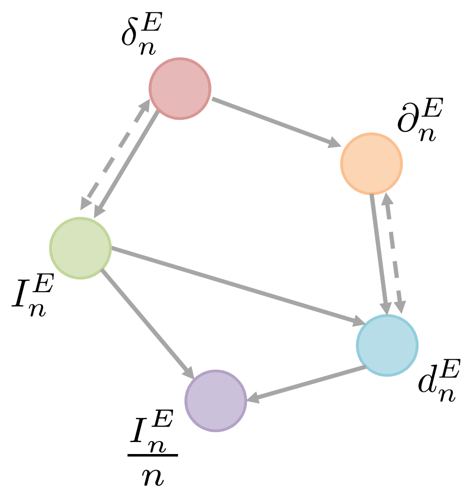

Figure 1: The implication scheme: The arrow $\alpha \longrightarrow \beta$ means that α is stronger than β; $\alpha \dashleftarrow\dashrightarrow \beta$ means that α coincides with β when $\pi_n = P_n$, where $I_n^E \to d_n^E$ is due to [15] and $d_n^E \to \frac{1}{n}I_n^E$ is due to [11]. In the finite alphabet case, exponential decay of d_n^E (with increasing n) implies that of I_n^E (cf. [12]).

We notice that all of ϵ_n^B, δ_n^E, ∂_n^E, d_n^E and I_n^E, $\frac{1}{n}I_n^E$ defined here are the measures *averaged* over the message set $\mathcal{M}_n$ with the uniform distribution. On the other hand, we can consider also the criteria *maximized* over the message set $\mathcal{M}_n$, which will be discussed later in Remark 3.8.

The secrecy capacities δ-C_s and ∂-C_s between Alice and Bob are defined to be the supremum of all δ-achievable rates and that of all ∂-achievable rates, respectively. Similarly, the secrecy capacity d-C_s with d-achievability, the secrecy capacity i-C_s with i-achievability as well as the secrecy capacity w-C_s with w-achievability can also be defined.

C. Concatenation

Let $\mathcal{V}$ be an arbitrary alphabet (not necessarily finite) and let V^n be an arbitrary auxiliary random variable with values in $\mathcal{V}^n$ such that $V^n \to X^n \to Y^n Z^n$ forms a Markov chain in this order, where X^n is an input variable for the wiretap channel (W_B^n, W_E^n); and Y^n, Z^n are the output variables of channels W_B^n, W_E^n due to the input X^n, respectively. Setting $\mathbf{V} = \{V^n\}_{n=1}^\infty, \mathbf{X} = \{X^n\}_{n=1}^\infty, \mathbf{Y} = \{Y^n\}_{n=1}^\infty, \mathbf{Z} = \{Z^n\}_{n=1}^\infty$, we denote this Markov chain by $\mathbf{V} \to \mathbf{X} \to \mathbf{YZ}$.

Definition 2.1 Given a general channel $W^n : \mathcal{X}^n \to \mathcal{Y}^n$, we define its *concatenated* channel $W^{n+} : \mathcal{V}^n \to \mathcal{Y}^n$ so that *

$$W^{n+}(\mathbf{y}|\mathbf{v}) = \sum_{\mathbf{x}\in\mathcal{X}^n} W^n(\mathbf{y}|\mathbf{x}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}), \tag{2.11}$$

where $P_{X^n|V^n} : \mathcal{V}^n \to \mathcal{X}^n$ is an arbitrary auxiliary channel. In particular, we say that a pair (W_B^{n+}, W_E^{n+}) is a *concatenation* of the wiretap channel (W_B^n, W_E^n), if

$$W_B^{n+}(\mathbf{y}|\mathbf{v}) = \sum_{\mathbf{x}\in\mathcal{X}^n} W_B^n(\mathbf{y}|\mathbf{x}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) \tag{2.12}$$

$$W_E^{n+}(\mathbf{z}|\mathbf{v}) = \sum_{\mathbf{x}\in\mathcal{X}^n} W_E^n(\mathbf{z}|\mathbf{x}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}). \tag{2.13}$$

with the auxiliary channel $P_{X^n|V^n}$. □

*We use the convention that, given random variables S and T, $P_S(\cdot)$ and $P_{S|T}(\cdot|\cdot)$ denote the probability distribution of S, and the conditional probability distribution of S given T, respectively

Definition 2.2 A wiretap channel (W_B^n, W_E^n) is said to be stationary and memoryless if, with some channels $W_B : \mathcal{X} \to \mathcal{Y}, W_E : \mathcal{X} \to \mathcal{Z}$, it holds that

$$W_B^n(\mathbf{y}|\mathbf{x}) = \prod_{k=1}^{n} W_B(y_k|x_k), \quad W_E^n(\mathbf{z}|\mathbf{x}) = \prod_{k=1}^{n} W_E(z_k|x_k), \tag{2.14}$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_n)$, $\mathbf{y} = (y_1, y_2, \cdots, y_n)$, $\mathbf{z} = (z_1, z_2, \cdots, z_n)$. □

Then, the secrecy capacity d-C_s for a general wiretap channel as well as for a stationary memoryless wiretap channel is given as follows, where formula (2.15) is due to Hayashi [10], and formula (2.16) is due to Csiszár [12].

Theorem 2.1 A general formula[†]

$$\begin{aligned} &\text{d-}C_s \\ &= \sup_{\mathbf{VX}} \left(\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^{n+}(Y^n|V^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^{n+}(Z^n|V^n)}{P_{Z^n}(Z^n)} \right) \end{aligned} \tag{2.15}$$

holds. Furthermore, if (W_B^n, W_E^n) is stationary and memoryless, then (2.15) reduces to

$$\text{d-}C_s = \sup_{VX}(I(V;Y) - I(V;Z)), \tag{2.16}$$

where $I(S;T)$ denotes the mutual information between S and T (cf. Cover and Thomas [9]); and X (taking values in $\mathcal{X}$) is any channel input; Y, Z (taking values in $\mathcal{Y}, \mathcal{Z}$) are the outputs via channels W_B, W_E due to X, respectively, whereas V is an arbitrary auxiliary random variable such that $V \to X \to YZ$ forms a Markov chain. □

D. More capable channel

Definition 2.3 If, for any $\mathbf{V} = \{V^n\}_{n=1}^{\infty}$ and $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$, it holds that

$$\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{W_B^{n+}(Y^n|V^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{W_E^{n+}(Z^n|V^n)} \geq 0, \tag{2.17}$$

then we say that channel W_B^n is *more capable* than channel W_E^n; and moreover, for simplicity, we also say that a wiretap channel (W_B^n, W_E^n) is more capable, with abuse of notation. □

[†]For a sequence $\{S_n\}_{n=1}^{\infty}$ of real-valued random variables, we define $\text{p-}\liminf_{n\to\infty} S_n \equiv \sup\{\alpha | \lim_{n\to\infty} \Pr\{S_n < \alpha\} = 0\}$ and $\text{p-}\limsup_{n\to\infty} S_n \equiv \inf\{\beta | \lim_{n\to\infty} \Pr\{S_n > \beta\} = 0\}$ (cf. Han [19]).

Remark 2.1 If a wiretap channel (W_B^n, W_E^n) is stationary and memoryless, and more capable, then condition (2.17) implies that

$$I(X;Y|V) \geq I(X;Z|V) \tag{2.18}$$

for any VX such that Y, Z are the outputs of channels $W_B : \mathcal{X} \to \mathcal{Y}, W_E : \mathcal{X} \to \mathcal{Z}$ due to input X, respectively ($V \to X \to YZ$). This is seen just by taking an i.i.d. $V^n X^n$ in (2.17) and invoking the weak law of large numbers. Moreover, noticing that (2.18) is written as

$$\sum_{v\in\mathcal{V}} I(X;Y|V=v)P_V(v) \geq \sum_{v\in\mathcal{V}} I(X;Z|V=v)P_V(v),$$

we see that (2.18) is equivalent to

$$I(X;Y) \geq I(X;Z) \quad \text{for all } X,$$

which is the original definition of "more capable" by Csiszár and Körner [13]. □

ThusDwe have the following theorem (also cf. Koga and Sato [8]):

Theorem 2.2 If a wiretap channel (W_B^n, W_E^n) is more capable, formula (2.15) is simplified as follows (the non-concatenated channel):

$$\begin{aligned} &\text{d-}C_s \\ &\quad = \sup_{\mathbf{X}} \left(\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n}(Z^n)} \right). \end{aligned} \tag{2.19}$$

Furthermore, if (W_B^n, W_E^n) is stationary and memoryless, then (2.19) reduces to

$$\text{d-}C_s = \sup_X (I(X;Y) - I(X;Z)). \tag{2.20}$$

Remark 2.2 A strengthening of (2.16) and (2.20) is given later in Theorem 3.5. □

Proof of Theorem 2.2:

Keeping the Markov property $\mathbf{V} \to \mathbf{X} \to \mathbf{YZ}$ in mind, we have

$$\begin{aligned} &\text{d-}C_s \\ &\quad = \sup_{\mathbf{VX}} \left(\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^{n+}(Y^n|V^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^{n+}(Z^n|V^n)}{P_{Z^n}(Z^n)} \right) \end{aligned}$$

$$
\begin{aligned}
&= \sup_{\mathbf{VX}} \Bigg(\text{p-}\liminf_{n\to\infty} \left[\frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n}(Y^n)} - \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n|V^n}(Y^n|V^n)} \right] \\
&\quad - \text{p-}\limsup_{n\to\infty} \left[\frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n}(Z^n)} - \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n|V^n}(Z^n|V^n)} \right] \Bigg) \\
&\le \sup_{\mathbf{VX}} \Bigg(\left[\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n}(Y^n)} - \text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n|V^n}(Y^n|V^n)} \right] \\
&\quad - \left[\text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n}(Z^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n|V^n}(Z^n|V^n)} \right] \Bigg) \\
&= \sup_{\mathbf{VX}} \Bigg(\left[\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n}(Z^n)} \right] \\
&\quad - \left[\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n|V^n}(Y^n|V^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n|V^n}(Z^n|V^n)} \right] \Bigg) \\
&\le \sup_{\mathbf{X}} \left[\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n}(Z^n)} \right],
\end{aligned} \tag{2.21}
$$

where the last inequality follows from condition (2.17). Finally, notice that inequality in (2.21) reduces to equality by choosing $\mathbf{V}$ as $V^n = X^n$, thereby completing the proof of (2.19). □

Next, let us give here an information spectrum proof of (2.20) relying on the general formula (2.19). Han [19] has shown that ‡

$$
\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n}(Y^n)} \le \liminf_{n\to\infty} \frac{1}{n} I(X^n;Y^n), \tag{2.22}
$$

$$
\text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n}(Z^n)} \ge \limsup_{n\to\infty} \frac{1}{n} I(X^n;Z^n). \tag{2.23}
$$

Hence, we have

$$
\begin{aligned}
&\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n}(Z^n)} \\
&\quad \le \liminf_{n\to\infty} \left(\frac{1}{n} I(X^n;Y^n) - \frac{1}{n} I(X^n;Z^n) \right) \\
&\quad \le \sup_X (I(X;Y) - I(X;Z)),
\end{aligned} \tag{2.24}
$$

where the last inequality follows in parallel with the argument employed in El Gamal and Kim [20, p.555] with X^n in place of M, taking account that the

‡Inequality (2.23) holds under an assumption such as the finiteness of alphabet $\mathcal{Z}$.

wiretap channel here in consideration is stationary and memoryless. Thus, the convere part has been established. The direct part is trivial: just consider any i.i.d. input X^n in (2.19). $\square$

3 Evaluation of reliability and security

In this section, with criterion (2.4) we are now interested in exponentially decreasing rates of ϵ_n^B, δ_n^E as n tends to ∞. To this end, define the following two functions: let $W^n(\mathbf{y}|\mathbf{v}) : \mathcal{V}^n \to \mathcal{Y}^n$, $W^n(\mathbf{z}|\mathbf{v}) : \mathcal{V}^n \to \mathcal{Z}^n$ be arbitrary channels and $Q(\mathbf{v})$ be an arbitrary auxiliary input distribution on $\mathcal{V}^n$, and set

$$\phi(\rho|W^n, Q) \equiv -\log \sum_{\mathbf{y}} \left(\sum_{\mathbf{v}} Q(\mathbf{v}) W^n(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \right)^{1+\rho}, \tag{3.1}$$

$$\psi(\rho|W^n, Q) \equiv -\log \sum_{\mathbf{z}} \left(\sum_{\mathbf{v}} Q(\mathbf{v}) W^n(\mathbf{z}|\mathbf{v})^{1+\rho} \right) W_Q^n(\mathbf{z})^{-\rho}, \tag{3.2}$$

where $W_Q^n(\mathbf{z}) = \sum_{\mathbf{v}} Q(\mathbf{v}) W^n(\mathbf{z}|\mathbf{v})$. Then, we have

Theorem 3.1 Let (W_B^n, W_E^n) be a general wiretap channel, and M_n, L_n be arbitrary positive integers, then there exists a pair (φ_n, ψ_n^B) of encoder and decoder such that

$$\epsilon_n^B \leq 2 \inf_{0 \leq \rho \leq 1} (M_n L_n)^\rho e^{-\phi(\rho|W_B^{n+}, Q)}, \tag{3.3}$$

$$\delta_n^E \leq 2 \inf_{0 < \rho \leq 1} \frac{e^{-\psi(\rho|W_E^{n+}, Q)}}{\rho L_n^\rho}, \tag{3.4}$$

where (W_B^{n+}, W_E^{n+}) is a concatenation of (W_B^n, W_E^n) (cf. Definition 2.1). $\square$

Remark 3.1 Later we define the rates $R_B = \frac{1}{n} \log M_n$ and $R_E = \frac{1}{n} \log L_n$, which is called the coding rate for Bob and the resolvability rate against Eve, respectively. Rate R_B is quite popular in channel coding, whereas rate R_E, roughly speaking, indicates the rate of a large dice with L_n faces to provide randomness needed to implement an efficient *stochastic* encoder φ_n to deceive Eve. $\square$

Remark 3.2 Formula (3.3) without concatenation is due to Gallager [16], while formula (3.4) without concatenation is due to Hayashi [7]. However, what was actually claimed in [7] is, instead of (3.4),

$$I_n^E \le 2 \inf_{0<\rho\le 1} \frac{e^{-\psi(\rho|W_E^{n+},Q)}}{\rho L_n^{\rho}} \tag{3.5}$$

(cf. (2.7) in Section 1.B). Since $I_n^E \le \delta_n^E$ holds for any π_n, formula (3.5) is weaker than formula (3.4) in general. This difference, though seemingly not substantial, is of practical significance from the standpoint of security performance, because we can control security against Eve by adaptationally resetting π_n in the *nonstationary* real process of communication. □

Proof of Theorem 3.1:

The proof of (3.4) for the case with the non-concatenated (W_B^n, W_E^n) is given in [7] from the viewpoint of universal$_2$ hashing problems. However, this proof is rather complicated in that the arguments there are not focused enough on the information-theoretic core of formula (3.4).

Indeed, from the information-spectrum point of view, we had already established the crux of Theorem 3.1 in the short proof of Theorem 12 in Han and Verdú [21, p.768]. Let us just repeat it here for the sake of the reader's convenience.

First, set $P_{V^n} = Q$ and generate a random code $\mathcal{C} = \{V_1^n, V_2^n, \cdots, V_{M_nL_n}^n\}$ of size M_nL_n, where $V_1^n, V_2^n, \cdots, V_{M_nL_n}^n$ are i.i.d. random variables with common distribution Q on $\mathcal{V}^n$, and divide the M_nL_n random codewords $V_1^n, V_2^n, \cdots,$ $V_{M_nL_n}^n$ into M_n subcodes of equal size L_n so that

$$\begin{aligned} \mathcal{C}_1 &= \{V_1^n, V_2^n, \cdots, V_{L_n}^n\}, \\ \mathcal{C}_2 &= \{V_{L_n+1}^n, \cdots, V_{2L_n}^n\}, \\ \cdots & \quad \cdots\cdots\cdots\cdots \\ \mathcal{C}_{M_n} &= \{V_{L_n(M_n-1)+1}^n, \cdots, V_{M_nL_n}^n\}. \end{aligned} \tag{3.6}$$

For each message $i \in \mathcal{M}_n \equiv \{1, 2, \cdots, M_n\}$, the stochastic encoder $\varphi_n : \mathcal{M}_n \to \mathcal{V}^n$ produces the uniform distribution over $\mathcal{C}_i$. The decoder $\psi_n^B : \mathcal{Y}^n \to \mathcal{M}_n$ tries to decode all of these M_nL_n codewords $V_1^n, V_2^n, \cdots, V_{M_nL_n}^n$. Then, the *reliability* formula for channel $W_B^{n+} : \mathcal{V}^n \to \mathcal{Y}^n$:

$$\mathrm{E}_{\mathcal{C}}[\epsilon_n^B] \le \inf_{0\le\rho\le 1} (M_nL_n)^{\rho} e^{-\phi(\rho|W_B^{n+},Q)} \tag{3.7}$$

immediately follows from Gallager [16] with maximum likelihood decoding, where $\mathrm{E}_{\mathcal{C}}$ denotes the expectation with respect to the random code $\mathcal{C}$.

Next, for each $i \in \mathcal{M}_n$ we use the subcode $\mathcal{C}_i$ to produce an output distribution on $\mathcal{Z}^n$ that approximates enough the target output distribution π_n on $\mathcal{Z}^n$ generated via channel $W_E^{n+} : \mathcal{V}^n \to \mathcal{Z}^n$ due to the input distribution P_{V^n} (the *resolvability* [§] problem). Let $\hat{V}_i^n$ be the random variable taking values uniformly in the subcode $\mathcal{C}_i$, and let $\hat{Z}_i^n$ be the output via channel W_E^{n+} due to the input $\hat{V}_i^n$ ($i = 1, 2, \cdots, M_n$), with the probability distribution of $\hat{Z}_i^n$ denoted by $P_n^{(i)}$. We now evaluate the degree of approximation in terms of the divergence $D(P_n^{(i)}||\pi_n)$. By symmetry of the subcodes, we can focus on the case $i = 1$ without loss of generality. For notational simplicity, set

$$i_{V^n W^n}(\mathbf{v}, \mathbf{z}) = \log \frac{W_E^{n+}(\mathbf{z}|\mathbf{v})}{P_{Z^n}(\mathbf{z})}.$$

Then, we have

$$\begin{aligned}
&\mathrm{E}_{\mathcal{C}_1} D(P_n^{(1)}||\pi_n) \\
&= \sum_{\mathbf{z}\in\mathcal{Z}^n} \sum_{\mathbf{c}_1\in\mathcal{V}^n} \cdots \sum_{\mathbf{c}_{L_n}\in\mathcal{V}^n} P_{V^n}(\mathbf{c}_1)\cdots P_{V^n}(\mathbf{c}_{L_n}) \\
&\quad \cdot \frac{1}{L_n}\sum_{j=1}^{L_n} W^n(\mathbf{z}|\mathbf{c}_j) \log\left(\frac{1}{L_n}\sum_{k=1}^{L_n} \exp i_{V^n W^n}(\mathbf{c}_k, \mathbf{z})\right) \\
&= \sum_{\mathbf{c}_1\in\mathcal{V}^n} \cdots \sum_{\mathbf{c}_{L_n}\in\mathcal{V}^n} P_{V^n}(\mathbf{c}_1)\cdots P_{V^n}(\mathbf{c}_{L_n}) \\
&\quad \cdot \sum_{\mathbf{z}\in\mathcal{Z}^n} W^n(\mathbf{z}|\mathbf{c}_1) \log\left(\frac{1}{L_n}\sum_{k=1}^{L_n} \exp i_{V^n W^n}(\mathbf{c}_k, \mathbf{z})\right) \\
&\le \sum_{\mathbf{c}_1\in\mathcal{V}^n} \sum_{\mathbf{z}\in\mathcal{Z}^n} W^n(\mathbf{z}|\mathbf{c}_1) P_{V^n}(\mathbf{c}_1) \\
&\quad \cdot \log\left(\frac{1}{L_n} \exp i_{V^n W^n}(\mathbf{c}_1, \mathbf{z}) \right.\\
&\qquad \left. + \frac{1}{L_n}\sum_{k=2}^{L_n} \mathrm{E} \exp i_{V^n W^n}(V_k^n, \mathbf{z})\right) \\
&\le \mathrm{E}\left[\log\left(1 + \frac{1}{L_n} \exp i_{V^n W^n}(V^n, Z^n)\right)\right], \qquad (3.8)
\end{aligned}$$

[§] Csiszár [12] is the first who has looked at the security problem with wiretap channels from the viewpoint of resolvability devised by Han and Verdú [21].

where the first inequality follows from the concavity of the logarithm and the second one is a result of

$$\mathrm{E}[\exp i_{V^nW^n}(V_k^n, \mathbf{z})] = 1$$

for all $\mathbf{z} \in \mathcal{Z}^n$ and $k = 1, 2, \cdots, L_n$. Now, apply a simple inequality with $0 < \rho \leq 1$ and $x \geq 0$ (cf. [7]):

$$\log(1+x) \leq \frac{\log(1+x)^\rho}{\rho} \leq \frac{\log(1+x^\rho)}{\rho} \leq \frac{x^\rho}{\rho}$$

to (3.8) to eventaully obtain

$$\mathrm{E}_{\mathcal{C}_1} D(P_n^{(1)}||\pi_n) \leq \inf_{0<\rho\leq 1} \frac{e^{-\psi(\rho|W_E^{n+}, Q)}}{\rho L_n^\rho},$$

from which it follows that

$$\mathrm{E}_{\mathcal{C}} \left[\frac{1}{M_n} \sum_{i=1}^{M_n} D(P_n^{(i)}||\pi_n) \right] \leq \inf_{0<\rho\leq 1} \frac{e^{-\psi(\rho|W_E^{n+}, Q)}}{\rho L_n^\rho},$$

that is,

$$\mathrm{E}_C[\delta_n^E] \leq \inf_{0<\rho\leq 1} \frac{e^{-\psi(\rho|W_E^{n+}, Q)}}{\rho L_n^\rho}. \tag{3.9}$$

Thus, in view of (3.7) and (3.9) with Markov inequality, we conclude that there exists at least one non-random pair (φ_n, ψ_n^B) of encoder and decoder satisfying (3.3), (3.4), thereby completing the proof of the theorem. $\square$

An immediate consequence of Theorem 3.1 is the following, where

$$\phi(-\rho|W^n, Q) \equiv -\log \sum_{\mathbf{z}} \left(\sum_{\mathbf{v}} Q(\mathbf{v}) W^n(\mathbf{z}|\mathbf{v})^{\frac{1}{1-\rho}} \right)^{1-\rho}. \tag{3.10}$$

Theorem 3.2 Let M_n and L_n be arbitrary positive integers, then there exists a pair (φ_n, ψ_n^B) of encoder and decoder such that

$$\epsilon_n^B \leq 2 \inf_{0\leq\rho\leq 1} (M_n L_n)^\rho e^{-\phi(\rho|W_B^{n+}, Q)}, \tag{3.11}$$

$$\delta_n^E \leq 2 \inf_{0<\rho<1} \frac{e^{-\phi(-\rho|W_E^{n+}, Q)}}{\rho L_n^\rho}. \tag{3.12}$$

Proof: A simple inequalitiy (due to Hölder's inequality)

$$\left(\sum_{\mathbf{v}} Q(\mathbf{v}) W_E^{n+}(\mathbf{z}|\mathbf{v})^{1+\rho}\right) W_Q^{n+}(\mathbf{z})^{-\rho} \le \left(\sum_{\mathbf{v}} Q(\mathbf{v}) W_E^{n+}(\mathbf{z}|\mathbf{v})^{\frac{1}{1-\rho}}\right)^{1-\rho} \tag{3.13}$$

for $0 \le \rho < 1$ was shown in [7], [18], which together with Theorem 3.1 yields Theorem 3.2. □

Remark 3.3 (Two upper bounds on security measure) Notice here that the same function $\phi(\rho|W^{n+}, Q)$ appears with different values of ρ and different channels W_B^{n+}, W_E^{n+} in (3.11) and (3.12), respectively. As was pointed out in the above, Hayashi [7] has addressed only *non-concatenated* wiretap channels (W_B^n, W_E^n) instead of concatenated ones (W_B^{n+}, W_E^{n+}) in Theorem 3.1 and Theorem 3.2. However, the introduction of (W_B^{n+}, W_E^{n+}) is necessitated and of great importance in studying reliability and security with the wiretap channel (cf. Remark 3.4). Although upper bound (3.5) is tighter than upper bound (3.12) in general, as will turn out in Section 5, upper bound (3.5) is quite intractable compared with upper bound (3.12) when dealing with the security problem with *cost constraint*. See Remark 5.2. □

Remark 3.4 (Tradeoff of reliability and security by concatenation) The reason why we have introduced the concatenated channel $W^{n+}(\mathbf{y}|\mathbf{v})$ instead of the non-concatenated channel $W^n(\mathbf{y}|\mathbf{x})$ is the following. The quantity $A_n \equiv e^{-\phi(\rho|W_B^{n+}, Q)}$ in (3.11) is rewritten, by concavity of the function $f(x) = x^{\frac{1}{1+\rho}}$, as

$$\begin{aligned} A_n &= \sum_{\mathbf{y}} \left(\sum_{\mathbf{v}} Q(\mathbf{v}) \left(\sum_{\mathbf{x}} P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) W_B^n(\mathbf{y}|\mathbf{x})\right)^{\frac{1}{1+\rho}}\right)^{1+\rho} && (3.14) \\ &\ge \sum_{\mathbf{y}} \left(\sum_{\mathbf{v}} \sum_{\mathbf{x}} Q(\mathbf{v}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) W_B^n(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}}\right)^{1+\rho} && (3.15) \\ &= \sum_{\mathbf{y}} \left(\sum_{\mathbf{x}} P(\mathbf{x}) W_B^n(\mathbf{y}|\mathbf{x})^{\frac{1}{1+\rho}}\right)^{1+\rho}, && (3.16) \end{aligned}$$

where $P(\mathbf{x}) = \sum_{\mathbf{v}} Q(\mathbf{v}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v})$. This implies that concatenation decreases reliability for the channel for Bob. On the other hand, however, the quantity $B_n \equiv e^{-\phi(-\rho|W_E^{n+}, Q)}$ in (3.12) is rewritten, by convexity of the function $g(x) = x^{\frac{1}{1-\rho}}$, as

$$B_n = \sum_{\mathbf{z}} \left(\sum_{\mathbf{v}} Q(\mathbf{v}) \left(\sum_{\mathbf{x}} P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) W_E^n(\mathbf{z}|\mathbf{x}) \right)^{\frac{1}{1-\rho}} \right)^{1-\rho} \tag{3.17}$$

$$\leq \sum_{\mathbf{z}} \left(\sum_{\mathbf{v}} \sum_{\mathbf{x}} Q(\mathbf{v}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) W_E^n(\mathbf{z}|\mathbf{x})^{\frac{1}{1-\rho}} \right)^{1-\rho} \tag{3.18}$$

$$= \sum_{\mathbf{z}} \left(\sum_{\mathbf{x}} P(\mathbf{x}) W_E^n(\mathbf{z}|\mathbf{x})^{\frac{1}{1-\rho}} \right)^{1-\rho}, \tag{3.19}$$

which implies that concatenation increases security against the channel for Eve. Thus, we can control the tradeoff between reliability and security (usually conflicting) by adequate choice of an auxiliary channel $P_{X^n|V^n}$ (e.g., see Fig.5 later for the case of stationary memoryless wiretap channels). Finally, it should be noted that $C_n \equiv \psi(\rho|W_E^{n+}, Q)$ in (3.5) does *not* have such a nice tradeoff property in general, because it lacks the convexity in $W_E^n(\mathbf{z}|\mathbf{x})$ unlike in the above. □

So far we have considered about performance of general wiretap channels. Now, to be specific, suppose that the channel (W_B^n, W_E^n) is stationary and memoryless, and

$$P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) = \prod_{i=1}^{n} P_{X|V}(x_i|v_i),$$

$$P_{V^n}(\mathbf{v}) = \prod_{i=1}^{n} P_V(v_i),$$

where $\mathbf{x} = (x_1, \cdots, x_n), \mathbf{v} = (v_1, \cdots, v_n)$. Moreover, set $M_n = e^{nR_B}, L_n = e^{nR_E}$ with rates R_B, R_E, which we call Bob's *coding rate* and Eve's *resolvability rate*, respectively. Accordingly, with $q = P_V$ let us define the *reliability function* $F(q, R_B, R_E, n)$ for Bob, and the *security functions* $G(q, R_E, n)$, $H(q, R_E, n)$ against Eve, by

$$F(q, R_B, R_E, n) \equiv \sup_{0 \leq \rho \leq 1} \left(\phi(\rho|W_B^+, q) - \rho(R_B + R_E) - \frac{\rho \log 3}{n} \right), \tag{3.20}$$

$$G(q, R_E, n) \equiv \sup_{0 < \rho \leq 1} \left(\psi(\rho|W_E^+, q) + \rho R_E + \frac{\log \rho}{n} \right), \tag{3.21}$$

$$H(q,R_E,n) \equiv \sup_{0<\rho<1}\left(\phi(-\rho|W_E^+,q)+\rho R_E+\frac{\log\rho}{n}\right), \tag{3.22}$$

where (W_B^+,W_E^+) is a concatenation of (W_B,W_E) (cf. Definition 2.1), that is,

$$W_B^+(y|v) = \sum_{x\in\mathcal{X}} W_B(y|x)P_{X|V}(x|v), \tag{3.23}$$

$$W_E^+(z|v) = \sum_{x\in\mathcal{X}} W_E(z|x)P_{X|V}(x|v); \tag{3.24}$$

and we have set

$$\phi(\rho|W_B^+,q) = -\log\sum_y\left(\sum_v q(v)W_B^+(y|v)^{\frac{1}{1+\rho}}\right)^{1+\rho}, \tag{3.25}$$

$$\psi(\rho|W_E^+,q) = -\log\sum_z\left(\sum_v q(v)W_E^+(z|v)^{1+\rho}\right)W_q^+(z)^{-\rho} \tag{3.26}$$

$$\phi(-\rho|W_E^+,q) = -\log\sum_z\left(\sum_v q(v)W_E^+(z|v)^{\frac{1}{1-\rho}}\right)^{1-\rho} \tag{3.27}$$

with $W_q^+(z)=\sum_v q(v)W_E^+(z|v)$. Then, Theorem 3.1 yields

Theorem 3.3 If a wiretap channel (W_B^n,W_E^n) is stationary and memoryless, we have

$$\epsilon_n^B \leq 2e^{-nF(q,R_B,R_E,+\infty)}, \tag{3.28}$$

$$\delta_n^E \leq 2e^{-nG(q,R_E,n)}. \tag{3.29}$$

Similarly, Theorem 3.2 yields

Theorem 3.4 If a wiretap channel (W_B^n,W_E^n) is stationary and memoryless, we have

$$\epsilon_n^B \leq 2e^{-nF(q,R_B,R_E,+\infty)}, \tag{3.30}$$

$$\delta_n^E \leq 2e^{-nH(q,R_E,n)}. \tag{3.31}$$

Remark 3.5 It should be noted that, on the right-hand sides of (3.20) $\sim$ (3.22), the terms of order $O(\frac{1}{n})$ approach zero as n tends to ∞, so that these terms do not effect the exponents. Also, the $+\infty$ in $F(q,R_B,R_E,+\infty)$ in (3.28) and (3.30) means that the term $\frac{\rho\log 3}{n}$ is unnecessary here, but it is necessary when we consider the maximum criterion instead of the average criterion (cf. Remark 3.8). $\square$

Remark 3.6 (Reliability and security functions) The function $F(q, R_B, R_E, n)$ specifies *performance of channel coding* (called the *random coding exponent* of Gallager [16]), whereas the functions $G(q, R_E, n)$, $H(q, R_E, n)$ specify *performance of channel resolvability* (cf. Han and Verdú [21], Han [19], Hayashi [10], Bloch and Laneman [11]). In particular, the function $H(q, R_E, +\infty)$ coincides with the converse channel coding exponent of Arimoto [17]. It is easy to check that $F(q, R_B, R_E, n)$ is monotone strictly decreasing and convex with respect to $R_B + R_E$, and $H(q, R_E, n)$, $G(q, R_E, n)$ are monotone strictly increasing and convex with respect to R_E. Also, it is not difficult to check that

$$F(q, R_B, R_E, +\infty) = 0 \text{ at } R_B + R_E = I(q, W_B^+);$$

$$H(q, R_E, +\infty) = G(q, R_E, +\infty) = 0 \text{ at } R_E = I(q, W_E^+),$$

where $I(q, W)$ indicates the mutual information between the input and the output induced by input q and channel W. Therefore, $F(q, R_B, R_E, +\infty)$ is positive for $R_B + R_E < I(q, W_B^+)$, whereas $H(q, R_E, +\infty)$, $G(q, R_E, +\infty)$ are positive for $R_E > I(q, W_E^+)$.
It should be also noted that if channel W_B^+ is strictly more capable than channel W_E^+ then we can take an input distribution q so that $I(q, W_B^+) > I(q, W_E^+)$. Thus, it is concluded that rate $I(q, W_B^+) - I(q, W_E^+)$ is δ-achievable in the sense of Section 2. This is because, for any small $\gamma > 0$, we can choose R_B, R_E so that $R_E = I(q, W_E^+) + \frac{\gamma}{2}$, $R_B + R_E = I(q, W_B^+) - \frac{\gamma}{2}$ and hence $R_B = I(q, W_B^+) - I(q, W_E^+) - \gamma$ for which $F(q, R_B, R_E, +\infty) > 0$, $H(q, R_E, +\infty) > 0, G(q, R_E, +\infty) > 0$ and therefore both of ϵ_n^B and δ_n^E exponentially decay with increasing n. Thus, summarizing these observations, we can depict Fig.3, Fig.5 and Fig.6 (cf. Remark 3.4 and Remark 3.7), where we are considering, as an example, wiretap channels consisting of a pair of two BSC's (Binary Symmetric Channels). It should be noted here that for any pair of BSC's one is degraded (and hence also is more capable) than the other one, so that in calculating the δ-secrecy capacity we can invoke formula (3.33) in Theorem 3.5 to follow. More specifically, let q indicate the input maximizing $I(q, W_B^+) - I(q, W_E^+)$, then this $I(q, W_B^+) - I(q, W_E^+)$ gives the δ-secrecy capacity δ-C_s. □

Remark 3.7 (Tradeoff of reliability and security by rate exchange) In Remark 3.4 we have shown how to control tradeoff between reliability and security by adequate choice of auxiliary channels (concatenation), yielding increase of the values of $H(q, R_E, n)$, $G(q, R_E, n)$ but decrease of the value of

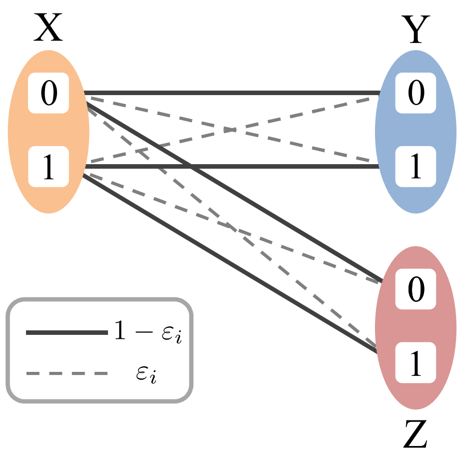

Figure 2: Non-concatenated BSC (Binary Symmetric Channel) with crossover probability ε_i $(i = y, z)$; ε_y, ε_z denote crossover probabilities for channel $P_{Y|X} : \mathcal{X} \to \mathcal{Y}$ and channel $P_{Z|X} : \mathcal{X} \to \mathcal{Z}$, respectively.

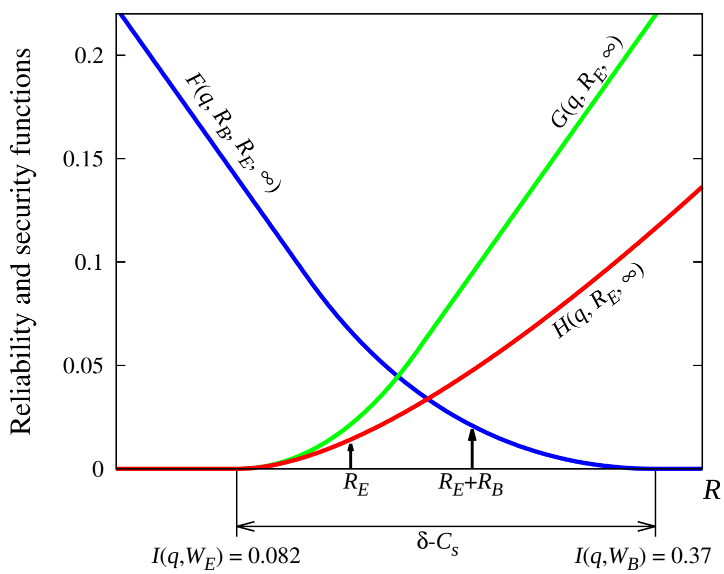

Figure 3: Reliability and security functions for non-concatenated BSC ($\varepsilon_y = 0.1$, $\varepsilon_z = 0.3$).

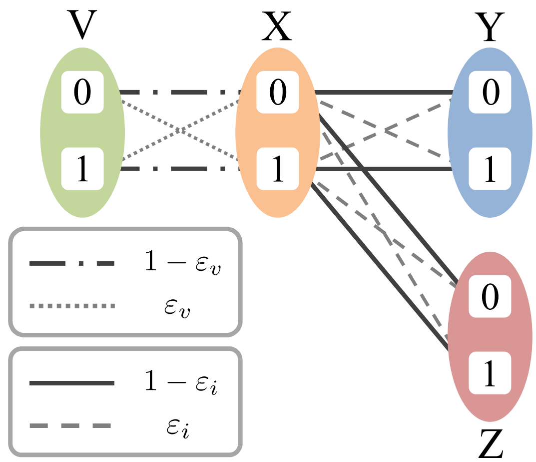

Figure 4: Concatenated BSC; ε_v denotes crossover probability for channel $P_{X|V} : \mathcal{V} \to \mathcal{X}$.

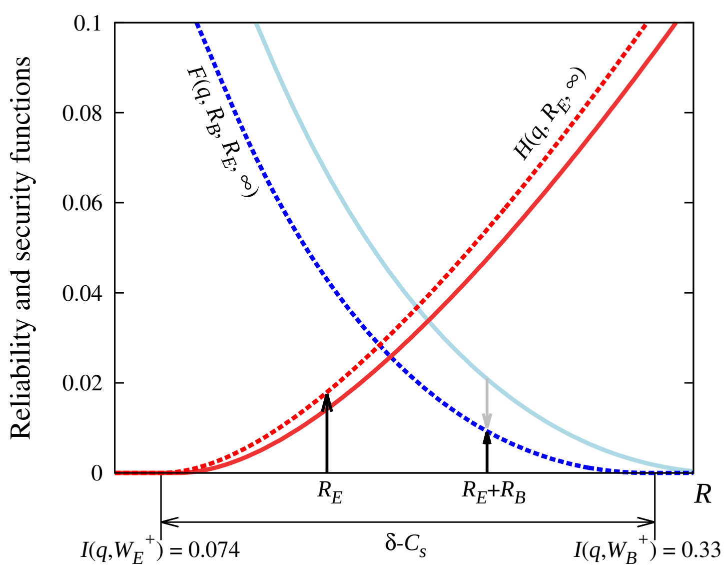

Figure 5: Tradeoff by concatenation: reliability and security functions for non-concatenated (solid lines) and concatenated (dashed lines, $\varepsilon_v = 0.025$) BSC ($\varepsilon_y = 0.1$, $\varepsilon_z = 0.3$), where reliability for Bob decreases but security against Eve increases.

$F(q, R_B, R_E, n)$, which implies that enhancement of security was attained at the expense of reliability with fixed rates R_B, R_E, and vice versa (see Fig. 5). One more way to control such a tradeoff is to handle rates R_B, R_E, where enhancement of security is attained at the expense of rate R_B but not at reliability: with the same functions $F(q, R_B, R_E, n)$, $H(q, R_E, n)$, $G(q, R_E, n)$ as above, we let R_E increase while keeping the sum $R_B + R_E$ unchanged, which implies decrease of rate R_B but no expense of reliability, because then the values of $H(q, R_E, n)$, $G(q, R_E, n)$ increase but that of $F(q, R_B, R_E, n)$ remains unchanged (e.g., see Fig.6). $\square$

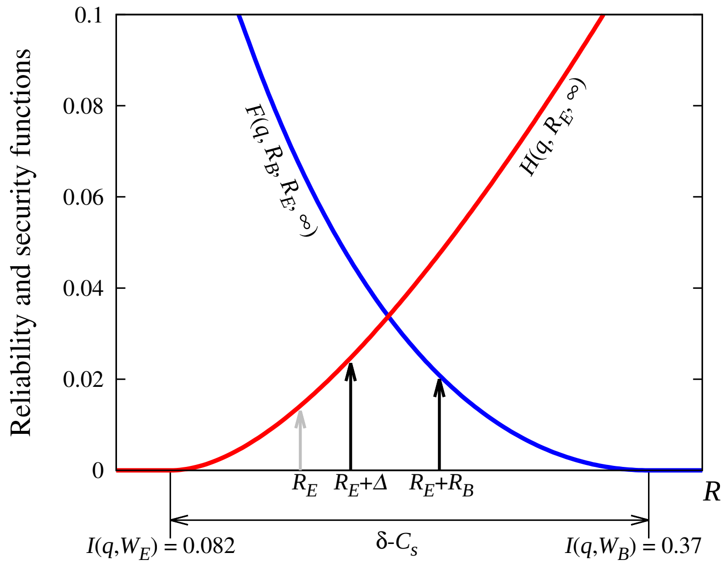

Figure 6: Tradeoff by rate exchange: let $R_E \to R_E + \Delta$ and $R_B \to R_B - \Delta$, then security against Eve increases by Δ and rate R_B decreases by Δ but at no expense of reliability for Bob.

An immediate consequence from the observation in Remark 3.6 is the following theorem, which is a strengthening of the latter parts in Theorem 2.1 and Theorem 2.2:

Theorem 3.5 If a wiretap channel (W_B^n, W_E^n) is stationary and memoryless, then it holds that

$$\delta\text{-}C_s = \sup_{VX}(I(V;Y) - I(V;Z)), \tag{3.32}$$

which means that the δ-secrecy capacity is attained in general with a concatenated channel. Furthermore, if (W_B^n, W_E^n) is more capable, then it holds that

$$\delta\text{-}C_s = \sup_X (I(X;Y) - I(X;Z)), \tag{3.33}$$

which means that in this case there is no need to consider concatenation. □

Remark 3.8 (Average vs. maximum criteria) Recall that upper bounds in (3.28), (3.29) and (3.30), (3.31) are based on the averaged criteria as mentioned in Section 1.B. Alternatively, instead of the averaged criteria ϵ_n^B and δ_n^E, we can define the maximum criteria $\text{m-}\epsilon_n^B$ and $\text{m-}\delta_n^E$ as follows.

$$\text{m-}\epsilon_n^B \equiv \max_{i\in\mathcal{M}_n} \Pr\{\psi_n^B(\varphi_n(i)) \neq i\}, \tag{3.34}$$

$$\text{m-}\delta_n^E \equiv \max_{i\in\mathcal{M}_n} D(P_n^{(i)}||\pi_n). \tag{3.35}$$

With these criteria, (3.28), (3.29) and (3.30), (3.31) are slightly modified, using Markov inequality, as

$$\text{m-}\epsilon_n^B \leq 6e^{-nF(q,R_B,R_E,n)}, \tag{3.36}$$

$$\text{m-}\delta_n^E \leq 6e^{-nG(q,R_E,n)}, \tag{3.37}$$

and

$$\text{m-}\epsilon_n^B \leq 6e^{-nF(q,R_B,R_E,n)}, \tag{3.38}$$

$$\text{m-}\delta_n^E \leq 6e^{-nH(q,R_E,n)}, \tag{3.39}$$

respectively. Bounds (3.36) (and (3.38)) are well known in channel coding (cf. Gallager [16]), whereas bounds (3.37) and (3.39) are taken into consideration for the first time in this paper.

In channel coding, which of the averaged ϵ_n^B or the maximum $\text{m-}\epsilon_n^B$ we should take would be rather a matter of preference or the context. On the other hand, however, which of the averaged δ_n^E or the maximum $\text{m-}\delta_n^E$ we should take is a serious matter from the viewpoint of security. This is because, even with small δ_n^E, we cannot exclude a possibility that the divergence distance $D(P_n^{(i)}||\pi_n)$ is very large for some particular $i \in \mathcal{M}_n$, and hence $\text{m-}\delta_n^E$ is also very large, which implies that the message i is not saved from a serious risk of successful decryption by Eve. On the other hand, with small $\text{m-}\delta_n^E$, every message $i \in \mathcal{M}_n$ is guaranteed to be kept highly confidential against Eve as well. Thus, we prefer the criterion $\text{m-}\delta_n^E$ as well as $\text{m-}\epsilon_n^B$ in this paper. □

4 Secrecy capacity with cost constraint

So far we have established the fundamental theorems/corollaries about the problem of secrecy capacities, and in particular, that of reliability and security functions in the context of wiretap channels *without* cost constraint. However, from the viewpoint of communication technologies, it is sometimes needed to impose *cost constraint* on channel inputs. In this section we address this problem.

For $n = 1, 2, \cdots$ fix a mapping $c_n : \mathcal{X}^n \to \mathbf{R}$ (the set of real numbers) arbitrarily. For $\mathbf{x} \in \mathcal{X}^n$ we call $c_n(\mathbf{x})$ the cost of $\mathbf{x}$ and $\frac{1}{n}c_n(\mathbf{x})$ the cost per letter. In the channel coding problem with cost constraint, we require the all the encoder outputs $\varphi_n(i) \in \mathcal{X}^n$ satisfy

$$\Pr\left\{\frac{1}{n}c_n(\varphi_n(i)) \le \Gamma\right\} = 1 \quad (i = 1, 2, \cdots, M_n) \tag{4.1}$$

for all $n = 1, 2, \cdots$, where Γ is an arbitrarily given constant, which we call *cost constraint* Γ. Notice here that the encoder φ_n may be *stochastic*.

We say that a rate R is Γ-achievable if there exists a pair (φ_n, ψ_n^B) of encoder and decoder satisfying (2.6), (2.9) and (4.1).

The secrecy capacity d-$C_s(\Gamma)$ is defined to be the supremum of all Γ-achievable rates. Similarly, we can define also w-$C_s(\Gamma)$, i-$C_s(\Gamma)$, ∂-$C_s(\Gamma)$ and δ-$C_s(\Gamma)$. To give the formula for d-$C_s(\Gamma)$ with a general wiretap channel, define

$$\mathcal{X}^n(\Gamma) = \left\{\mathbf{x} \in \mathcal{X}^n \,\middle|\, \frac{1}{n}c_n(\mathbf{x}) \le \Gamma\right\} \tag{4.2}$$

and let $\mathcal{S}_\Gamma$ denote the set of all input processes $\mathbf{X} = \{X^n\}_{n=1}^\infty$ satisfying

$$\Pr\{X^n \in \mathcal{X}^n(\Gamma)\} = 1 \tag{4.3}$$

for all $n = 1, 2, \cdots$. Then, we have the following two theorems due to Bloch and Laneman [11].

Theorem 4.1 A general formula

$$\text{d-}C_s(\Gamma) = \sup_{\mathbf{VX}:\mathbf{X}\in\mathcal{S}_\Gamma}\left(\text{p-}\liminf_{n\to\infty}\frac{1}{n}\log\frac{W_B^{n+}(Y^n|V^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty}\frac{1}{n}\log\frac{W_E^{n+}(Z^n|V^n)}{P_{Z^n}(Z^n)}\right) \tag{4.4}$$

holds, where $\mathbf{V} \to \mathbf{X} \to \mathbf{YZ}$ forms a Markov chain. $\square$

As a special but important case, let us here consider the case in which the channel (W_B^n, W_E^n) is stationary and memoryless with *additive cost* $c : \mathcal{X} \to \mathbf{R}$ in the sense that $c_n(\mathbf{x}) = \sum_{i=1}^n c(x_i)$ where $\mathbf{x} = (x_1, \cdots, x_n)$. Then, we obtain

Theorem 4.2 If a wiretap channel (W_B^n, W_E^n) is stationary and memoryless, then (4.4) reduces to

$$\text{d-}C_s(\Gamma) = \sup_{VX:\mathrm{E}c(X)\le\Gamma} (I(V;Y) - I(V;Z)), \tag{4.5}$$

where X is any channel input satisfying $\mathrm{E}c(X) \le \Gamma$; and Y, Z are the outputs via channels W_B, W_E due to X, respectively, whereas V is an arbitrary auxiliary random variable such that $V \to X \to YZ$ forms a Markov chain. □

Remark 4.1 Actually, it can be shown that d-$C_s(\Gamma)$ = i-$C_s(\Gamma)$ = δ-$C_s(\Gamma)$ if all alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are *finite*; the first equality is shown in [11] along with [12], and the second equality follows as a special case from Theorem 5.3 later in Section 5 on reliability and security with cost constraint (see, also Remark 3.6). □

Proof of Theorem 4.2:

Although the proof of the achievability part is given in [11], which carries out direct calculation to derive a resolvability result but does not rely on Theorem 4.1, we will give here another quite different proof relying solely on Theorem 4.1, since it itself is of independent interest from the viewpoint of information spectrum methods. As for the detailed proof, see Appendix I. □

In passing this section, let us address the problem of a general *more capable* wiretap channel (W_B^n, W_E^n) with cost constraint (cf. Definition 2.3). In this respect, we have

Theorem 4.3 If a wiretap channel (W_B^n, W_E^n) is more capable, formula (4.4) is simplified as follows (non-concatenated channel):

$$\begin{aligned}&\text{d-}C_s(\Gamma)\\ &= \sup_{\mathbf{X}\in\mathcal{S}_\Gamma} \left(\text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{W_B^n(Y^n|X^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{W_E^n(Z^n|X^n)}{P_{Z^n}(Z^n)} \right).\end{aligned} \tag{4.6}$$

Furthermore, if (W_B^n, W_E^n) is stationary and memoryless, then (4.6) reduces to

$$\text{d-}C_s(\Gamma) = \sup_{X:\mathrm{E}c(X)\le\Gamma} (I(X;Y) - I(X;Z)). \tag{4.7}$$

Remark 4.2 Like in Remark 4.1, it can be shown also in (4.7) that d-$C_s(\Gamma) =$ i-$C_s(\Gamma) = \delta$-$C_s(\Gamma)$ if all alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are *finite*. This equivalence follows as a special case from Theorem 5.4 in Section 5. $\square$

Proof of Theorem 4.3:

Keeping the cost constraint here in mind, formula (4.6) follows from Theorem 4.1 in the entirely same way that Theorem 2.2 was derived from Theorem 2.1, whereas (4.7) follows from the argument as in the proof of Theorem 4.2. $\square$

5 Reliability and security with cost constraint

Having dealt with the problem of reliability and security *without* cost constraint in Section 3, the problem of reliability and security *with* cost constraint is considered in this section.

Let (W_B^n, W_E^n) be a general wiretap channel and $c_n : \mathcal{X}^n \to \mathbf{R}$ be a cost function as defined in Section 4. If a Markov chain $V^n \to X^n \to Y^n Z^n$ for (W_B^n, W_E^n) defined as in Section 2 satisfies cost constraint

$$\Pr\{X^n \in \mathcal{X}^n(\Gamma)\} = 1, \tag{5.1}$$

where $\mathcal{X}^n(\Gamma)$ is defined as in (4.2), that is,

$$\mathcal{X}^n(\Gamma) = \left\{ \mathbf{x} \in \mathcal{X}^n \left| \frac{1}{n} c_n(\mathbf{x}) \leq \Gamma \right.\right\}, \tag{5.2}$$

then, we call (W_B^n, W_E^n) a wiretap channel with *cost constraint* Γ. Notice here that we want to impose cost constraint on X^n but not on V^n, which causes some inconvenient subtleties.

Thus, paralleling with the arguments that have derived Theorem 3.1 and Theorem 3.2, we have the following theorem:

Theorem 5.1 Let (W_B^n, W_E^n) be a general wiretap channel with cost constraint Γ, and M_n, L_n be arbitrary positive integers, then the random code $\mathcal{C}$ as used in the proof of Theorem 3.1 has the performance with $0 < \rho < 1$ as

$$\mathrm{E}_{\mathcal{C}}[\epsilon_n^B] \quad \leq \quad (M_n L_n)^\rho \sum_{\mathbf{y}} \left(\sum_{\mathbf{v}} Q(\mathbf{v}) W_B^{n+}(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}} \right)^{1+\rho}, \tag{5.3}$$

$$\mathrm{E}_{\mathcal{C}}[\delta_n^E] \leq \frac{1}{\rho L_n^\rho}\sum_{\mathbf{z}}\left(\sum_{\mathbf{v}} Q(\mathbf{v}) W_E^{n+}(\mathbf{z}|\mathbf{v})^{\frac{1}{1-\rho}}\right)^{1-\rho}, \tag{5.4}$$

$$\mathrm{E}_{\mathcal{C}}[\delta_n^E] \leq \frac{1}{\rho L_n^\rho}\sum_{\mathbf{z}}\left(\sum_{\mathbf{v}} Q(\mathbf{v}) W_E^{n+}(\mathbf{z}|\mathbf{v})^{1+\rho}\right) W_Q^n(\mathbf{z})^{-\rho}, \tag{5.5}$$

where

$$\sum_{\mathbf{x}\in\mathcal{X}^n(\Gamma)} P_{X^n}(\mathbf{x}) = 1, \tag{5.6}$$

$$W_Q^n(\mathbf{z}) = \sum_{\mathbf{v}} Q(\mathbf{v}) W_E^{n+}(\mathbf{z}|\mathbf{v}). \tag{5.7}$$

Proof:

Keeping (5.6) in mind, it suffices to invoke the arguments as used in the proofs of Theorem 3.1 and Theorem 3.2 (see, Han [19, p.423]). $\square$

Remark 5.1 With condition (5.6), $W_B^{n+}(\mathbf{y}|\mathbf{v}), W_E^{n+}(\mathbf{z}|\mathbf{v})$ can be written as

$$W_B^{n+}(\mathbf{y}|\mathbf{v}) = \sum_{\mathbf{x}\in\mathcal{X}^n(\Gamma)} W_B^n(\mathbf{y}|\mathbf{x}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}), \tag{5.8}$$

$$W_E^{n+}(\mathbf{z}|\mathbf{v}) = \sum_{\mathbf{x}\in\mathcal{X}^n(\Gamma)} W_E^n(\mathbf{z}|\mathbf{x}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}). \tag{5.9}$$

$\square$

Suppose now that we are given a stationary and memoryless wiretap channel (W_B^n, W_E^n) with *additive* cost constraint Γ. With this important class of channels, we attempt to bring out specific useful insights on the basis of these bounds (5.3) and (5.4). To do so, let us consider the case in which $V^nX^n = (V_1X_1, \cdots, V_nX_n)$ are i.i.d. variables with common joint distribution

$$P_{XV}(x, v) \quad ((v, x) \in \mathcal{V}\times\mathcal{X}), \tag{5.10}$$

then, the probabilities of X^n and V^n, and the conditional probability of X^n given V^n are written as

$$P_{X^n}(\mathbf{x}) = \prod_{i=1}^n P_X(x_i), \tag{5.11}$$

$$P_{V^n}(\mathbf{v}) = \prod_{i=1}^n P_V(v_i), \tag{5.12}$$

$$P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) = \prod_{i=1}^{n} P_{X|V}(x_i|v_i), \tag{5.13}$$

respectively, where

$$\mathbf{x} = (x_1, \cdots, x_n), \quad \mathbf{v} = (v_1, \cdots, v_n).$$

It should be noted here that V^n indicates a channel input for (W_B^{n+}, W_E^{n+}), and X^n indicates a channel input for (W_B^n, W_E^n). We assume that the probability distribution P_X on $\mathcal{X}$ satisfies

$$\sum_{x\in\mathcal{X}} P_X(x)c(x) \le \Gamma - \tau, \tag{5.14}$$

where $c(\cdot)$ is an additive cost function and $\tau > 0$ is an arbitrarily small number, and define

$$\chi(\mathbf{x}) = \begin{cases} 1 & \text{for } \sum_{i=1}^n c(x_i) \le n\Gamma, \\ 0 & \text{otherwise}; \end{cases} \tag{5.15}$$

$$\mu_n = \sum_{\mathbf{x}} \chi(\mathbf{x}) \prod_{i=1}^{n} P_X(x_i). \tag{5.16}$$

It is easy to see that $\mu_n \to 1$ as $n \to +\infty$ by means of the weak law of large numbers. We rewrite μ_n as follows:

$$\begin{aligned} \mu_n &= P_{X^n}(\mathcal{X}^n(\Gamma)) \\ &= \sum_{\mathbf{v}\in\mathcal{V}^n} P_{X^n|V^n}(\mathcal{X}^n(\Gamma)|\mathbf{v}) P_{V^n}(\mathbf{v}), \end{aligned} \tag{5.17}$$

then, by means of Markov inequality there exists a subset $\mathcal{T}_1 \subset \mathcal{V}^n$ such taht

$$\alpha_n \stackrel{\Delta}{=} P_{V^n}(\mathcal{T}_1) \ \ge \ 1 - \sqrt{1-\mu_n} \stackrel{\Delta}{=} \beta_n, \tag{5.18}$$

$$\gamma_n(\mathbf{v}) \stackrel{\Delta}{=} P_{X^n|V^n}(\mathcal{X}^n(\Gamma)|\mathbf{v}) \ \ge \ \beta_n \quad \text{for all } \mathbf{v} \in \mathcal{T}_1. \tag{5.19}$$

Obviously, $\alpha_n, \beta_n, \gamma_n(\mathbf{v}) \to 1$ as $n \to +\infty$. Define

$$\tilde{P}_{V^n}(\mathbf{v}) = \frac{P_{V^n}(\mathbf{v})}{\alpha_n} \quad (\mathbf{v} \in \mathcal{T}_1), \tag{5.20}$$

$$\tilde{P}_{X^n|V^n}(\mathbf{x}|\mathbf{v}) = \frac{P_{X^n|V^n}(\mathbf{x}|\mathbf{v})}{\gamma_n(\mathbf{v})} \quad (\mathbf{x} \in \mathcal{X}^n(\Gamma), \mathbf{v} \in \mathcal{T}_1), \tag{5.21}$$

which are obviously a probability distribution and a conditional probability distribution. On the other hand, notice that $\chi(\mathbf{x})$ can be upper bounded (for all $\mathbf{x} \in \mathcal{X}^n$) as

$$\chi(\mathbf{x}) \le \exp\left[(1+\rho)r\left(n\Gamma - \sum_{i=1}^{n} c(x_i)\right)\right], \tag{5.22}$$

where $r \ge 0$ is an arbitrary number. Now substitute $\tilde{P}_{V^n}(\mathbf{v})$ and $\tilde{P}_{X^n|V^n}(\mathbf{x}|\mathbf{v})$ into $Q(\mathbf{v})$ in (5.3) and $P_{X^n|V^n}(\mathbf{x}|\mathbf{v})$ in (5.8), respectively, to obtain

$$\begin{aligned}
&W_B^{n+}(\mathbf{y}|\mathbf{v}) \\
&\quad = \frac{1}{\gamma_n(\mathbf{v})} \sum_{\mathbf{x} \in \mathcal{X}^n(\Gamma)} W_B^n(\mathbf{y}|\mathbf{x}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) \\
&\quad = \frac{1}{\gamma_n(\mathbf{v})} \sum_{\mathbf{x} \in \mathcal{X}^n} W_B^n(\mathbf{y}|\mathbf{x}) \chi(\mathbf{x}) P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) \\
&\quad \le \frac{1}{\gamma_n(\mathbf{v})} \sum_{\mathbf{x} \in \mathcal{X}^n} W_B^n(\mathbf{y}|\mathbf{x}) \exp\left[(1+\rho)r\left(n\Gamma - \sum_{i=1}^{n} c(x_i)\right)\right] P_{X^n|V^n}(\mathbf{x}|\mathbf{v}) \\
&\quad \le \frac{1}{\beta_n} \sum_{\mathbf{x} \in \mathcal{X}^n} W_B^n(\mathbf{y}|\mathbf{x}) \exp\left[(1+\rho)r\left(n\Gamma - \sum_{i=1}^{n} c(x_i)\right)\right] P_{X^n|V^n}(\mathbf{x}|\mathbf{v})
\end{aligned} \tag{5.23}$$

and

$$\mathrm{E}_{\mathcal{C}}[\epsilon_n^B] \le \frac{1}{\alpha_n^{1+\rho}} (M_n L_n)^\rho \sum_{\mathbf{y}} \left(\sum_{\mathbf{v}} P_{V^n}(\mathbf{v}) W_B^{n+}(\mathbf{y}|\mathbf{v})^{\frac{1}{1+\rho}}\right)^{1+\rho}, \tag{5.24}$$

which together with (5.11) $\sim$ (5.13) yields, with $0 \le \rho \le 1$,

$$\begin{aligned}
&\mathrm{E}_{\mathcal{C}}[\epsilon_n^B] \\
&\quad \le \frac{1}{\alpha_n^{1+\rho}\beta_n} (M_n L_n)^\rho \\
&\qquad \cdot \left[\sum_{y \in \mathcal{Y}} \left(\sum_{v \in \mathcal{V}} q(v) \left[\sum_{x \in \mathcal{X}} W_B(y|x) P_{X|V}(x|v) e^{(1+\rho)r[\Gamma - c(x)]}\right]^{\frac{1}{1+\rho}}\right)^{1+\rho}\right]^n,
\end{aligned} \tag{5.25}$$

where we have put $q(v) = P_V(v)$ for simplicity.

Next, let us evaluate upper bound (5.4). In the way similar to the argument above with $-\rho$ in place of ρ, we obtain with $0 < \rho < 1$:

$$
\begin{aligned}
&\mathrm{E}_{\mathcal{C}}[\delta_n^E] \\
&\quad\le \frac{1}{\alpha_n^{1-\rho}\beta_n}\frac{1}{\rho L_n^{\rho}} \\
&\qquad \cdot \left[\sum_{z\in\mathcal{Z}}\left(\sum_{v\in\mathcal{V}} q(v)\left[\sum_{x\in\mathcal{X}} W_E(z|x)P_{X|V}(x|v)e^{(1-\rho)r[\Gamma-c(x)]}\right]^{\frac{1}{1-\rho}}\right)^{1-\rho}\right]^n .
\end{aligned}
\tag{5.26}
$$

Remark 5.2 (Two security functions) So far, we have established evaluation of upper bounds (5.3) and (5.4) when the channel (W_B^n, W_E^n) is stationary and memoryless. It should be noted, however, that we did not evaluate upper bound (5.5). This is because (5.5) contains the term $W_Q^n(\mathbf{z})$ with negative power $-\rho$, and hence upper bounding for (5.5) does not work. Thus, we prefer bound (5.4) rather than bound (5.5). □

Remark 5.3 As was shown, $\lim_{n\to\infty}\mu_n = 1$ and hence $\lim_{n\to\infty}\alpha_n = \lim_{n\to\infty}\beta_n = 1$, so that the normalizing probability μ_n does not effect the exponentially decreasing rates of ϵ_n^B, δ_n^E with increasing n. On the other hand, Gallager [16] used, instead of (5.22), upper bound

$$
\chi(\mathbf{x}) \le \exp\left[(1+\rho)r\left(\sum_{i=1}^{n} c(x_i) - n\Gamma + \delta\right)\right], \tag{5.27}
$$

where $\delta > 0$ is an arbitrary constant, and showed that, with this choice of upper bound, the normalizing probability μ_n satisfies $\lim_{n\to\infty}\sqrt{n}\mu_n = c$ (constant), and then here too μ_n does not effect the exponentially decreasing rates. We prefer upper bound (5.22) in this paper (except for in Theorems 9.2 and 9.4 later in Section 9), because it provides us with reasonable evaluation of the reliability and security functions for binary symmetric wiretap channels, Gaussian wiretap channels and also for Poisson wiretap channels to be treated in this section and in Sections 7, 9 and 10. □

Let us now give more compact forms to (5.25) and (5.26). To do so, define the *reliability function* $F_c(q, R_B, R_E, n)$ and the *security function* $H_c(q, R_E, n)$, with cost constraint, as

$$
F_c(q, R_B, R_E, n)
$$

$$\equiv \sup_{0\le\rho\le 1}\sup_{r\ge 0}\left(\phi(\rho|W_B,q,r)-\rho(R_B+R_E)+\frac{\log(\alpha_n\beta_n^{1+\rho})-\rho\log 3}{n}\right), \tag{5.28}$$

$$\begin{aligned} & H_c(q,R_E,n) \\ & \equiv \sup_{0<\rho<1}\sup_{r\ge 0}\left(\phi(-\rho|W_E,q,r)+\rho R_E+\frac{\log(\alpha_n\beta_n^{1-\rho})+\log\rho}{n}\right), \end{aligned} \tag{5.29}$$

where for fixed rates R_B, R_E we have set $M_n = e^{nR_B}, L_n = e^{nR_E}$, and

$$\begin{aligned} & \phi(\rho|W_B,q,r) \\ & = -\log\left[\sum_{y\in\mathcal{Y}}\left(\sum_{v\in\mathcal{V}}q(v)\left[\sum_{x\in\mathcal{X}}W_B(y|x)P_{X|V}(x|v)e^{(1+\rho)r[\Gamma-c(x)]}\right]^{\frac{1}{1+\rho}}\right)^{1+\rho}\right], \end{aligned} \tag{5.30}$$

$$\begin{aligned} & \phi(-\rho|W_E,q,r) \\ & = -\log\left[\sum_{z\in\mathcal{Z}}\left(\sum_{v\in\mathcal{V}}q(v)\left[\sum_{x\in\mathcal{X}}W_E(z|x)P_{X|V}(x|v)e^{(1-\rho)r[\Gamma-c(x)]}\right]^{\frac{1}{1-\rho}}\right)^{1-\rho}\right]. \end{aligned} \tag{5.31}$$

Thus, after some manipulations repeatedly using Markov inequality

$$\Pr\{W > 2\mathrm{E}(W)\} < 1/2, \quad \Pr\{W > 3\mathrm{E}(W)\} < 1/3,$$

we reach the following one of the main theorems:

Theorem 5.2 s Let (W_B^n, W_E^n) be stationary and memoryless, then there exists a pair (φ_n, ψ_n^B) of encoder and decoder such that

$$\mathrm{m}\text{-}\epsilon_n^B \le 6e^{-nF_c(q,R_B,R_E,n)}, \tag{5.32}$$
$$\mathrm{m}\text{-}\delta_n^E \le 6e^{-nH_c(q,R_E,n)}, \tag{5.33}$$

where $\sum_{x\in\mathcal{X}} P_X(x)c(x) \le \Gamma$ and $\mathrm{m}\text{-}\epsilon_n^B$, $\mathrm{m}\text{-}\delta_n^E$ are the maximum criteria as defined in Remark 3.8. □

Remark 5.4 Notice here that the cost constraint imposed on P_X was originally of the form $\sum_{x\in\mathcal{X}} P_X(x)c(x) \le \Gamma - \tau$ with an arbitrarily small number $\tau > 0$, which, on one hand, is rewritten equivalently as the cost constraint on $q = P_V$:

$$\sum_{v\in\mathcal{V}} \left[\sum_{x\in\mathcal{X}} c(x)P_{X|V}(x|v)\right] q(v) \le \Gamma - \tau \tag{5.34}$$

with fixed $P_{X|V}$. This implies that the cost constraint is linear and hence continuous in $q(v)$. On the other hand, the right-hand sides of (5.30) and (5.31) are also continuous in $q(x)$, e.g., under the L_a metric ($a > 1$). Therefore, we can let $\tau \to 0$ in (5.34) to yield the cost constraint $\sum_{x\in\mathcal{X}} P_X(x)c(x) \le \Gamma$, in that the functions $F_c(q, R_B, R_E, n)$ of $R_B + R_E$ and $H_c(q, R_E, n)$ of R_E are both convex closed (cf. Rockafeller [22]). □

Remark 5.5 In view of Remark 5.3 it should be noted that, in the brackets on the right-hand sides of (5.28) and (5.29), the third terms are of order $O(\frac{1}{n})$, approaching zero as n tends to ∞, so that these terms do not effect the exponents. □

Remark 5.6 (Reliability and security functions) Keeping Remark 3.6 in mind, we conclude that the reliability function $F_c(q, R_B, R_E, n)$ is monotone strictly decreasing and convex in $R_B + R_E$, whereas the security function $H_c(q, R_E, n)$ is monotone strictly increasing and convex in R_E. Moreover, as will be seen from the proof of Theorem 5.3, $F_c(q, R_B, R_E, +\infty)$ is nonnegative at $R_B + R_E = I(q, W_B^+)$, and $H_c(q, R_E, +\infty)$ is nonnegative at $R_E = I(q, W_E^+)$. Thus, $F_c(q, R_B, R_E, +\infty)$ is positive for $R_B+R_E < I(q, W_B^+)$, and $H_c(q, R_E, +\infty)$ is positive for $R_E > I(q, W_E^+)$. See Fig.7∼ Fig.11. Here too, we are considering, as an example, binary symmetric wiretap channels. As for the more capability of these wiretap channels, refer to the latter part of Remark 3.6.
Thus, it is concluded that rate $I(q, W_B^+) - I(q, W_E^+)$ is δ-achievable in the sense of Section 2 under cost constraint. This is because, for any small $\gamma > 0$, we can choose R_B, R_E so that $R_E = I(q, W_E^+) + \frac{\gamma}{2}$, $R_B + R_E = I(q, W_B^+) - \frac{\gamma}{2}$ and hence $R_B = I(q, W_B^+) - I(q, W_E^+) - \gamma$ for which $F_c(q, R_B, R_E, +\infty) > 0$, $H_c(q, R_E, +\infty) > 0$ and therefore both of ϵ_n^B and δ_n^E exponentially decay with increasing n. Furthermore, let q indicate the input maximizing $I(q, W_B^+) - I(q, W_E^+)$ under condition $\mathrm{E}c(X) \le \Gamma$, then we see with this q that $R_B = I(q, W_B^+) - I(q, W_E^+)$ gives the δ- secrecy capacity δ-$C_s(\Gamma)$ under cost constraint Γ. Here, in calculating the δ- secrecy capacity we invoke formula (5.44) in Theorem 5.4. □

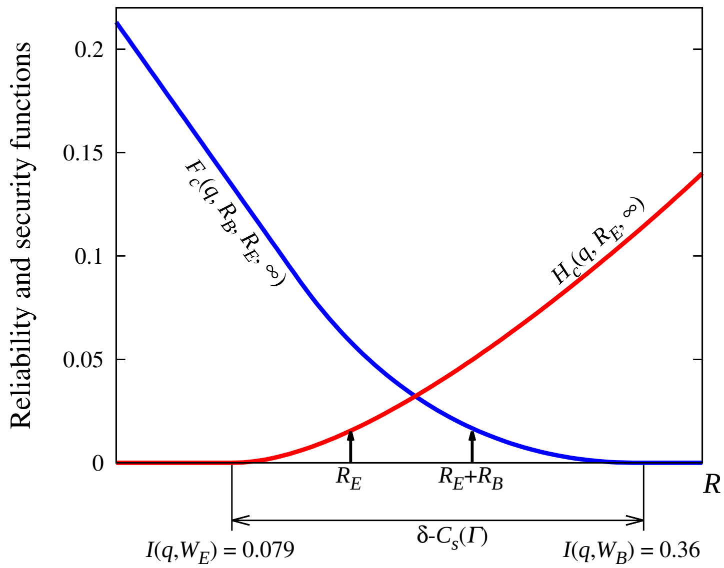

Figure 7: Reliability and security functions with cost constraint for non-concatenated BSC ($\varepsilon_y = 0.1$, $\varepsilon_z = 0.3$, $c(0) = 1$, $c(1) = 2$, $\Gamma = 1.4$).

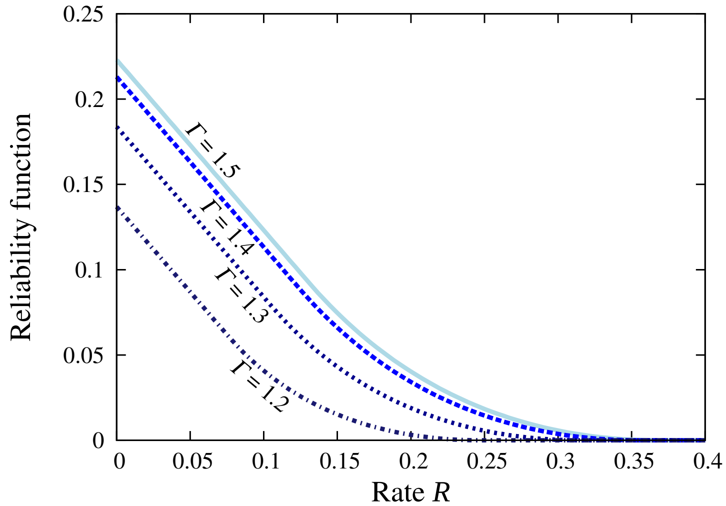

Figure 8: Reliability function for non-concatenated BSC with varied cost constraint Γ. Function curve moves upward as allowed cost Γ becomes large ($\varepsilon_y = 0.1$, $c(0) = 1$, $c(1) = 2$).

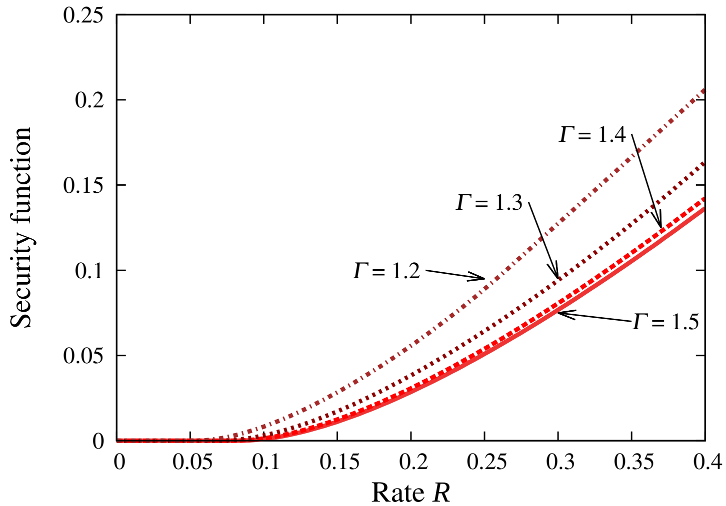

Figure 9: Security function for non-concatenated BSC with varied cost constraint Γ. Function curve moves downward as allowed cost Γ becomes large ($\varepsilon_z = 0.3$, $c(0) = 1$, $c(1) = 2$).

Remark 5.7 (Non-concatenation) It is sometimes useful to consider the special case with $V \equiv X$ as random variables over $\mathcal{V} = \mathcal{X}$. In this case the above quantities $\phi(\rho|W_B, q, r)$, $\phi(-\rho|W_E, q, r)$ reduce to

$$\phi(\rho|W_B, q, r) = -\log\left[\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\mathcal{X}} q(x)W_B(y|x)^{\frac{1}{1+\rho}}e^{r[\Gamma-c(x)]}\right)^{1+\rho}\right], \tag{5.35}$$

$$\phi(-\rho|W_E, q, r) = -\log\left[\sum_{z\in\mathcal{Z}}\left(\sum_{x\in\mathcal{X}} q(x)W_E(z|x)^{\frac{1}{1-\rho}}e^{r[\Gamma-c(x)]}\right)^{1-\rho}\right], \tag{5.36}$$

where the reliability functions (5.35) with $c(x)-\Gamma$ instead of $\Gamma-c(x)$ is earlier found in Gallager [16] and (5.35) with $c(x)-\Gamma$ instead of $\Gamma-c(x)$ applied to Poisson channels is found in Wyner [16], while the security function (5.36) intervenes for the first time in this paper. $\square$

We now can draw several preliminary insights into secrecy capacities with cost constraint. Some of them are the following two theorems. The first one is a strengthening of Theorem 4.2 (cf. also Remark 4.1):

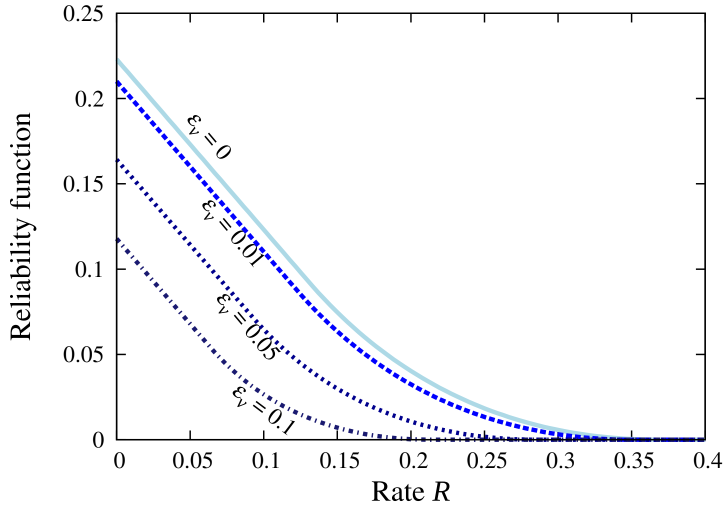

Figure 10: Reliability function with cost constraint for concatenated BSC with varied crossover probability ε_v. Function curve moves upward as ε_v becomes small ($\varepsilon_y = 0.1$, $c(0) = 1$, $c(1) = 2$, $\Gamma = 1.5$).

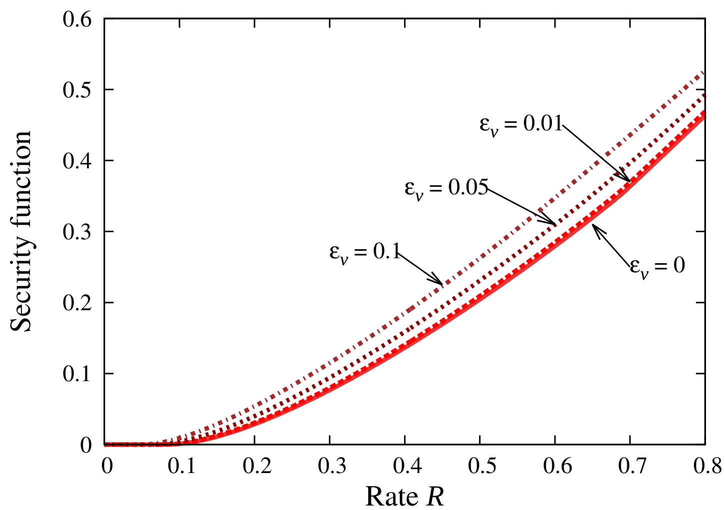

Figure 11: Security function with cost constraint for concatenated BSC with varied crossover probability ε_v. Function curve moves downward as ε_v becomes small ($\varepsilon_z = 0.3$, $c(0) = 1$, $c(1) = 2$, $\Gamma = 1.5$).

Theorem 5.3 If a wiretap channel (W_B^n, W_E^n) is stationary memoryless with *arbitrary* alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ (not necessarily finite), then

$$\delta\text{-}C_s(\Gamma) = \sup_{VX:\mathrm{E}c(X)\le\Gamma} (I(V;Y) - I(V;Z)) \tag{5.37}$$

under the maximum criterion ($\mathrm{m}\text{-}\epsilon_n^B, \mathrm{m}\text{-}\delta_n^E$). □

Proof:

Instead of (5.28) and (5.29), consider

$$\begin{aligned} F_{c0}(q, R_B, R_E, r) &\equiv \sup_{0\le\rho\le1} \left(\phi(\rho|W_B, q, r) - \rho(R_B + R_E)\right), &(5.38)\\ H_{c0}(q, R_E, r) &\equiv \sup_{0<\rho<1} \left(\phi(-\rho|W_E, q, r) + \rho R_E\right), &(5.39)\end{aligned}$$

then

$$\begin{aligned} F_{c0}(q, R_B, R_E, r=0) &\equiv \sup_{0\le\rho\le1} \left(\phi(\rho|W_B, q, r=0) - \rho(R_B + R_E)\right),\\ &= \sup_{0\le\rho\le1} \left(\phi(\rho|W_B^+, q) - \rho(R_B + R_E),\right) &(5.40)\\ H_{c0}(q, R_E, r=0) &\equiv \sup_{0<\rho<1} \left(\phi(-\rho|W_E, q, r=0) + \rho R_E\right)\\ &= \sup_{0<\rho<1} \left(\phi(-\rho|W_E^+, q) + \rho R_E\right), &(5.41)\end{aligned}$$

which are nothing but $F(q, R_B, R_E, +\infty)$ in (3.20) and $H(q, R_E, +\infty)$ in (3.22). Therefore,

$$\begin{aligned} F_c(q, R_B, R_E, +\infty) &\ge F_{c0}(q, R_B, R_E, r=0) = F(q, R_B, R_E, +\infty), &(5.42)\\ H_c(q, R_E, +\infty) &\ge H_{c0}(q, R_E, r=0) = H(q, R_E, +\infty). &(5.43)\end{aligned}$$

Now let P_{VX} denote the probability distribution of VX to achieve the supremum on the right-hand side of (5.37) and set $q = P_V$. Then, as was shown in Remark 3.6, it turns out that

$$I(q, W_B^+) - I(q, W_E^+) = \sup_{VX:\mathrm{E}c(X)\le\Gamma} (I(V;Y) - I(V;Z))$$

is δ-achievable in the sense of Section 2, where it is obvious that

$$\sum_{x\in\mathcal{X}} P_X(x)c(x) \le \Gamma.$$

Theorem 5.2 along with (5.42), (5.43) guarantees that this is δ-achievable under the maximum criterion $(\text{m-}\epsilon_n^B, \text{m-}\delta_n^E)$. $\square$

The second one is a strengthening of the latter part of Theorem 4.3:

Theorem 5.4 If a wiretap channel (W_B^n, W_E^n) is stationary memoryless with *arbitrary* alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ (not necessarily finite) and more capable, then

$$\delta\text{-}C_s(\Gamma) = \sup_{X:\mathrm{E}c(X)\le\Gamma} (I(X;Y) - I(X;Z)) \tag{5.44}$$

under the maximum criterion $(\text{m-}\epsilon_n^B, \text{m-}\delta_n^E)$. $\square$

Proof:

Consider the case with $V \equiv X$, then, by means of Remark 5.7, instead of (5.40) and (5.39) we have

$$\begin{aligned} F_{c0}(q, R_B, R_E, r=0) &\equiv \sup_{0\le\rho\le1} \left(\phi(\rho|W_B, q, r=0) - \rho(R_B + R_E)\right), \\ &= \sup_{0\le\rho\le1} \left(\phi(\rho|W_B, q) - \rho(R_B + R_E),\right) && (5.45) \\ H_{c0}(q, R_E, r=0) &\equiv \sup_{0<\rho<1} \left(\phi(-\rho|W_E, q, r=0) + \rho R_E\right) \\ &= \sup_{0<\rho<1} \left(\phi(-\rho|W_E, q) + \rho R_E\right), && (5.46) \end{aligned}$$

where

$$\begin{aligned} \phi(\rho|W_B, q) &= -\log\left[\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\mathcal{X}} q(x) W_B(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho}\right], && (5.47) \\ \phi(-\rho|W_E, q) &= -\log\left[\sum_{z\in\mathcal{Z}}\left(\sum_{x\in\mathcal{X}} q(x) W_E(z|x)^{\frac{1}{1-\rho}}\right)^{1-\rho}\right]. && (5.48) \end{aligned}$$

The rest is similar to the argument in the proof of Theorem 5.3. $\square$

Remark 5.8 (Equivalence of two cost constraints) As was discussed in Remark 5.4, since it holds that

$$\sum_{x\in\mathcal{X}} P_X(x)c(x) = \sum_{v\in\mathcal{V}} P_V(v)\overline{c}(v), \tag{5.49}$$

where

$$\overline{c}(v) = \sum_{x\in\mathcal{X}} c(x) P_{X|V}(x|v), \tag{5.50}$$

we see that $\mathrm{E}c(X) = \mathrm{E}\overline{c}(V)$, and hence $\sum_{x\in\mathcal{X}} P_X(x)c(x) \le \Gamma - \tau$ is equivalent to $\sum_{v\in\mathcal{V}} P_V(v)\overline{c}(v) \le \Gamma - \tau$, where $\tau > 0$ is an arbitrarily small number. Therefore, with condition (5.14), it is concluded again by virtue of the weak law of large numbers that, according to (5.16),

$$\overline{\mu}_n \stackrel{\Delta}{=} \sum_{\mathbf{v}} \overline{\chi}(\mathbf{v}) \prod_{i=1}^{n} P_V(v_i) \to 1 \quad (n \to +\infty),$$

where $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ and

$$\overline{\chi}(\mathbf{v}) = \begin{cases} 1 & \text{for } \sum_{i=1}^{n} \overline{c}(v_i) \le n\Gamma, \\ \\ 0 & \text{otherwise,} \end{cases} \tag{5.51}$$

so that $\mathcal{T}_1$ in the above can be replaced by $\mathcal{T}_1 \cap \{\mathbf{v} \in \mathcal{V}^n | \overline{\chi}(\mathbf{v}) = 1\}$ without any change of the conclusions (5.25) and (5.26). This observation means that cost constraint Γ (with cost $c(x)$) on X^n of the concatenated channel (W_B^{n+}, W_E^{n+}) is tantamount to cost constraint Γ (with cost $\overline{c}(v)$) on the input V^n of (W_B^{n+}, W_E^{n+}), and vice versa. Thus, introducing the upper bound

$$\overline{\chi}(\mathbf{v}) \le \exp\left[s \left(n\Gamma - \sum_{i=1}^{n} \overline{c}(v_i) \right) \right], \tag{5.52}$$

where $s \ge 0$ is an arbitrary number, we can strengthen upper bounds (5.25) and (5.26) as follows:

$$\begin{aligned} &\mathrm{E}_{\mathcal{C}}[\epsilon_n^B] \\ &\quad \le \frac{1}{\alpha_n^{1+\rho}\beta_n} (M_n L_n)^{\rho} \\ &\qquad \cdot \left[\sum_{y\in\mathcal{Y}} \left(\sum_{v\in\mathcal{V}} q(v) e^{s[\Gamma - \overline{c}(v)]} \left[\sum_{x\in\mathcal{X}} W_B(y|x) P_{X|V}(x|v) e^{(1+\rho)r[\Gamma - c(x)]} \right]^{\frac{1}{1+\rho}} \right)^{1+\rho} \right]^n, \end{aligned} \tag{5.53}$$

$$\begin{aligned} &\mathrm{E}_{\mathcal{C}}[\delta_n^E] \\ &\quad \le \frac{1}{\alpha_n^{1-\rho}\beta_n} \frac{1}{\rho L_n^{\rho}} \\ &\qquad \cdot \left[\sum_{z\in\mathcal{Z}} \left(\sum_{v\in\mathcal{V}} q(v) e^{s[\Gamma - \overline{c}(v)]} \left[\sum_{x\in\mathcal{X}} W_E(z|x) P_{X|V}(x|v) e^{(1-\rho)r[\Gamma - c(x)]} \right]^{\frac{1}{1-\rho}} \right)^{1-\rho} \right]^n, \end{aligned} \tag{5.54}$$

respectively. Accordingly, (5.30) and (5.31) are replaced by

$$\begin{aligned}
&\phi(\rho|W_B, q, s, r) \\
&= -\log\left[\sum_{y\in\mathcal{Y}}\left(\sum_{v\in\mathcal{V}} q(v)e^{s[\Gamma-\overline{c}(v)]}\left[\sum_{x\in\mathcal{X}} W_B(y|x)P_{X|V}(x|v)e^{(1+\rho)r[\Gamma-c(x)]}\right]^{\frac{1}{1+\rho}}\right)^{1+\rho}\right],
\end{aligned} \tag{5.55}$$

$$\begin{aligned}
&\phi(-\rho|W_E, q, s, r) \\
&= -\log\left[\sum_{z\in\mathcal{Z}}\left(\sum_{v\in\mathcal{V}} q(v)e^{s[\Gamma-\overline{c}(v)]}\left[\sum_{x\in\mathcal{X}} W_E(z|x)P_{X|V}(x|v)e^{(1-\rho)r[\Gamma-c(x)]}\right]^{\frac{1}{1-\rho}}\right)^{1-\rho}\right].
\end{aligned} \tag{5.56}$$

Although Theorem 5.2 with (5.28), (5.29) (where $\sup_{r\geq 0}$ is replaced by $\sup_{s\geq 0, r\geq 0}$) and with (5.55), (5.56) gives the performance better than or equal to either of the original version without the term $e^{s[\Gamma-\overline{c}(v)]}$ and the version without the terms $e^{(1+\rho)r[\Gamma-c(x)]}$ and $e^{(1-\rho)r[\Gamma-c(x)]}$, we do not go into the details here. However, it should be noted that these two versions as well as the version with both terms give the same reliability and security functions with the maximizing parameters $s = r = 0$ in either case of symmetrically concatenated (i.e., with $P_{X|V}(1|0) = P_{X|V}(0|1)$) BSC's (see Fig.10 and Fig.11) and concatenated Poisson channels (see Section 10). On the other hand, numerical examples show that in the case of asymmetrically concatenated (i.e., with $P_{X|V}(1|0) \neq P_{X|V}(0|1)$) BSC's, the reliability and security functions for the version with both terms allowed actually reduces (with the maximizing parameters that may be $s > 0 \wedge r = 0$ or $r > 0 \wedge s = 0$; e.g., see Fig.12) to either those for the version only with the term $e^{s[\Gamma-\overline{c}(v)]}$ or those for the version only with the term $e^{(1\pm\rho)r[\Gamma-c(x)]}$, depending on the values of transition probabilities of the concatenated channel. □

6 Secrecy capacity of Poisson wiretap channel

In this section, we consider application of Theorem 5.4 to the Poisson wiretap channel to determine its secrecy capacity. First of all, let us define the Poisson wiretap channel (cf. [23], [25], [6]). The input process to the Poisson channel is a waveform denoted by X_t $(0 \leq t \leq T)$ satisfying $X_t \geq 0$ for all t, where T

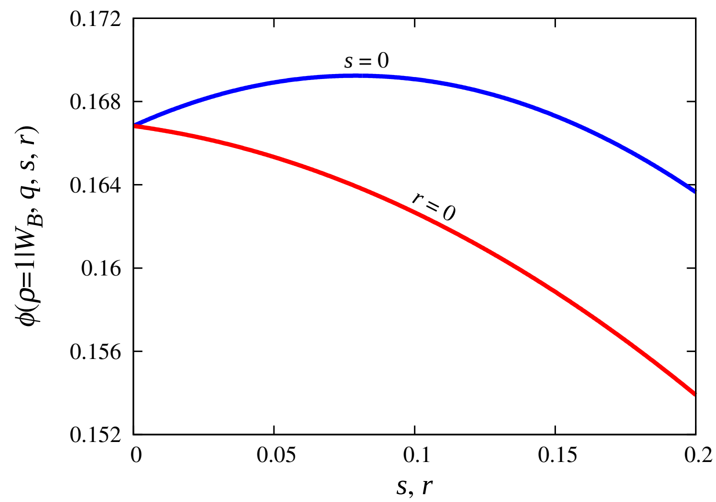

Figure 12: Reliability exponent for asymmetrically concatenated BSC with varied s, r; $P_{X|V}(1|0) = 0.1, P_{X|V}(0|1) = 0$; $P_{Y|X}(1|0) = P_{Y|X}(0|1) = 0.1$.

is an arbitrarily large time span. We assume that the input process is not only peak power limited, i.e., $0 \leq X_t \leq 1$ for all t but also average power limited, i.e.,

$$\frac{1}{T}\int_0^T X_t dt \leq \Gamma \quad (0 \leq \Gamma \leq 1). \tag{6.1}$$

The output signal to be received by the legitimate receiver Bob is a Poisson counting process Y_t $(0 \leq t \leq T)$ with instantaneous rate $A_y X_t + \lambda_y$ ($\lambda_y \geq 0$ is the dark current, and $A_y > 0$ specifies attenuation of signal) such that

$$Y_{t=0} = 0, \tag{6.2}$$

and, for $0 \leq t, t+\tau \leq T$ $(\tau > 0)$,

$$\Pr\{Y_{t+\tau} - Y_t = j\} = \frac{e^{-\Lambda}\Lambda^j}{j!} \quad (j = 0, 1, 2, \cdots), \tag{6.3}$$

where

$$\Lambda = \int_t^{t+\tau} (A_y X_t + \lambda_y) dt. \tag{6.4}$$

Similarly, the output signal to be received by the eavesdropper Eve is a Poisson counting process Z_t $(0 \leq t \leq T)$ with instantaneous rate $A_z X_t + \lambda_z$.

We now want to discretize the continuous time process like this into a discrete time process in order to make the problem more tractable with asymptotically negligible loss of performance. To do so, we follow the way that Wyner [23] has demonstrated, and for the reader's convenience we review here his formulation to be exact. Let $\Delta > 0$ be an arbitrary very small constant. Then, we assume the following.

a) The channel input X_t is constant for $(i-1)\Delta < t \le i\Delta$ $(i = 1, 2, \cdots)$, and X_t takes only the values 0 or 1. For $i = 1, 2, \cdots$, define as $x_i = 0$ or 1 according as $X_t = 0$ or 1 in the interval $((i-1)\Delta, i\Delta]$.
b) Bob observes only the samples $Y_{i\Delta}$ $(i = 1, 2, \cdots)$, and define as $y_i = 1$ if $Y_{i\Delta} - Y_{(i-1)\Delta} = 1$; $y_i = 0$ otherwise.
c) Eve observes only the samples $Z_{i\Delta}$ $(i = 1, 2, \cdots)$, and define as $z_i = 1$ if $Z_{i\Delta} - Z_{(i-1)\Delta} = 1$; $z_i = 0$ otherwise.

Owing to the discretization under assumptions a), b), c), we have two channels W_B, W_E for Bob and Eve, respectively, i.e., two-input two-output stationary memoryless discrete channels (DMC) such as $W_B : x_i \to y_i$ and $W_E : x_i \to z_i$, whose transition probabilities are given, up to the first order, as

$$\begin{aligned} W_B(1|0) &= \lambda_y \Delta e^{-\lambda_y \Delta} \\ &\simeq \lambda_y \Delta = s_y A_y \Delta, \qquad (6.5) \\ W_B(1|1) &= (A_y + \lambda_y)\Delta e^{-(A_y+\lambda_y)\Delta} \\ &\simeq (A_y + \lambda_y)\Delta = A_y(1+s_y)\Delta; \qquad (6.6) \end{aligned}$$

$$\begin{aligned} W_E(1|0) &= \lambda_z \Delta e^{-\lambda_z \Delta} \\ &\simeq \lambda_z \Delta = s_z A_z \Delta, \qquad (6.7) \\ W_E(1|1) &= (A_z + \lambda_z)\Delta e^{-(A_z+\lambda_z)\Delta} \\ &\simeq (A_z + \lambda_z)\Delta = A_z(1+s_z)\Delta, \qquad (6.8) \end{aligned}$$

where we have put

$$s_y = \frac{\lambda_y}{A_y}, \quad s_z = \frac{\lambda_z}{A_z}. \qquad (6.9)$$

Furthermore, a given fixed constant $\Delta > 0$ small enough, define the whole time interval $T = n\Delta$, where n denotes the block length of the DMC. Then, the power constraint (6.1) is equivalent to

$$\frac{1}{n}\sum_{i=1}^{n} c(x_i) \le \Gamma, \qquad (6.10)$$

where the additive cost $c(x)$ is defined as $c(x) = x$ for $x = 0, 1$. We are now almost ready to apply Theorem 5.2 and Theorem 5.4 to find secrecy capacities and reliability/security functions.

However, since Theorem 5.4 holds only for more capable channels, we need to impose some restriction on the class of wiretap channels as above formulated. In this connection, we introduce the concept of degradedness of channels as follows:

Definition 6.1 ([14]) A wiretap channel (W_B, W_E) is said to be (statistically) *degraded* ¶ if there exists an auxiliary channel $V : \mathcal{Y} \to \mathcal{Z}$ such that

$$W_E(z|x) = \sum_{y \in \mathcal{Y}} W_B(y|x) V(z|y). \tag{6.11}$$

In this connection, we have the following theorems:

Theorem 6.1 ([25], [6]) A wiretap channel is degraded if

$$A_y \geq A_z \tag{6.12}$$

and

$$\frac{\lambda_y}{A_y} \leq \frac{\lambda_z}{A_z}. \tag{6.13}$$

Theorem 6.2 ([14]) A wiretap channel is more capable if it is degraded. □

Thus, in the sequel, we confine ourselves to the class of wiretap channels satisfying (6.12) and (6.13) to guarantee application of Theorem 5.4, where we assume that at least one of them holds with strict inequality; otherwise the problem is trivial.

With these preparations, we now go to the problem of determining the secrecy capacity. Let X be a channel input, and Y, Z be the channel output via W_B, W_E, respectively, due to X. Following Wyner [23] with $q = \Pr\{X = 1\}$, we directly compute the mutual informations to have

$$\begin{aligned} I(X;Y) &= \Delta A_y[-(q+s_y)\log(q+s_y) + q(1+s_y)\log(1+s_y) \\ &\quad +(1-q)s_y \log s_y] \stackrel{\Delta}{=} f(q), \end{aligned} \tag{6.14}$$

¶More exactly, we should say that channel W_E is degraded than channel W_B. Here, with abuse of notation, we simply say that (W_B, W_E) is degraded.

$$I(X;Z) = \Delta A_z[-(q+s_z)\log(q+s_z) + q(1+s_z)\log(1+s_z) + (1-q)s_z\log s_z] \triangleq g(q), \quad (6.15)$$

$$\sigma(q) \triangleq f(q) - g(q). \quad (6.16)$$

Then, it is evident that

$$\sigma(0) = \sigma(1) = 0. \quad (6.17)$$

Moreover,

$$\sigma''(q) = -\frac{\Delta A_y}{q+s_y} + \frac{\Delta A_z}{q+s_z} < 0, \quad (6.18)$$

where the inequality follows from (6.12) and (6.13). Therefore, $\sigma(q)$ is strictly concave and takes the maximum value at the unique $q = q^*$ in the interval $(0,1)$ with $\sigma'(q^*) = 0$. Thus, we have the following one of the main results.

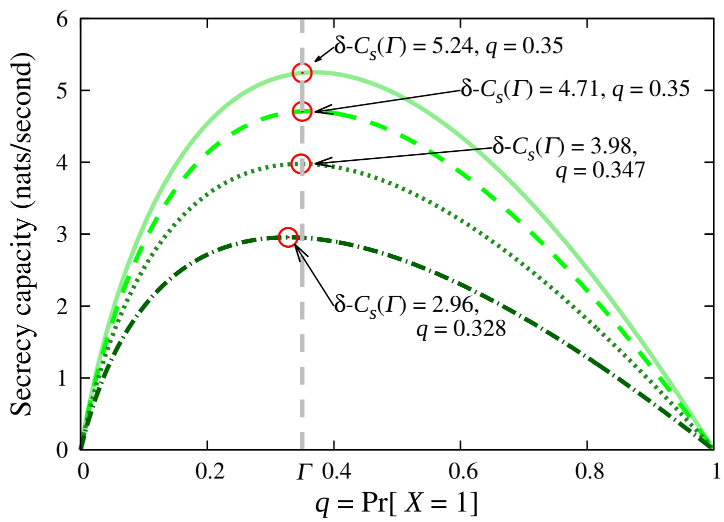

Figure 13: Secrecy capacity for Poisson channel with varied Eve's channel; solid line corresponds to $A_z = 12$, $\lambda_z = 10$; dashed line corresponds to $A_z = 14$, $\lambda_z = 8$; dotted line corresponds to $A_z = 16$, $\lambda_z = 6$; single dotted line corresponds to $A_z = 18$, $\lambda_z = 4$ ($A_y = 20$, $\lambda_y = 2$, $\Gamma = 0.5$).

Theorem 6.3 The δ-secrecy capacity with cost constraint δ-$C_s(\Gamma)$ per second of the wiretap channel (W_B, W_E) is given by

$$\delta\text{-}C_s(\Gamma)$$

$$
\begin{aligned}
= \quad & \log \frac{(q^*_\Gamma + s_z)^{(q^*_\Gamma + s_z)A_z}}{(q^*_\Gamma + s_y)^{(q^*_\Gamma + s_y)A_y}} + \log \frac{s_y^{s_y A_y}}{s_z^{s_z A_z}} \\
& + q^*_\Gamma \left(\log \frac{(q^* + s_y)^{A_y}}{(q^* + s_z)^{A_z}} + A_y - A_z \right) \qquad (6.19)
\end{aligned}
$$

under the maximum criterion (m-ϵ_n^B, m-δ_n^E), where $q = q^*$ is the unique solution in $(0, 1)$ of the equation:

$$
\frac{(A_y q^* + \lambda_y)^{A_y}}{(A_z q^* + \lambda_z)^{A_z}} = e^{A_z - A_y} \frac{(A_y + \lambda_y)^{A_y + \lambda_y}}{(A_z + \lambda_z)^{A_z + \lambda_z}} \frac{\lambda_z^{\lambda_z}}{\lambda_y^{\lambda_y}}, \qquad (6.20)
$$

and

$$
q^*_\Gamma = \min(q^*, \Gamma). \qquad (6.21)
$$

Proof :

We develop $\sigma(q)$ in (6.16) as follows:

$$
\begin{aligned}
\sigma(q) \quad = \quad & \Delta A_y[-(q + s_y)\log(q + s_y) \\
& + q(1 + s_y)\log(1 + s_y) + (1 - q)s_y \log s_y] \\
& + \Delta A_z[(q + s_z)\log(q + s_z) \\
& - q(1 + s_z)\log(1 + s_z) - (1 - q)s_z \log s_z]. \qquad (6.22)
\end{aligned}
$$

Then, a direct computation shows that

$$
\begin{aligned}
\sigma'(q) \quad = \quad & \Delta A_y[-\log(q + s_y) - 1 \\
& + (1 + s_y)\log(1 + s_y) - s_y \log s_y] \\
& + \Delta A_z[\log(q + s_z) + 1 \\
& - (1 + s_z)\log(1 + s_z) + s_z \log s_z] \\
= \quad & \Delta \left[(A_z - A_y) - \log \frac{(q + s_y)^{A_y}}{(q + s_z)^{A_z}} \right. \\
& \left. + \log \frac{(1 + s_y)^{(1+s_y)A_y}}{(1 + s_z)^{(1+s_z)A_z}} - \log \frac{s_y^{s_y A_y}}{s_z^{s_z A_z}} \right]. \qquad (6.23)
\end{aligned}
$$

Hence, the solution $q = q^*$ of the equation $\sigma'(q) = 0$ is given by

$$
\begin{aligned}
\log \frac{(q^* + s_y)^{A_y}}{(q^* + s_z)^{A_z}} \quad = \quad & (A_z - A_y) + \log \frac{(1 + s_y)^{(1+s_y)A_y}}{(1 + s_z)^{(1+s_z)A_z}} \\
& - \log \frac{s_y^{s_y A_y}}{s_z^{s_z A_z}}, \qquad (6.24)
\end{aligned}
$$

which is equivalent to

$$\frac{(A_y q^* + \lambda_y)^{A_y}}{(A_z q^* + \lambda_z)^{A_z}} = e^{A_z - A_y} \frac{(A_y + \lambda_y)^{A_y + \lambda_y}}{(A_z + \lambda_z)^{A_z + \lambda_z}} \frac{\lambda_z^{\lambda_z}}{\lambda_y^{\lambda_y}}. \tag{6.25}$$

On the other hand,

$$\begin{aligned}
\sigma(q) &= \Delta \log \frac{(q + s_z)^{(q+s_z)A_z}}{(q + s_y)^{(q+s_y)A_y}} \\
&\quad + \Delta q \log \frac{(1 + s_y)^{(1+s_y)A_y}}{(1 + s_z)^{(1+s_z)A_z}} \\
&\quad + \Delta (1 - q) \log \frac{s_y^{s_y A_y}}{s_z^{s_z A_z}} \\
&= \Delta \log \frac{(q + s_z)^{(q+s_z)A_z}}{(q + s_y)^{(q+s_y)A_y}} + \Delta \log \frac{s_y^{s_y A_y}}{s_z^{s_z A_z}} \\
&\quad + \Delta q \left(\log \frac{(1 + s_y)^{(1+s_y)A_y}}{(1 + s_z)^{(1+s_z)A_z}} - \log \frac{s_y^{s_y A_y}}{s_z^{s_z A_z}} \right) \\
&= \Delta \log \frac{(q + s_z)^{(q+s_z)A_z}}{(q + s_y)^{(q+s_y)A_y}} + \Delta \log \frac{s_y^{s_y A_y}}{s_z^{s_z A_z}} \\
&\quad + \Delta q \left(\log \frac{(q^* + s_y)^{A_y}}{(q^* + s_z)^{A_z}} + A_y - A_z \right),
\end{aligned} \tag{6.26}$$

where we used (6.24) in the last step. Consequently, with $q_\Gamma^* = \min(q^*, \Gamma)$,

$$\begin{aligned}
&\max_{X : \mathrm{E}c(X) \le \Gamma} (I(X;Y) - I(X;Z)) \\
&= \max_{0 \le q \le \Gamma} (I(X;Y) - I(X;Z)) \\
&= \Delta \log \frac{(q_\Gamma^* + s_z)^{(q_\Gamma^* + s_z)A_z}}{(q_\Gamma^* + s_y)^{(q_\Gamma^* + s_y)A_y}} + \Delta \log \frac{s_y^{s_y A_y}}{s_z^{s_z A_z}} \\
&\quad + \Delta q_\Gamma^* \left(\log \frac{(q^* + s_y)^{A_y}}{(q^* + s_z)^{A_z}} + A_y - A_z \right).
\end{aligned} \tag{6.27}$$

Since Theorem 5.4 claims that the left-hand side of (6.27) gives the δ-secrecy capacity per channel use, it is concluded that the δ-secrecy capacity δ-$C_s(\Gamma)$ per second is given by (6.19). □

Remark 6.1 It is easy to check that, in the special case *without* cost constraint (i.e., $\Gamma = 1$ and hence $q^*_\Gamma = q^*$), (6.19) reduces to

$$\begin{aligned} \delta\text{-}C_s(1) &= q^*(A_y - A_z) + \log \frac{\lambda_y^{\lambda_y}}{\lambda_z^{\lambda_z}} \\ &\quad + \log \frac{(A_z q^* + \lambda_z)^{\lambda_z}}{(A_y q^* + \lambda_y)^{\lambda_y}}. \end{aligned} \tag{6.28}$$

A similar but weak version w-C_s *without* cost constraint Γ of this formula is developed in Laourine and Wagner [6], that is.

$$\begin{aligned} \text{w-}C_s &= q^*(A_y - A_z) + \log \frac{\lambda_y^{\lambda_y}}{\lambda_z^{\lambda_z}} \\ &\quad + \log \frac{(A_z q^* + \lambda_z)^{\lambda_z}}{(A_y q^* + \lambda_y)^{\lambda_y}} \end{aligned} \tag{6.29}$$

with the same equation as (6.20). As for the definition of δ-$C_s(\Gamma)$ and w-C_s, see Sections 2 and 4. It should be remarked here that, from the standpoint of security, formula (6.28) is much stronger than formula (6.29) as was discussed in Section 1.B. Indeed, what we wanted to obtain in this paper is just of the type (6.28) but not of the type (6.29). □

Example 6.1 Let us quote here the *worst case scenario* as demonstrated in [6], which is a particularly insightful case (due to thinning by Eve) specified by

$$\frac{\lambda_y}{A_y} = \frac{\lambda_z}{A_z} = s.$$

In this case, it is easy to see (cf. [6]) that q^* is given by

$$q^* = \frac{(1+s)^{1+s}}{es^s} - s. \tag{6.30}$$

It is then also easy to verify that (6.19) reduces to

$$\delta\text{-}C_s(\Gamma) = (A_y - A_z)\left[\begin{array}{r} -(q_\Gamma + s)\log(q_\Gamma + s) + s\log s \\ +q_\Gamma\left[(1+s)\log(1+s) - s\log s\right] \end{array}\right], \tag{6.31}$$

where

$$q_\Gamma = \min\left(\frac{(1+s)^{1+s}}{es^s} - s, \Gamma\right).$$

Formula (6.31) with $\Gamma = 1$ (no cost constraint) yields a strong secrecy version of the formula given in [6], i.e.,

$$\delta\text{-}C_s(1) = (\lambda_y - \lambda_z)\left(\frac{1}{e}\left(1 + \frac{1}{s}\right)^{1+s} - (1+s)\log\left(1 + \frac{1}{s}\right)\right).$$

Moreover, in the particular case with $s = 0$ (no dark current), (6.31) reduces to

$$\delta\text{-}C_s(\Gamma) = -(A_y - A_z)q_\Gamma \log q_\Gamma, \tag{6.32}$$

where

$$q_\Gamma = \min\left(\frac{1}{e}, \Gamma\right),$$

from which it follows that

$$\delta\text{-}C_s(1) = \frac{A_y - A_z}{e},$$

which was shown in [6] with the weak version w-C_s instead of δ-$C_s(1)$. Formulas (6.31) and (6.32) with $A_z = 0$ (absence of the eavesdropper) are found earlier in Wyner [23]. □

7 Reliability and security functions of Poisson wiretap channel

In this section, we consider application of Theorem 5.2 to the Poisson wiretap channel to evaluate its reliability and security functions. Here too, as in the previous section, we use the same two-input two-output stationary memoryless channel model specified with the transition probabilities and the cost constraint with parameters (6.5) $\sim$ (6.10). In this section we focus on wiretap channels *without* concatenation (i.e., $V \equiv X$; cf. Remark 5.7), and later in Section 10 extend it to the case of wiretap channels *with* concatenation. Also, we assume that the conditions for degradedness (6.12) and (6.13) in Theorem 6.1 are satisfied.

A. Reliability function

The first concern in this section is on the behavior of the reliability function for Bob. Formula (5.32) of Theorem 5.2 is written as

$$\begin{aligned}
\text{m-}\epsilon_n^B &\leq 6e^{-nF_c(q,R_{B0},R_{E0},n)} \\
&= \exp[-n \sup_{0\leq\rho\leq 1} \sup_{r\geq 0} (E_{B0}(\rho,q,r) - \rho(R_{B0}+R_{E0}) + O(1/n))] \\
&= \exp[-n \sup_{0\leq\rho\leq 1} \sup_{r\geq 0} (E_{B0}(\rho,q,r) - \rho(R_{B0}+R_{E0})) + O(1)], \quad (7.1)
\end{aligned}$$

where we have set $E_{B0}(\rho, q, r) = \phi(\rho|W_B, q, r)$. Let us first evaluate $E_{B0}(\rho, q, r)$. Taking account of (5.35), we have

$$
\begin{aligned}
E_{B0}(\rho, q, r) &= -\log\left[\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\mathcal{X}} q(x)W_B(y|x)^{\frac{1}{1+\rho}} e^{r[\Gamma - c(x)]}\right)^{1+\rho}\right] \\
&= -\log\left[\sum_{y=0}^{1}\left(\sum_{x=0}^{1} q(x)W_B(y|x)^{\frac{1}{1+\rho}} e^{r[\Gamma - c(x)]}\right)^{1+\rho}\right] \\
&= -\log\sum_{y=0}^{1} V_y^{1+\rho} - r(1+\rho)\Gamma, \qquad (7.2)
\end{aligned}
$$

where

$$
V_y = \sum_{x=0}^{1} q(x)W_B(y|x)^{\frac{1}{1+\rho}} e^{-rx} \quad (y = 0, 1).
$$

(It should be noted here that in evaluation of (7.2) Wyner [23] used $c(x) - \Gamma$ instead of $\Gamma - c(x)$, which causes some subtle irrelevance.) With $q = q(1)$, an elementary caluculation using (6.5) and (6.6) leads, up to the first order, to

$$
\begin{aligned}
E_{B0}(\rho, q, r) &= -r(1+\rho)\Gamma - (1+\rho)\log(1 - q + qe^{-r}) \\
&\quad +\Delta A_y\left[\frac{(1-q)s_y + qe^{-r}(1+s_y)}{1-q+qe^{-r}}\right] \\
&\quad -\Delta A_y\left[\frac{(1-q)s_y^{\frac{1}{1+\rho}} + qe^{-r}(1+s_y)^{\frac{1}{1+\rho}}}{1-q+qe^{-r}}\right]^{1+\rho}. \qquad (7.3)
\end{aligned}
$$

First, in order to maximize $E_{B0}(\rho, q, r)$ with respect to r, set

$$
g(r) = -r(1+\rho)\Gamma - (1+\rho)\log(1 - q + qe^{-r}).
$$

Then,

$$
g'(r) = -(1+\rho)\Gamma + (1+\rho)\frac{qe^{-r}}{1-q+qe^{-r}},
$$

$$
g''(r) = -(1+\rho)\frac{q^2e^{-r}(1-q)}{(1-q+qe^{-r})^2} < 0,
$$

which means that $g(r)$ is strictly concave. It is evident that

$$g(0) = 0, \quad g'(0) = -(1+\rho)(\Gamma - q) \le 0,$$

where we have used that cost constraint $\mathrm{E}c(X) \le \Gamma$ is written as $q \le \Gamma$. Consequently, we have

$$\max_{r \ge 0} g(r) = g(0) = 0, \tag{7.4}$$

and hence

$$\begin{aligned} E_{B0}(\rho, q) &\stackrel{\Delta}{=} \max_{r\ge 0} E_{B0}(\rho, q, r) \\ &= \Delta A_y \left[(1-q)s_y + (1+s_y)q - \left[(1-q)s_y^{\frac{1}{1+\rho}} + q(1+s_y)^{\frac{1}{1+\rho}} \right]^{1+\rho} \right]. \\ &= \Delta A_y \left[q + s_y - s_y(1+\tau_y q)^{1+\rho} \right], \end{aligned} \tag{7.5}$$

where

$$\tau_y = \left(1 + \frac{1}{s_y}\right)^{\frac{1}{1+\rho}} - 1. \tag{7.6}$$

On the other hand, (7.1) is rewritten as

$$\text{m-}\epsilon_n^B \le \exp[-n \sup_{0\le\rho\le 1} (E_{B0}(\rho, q) - \rho(R_{B0} + R_{E0})) + O(1)]. \tag{7.7}$$

Notice here that $E_{B0}(\rho, q) - \rho(R_{B0} + R_{E0})$ in (7.7) is the exponent per channel use, so that

$$\frac{E_{B0}(\rho, q) - \rho(R_{B0} + R_{E0})}{\Delta}$$

gives the exponent per second. Therefore,

$$E_B(\rho, q) = \frac{E_{B0}(\rho, q)}{\Delta}, \quad R_B = \frac{R_B B0}{\Delta}, \quad R_E = \frac{R_{E0}}{\Delta}$$

gives the exponents per second. Thus, taking account of $T = n\Delta$, it turns out that (7.7) is equivalent to

$$\text{m-}\epsilon_n^B \le \exp[-T \sup_{0\le\rho\le 1} (E_B(\rho, q) - \rho(R_B + R_E)) + O(1)], \tag{7.8}$$

where

$$E_B(\rho, q) = A_y \left[q + s_y - s_y(1+\tau_y q)^{1+\rho} \right]. \tag{7.9}$$

Since $E_B(\rho, q)$ is concave in ρ (cf. Gallager [16]), the supremum

$$\sup_{0 \le \rho \le 1} \left(E_B(\rho, q) - \rho(R_B + R_E) \right)$$

is specified by the equation:

$$\frac{dE_B(\rho, q)}{d\rho} = R_B + R_E \quad (0 \le \rho \le 1). \tag{7.10}$$

Carrying out a direct calculation of the left-hand side of (7.10), it follows that

$$\begin{aligned} R_B + R_E &= \frac{dE_B(\rho, q)}{d\rho} \\ &= A_y s_y \left[q \left(1 + \frac{1}{s_y} \right)^{\frac{1}{1+\rho}} \frac{(1 + \tau_y q)^\rho}{1 + \rho} \log \left(1 + \frac{1}{s_y} \right) - (1 + \tau_y q)^{1+\rho} \log(1 + \tau_y q) \right], \end{aligned} \tag{7.11}$$

which together with (7.6) and (7.9) gives the parametric representation of the reliability function under the maximum criterion $\text{m-}\epsilon_n^B$ with parameter ρ.

Remark 7.1 The function

$$f_B(R, q) \stackrel{\Delta}{=} \sup_{0 \le \rho \le 1} \left(E_B(\rho, q) - \rho R \right) \quad (R = R_B + R_E) \tag{7.12}$$

can be derived by eliminating ρ from (7.9) using (7.11), and is zero at

$$\begin{aligned} & R_B + R_E = \\ &= A_y s_y \left[q \left(1 + \frac{1}{s_y} \right) \log \left(1 + \frac{1}{s_y} \right) - \left(1 + \frac{q}{s_y} \right) \log \left(1 + \frac{q}{s_y} \right) \right] \\ &\stackrel{\Delta}{=} h_B(q) = I(q, W_B)/\Delta, \end{aligned} \tag{7.13}$$

and $f_B(R_B + R_E, q)$ is convex and positive in the range: $R_B + R_E < h_B(q)$. $\square$

B. Security function

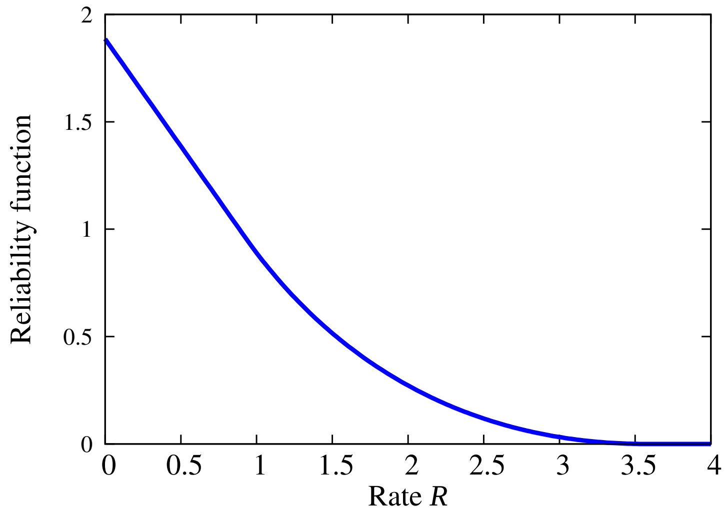

Figure 14: Reliability function for Poisson channel ($A_y = 12$, $\lambda_y = 0.5$, $\Gamma = 0.5$).

Let us now turn to the problem of evaluating the security function against Eve. We proceed in parallel with the above case of reliability function. Formula (5.33) of Theorem 5.2 is written as

$$
\begin{aligned}
\text{m-}\delta_n^E &\leq 6e^{-nH_c(q,R_{E0},n)} \\
&= \exp[-n \sup_{0<\rho<1} \sup_{r\geq 0} (E_{E0}(\rho,q,r) + \rho R_{E0} + O(1/n))] \\
&= \exp[-n \sup_{0<\rho<1} \sup_{r\geq 0} (E_{E0}(\rho,q,r) + \rho R_{E0}) + O(1)], \qquad (7.14)
\end{aligned}
$$

where we have set $E_{E0}(\rho,q,r) = \phi(-\rho|W_E,q,r)$. Let us evaluate $E_{E0}(\rho,q,r)$. Taking account of (5.36), we have

$$
\begin{aligned}
E_{E0}(\rho,q,r) &= -\log\left[\sum_{z\in\mathcal{Z}}\left(\sum_{x\in\mathcal{X}} q(x)W_E(z|x)^{\frac{1}{1-\rho}} e^{r[\Gamma-c(x)]}\right)^{1-\rho}\right] \\
&= -\log\left[\sum_{z=0}^{1}\left(\sum_{x=0}^{1} q(x)W_E(z|x)^{\frac{1}{1-\rho}} e^{r[\Gamma-c(x)]}\right)^{1-\rho}\right] \\
&= -\log\sum_{z=0}^{1} V_z^{1-\rho} - r(1-\rho)\Gamma, \qquad (7.15)
\end{aligned}
$$

where

$$V_z = \sum_{x=0}^{1} q(x) W_E(z|x)^{\frac{1}{1-\rho}} e^{-rx} \quad (z = 0, 1).$$

With $q = q(1)$, an elementary caluculation using (6.7) and (6.8) leads, up to the first order, to

$$\begin{aligned} E_{E0}(\rho, q, r) &= -r(1-\rho)\Gamma - (1-\rho)\log(1 - q + qe^{-r}) \\ &\quad + \Delta A_z \left[\frac{(1-q)s_z + qe^{-r}(1+s_z)}{1 - q + qe^{-r}} \right] \\ &\quad - \Delta A_z \left[\frac{(1-q)s_z^{\frac{1}{1-\rho}} + qe^{-r}(1+s_z)^{\frac{1}{1-\rho}}}{1 - q + qe^{-r}} \right]^{1-\rho}. \end{aligned} \tag{7.16}$$

In order to first maximize $E_{E0}(\rho, q, r)$ with respect to r, set

$$h(r) = -r(1-\rho)\Gamma - (1-\rho)\log(1 - q + qe^{-r}).$$

Then,

$$h'(r) = -(1-\rho)\Gamma + (1-\rho)\frac{qe^{-r}}{1 - q + qe^{-r}},$$

$$h''(r) = -(1-\rho)\frac{q^2 e^{-r}(1-q)}{(1 - q + qe^{-r})^2} < 0,$$

which means that $h(r)$ is strictly concave. It is evident that

$$h(0) = 0, \quad h'(0) = -(1-\rho)(\Gamma - q) \leq 0.$$

Consequently, we have

$$\max_{r \geq 0} h(r) = h(0) = 0, \tag{7.17}$$

and hence

$$\begin{aligned} E_{E0}(\rho, q) &\stackrel{\Delta}{=} \max_{r \geq 0} E_{E0}(\rho, q, r) \\ &= \Delta A_z \left[(1-q)s_z + (1+s_z)q - \left[(1-q)s_z^{\frac{1}{1-\rho}} + q(1+s_z)^{\frac{1}{1-\rho}} \right]^{1-\rho} \right]. \\ &= \Delta A_z \left[q + s_z - s_z(1 + \tau_z q)^{1-\rho} \right], \end{aligned} \tag{7.18}$$

where

$$\tau_z = \left(1 + \frac{1}{s_z}\right)^{\frac{1}{1-\rho}} - 1. \tag{7.19}$$

On the other hand, (7.14) is rewritten as

$$\text{m-}\delta_n^E \leq \exp[-n \sup_{0<\rho<1} (E_{E0}(\rho, q) + \rho R_{E0}) + O(1)]. \tag{7.20}$$

Notice here that $E_{E0}(\rho, q) + \rho R_{E0}$ in (7.20) is the exponent per channel use, so that

$$\frac{E_{E0}(\rho, q) + \rho R_{E0}}{\Delta}$$

gives the exponent per second. Therefore,

$$E_E(\rho, q) = \frac{E_{E0}(\rho, q)}{\Delta}, \quad R_E = \frac{R_{E0}}{\Delta}$$

gives the exponents per second. Thus, taking account of $T = n\Delta$, it turns out that (7.20) is equivalent to

$$\text{m-}\delta_n^E \leq \exp[-T \sup_{0<\rho<1} (E_E(\rho, q) + \rho R_E) + O(1)], \tag{7.21}$$

where

$$E_E(\rho, q) = A_z \left[q + s_z - s_z(1 + \tau_z q)^{1-\rho}\right]. \tag{7.22}$$

Since $E_E(\rho, q)$ is concave in ρ, the supremum

$$\sup_{0<\rho<1} (E_E(\rho, q) + \rho R_E)$$

is specified by the equation;

$$-\frac{dE_E(\rho, q)}{d\rho} = R_E \quad (0 < \rho < 1). \tag{7.23}$$

Carrying out a direct calculation of the left-hand side of (7.23), it follows that

$$\begin{aligned} R_E &= -\frac{dE_E(\rho, q)}{d\rho} \\ &= A_z s_z \left[q\left(1 + \frac{1}{s_z}\right)^{\frac{1}{1-\rho}} \frac{(1 + \tau_z q)^{-\rho}}{1-\rho} \log\left(1 + \frac{1}{s_z}\right) - (1 + \tau_z q)^{1-\rho} \log(1 + \tau_z q)\right], \end{aligned} \tag{7.24}$$

which together with (7.19) and (7.22) gives the parametric representation of the security function under the maximum criterion $\text{m-}\delta_n^E$ with parameter ρ.

Remark 7.2 The function

$$f_E(R,q) \stackrel{\Delta}{=} \sup_{0<\rho<1} \left(E_E(\rho,q)+\rho R\right) \quad (R=R_E) \tag{7.25}$$

can be derived by eliminating ρ from (7.22) using (7.24), and is zero at

$$\begin{aligned} R_E &= A_z s_z \left[q\left(1+\frac{1}{s_z}\right)\log\left(1+\frac{1}{s_z}\right) - \left(1+\frac{q}{s_z}\right)\log\left(1+\frac{q}{s_z}\right)\right] \\ &\stackrel{\Delta}{=} h_E(q) = I(q,W_E)/\Delta, \end{aligned} \tag{7.26}$$

and $f_E(R_E,q)$ is convex and positive in the range: $R_E > h_E(q)$. It should be noted here that the form of the function $f_E(R,q)$ is the same as that of $f_B(R,q)$ in (7.12) of Remark 7.1, while they are positive in the opposite directions, i.e., (7.26) and $R_E > h_E(q)$ correspond to (7.13) and $R_B + R_E < h_B(q)$, respectively. $\square$

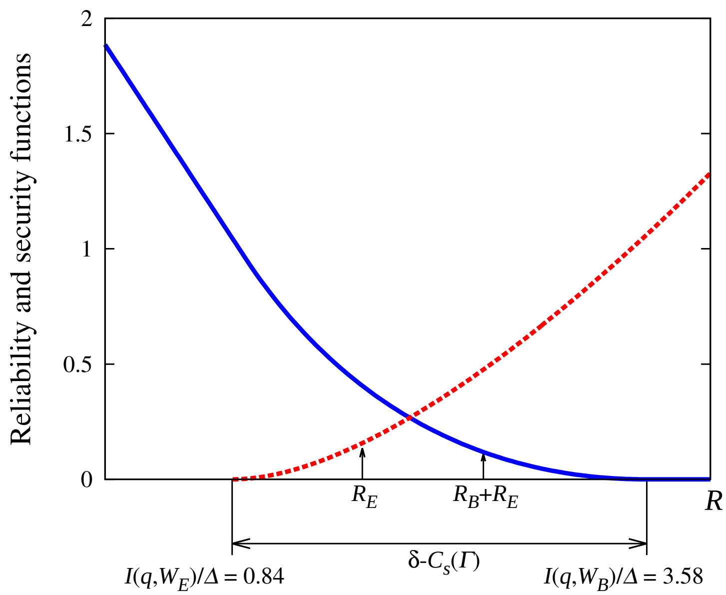

Figure 15: Reliability and security functions for Poisson channel ($A_y = 12$, $\lambda_y = 0.5$, $A_z = 5$, $\lambda_z = 1.5$, $\Gamma = 0.5$).

Remark 7.3 As was stated in the previous section, *degradedness* implies *more capability*, so that it holds that $I(q; W_B) > I(q; W_E)$ owing to the assumed degradedness, which guarantees that the security function curve crosses the reliability function curve. This property enables us to control tradeoff between reliability and security by the operation of rate exchange (cf. Remark 3.7). It should be noted here that in the above arguments the common input probability q is shared by both the reliability function and the security function. This implies that maximization over q should not be taken separately for the reliability function and the security function, but should be taken for $I(q; W_B) - I(q; W_E)$ to achieve the δ-secrecy capacity δ-$C_s(\Gamma)$ of the wiretap channel, as long as q satisfies the cost constraint $q \le \Gamma$. $\square$

8 Secrecy capacity of Gaussian wiretap channel

In this section, we first consider application of Theorem 5.4 to the discrete time stationary memoryless Gaussian wiretap channel to determine the δ-secrecy capacity. Let the Gaussian wiretap channel be denoted by (W_B, W_E) and the input by X, and let Y, Z be the outputs via channels W_B, W_E, respectively, due to the input X, i.e.,

$$Y = A_y X + N_y, \tag{8.1}$$
$$Z = A_z X + N_z, \tag{8.2}$$

where $A_y > 0, A_z > 0$ are positive constants specifying attenuation of signal, and N_y, N_z are Gaussian *additive* noises with variances σ_y^2, σ_z^2, respectively. Here, we have an analogue of Theorem 6.1:

Theorem 8.1 A Gaussian wiretap channel is (statistically) degraded if

$$A_y \ge A_z \tag{8.3}$$

and

$$\frac{\sigma_y}{A_y} \le \frac{\sigma_z}{A_z}. \tag{8.4}$$

Proof: Set

$$\tilde{\sigma}^2 = \frac{A_y^2}{A_z^2}\sigma_z^2 - \sigma_y^2,$$

where $\tilde{\sigma} \geq 0$ follows from (8.4). Then, there exists a fictitious Gaussian noise $\tilde{N}$ with variance $\tilde{\sigma}^2$ that is independent from N_y such that

$$A_y X + N_y + \tilde{N} \simeq A_y X + \frac{A_y}{A_z} N_z = \frac{A_y}{A_z}(A_z X + N_z),$$

where "$U \simeq V$" means that U and V are subject to the same statistics. In view of (8.3), this means that $A_z X + N_z$ can be obtained by adding the fictitious noise $\tilde{N}$ and attenuating $A_y X + N_y + \tilde{N}$. $\square$

Hereafter, we assume that conditions (8.3) and (8.4) are satisfied. Since degradedness implies more capability (cf. Theorem 6.2), we can invoke Theorem 5.4 with cost function $c(x) = x^2$ and cost constraint $\mathrm{E}c(X) \leq \Gamma$ to have

Theorem 8.2 The δ-secrecy capacity δ-$C_s(\Gamma)$ of a Gaussian wiretap channel under cost constraint Γ is given by

$$\delta\text{-}C_s(\Gamma) = \frac{1}{2}\log\left(1 + \frac{A_y^2\Gamma}{\sigma_y^2}\right) - \frac{1}{2}\log\left(1 + \frac{A_z^2\Gamma}{\sigma_z^2}\right), \tag{8.5}$$

under the maximum criterion (m-ϵ_n^B, m-δ_n^E).

Remark 8.1 A weak secrecy version of formula (8.5) with $A_y = A_z = 1$ is found earlier in Cheong and Hellman [26]:

$$\text{w-}C_s(\Gamma) = \frac{1}{2}\log\left(1 + \frac{\Gamma}{\sigma_y^2}\right) - \frac{1}{2}\log\left(1 + \frac{\Gamma}{\sigma_z^2}\right). \tag{8.6}$$

Proof of Theorem 8.2:

Define the differential entropy for probability density function $f(u)$ by

$$h(f) = -\int f(u)\log f(u)du.$$

Then,

$$\begin{aligned}
&I(X;Y) - I(X;Z) \\
&\quad= h(Y) - h(Z) - h(Y|X) + h(Z|X) \\
&\quad= h(Y) - h(Z) - h(N_y) + h(N_z) \\
&\quad= h(Y) - h(Z) - \frac{1}{2}\log\sigma_y^2 + \frac{1}{2}\log\sigma_z^2 \\
&\quad= h\left(\frac{Y}{A_y}\right) - h\left(\frac{Z}{A_z}\right) + \log(A_y/A_z) - \frac{1}{2}\log\sigma_y^2 + \frac{1}{2}\log\sigma_z^2. \qquad (8.7)
\end{aligned}$$

We now observe the following equivalence:

$$\max_{X:\mathrm{E}c(X)\le\Gamma}(I(X;Y)-I(X:Z)) \quad\Longleftrightarrow \max_{X:\mathrm{E}c(X)\le\Gamma}\left(h\left(\frac{Y}{A_y}\right)-h\left(\frac{Z}{A_z}\right)\right). \tag{8.8}$$

On the other hand, Liu and Viswanath [27] qurantees that the maximization on the right-hand side is attained by a Gaussian density P_X with variance σ^2. It is then easy to check that

$$\begin{aligned} g(\Gamma) &\triangleq \max_{X:\mathrm{E}c(X)\le\Gamma}\left(h\left(\frac{Y}{A_y}\right)-h\left(\frac{Z}{A_z}\right)\right) \\ &= \frac{1}{2}\max_{\sigma^2\le\Gamma}\left[\log\left(\sigma^2+\frac{\sigma_y^2}{A_y^2}\right)-\log\left(\sigma^2+\frac{\sigma_z^2}{A_z^2}\right)\right] \\ &= \log\left(\Gamma+\frac{\sigma_y^2}{A_y^2}\right)-\log\left(\Gamma+\frac{\sigma_z^2}{A_z^2}\right), \end{aligned} \tag{8.9}$$

where in the last step we have used (8.4). Susbtituting (8.9) into (8.7) and rearranging it, we eventually obtain

$$\max_{X:\mathrm{E}c(X)\le\Gamma}(I(X;Y)-I(X:Z)) = \frac{1}{2}\log\left(1+\frac{A_y^2\Gamma}{\sigma_y^2}\right)-\frac{1}{2}\log\left(1+\frac{A_z^2\Gamma}{\sigma_z^2}\right), \tag{8.10}$$

which together with Theorem 5.4 concludes Theorem 8.2. $\square$

9 Reliability and security functions of Gaussian wiretap channel

In this section, we consider application of Theorem 5.2 to the Gaussian wiretap channel to evaluate its reliability and security functions. To this end, it is convenient here to use, according to (5.32) and (5.33), formulas

$$\text{m-}\epsilon_n^B \le \exp[-n\sup_{0\le\rho\le1}\sup_{s\ge0}(E_B(\rho,q,s)-\rho(R_B+R_E))+O(1)], \tag{9.1}$$

$$\text{m-}\delta_n^E \le \exp[-n\sup_{0<\rho<1}\sup_{s\ge0}(E_E(\rho,q,s)+\rho R_E)+O(1)], \tag{9.2}$$

where

$$E_B(\rho,q,s) = -\log\left[\int_y\left(\int_x q(x)W_B(y|x)^{\frac{1}{1+\rho}}e^{s[\Gamma-c(x)]}dx\right)^{1+\rho}dy\right], \tag{9.3}$$

$$E_E(\rho,q,s) \quad = \quad -\log\left[\int_y\left(\int_x q(x)W_E(y|x)^{\frac{1}{1-\rho}}e^{s[\Gamma-c(x)]}dx\right)^{1-\rho}dy\right]. \quad (9.4)$$

Remark 9.1 These formulas (9.1) $\sim$ (9.4) are the continuous alphabet non-concatenated versions of (5.32) and (5.33) in Theorem 5.2 (cf. Remark 5.7). $\square$

A. Reliability function

We first insert

$$W_B(y|x) = \frac{1}{\sqrt{2\pi\sigma_B^2}}\exp\left[-\frac{(y-x)^2}{2\sigma_B^2}\right] \tag{9.5}$$

and

$$q(x) = \frac{1}{\sqrt{2\pi A_y^2\Gamma}}\exp\left[-\frac{x^2}{2A_y^2\Gamma}\right] \tag{9.6}$$

into (9.3) to have

$$\begin{aligned} &E_B(\rho,q,s)\\ &= \quad -s(1+\rho)A_y^2\Gamma + \frac{1}{2}\log(1+2sA_y^2\Gamma) + \frac{\rho}{2}\log\left[1+2sA_y^2\Gamma+\frac{A_y^2\Gamma}{(1+\rho)\sigma_B^2}\right], \end{aligned} \tag{9.7}$$

where $s \geq 0$ is an arbitrary constant. Set

$$A_B \quad = \quad \frac{A_y^2\Gamma}{\sigma_B^2}, \tag{9.8}$$

$$\beta_B \quad = \quad 1+2sA_y^2\Gamma+\frac{A_B}{(1+\rho)}, \tag{9.9}$$

where β_B ranges as

$$1+\frac{A_B}{1+\rho} \leq \beta_B < +\infty. \tag{9.10}$$

Use (9.8) and (9.9) to eliminate Γ, σ_B^2 and s from (9.7), and consider $E_B(\rho,q,s)$ as a function $E_B(A_B,\beta_B,\rho)$ of A_B,β_B,ρ, then

$$\begin{aligned} &E_B(A_B,\beta_B,\rho)\\ &= \quad \frac{1}{2}\left[(1-\beta_B)(1+\rho)+A_B+\log\left(\beta_B-\frac{A_B}{1+\rho}\right)+\rho\log\beta_B\right]. \end{aligned} \tag{9.11}$$

Hence,

$$\frac{dE_B}{d\beta_B} = \frac{1}{2}\left[-(1+\rho) + \frac{1+\rho}{\beta_B(1+\rho) - A_B} + \frac{\rho}{\beta_B}\right]. \tag{9.12}$$

Notice that the right-hand side of (9.12) is decreasing in β_B because $\beta_B(1+\rho) - A_B > 0$ owing to (9.10) and that

$$\frac{dE_B}{d\beta_B} < 0 \text{ at } \beta_B = 1 + \frac{A_B}{1+\rho} \quad \text{and} \quad \frac{dE_B}{d\beta_B} < 0 \text{ when } \beta_B \to +\infty.$$

Therefore, E_B has the maximum value at $\beta_B = 1 + \frac{A_B}{1+\rho}$, i.e.,

$$E_B(\rho) \stackrel{\Delta}{=} \max_{\beta_B \geq 1 + \frac{A_B}{1+\rho}} E_B(A_B, \beta_B, \rho) = \frac{\rho}{2}\log\left(1 + \frac{A_B}{1+\rho}\right) \quad (0 \leq \rho \leq 1). \tag{9.13}$$

On the other hand, $E_B(\rho) - \rho(R_B + R_E)$ has a stationary point with respect to ρ, i.e.,

$$\begin{aligned} & \frac{\partial(E_B(\rho) - \rho(R_B + R_E))}{\partial\rho} \\ = \ & \frac{1}{2}\log\left(1 + \frac{A_B}{1+\rho}\right) - \frac{\rho A_B}{2(1+\rho)(1+\rho+A_B)} - (R_B + R_E) = 0. \end{aligned} \tag{9.14}$$

Hence,

$$R_B + R_E = \frac{1}{2}\log\left(1 + \frac{A_B}{1+\rho}\right) - \frac{\rho A_B}{2(1+\rho)(1+\rho+A_B)}. \tag{9.15}$$

As a consequence, by means of (9.13) and (9.15), we obtain

$$\begin{aligned} E_B(R_B, R_E) & \stackrel{\Delta}{=} E_B(\rho) - \rho(R_B + R_E) \\ & = \frac{\rho^2 A_B}{2(1+\rho)(1+\rho+A_B)}. \end{aligned} \tag{9.16}$$

Thus, we have

Theorem 9.1 (Reliability function) The reliability function $E_B(R_B, R_E)$ of a Gaussian wiretap channel under the maximum criterion m-ϵ_n^B is given by the following parametric representation with $0 \leq \rho \leq 1$:

$$E_B(R_B, R_E) = \frac{\rho^2 A_B}{2(1+\rho)(1+\rho+A_B)}, \tag{9.17}$$

$$R_B + R_E = \frac{1}{2}\log\left(1 + \frac{A_B}{1+\rho}\right) - \frac{\rho A_B}{2(1+\rho)(1+\rho+A_B)} \tag{9.18}$$

for $R_c \le R_B + R_E \le \frac{1}{2}\log(1 + A_B)$ with

$$R_c = \frac{1}{2}\log\left(1 + \frac{A_B}{2}\right) - \frac{A_B}{4(2 + A_B)},$$

whereas, for $0 \le R_B + R_E \le R_c$,

$$E_B(R_B, R_E) = \frac{1}{2}\log\left(1 + \frac{A_B}{2}\right) - (R_B + R_E). \tag{9.19}$$

□

So far we have established the formula for reliability function based on upper bound (5.22). In contrast with Theorem 9.1, Gallager [16] has derived another reliability function based on upper bound (5.27), leading to the exponent formula

$$E_B(\rho, q, s) = -\log\left[\int_y \left(\int_x q(x) W_B(y|x)^{\frac{1}{1+\rho}} e^{s[c(x-\Gamma)]} dx\right)^{1+\rho} dy\right] \tag{9.20}$$

instead of (9.3). It should be noted here that in (9.20) $c(x) - \Gamma$ appears instead of $\Gamma - c(x)$. Then, we have

Theorem 9.2 (Reliability function: Gallager) The reliability function of a Gaussian wiretap channel under the maximum criterion m-ϵ_n^B is given by

$$\begin{aligned} &E_B(R_B, R_E) \\ &= \frac{A_B}{4\beta_B}\left[(\beta_B + 1) - (\beta_B - 1)\sqrt{1 + \frac{4\beta_B}{A_B(\beta_B - 1)}}\right] \\ &\quad + \frac{1}{2}\log\left[\beta_B - \frac{A_B(\beta_B - 1)}{2}\left(\sqrt{1 + \frac{4\beta_B}{A_B(\beta_B - 1)}} - 1\right)\right] \end{aligned} \tag{9.21}$$

with $\beta_B = e^{2(R_B + R_E)}$. Formula (9.21) is valid in the range of $R = R_B + R_E$ as follows:

$$\frac{1}{2}\log\left[\frac{1}{2} + \frac{A_B}{4} + \frac{1}{2}\sqrt{1 + \frac{A_B}{4}}\right] \le R \le \frac{1}{2}\log(1 + A_B). \tag{9.22}$$

For R less than the left-hand side of (9.22), we must choose $\rho = 1$ yielding

$$E_B(R_B, R_E) = 1 - \beta_B + \frac{A_B}{2} + \frac{1}{2}\log\left(\beta_B - \frac{A_B}{2}\right) + \frac{1}{2}\log\beta_B - (R_B + R_E), \tag{9.23}$$

where

$$\beta_B = \frac{1}{2}\left[1 + \frac{A_B}{2} + \sqrt{1 + \frac{A_B^2}{4}}\right].$$

□

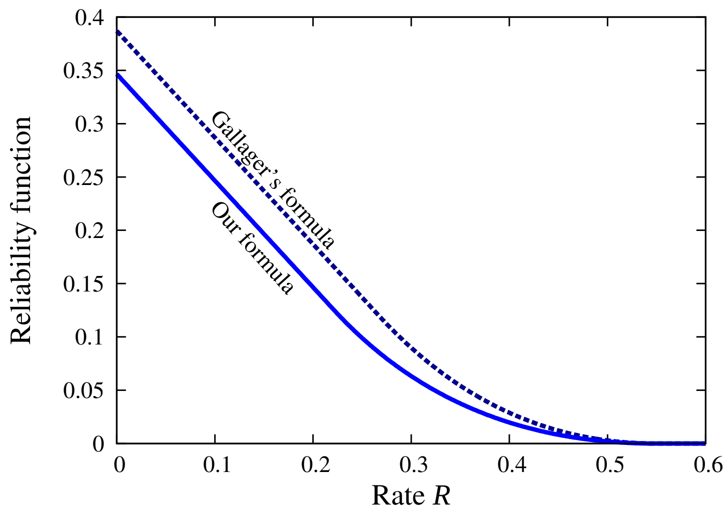

Figure 16: Comparison of reliability function for Gaussian channel derived in this paper (solid line) and derived by Gallager (dashed line) ($A_y = 1$, $\sigma_y = 0.5$, $\Gamma = 0.5$).

B. Security function

In this subsection, we evaluate the right-hand side of (9.2) on the security function. The arguments here proceed in parallel with those in the previous subsection with due modifications and $-\rho$ instead of ρ. Here too, we insert

$$W_E(y|x) = \frac{1}{\sqrt{2\pi\sigma_E^2}} \exp\left[-\frac{(y-x)^2}{2\sigma_E^2}\right] \tag{9.24}$$

and

$$q(x) = \frac{1}{\sqrt{2\pi A_z^2 \Gamma}} \exp\left[-\frac{x^2}{2A_z^2\Gamma}\right] \tag{9.25}$$

into (9.4) to have

$E_E(\rho, q, s)$

$$= \quad -s(1-\rho)A_z^2\Gamma + \frac{1}{2}\log(1+2sA_z^2\Gamma) - \frac{\rho}{2}\log\left[1+2sA_z^2\Gamma + \frac{A_z^2\Gamma}{(1-\rho)\sigma_E^2}\right], \tag{9.26}$$

where $s \geq 0$ is an arbitrary constant. Here we set

$$A_E \quad = \quad \frac{A_z^2\Gamma}{\sigma_E^2}, \tag{9.27}$$

$$\beta_E \quad = \quad 1 + 2sA_z^2\Gamma + \frac{A_E}{(1-\rho)}, \tag{9.28}$$

then β_E ranges as

$$1 + \frac{A_E}{1-\rho} \leq \beta_E < +\infty. \tag{9.29}$$

Use (9.27) and (9.28) to eliminate Γ, σ_E^2 and s from (9.26), and consider $E_E(\rho, q, s)$ as a function $E_B(A_E, \beta_E, \rho)$ of A_B, β_E, ρ, then

$$\begin{aligned} & E_E(A_E, \beta_E, \rho) \\ = \quad & \frac{1}{2}\left[(1-\beta_E)(1-\rho) + A_E + \log\left(\beta_E - \frac{A_E}{1-\rho}\right) - \rho\log\beta_E\right]. \end{aligned} \tag{9.30}$$

A stationary point with respect to β_E (and hence also with respect to s) is specified by

$$\frac{dE_E}{d\beta_E} = \frac{1}{2}\left[-(1-\rho) + \frac{1-\rho}{\beta_E(1-\rho)-A_E} - \frac{\rho}{\beta_E}\right] = 0. \tag{9.31}$$

Notice that the right-hand side of (9.31) is decreasing in β_E because $\beta_E(1-\rho) - A_E \geq 1-\rho$ owing to (9.29) and that

$$\frac{dE_E}{d\beta_E} > 0 \text{ at } \beta_E = 1 + \frac{A_E}{1-\rho} \quad \text{and} \quad \frac{dE_E}{d\beta_E} < 0 \text{ when } \beta_E \to +\infty.$$

Therefore, equation (9.31) has the unique solution for β_E, i.e.,

$$\beta_E = \frac{1}{2}\left(1 + \frac{A_E}{1-\rho}\right)\left[1 + \sqrt{1 + \frac{4A_E\rho}{(1-\rho+A_E)^2}}\right]. \tag{9.32}$$

On the other hand, $E_E + \rho R_E$ has a stationary point with respect to ρ, i.e.,

$$\begin{aligned} & \frac{\partial(E_E + \rho R_E)}{\partial\rho} \\ = \quad & -\frac{1}{2}\left[1 - \beta_E + \frac{\beta_E}{\beta_E(1-\rho)-A_E} - \frac{1}{1-\rho} + \log\beta_E\right] + R_E = 0. \end{aligned} \tag{9.33}$$

From (9.31) and (9.33), it follows that

$$R_E = \frac{1}{2}\log \beta_E. \tag{9.34}$$

Furthermore, combining (9.30) with (9.34), we obtain

$$\begin{aligned} E_E(R_E) &\stackrel{\Delta}{=} E_E + \rho R_E \\ &= \frac{1}{2}\left[(1-\beta_E)(1-\rho) + A_E + \log\left(\beta_E - \frac{A_E}{1-\rho}\right)\right]. \end{aligned} \tag{9.35}$$

On the other hand, equation (9.31) can be solved for ρ as follows:

$$1-\rho = \frac{A_E}{2\beta_E}\left[1+\sqrt{1+\frac{4\beta_E}{A_E(\beta_E-1)}}\right], \tag{9.36}$$

which inserted into (9.35) yields the following theorem:

Theorem 9.3 (Security function) The security function of a Gaussian wire-tap channel under the maximum criterion $\text{m-}\delta_n^E$ is given by

$$\begin{aligned} &E_E(R_E) \\ &= \frac{A_E}{4\beta_E}\left[(\beta_E+1) - (\beta_E-1)\sqrt{1+\frac{4\beta_E}{A_E(\beta_E-1)}}\right] \\ &\quad + \frac{1}{2}\log\left[\beta_E - \frac{A_E(\beta_E-1)}{2}\left(\sqrt{1+\frac{4\beta_E}{A_E(\beta_E-1)}}-1\right)\right] \end{aligned} \tag{9.37}$$

with $\beta_E = e^{2R_E}$. □

Remark 9.2 It should be noted that the form of the function in (9.21) is the same as that in (9.37), but the ranges where they are valid are opposite, i.e., formula (9.37) is valid in the range of R_E:

$$R_E \geq \frac{1}{2}\log(1+A_E), \tag{9.38}$$

where parameter $\rho = 0$ corresponds to $R_E = \frac{1}{2}\log(1+A_E)$ and $\rho \to 1$ corresponds to $R_E \to +\infty$, whereas (9.21) along with (9.23) is valid when $R_B + R_E \leq \frac{1}{2}\log(1+A_B)$. □

Now in view of Theorem 9.2, one may be tempted to derive the security function based on upper bound (5.27), leading to the exponent formula

$$E_E(\rho, q, s) = -\log\left[\int_y\left(\int_x q(x)W_E(y|x)^{\frac{1}{1-\rho}}e^{s[c(x)-\Gamma]}dx\right)^{1-\rho}dy\right] \qquad (9.39)$$

instead of (9.4). It should be noted here that in (9.39) $c(x)-\Gamma$ appears instead of $\Gamma - c(x)$. Let us see what happens in this case. It is first straightforward to check that (9.39) is developed as

$$\begin{aligned} & E_E(\rho, q, s) \\ = \;& s(1-\rho)A_z^2\Gamma + \frac{1}{2}\log(1-2sA_z^2\Gamma) - \frac{\rho}{2}\log\left[1-2sA_z^2\Gamma + \frac{A_z^2\Gamma}{(1-\rho)\sigma_E^2}\right] \end{aligned} \qquad (9.40)$$

with

$$A_E = \frac{A_z^2\Gamma}{\sigma_E^2}, \qquad (9.41)$$

$$\beta_E = 1 - 2sA_z^2\Gamma + \frac{A_E}{1-\rho}, \qquad (9.42)$$

where it is evident that β_E ranges as

$$\frac{A_E}{1-\rho} < \beta_E \le 1 + \frac{A_E}{1-\rho}.$$

As was shown in the proof of Theorem 9.1, (9.39) is rewritten as a function of A_E, β_E, ρ as follows:

$$\begin{aligned} & E_E(A_E, \beta_E, \rho) \\ = \;& \frac{1}{2}\left[(1-\beta_E)(1-\rho) + A_E + \log\left(\beta_E - \frac{A_E}{1-\rho}\right) - \rho\log\beta_E\right]. \end{aligned} \qquad (9.43)$$

Then, it is not difficult to verify that

$$\begin{aligned} E_E(\rho) &\triangleq \max_{\frac{A_E}{1-\rho} < \beta_E \le 1 + \frac{A_E}{1-\rho}} E_E(A_E, \beta_E, \rho) \\ &= E_E(A_E, 1 + \frac{A_E}{1-\rho}, \rho) \\ &= -\frac{\rho}{2}\log\left(1 + \frac{A_E}{1-\rho}\right). \end{aligned} \qquad (9.44)$$

Moreover, the equation

$$\frac{\partial(E_E(\rho)+\rho R_E)}{\partial\rho}=0$$

yields

$$R_E=\frac{1}{2}\log\left(1+\frac{A_E}{1-\rho}\right)+\frac{\rho A_E}{2(1-\rho)(1-\rho+A_E)}. \tag{9.45}$$

Then, from (9.44) and (9.45) it follows that

$$E_E(R_E)\stackrel{\Delta}{=}E_E(\rho)+\rho R_E=\frac{\rho^2 A_E}{2(1-\rho)(1-\rho+A_E)}. \tag{9.46}$$

Thus, we have

Theorem 9.4 (Security function: Gallager-type) The security function $E_B(R_B, R_E)$ of a Gaussian wiretap channel under the maximum criterion $\mathrm{m}\text{-}\delta_n^E$ is given by the following parametric representation with $0\le\rho\le 1$:

$$E_E(R_E) = \frac{\rho^2 A_E}{2(1-\rho)(1-\rho+A_E)}, \tag{9.47}$$

$$R_E = \frac{1}{2}\log\left(1+\frac{A_E}{1-\rho}\right)+\frac{\rho A_E}{2(1-\rho)(1-\rho+A_E)} \tag{9.48}$$

for $R_E\ge\frac{1}{2}\log(1+A_E)$. □

Remark 9.3 In Fig.19 we see that as for the reliability function Gallager bound outperforms our bound, whereas as for the security function our bound outperforms Gallager-type bound. It is interesting to observe a kind of dualities holding among Theorem 9.1 $\sim$ Theorem 9.4, which is illustrated in Fig. 17. □

10 Concatenation for Poisson wiretap channel

In this section, we investigate the effects of *concatenation* for performance of Poisson wiretap channels. We first observe a basic property (*invariance*) of Poisson wiretap channel under concatenation (on the basis of Theorem 5.2 and Theorem 5.4). Here too, we use the notation as used in Sections 6 and 7.

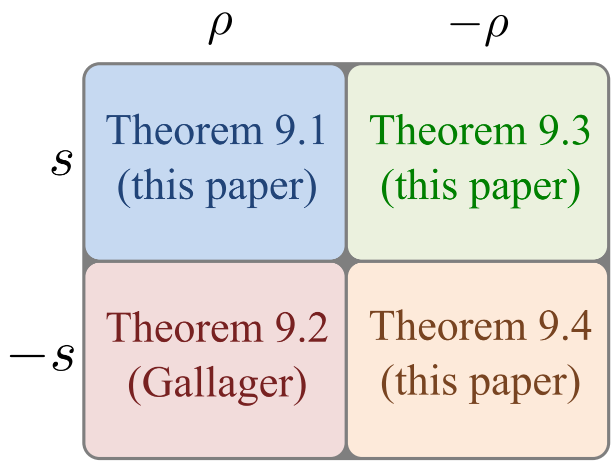

Figure 17: The reliability function in Theorem 9.1 (this paper) has the same form as that of the security function in Theorem 9.4 (this paper: Gallager-type), whereas the security function in Thoerm 9.3 (this paper) has the same form as that of the reliability function in Theorem 9.2 (Gallager).

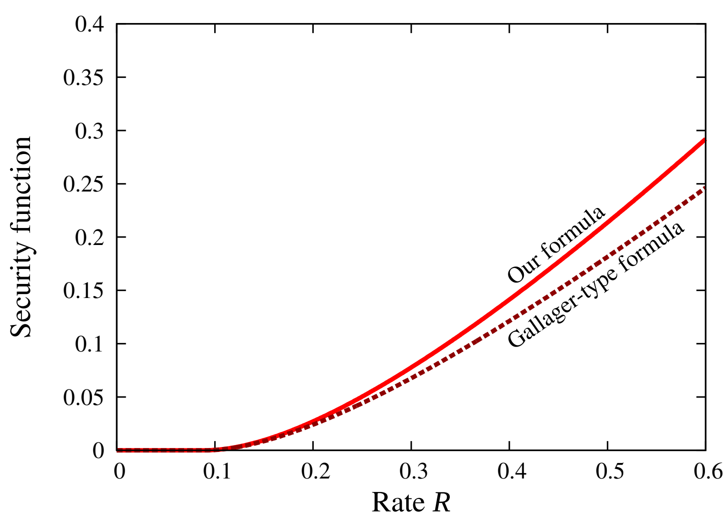

Figure 18: Comparison of security function for Gaussian channel with formula derived in this paper (solid line) and Gallager-type formula (dashed line) ($A_z = 0.5$, $\sigma_z = 0.8$, $\Gamma = 0.5$).

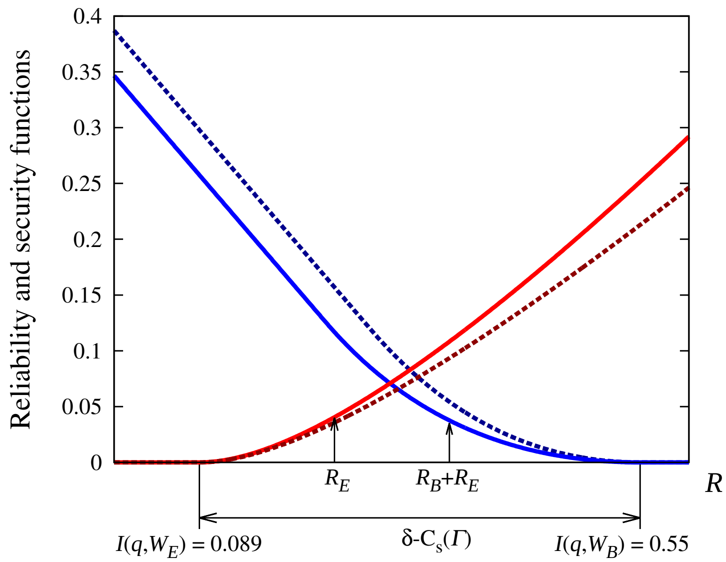

Figure 19: Reliability and security functions: comparison of Gallager-type formula (dashed line) and our forumla (solid line) for Gaussian channel ($A_y = 1$, $\sigma_y = 0.5$, $A_z = 0.5$, $\sigma_z = 0.8$, $\Gamma = 0.5$).

Theorem 10.1 Conditions (6.12) and (6.13) are preserved under concatenation. □

Proof:

Set the transition probabilities $P_{X|V}$ of the auxiliary binary channel as

$$a = P_{X|V}(1|1), \quad 1 - a = P_{X|V}(0|1); \tag{10.1}$$

$$b = P_{X|V}(1|0), \quad 1 - b = P_{X|V}(0|0), \tag{10.2}$$

where we assume that $a > b$. Then, the transition probabilities of the concatenated channel (W_B^+, W_E^+) are given by

$$W_B^+(1|1) = [a(A_y + \lambda_y) + (1-a)\lambda_y]\,\Delta = [aA_y + \lambda_y]\Delta, \tag{10.3}$$

$$W_B^+(1|0) = [b(A_y + \lambda_y) + (1-b)\lambda_y]\,\Delta = [bA_y + \lambda_y]\Delta; \tag{10.4}$$

$$W_E^+(1|1) = [a(A_z + \lambda_z) + (1-a)\lambda_z]\,\Delta = [aA_z + \lambda_z]\Delta, \tag{10.5}$$

$$W_E^+(1|0) = [b(A_z + \lambda_z) + (1-b)\lambda_z]\,\Delta = [bA_z + \lambda_z]\Delta. \tag{10.6}$$

Notice that the concatenated channel is also a Poisson wiretap channel, and let the peak power and dark currents of the concatenated channel be denoted by A_y^+, A_z^+, λ_y^+, λ_z^+, respectively, then we obtain

$$\begin{aligned} \lambda_y^+ &= W_B^+(1|0)/\Delta = bA_y + \lambda_y, && (10.7)\\ A_y^+ &= (W_B^+(1|1) - W_B^+(1|0))/\Delta = (a-b)A_y, && (10.8) \end{aligned}$$

$$\begin{aligned} \lambda_z^+ &= W_E^+(1|0)/\Delta = bA_z + \lambda_z, && (10.9)\\ A_z^+ &= (W_E^+(1|1) - W_E^+(1|0))/\Delta = (a-b)A_z, && (10.10) \end{aligned}$$

which means that concatenation has the effect of not only attenuating peak powers to a factor of $a-b$ but also augmenting a factor of b to dark currents. Recall that we have set as

$$s_y = \frac{\lambda_y}{A_y}, \quad s_z = \frac{\lambda_z}{A_z}. \tag{10.11}$$

According to (10.11), set

$$s_y^+ = \frac{\lambda_y^+}{A_y^+}, \quad s_z^+ = \frac{\lambda_z^+}{A_z^+}, \tag{10.12}$$

then,

$$\begin{aligned} s_y^+ &= \frac{bA_y + \lambda_y}{(a-b)A_y} = \frac{b + s_y}{a-b}, && (10.13)\\ s_z^+ &= \frac{bA_z + \lambda_y}{(a-b)A_z} = \frac{b + s_z}{a-b}. && (10.14) \end{aligned}$$

from which it follows that

$$s_y \le s_z \Longleftrightarrow s_y^+ \le s_z^+. \tag{10.15}$$

Moreover, from (10.8) and (10.10) it follows that

$$A_y \ge A_z \Longleftrightarrow A_y^+ \ge A_z^+, \tag{10.16}$$

which completes the proof. □

Since we are considering the case where the non-concatenated channel (W_B, W_E) satisfies conditions (6.12) and (6.13), Theorem 10.1 ensures that the concatenated channel (W_B^+, W_E^+) also satisfies these conditions as well. Therefore, in view of Theorem 6.1 and Theorem 6.2, (W_B^+, W_E^+) is more capable, so that we can use the same arguments as were developed in Section 6. On the other hand, $p = \Pr\{X = 1\}$ is given as

$$p = qa + (1-q)b \tag{10.17}$$

with $q = \Pr\{V = 1\}$. Therefore, solving (10.17) with respect to q, we see that the problem with cost constraint $p \le \Gamma$ $(c(x) = x)$ on X^n of the concatenated channel (W_B^{n+}, W_E^{n+}) equivalently boils down to that with cost constraint Γ^+ $(\overline{c}(v) = v)$ on V^n of (W_B^{n+}, W_E^{n+}) with

$$q \le \frac{\Gamma - b}{a - b} \triangleq \Gamma^+, \tag{10.18}$$

where $\Gamma \ge b$ is assumed (cf. Remark 5.8). Thus, based on (10.15) $\sim$ (10.18), we can develop the same arguments on secrecy capacity as well as reliability/security functions as in Sections 6 and 7, which will be briefly summarized in the sequel.

A. Secrecy capacity

The following theorem is the concatenation counterpart of Theorem 6.3 without concatenation.

Theorem 10.2 Let $a > b$ and $\Gamma \ge b$. Then, the δ-secrecy capacity with cost constraint $\delta\text{-}C_s^+(\Gamma)$ per second of the concatenated wiretap channel (W_B^+, W_E^+) is given by

$$\begin{aligned} &\delta\text{-}C_s^+(\Gamma) \\ &\quad = \log \frac{(q_\Gamma^* + s_z^+)^{(q_\Gamma^*+s_z^+)A_z^+}}{(q_\Gamma^* + s_y^+)^{(q_\Gamma^*+s_y^+)A_y^+}} + \log \frac{(s_y^+)^{s_y^+ A_y^+}}{(s_z^+)^{s_z^+ A_z^+}} \\ &\qquad + q_\Gamma^* \left(\log \frac{(q^* + s_y^+)^{A_y^+}}{(q^* + s_z^+)^{A_z^+}} + A_y^+ - A_z^+ \right) \end{aligned} \tag{10.19}$$

under the maximum criterion (m-ϵ_n^B, m-δ_n^E), where $q = q^*$ is the unique solution in $(0,1)$ of the equation:

$$\frac{(A_y^+ q^* + \lambda_y^+)^{A_y^+}}{(A_z^+ q^* + \lambda_z^+)^{A_z^+}} = e^{A_z^+ - A_y^+} \frac{(A_y^+ + \lambda_y^+)^{A_y^+ + \lambda_y^+}}{(A_z^+ + \lambda_z^+)^{A_z^+ + \lambda_z^+}} \frac{(\lambda_z^+)^{\lambda_z^+}}{(\lambda_y^+)^{\lambda_y^+}}, \tag{10.20}$$

and

$$q_\Gamma^* = \min(q^*, \Gamma^+). \tag{10.21}$$

Proof:

It is not difficult to check that $0 < q* < 1$ as was shown in Section 6. Then, it suffices to proceed in parallel with the proof of Theorem 6.3. $\square$

Remark 10.1 It is easy to check that, in the special case *without* cost constraint (i.e., $\Gamma = a$ and hence $q^*_\Gamma = q^*$), (10.19) reduces to

$$\begin{aligned} \delta\text{-}C_s^+(a) \quad &= \quad q^*(A_y^+ - A_z^+) + \log \frac{(\lambda_y^+)^{\lambda_y^+}}{(\lambda_z^+)^{\lambda_z^+}} \\ &\quad + \log \frac{(A_z^+ q^* + \lambda_z^+)^{\lambda_z^+}}{(A_y^+ q^* + \lambda_y^+)^{\lambda_y^+}} \qquad (10.22) \end{aligned}$$

with equation (10.20). Similarly for the worst case scenario (cf. Example 6.1). $\square$

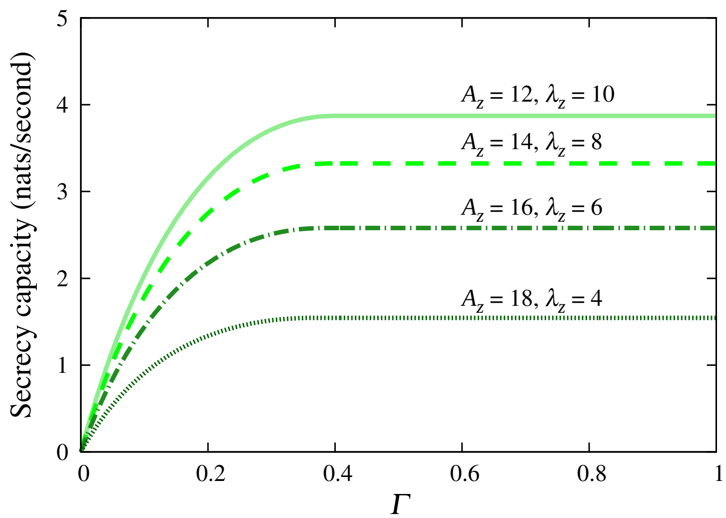

Figure 20: Secresy capacity for non-concatenated Poisson channel with varied Eve's channel ($A_y = 20$, $\lambda_y = 2$).

Remark 10.2 In view of Thorems 5.3 and 5.4 and from the definition of the concatenated channel (W_B^+, W_E^+), we see that concatenation decreases secrecy capacity and mutual information, i.e.,

$$\begin{aligned} \delta\text{-}C_s^+(\Gamma) &\leq \delta\text{-}C_s(\Gamma), \\ I(q, W_B^+) &\leq I(p, W_B), \\ I(q, W_E^+) &\leq I(p, W_E). \end{aligned}$$

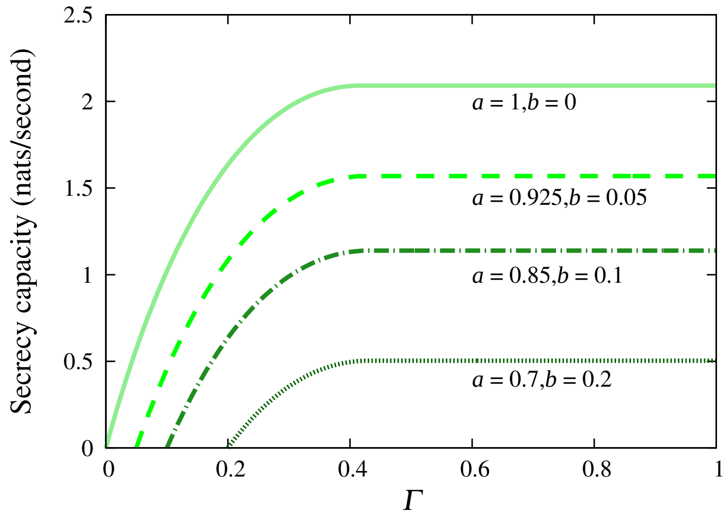

Figure 21: Secresy capacity for concatenated Poisson channel with varied a, b ($A_y = 12$, $\lambda_y = 2$, $A_z = 5$, $\lambda_y = 4$).

However, at the expense of this decrease, the security against Eve increases (cf. Remark 3.4). $\square$

B. Reliability function

Theorem 10.3 The maximum error probability $\mathrm{m}\text{-}\epsilon_n^B$ for (W_B^+, W_E^+) is upper bounded (with $0 \le q \le \Gamma^+$) as

$$\mathrm{m}\text{-}\epsilon_n^B \le \exp[-T \sup_{0\le\rho\le 1} \left(E_B^+(\rho, q) - \rho(R_B + R_E)\right) + O(1)], \tag{10.23}$$

where

$$E_B^+(\rho, q) = A_y^+ \left[q + s_y^+ - s_y^+(1 + \tau_y^+ q)^{1+\rho}\right], \tag{10.24}$$

$$\tau_y^+ = \left(1 + \frac{1}{s_y^+}\right)^{\frac{1}{1+\rho}} - 1. \tag{10.25}$$

Furthermore, the ρ to attain the supremum in (10.23) is specified by

$$R_B + R_E = \frac{dE_B^+(\rho, q)}{d\rho}$$
$$= A_y^+ s_y^+ \left[q \left(1 + \frac{1}{s_y^+}\right)^{\frac{1}{1+\rho}} \frac{(1 + \tau_y^+ q)^\rho}{1 + \rho} \log\left(1 + \frac{1}{s_y^+}\right)\right.$$

$$-(1+\tau_y^+ q)^{1+\rho}\log(1+\tau_y^+ q)\Bigg], \tag{10.26}$$

which together with (10.24) and (10.25) gives the parametric representation of the reliability function with parameter ρ. $\square$

Proof:

Since the concatenated channel (W_B^+, W_E^+) also satisfies conditions (6.12) and (6.13), it suffices to replace A_y, λ_y, τ_y, $E_B(\rho, q)$ in (7.6), (7.9) and (7.11) by $A_y^+, \lambda_y^+, \tau_y^+$, $E_B^+(\rho, q)$, respectively. $\square$

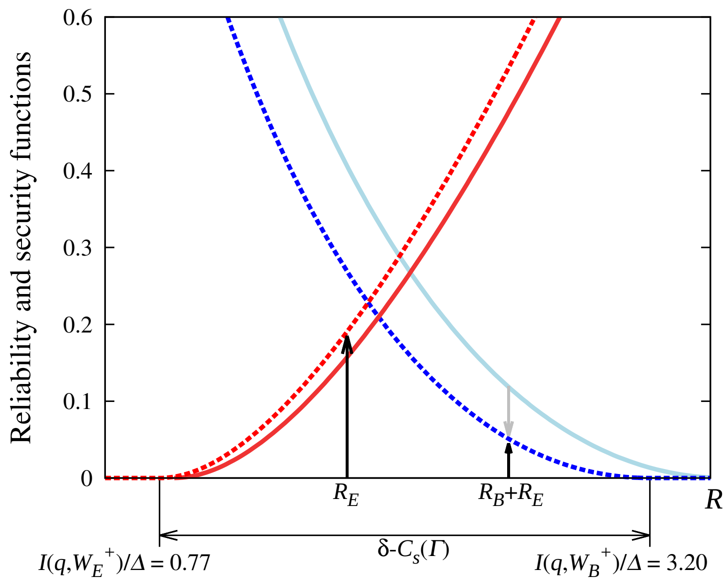

Figure 22: Comparison of reliability and security function for non-concatenated (solid line) and concatenated (dash line, a=0.98, $b = 0.02$) Poisson channel ($A_y = 12$, $\lambda_y = 0.5$, $A_z = 5$, $\lambda_z = 1.5$, $\Gamma = 0.5$).

Another proof of Theorem 10.3:

So far we have shown Theorem 10.3 by equivalently transforming the problem with cost constraint at the input of channel (W_B, W_E) to that with cost constraint at the input of channel (W_B^+, W_E^+). However, it is also possible to establish Theorem 10.3 by making a direct but somewhat tedious calculation of $F_c(q, R_B, R_E, n)$ on the right-hand side of (5.32) in Theorem 5.2 using (5.28) and (5.30). With probability prameters a, b as in (10.1) and (10.2), the computation is actually carried out in Appendix II. $\square$

Theorem 10.4 The maximum divergence $_{\mathrm{m}}\text{-}\delta_n^E$ for (W_B^+, W_E^+) is upper bounded (with $0 \le q \le \Gamma^+$) as

$$_{\mathrm{m}}\text{-}\delta_n^E \le \exp[-T \sup_{0<\rho<1} \left(E_E^+(\rho, q) + \rho R_E\right) + O(1)], \tag{10.27}$$

where

$$E_E^+(\rho, q) = A_z^+ \left[q + s_z^+ - s_z^+(1 + \tau_z^+ q)^{1-\rho}\right], \tag{10.28}$$

$$\tau_z^+ = \left(1 + \frac{1}{s_z^+}\right)^{\frac{1}{1-\rho}} - 1. \tag{10.29}$$

Furthermore, the ρ to attain the supremum in (10.27) is specified by

$$\begin{aligned} R_E &= -\frac{dE_E^+(\rho, q)}{d\rho} \\ &= A_z^+ s_z^+ \left[q\left(1 + \frac{1}{s_z^+}\right)^{\frac{1}{1-\rho}} \frac{(1 + \tau_z^+ q)^{-\rho}}{1-\rho} \log\left(1 + \frac{1}{s_z^+}\right) \right. \\ &\quad \left. -(1 + \tau_z^+ q)^{1-\rho} \log(1 + \tau_z^+ q) \right], \end{aligned} \tag{10.30}$$

which together with (10.28) and (10.29) gives the parametric representation of the security function with parameter ρ. □

Proof:

Since the concatenated channel (W_B^+, W_E^+) also satisfies conditions (6.12) and (6.13), it suffices to replace A_z, λ_z, τ_z, $E_E(\rho, q)$ in (7.19), (7.22) and (7.24) by $A_z^+, \lambda_z^+, \tau_z^+$, $E_E^+(\rho, q)$, respectively. □

Another proof of Theorem 10.4:

The proof fairly parallels the another proof of Theorem 10.3 with $-\rho$ in place of ρ on the basis of (5.29), (5.31) and (5.33). The details are given in Appendix III. □

11 Concluding remarks

So far we have established the δ-secrecy capacity with cost constraint (in the strongest and maximized security sense) as well as a pair of reliability and security functions for a general wiretap channel, and also for Binary symmetric wiretap channels (BSC), Gaussian wiretap channels and Poisson wiretap-channels. The key concept of a pair of reliability function and security function has played the crucial role throughout in this paper.

On the other hand, we have introduced five measures of security (including a very typical mutual information), where the relation among these measures was demonstrated in Fig.1. The secrecy capacities according to these five measures were also defined. Among others, the formula for the δ-criterion is the strongest one, which was invoked everywhere in this paper when cost constraint is considered with the strongest security measure. Incidentally, superiority of the maximum security criterion to the average security criterion was demonstrated.

First, we have given a formula for reliability and security functions under *cost constraint* for a general wiretap channel, which was then particularized to that for a stationary memoryless channel to obtain further specific insights into the dual structure of a pair of reliability and security functions, quantitatively specified by the pair $(\phi(\rho|W_B, q, r),\ \phi(-\rho|W_E, q, r))$.

Next, we have investigated in details one of typically important channels: the Poisson wiretap channel, whose security-theoretic features have been clarified again from the viewpoint of a pair of reliability and security functions, where the formula for δ-secrecy capacity also naturally followed from the same point of view.

Similarly, also for the Gaussian wiretap channel it was possible to establish the δ-secrecy capacity and a pair of reliability and security functions as well, where we had four formulas for reliability and security functions depending on different upper bounding techniques on the characteristic function $\chi(\mathbf{x})$ to ensure the cost constraint: one of them is due to Gallager [16] and the other three are demonstrated for the first time in this paper. These were shown to have the two-folded dualities (cf. Fig. 17). Here, our δ-secrecy capacity formula under the maximum criterion with power constraint is stronger than that of Cheong and Hellman [26] from the viewpoint of security.

Moreover, we have introduced the concept of *concatenation* in order to enhance performance of the wiretap channel. Two ways to control the tradeoff between reliability and security were shown to be possible on the basis of *concatenation* and *rate exchange*, respectively.

Interestingly enough, it turned out that cost constraint Γ (with cost $c(x)$)

on X^n of the concatenated channel (W_B^{n+}, W_E^{n+}) is equivalent to cost constraint Γ (with cost $\overline{c}(v)$) on the input V^n of (W_B^{n+}, W_E^{n+}), where $\overline{c}(v) = \sum_{x\in\mathcal{X}} c(x)P_{X|V}(x|v)$. This principle has enabled us to drastically simplify the performance analysis of concatenated Poisson wiretap channels. Otherwise, a lot of tedious involved computations are needed as was shown in Appendix II and III.

Finally, it should be noted that *degradedness* of a Poisson wiretap channel as well as a Gaussian wiretap channel was assumed to guarantee the *more capability* as defined in Section 2, which enabled us to dispense with concatenation in calculating the secrecy capacity formula alone.

Appendix

Appendix I: *Proof of Theorem 4.2*

1) First let us give an information spectrum proof for the converse relying on the general formula (4.4). As in the proof of (2.20) in Theorem 2.2, we have

$$\text{p-}\liminf_{n\to\infty} \frac{1}{n}\log\frac{W_B^{n+}(Y^n|V^n)}{P_{Y^n}(Y^n)} \leq \liminf_{n\to\infty}\frac{1}{n}I(V^n;Y^n), \tag{A.1}$$

$$\text{p-}\limsup_{n\to\infty} \frac{1}{n}\log\frac{W_E^{n+}(Z^n|V^n)}{P_{Z^n}(Z^n)} \geq \limsup_{n\to\infty}\frac{1}{n}I(V^n;Z^n). \tag{A.2}$$

Hence, with any Markov chain $\mathbf{V}\to\mathbf{X}\to\mathbf{YZ}$ satisfying $\mathbf{X}\in\mathcal{S}_\Gamma$ as in Theorem 4.1 we have

$$\begin{aligned}
&\text{p-}\liminf_{n\to\infty} \frac{1}{n}\log\frac{W_B^{n+}(Y^n|V^n)}{P_{Y^n}(Y^n)} - \text{p-}\limsup_{n\to\infty} \frac{1}{n}\log\frac{W_E^{n+}(Z^n|V^n)}{P_{Z^n}(Z^n)}\\
&\leq \liminf_{n\to\infty}\left(\frac{1}{n}I(V^n;Y^n)-\frac{1}{n}I(V^n;Z^n)\right)\\
&\leq \sup_{VX:\mathrm{E}c(X)\leq\Gamma}(I(V;Y)-I(V;Z)),
\end{aligned} \tag{A.3}$$

where the last inequality follows in parallel with the argument employed in El Gamal and Kim [20, p.555] with V^n in place of M, yielding Markov chain $V\to X\to YZ$ in view that the wiretap channel in consideration is stationary and memoryless, and also it is not difficult to check that X here satisfies input constraint $\mathrm{E}c(X)\leq\Gamma$. Since $\mathbf{V}\to\mathbf{X}\to\mathbf{YZ}$ is arbitrary as far as $\mathbf{X}\in\mathcal{S}_\Gamma$, Theorem 4.1 concludes that

$$\text{d-}C_s(\Gamma) \le \sup_{VX:\mathrm{E}c(X)\le\Gamma} (I(V;Y) - I(V;Z)). \tag{A.4}$$

Thus, the convese part has been established.

2) The direct part is *not* trivial, which is shown by proceeding in parallel with the information spectrum proof of Han [19, Theorem 3.6.2]. Letting $\delta > 0$ be an arbitrarily small constant, define $\overline{V} \to \overline{X} \to \overline{YZ}$ as $V \to X \to YZ$ attaining the supremum

$$\sup_{VX:\mathrm{E}c(X)\le\Gamma-\delta} (I(V;Y) - I(V;Z)).$$

Let

$$\begin{aligned} \overline{V}^n &= (\overline{V}_1, \overline{V}_2, \cdots, \overline{V}_n), \\ \overline{X}^n &= (\overline{X}_1, \overline{X}_2, \cdots, \overline{X}_n), \\ \overline{Y}^n &= (\overline{Y}_1, \overline{Y}_2, \cdots, \overline{Y}_n), \\ \overline{Z}^n &= (\overline{Z}_1, \overline{Z}_2, \cdots, \overline{Z}_n) \end{aligned}$$

be stationary memoryless sequences subject to $P_{\overline{V}}, P_{\overline{X}}, P_{\overline{Y}}, P_{\overline{Z}}$, respectively, and set $\overline{\mathbf{V}} = \{\overline{V}^n\}_{n=1}^{\infty}$, $\overline{\mathbf{X}} = \{\overline{X}^n\}_{n=1}^{\infty}$, $\overline{\mathbf{Y}} = \{\overline{Y}^n\}_{n=1}^{\infty}$, $\overline{\mathbf{Z}} = \{\overline{Z}^n\}_{n=1}^{\infty}$. Then, from Khintchin's law of large numbers it holds that

$$\underline{I}(\overline{\mathbf{V}};\overline{\mathbf{Y}}) = I(\overline{V};\overline{Y}), \quad \overline{I}(\overline{\mathbf{V}};\overline{\mathbf{Z}}) = I(\overline{V};\overline{Z}), \tag{A.5}$$

where we use the notation that, for any general sources $\mathbf{S} = \{S^n\}_{n=1}^{\infty}, \mathbf{T} = \{T^n\}_{n=1}^{\infty}$,

$$\underline{I}(\mathbf{S};\mathbf{T}) = \text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{P_{T^n|S^n}(T^n|S^n)}{P_{T^n}(T^n)},$$

$$\overline{I}(\mathbf{S};\mathbf{T}) = \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{P_{T^n|S^n}(T^n|S^n)}{P_{T^n}(T^n)}.$$

Notice here that $\overline{\mathbf{X}} \in \mathcal{S}_\Gamma$ is not guaranteed. We construct $\mathbf{X}_1$ satisfying $\mathbf{X}_1 \in \mathcal{S}_\Gamma$ from $\overline{\mathbf{X}}$ in the following way. First, we note that $c(\overline{X}_i)$ $(i = 1, 2, \cdots, n)$ are independent and identically distributed. By taking

$$\frac{1}{n} \sum_{i=1}^{n} \mathrm{E}c(\overline{X}_i) \le \Gamma - \delta$$

into consideration, we obtain

$$\Pr\left\{\frac{1}{n}\sum_{i=1}^{n} c(\overline{X}_i) \le \Gamma\right\} \to 1 \quad \text{as } n \to \infty$$

due to Khinchin's law of large numbers again. Hence, setting

$$A_n \equiv \left\{\mathbf{x} \in \mathcal{X}^n \,\middle|\, \frac{1}{n}\sum_{i=1}^{n} c(x_i) \le \Gamma\right\},$$

it follows that

$$a_n \equiv P_{\overline{X}^n}(A_n) \to 1 \quad \text{as } n \to \infty. \tag{A.6}$$

If we define the input process $\mathbf{X}_1 = \{X_1^n\}_{n=1}^{\infty}$ by

$$P_{X_1^n}(\mathbf{x}) = \begin{cases} \dfrac{1}{a_n} P_{\overline{X}^n}(\mathbf{x}) & \text{for} \quad \mathbf{x} \in A_n, \\ \quad 0 & \text{for} \quad \mathbf{x} \notin A_n, \end{cases}$$

then it is evident that $\mathbf{X}_1 \in \mathcal{S}_\Gamma$. Let us here define Markov chain $V_1^n \to X_1^n \to Y_1^n Z_1^n$ so that $P_{Y_1^n|X_1^n}(\mathbf{y}|\mathbf{x}) = W_B^n(\mathbf{y}|\mathbf{x})$, $P_{Z_1^n|X_1^n}(\mathbf{z}|\mathbf{x}) = W_E^n(\mathbf{z}|\mathbf{x})$, $P_{V_1^n|X_1^n}(\mathbf{v}|\mathbf{x}) = P_{V^n|X^n}(\mathbf{v}|\mathbf{x})$, and set $\mathbf{V}_1 = \{V_1^n\}_{n=1}^{\infty}$, $\mathbf{Y}_1 = \{Y_1^n\}_{n=1}^{\infty}$, $\mathbf{Z}_1 = \{Z_1^n\}_{n=1}^{\infty}$. Then, for any $\mathbf{y} \in \mathcal{Y}^n$ it holds that

$$\begin{aligned} P_{\overline{Y}^n}(\mathbf{y}) &= \sum_{\mathbf{x}\in\mathcal{X}^n} W^n(\mathbf{y}|\mathbf{x}) P_{\overline{X}^n}(\mathbf{x}) \\ &\ge \sum_{\mathbf{x}\in A_n} W^n(\mathbf{y}|\mathbf{x}) P_{\overline{X}^n}(\mathbf{x}) \\ &= a_n \sum_{\mathbf{x}\in A_n} W^n(\mathbf{y}|\mathbf{x}) P_{X_1^n}(\mathbf{x}) \\ &= a_n P_{Y_1^n}(\mathbf{y}). \end{aligned}$$

Hence,

$$\frac{1}{n}\log\frac{1}{P_{\overline{Y}^n}(Y_1^n)} \le \frac{1}{n}\log\frac{1}{P_{Y_1^n}(Y_1^n)} + \frac{1}{n}\log\frac{1}{a_n}. \tag{A.7}$$

Similarly,

$$\frac{1}{n}\log\frac{1}{P_{\overline{V}^n}(V_1^n)} \le \frac{1}{n}\log\frac{1}{P_{V_1^n}(V_1^n)} + \frac{1}{n}\log\frac{1}{a_n}. \tag{A.8}$$

On the other hand, as was shown in the proof of Han [19, Lemma 1.4.1], it holds that

$$\Pr\left\{\frac{1}{n}\log\frac{1}{P_{V_1^n Y_1^n}(V_1^n Y_1^n)} \le \frac{1}{n}\log\frac{1}{P_{\overline{V}^n\overline{Y}^n}(V_1^n Y_1^n)} + \gamma_n\right\} \ge 1 - e^{-n\gamma_n}, \tag{A.9}$$

where

$$\gamma_n \to 0,\ n\gamma_n \to \infty \quad \text{as } n \to \infty.$$

Take an arbitrary constant α such that $\alpha < \underline{I}(\overline{\mathbf{V}};\overline{\mathbf{Y}})$, then summarizing (A.7), (A.8) and (A.9) with $\overline{W}_B^{n+}(\mathbf{y}|\mathbf{v}) = P_{\overline{Y}^n|\overline{V}^n}(\mathbf{y}|\mathbf{v})$ yields

$$\begin{aligned}
&\Pr\left\{\frac{1}{n}\log\frac{\overline{W}_B^{n+}(Y_1^n|V_1^n)}{P_{\overline{Y}^n}(Y_1^n)} < \alpha\right\} \\
&= \Pr\left\{\frac{1}{n}\log\frac{P_{\overline{V}^n\overline{Y}^n}(V_1^nY_1^n)}{P_{\overline{V}^n}(V_1^n)P_{\overline{Y}^n}(Y_1^n)} < \alpha\right\} \\
&\geq \Pr\left\{\frac{1}{n}\log\frac{P_{V_1^nY_1^n}(V_1^nY_1^n)}{P_{V_1^n}(V_1^n)P_{Y_1^n}(Y_1^n)} < \alpha - \frac{2}{n}\log\frac{1}{a_n} - \gamma_n\right\} - e^{-n\gamma_n} \\
&= \Pr\left\{\frac{1}{n}\log\frac{W_B^{n+}(Y_1^n|V_1^n)}{P_{Y_1^n}(Y_1^n)} < \alpha - \frac{2}{n}\log\frac{1}{a_n} - \gamma_n\right\} - e^{-n\gamma_n}. \qquad \text{(A.10)}
\end{aligned}$$

Moreover, noting that

$$\begin{aligned}
&\Pr\left\{\frac{1}{n}\log\frac{\overline{W}_B^{n+}(\overline{Y}^n|\overline{V}^n)}{P_{\overline{Y}^n}(\overline{Y}^n)} < \alpha\right\} \\
&\geq a_n \Pr\left\{\frac{1}{n}\log\frac{\overline{W}_B^{n+}(Y_1^n|V_1^n)}{P_{\overline{Y}^n}(Y_1^n)} < \alpha\right\} \qquad \text{(A.11)}
\end{aligned}$$

and substituting this inequality into (A.10) yields

$$\begin{aligned}
&\Pr\left\{\frac{1}{n}\log\frac{\overline{W}_B^{n+}(\overline{Y}^n|\overline{V}^n)}{P_{\overline{Y}^n}(\overline{Y}^n)} < \alpha\right\} \\
&\geq a_n \Pr\left\{\frac{1}{n}\log\frac{W_B^{n+}(Y_1^n|V_1^n)}{P_{Y_1^n}(Y_1^n)} < \alpha - \frac{2}{n}\log\frac{1}{a_n} - \gamma_n\right\} - a_n e^{-n\gamma_n}.
\end{aligned}$$

By noticing $\alpha < \underline{I}(\overline{\mathbf{X}};\overline{\mathbf{Y}})$ and $a_n \to 1$ as $n \to \infty$, we obtain

$$\lim_{n\to\infty}\Pr\left\{\frac{1}{n}\log\frac{W_B^{n+}(Y_1^n|V_1^n)}{P_{Y_1^n}(Y_1^n)} < \alpha - \frac{2}{n}\log\frac{1}{a_n} - \gamma_n\right\} = 0. \qquad \text{(A.12)}$$

Since α is arbitrary as far as it satisfies $\alpha < \underline{I}(\overline{\mathbf{V}};\overline{\mathbf{Y}})$, (A.12) implies

$$\underline{I}(\overline{\mathbf{V}};\overline{\mathbf{Y}}) \leq \underline{I}(\mathbf{V}_1;\mathbf{Y}_1). \qquad \text{(A.13)}$$

Next, we will show

$$\overline{I}(\overline{\mathbf{V}};\overline{\mathbf{Z}}) \geq \overline{I}(\mathbf{V}_1;\mathbf{Z}_1). \qquad \text{(A.14)}$$

To do so, we use, instead of inequalities (A.7), (A.8), (A.9) and (A.11), the following inequalities

$$\Pr\left\{\frac{1}{n}\log\frac{1}{P_{Z_1^n}(Z_1^n)} \le \frac{1}{n}\log\frac{1}{P_{\overline{Z}^n}(Z_1^n)} + \gamma_n\right\} \ge 1 - e^{-n\gamma_n}, \tag{A.15}$$

$$\Pr\left\{\frac{1}{n}\log\frac{1}{P_{V_1^n}(V_1^n)} \le \frac{1}{n}\log\frac{1}{P_{\overline{V}^n}(V_1^n)} + \gamma_n\right\} \ge 1 - e^{-n\gamma_n}, \tag{A.16}$$

$$\frac{1}{n}\log\frac{1}{P_{\overline{V}^n\overline{Z}^n}(V_1^n Z_1^n)} \le \frac{1}{n}\log\frac{1}{P_{V_1^n Z_1^n}(V_1^n Z_1^n)} + \frac{1}{n}\log\frac{1}{a_n}, \tag{A.17}$$

and

$$\begin{aligned}
&\Pr\left\{\frac{1}{n}\log\frac{\overline{W}_E^{n+}(\overline{Z}^n|\overline{V}^n)}{P_{\overline{Z}^n}(\overline{Z}^n)} > \beta\right\} \\
&\quad\ge\quad a_n \Pr\left\{\frac{1}{n}\log\frac{\overline{W}_E^{n+}(Z_1^n|V_1^n)}{P_{\overline{Z}^n}(Z_1^n)} > \beta\right\},
\end{aligned} \tag{A.18}$$

where β is any constant such that $\beta > \overline{I}(\mathbf{\overline{V}};\mathbf{\overline{Z}})$. Summarizing (A.15), (A.16), (A.17) and proceeding in parallel with the above argument, we obtain

$$\begin{aligned}
&\Pr\left\{\frac{1}{n}\log\frac{\overline{W}_E^{n+}(Z_1^n|V_1^n)}{P_{\overline{Z}^n}(Z_1^n)} > \beta\right\} \\
&= \Pr\left\{\frac{1}{n}\log\frac{P_{\overline{V}^n\overline{Z}^n}(V_1^n Z_1^n)}{P_{\overline{V}^n}(V_1^n)P_{\overline{Z}^n}(Z_1^n)} > \beta\right\} \\
&\ge \Pr\left\{\frac{1}{n}\log\frac{P_{V_1^n Z_1^n}(V_1^n Z_1^n)}{P_{V_1^n}(V_1^n)P_{Z_1^n}(Z_1^n)} > \beta + \frac{1}{n}\log\frac{1}{a_n} + 2\gamma_n\right\} - 2e^{-n\gamma_n} \\
&= \Pr\left\{\frac{1}{n}\log\frac{W_E^{n+}(Z_1^n|V_1^n)}{P_{Z_1^n}(Z_1^n)} > \beta + \frac{1}{n}\log\frac{1}{a_n} + 2\gamma_n\right\} - 2e^{-n\gamma_n}
\end{aligned} \tag{A.19}$$

from which, together with (A.18), it follows that

$$\begin{aligned}
&\Pr\left\{\frac{1}{n}\log\frac{\overline{W}_E^{n+}(\overline{Z}^n|\overline{V}^n)}{P_{\overline{Z}^n}(\overline{Z}^n)} > \beta\right\} \\
&\ge a_n \Pr\left\{\frac{1}{n}\log\frac{W_E^{n+}(Z_1^n|V_1^n)}{P_{Z_1^n}(Z_1^n)} > \beta + \frac{1}{n}\log\frac{1}{a_n} + 2\gamma_n\right\} - 2a_n e^{-n\gamma_n}.
\end{aligned}$$

As a consequence,

$$\lim_{n\to\infty}\Pr\left\{\frac{1}{n}\log\frac{W_E^{n+}(Z_1^n|V_1^n)}{P_{Z_1^n}(Z_1^n)} > \beta + \frac{1}{n}\log\frac{1}{a_n} + 2\gamma_n\right\} = 0. \tag{A.20}$$

Since β is arbitrary as far as it satisfies $\beta > \overline{I}(\overline{\mathbf{V}};\overline{\mathbf{Z}})$, (A.20) implies (A.14). From (A.5), (A.13) and (A.14) it follows that

$$I(\overline{V};\overline{Y}) \le \underline{I}(\mathbf{V}_1;\mathbf{Y}_1), \quad I(\overline{V};\overline{Z}) \ge \overline{I}(\mathbf{V}_1;\mathbf{Z}_1).$$

Hence,

$$I(\overline{V};\overline{Y}) - I(\overline{V};\overline{Z}) \le \underline{I}(\mathbf{V}_1;\mathbf{Y}_1) - \overline{I}(\mathbf{V}_1;\mathbf{Z}_1).$$

Therefore, in view of $\mathbf{X}_1 \in \mathcal{S}_\Gamma$ and by the definition of $\overline{VXYZ}$, we have

$$\begin{aligned}
&\sup_{VX:\mathrm{E}c(X)\le\Gamma-\delta} (I(V;Y) - I(V;Z)) \\
&\le \sup_{\mathbf{VX}:\mathbf{X}\in\mathcal{S}_\Gamma} (\underline{I}(\mathbf{V};\mathbf{Y}) - \overline{I}(\mathbf{V};\mathbf{Z})) = \text{d-}C_s(\Gamma). \qquad \text{(A.21)}
\end{aligned}$$

On the other hand, notice that $I(V;Y)$, $I(V;Z)$, $\mathrm{E}c(X)$ are all continuous in P_{VXY} and P_{VXZ} under the L_1 metric, then it is not difficult to verify by the contradiction argument that

$$h(\Gamma) \stackrel{\Delta}{=} \sup_{VX:\mathrm{E}c(X)\le\Gamma} (I(V;Y) - I(V;Z)) \qquad \text{(A.22)}$$

is left-continuous in Γ. Thus, letting $\delta \to 0$ in (A.21) we obtain

$$\sup_{VX:\mathrm{E}c(X)\le\Gamma} (I(V;Y) - I(V;Z)) \le \text{d-}C_s(\Gamma).$$

Appendix II: *Another proof of Theorem 10.3*

We begin with

$$\begin{aligned}
&\phi(\rho|W_B,q,r) \\
&= -\log\left[\sum_{y\in\mathcal{Y}}\left(\sum_{v\in\mathcal{V}} q(v)\left[\sum_{x\in\mathcal{X}} W_B(y|x)P_{X|V}(x|v)e^{(1+\rho)r[\Gamma-c(x)]}\right]^{\frac{1}{1+\rho}}\right)^{1+\rho}\right], \\
&= -\log\left[\sum_{y=0}^{1}\left(\sum_{v=0}^{1} q(v)\left[\sum_{x=0}^{1} W_B(y|x)P_{X|V}(x|v)e^{(1+\rho)r[\Gamma-x]}\right]^{\frac{1}{1+\rho}}\right)^{1+\rho}\right], \\
&= -\Gamma(1+\rho)r - \log\sum_{y=0}^{1} V_y^{1+\rho}, \qquad \text{(A.23)}
\end{aligned}$$

where

$$V_y = \sum_{v=0}^{1} q(v) \left[\sum_{x=0}^{1} W_B(y|x) P_{X|V}(x|v) e^{-(1+\rho)rx} \right]^{\frac{1}{1+\rho}} \quad (y=0,1), \qquad \text{(A.24)}$$

which are computed, up to the first order, as follows.

$$\begin{aligned}
V_0 & \\
= \;& q\Big[(1-a)(1-s_yA_y\Delta) + a(1-(1+s_y)A_y\Delta)e^{-(1+\rho)r}\Big]^{\frac{1}{1+\rho}} \\
& +(1-q)\Big[(1-b)(1-s_yA_y\Delta) + b(1-(1+s_y)A_y\Delta)e^{-(1+\rho)r}\Big]^{\frac{1}{1+\rho}} \\
= \;& q\Big[(1-a+ae^{-(1+\rho)r}) - ((1-a)s_yA_y\Delta + a(1+s_y)A_y\Delta e^{-(1+\rho)r})\Big]^{\frac{1}{1+\rho}} \\
& +(1-q)\Big[(1-b+be^{-(1+\rho)r}) - ((1-b)s_yA_y\Delta + b(1+s_y)A_y\Delta e^{-(1+\rho)r})\Big]^{\frac{1}{1+\rho}} \\
\simeq \;& q\alpha^{\frac{1}{1+\rho}}\left[1 - \frac{A_y\Delta\alpha_s}{(1+\rho)\alpha}\right] + (1-q)\beta^{\frac{1}{1+\rho}}\left[1 - \frac{A_y\Delta\beta_s}{(1+\rho)\beta}\right] \\
= \;& \left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right)\left[1 - \frac{A_y\Delta(q\alpha_s\alpha^{-1} + (1-q)\beta_s\beta^{-1})}{(1+\rho)(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}})}\right], \qquad \text{(A.25)}
\end{aligned}$$

where we have set

$$\begin{aligned}
\alpha &= 1-a+ae^{-(1+\rho)r}, \\
\beta &= 1-b+be^{-(1+\rho)r}, \\
\alpha_s &= (1-a)s_y + a(1+s_y)e^{-(1+\rho)r}, \\
\beta_s &= (1-b)s_y + b(1+s_y)e^{-(1+\rho)r}.
\end{aligned}$$

Hence,

$$V_0^{1+\rho} \simeq \left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right)^{1+\rho}\left[1 - \frac{A_y\Delta(q\alpha_s\alpha^{-1} + (1-q)\beta_s\beta^{-1})}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}}\right]. \qquad \text{(A.26)}$$

Similarly,

$$\begin{aligned}
V_1 & \\
= \;& q\Big[(1-a)s_yA_y\Delta + a(1+s_y)A_y\Delta e^{-(1+\rho)r}\Big]^{\frac{1}{1+\rho}}
\end{aligned}$$

$$
\begin{aligned}
&+(1-q)\Big[(1-b)s_y A_y\Delta + b(1+s_y)A_y\Delta e^{-(1+\rho)r}\Big]^{\frac{1}{1+\rho}} \\
= \;& (A_y\Delta)^{\frac{1}{1+\rho}}\left[q\alpha_s^{\frac{1}{1+\rho}} + (1-q)\beta_s^{\frac{1}{1+\rho}}\right] \\
= \;& (A_y\Delta)^{\frac{1}{1+\rho}}\left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right)\left[\frac{q\alpha_s^{\frac{1}{1+\rho}} + (1-q)\beta_s^{\frac{1}{1+\rho}}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}}\right].
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&V_1^{1+\rho} \\
= \;& (A_y\Delta)\left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right)^{1+\rho}\left[\frac{q\alpha_s^{\frac{1}{1+\rho}} + (1-q)\beta_s^{\frac{1}{1+\rho}}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}}\right]^{1+\rho}.
\end{aligned}
\tag{A.27}
$$

As a consequence, substituting (A.26) and (A.27) into (A.23) and rearranging it, we obtain

$$
\begin{aligned}
&\phi(\rho|W_B, q, r) \\
= \;& -(1+\rho)\log\left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right) - r(1+\rho)\Gamma \\
& -\log\left(1 - A_y\Delta\left[\frac{q\alpha_s\alpha^{-1} + (1-q)\beta_s\beta^{-1}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}}\right]\right. \\
& \left. +A_y\Delta\left[\frac{q\alpha_s^{\frac{1}{1+\rho}} + (1-q)\beta_s^{\frac{1}{1+\rho}}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}}\right]^{1+\rho}\right) \\
\simeq \;& -(1+\rho)\log\left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right) - r(1+\rho)\Gamma \\
& +A_y\Delta\left(\frac{q\alpha_s\alpha^{-1} + (1-q)\beta_s\beta^{-1}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}} - \left[\frac{q\alpha_s^{\frac{1}{1+\rho}} + (1-q)\beta_s^{\frac{1}{1+\rho}}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}}\right]^{1+\rho}\right).
\end{aligned}
\tag{A.28}
$$

Set

$$
g(r) = -(1+\rho)\log\left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right) - r(1+\rho)\Gamma,
$$

then

$$
g'(r) = -(1+\rho)\Gamma + (1+\rho)\left[\frac{aq\alpha^{\frac{1}{1+\rho}-1} + b(1-q)\beta^{\frac{1}{1+\rho}-1}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}}\right]e^{-(1+\rho)r}.
$$

Therefore, it is obvious that $g(0) = 0$ and moreover,

$$\begin{aligned} g'(0) &= -(1+\rho)\Gamma + (1+\rho)(aq + b(1-q)) \\ &= -(1+\rho)(\Gamma - (aq + b(1-q))) \\ &\leq 0, \end{aligned}$$

where the last step follows from cost constraint (10.18), in that $aq+b(1-q) \triangleq p \leq \Gamma$ is equivalent to $q \leq \Gamma^+$ (cf. also (10.17)). We now want to show that $g(r)$ is concave with respect to r. To this end, it suffices to show that

$$h(r) = -(1+\rho)\log\left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right)$$

is concave with respect to r. First, we see that

$$h'(r) = (1+\rho)\frac{aq\alpha^{\frac{1}{1+\rho}-1} + b(1-q)\beta^{\frac{1}{1+\rho}-1}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}}e^{-(1+\rho)r}.$$

Then,

$$\begin{aligned} \log h'(r) &= \log\left(aq\alpha^{\frac{1}{1+\rho}-1} + b(1-q)\beta^{\frac{1}{1+\rho}-1}\right) \\ &\quad -\log\left(q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}\right) \\ &\quad -(1+\rho)r + \log(1+\rho). \end{aligned}$$

Hence,

$$\begin{aligned} \left(\log h'(r)\right)' &= \frac{\rho\left[a^2q\alpha^{\frac{1}{1+\rho}-2} + b^2(1-q)\beta^{\frac{1}{1+\rho}-2}\right]e^{-(1+\rho)r}}{aq\alpha^{\frac{1}{1+\rho}-1} + b(1-q)\beta^{\frac{1}{1+\rho}-1}} \\ &\quad + \frac{\left[aq\alpha^{\frac{1}{1+\rho}-1} + b(1-q)\beta^{\frac{1}{1+\rho}-1}\right]e^{-(1+\rho)r}}{q\alpha^{\frac{1}{1+\rho}} + (1-q)\beta^{\frac{1}{1+\rho}}} \\ &\quad -(1+\rho). \end{aligned} \tag{A.29}$$

We then observe that

$$\frac{ae^{-(1+\rho)r}}{\alpha} = \frac{ae^{-(1+\rho)r}}{1-a+ae^{-(1+\rho)r}} \leq 1,$$

$$\frac{be^{-(1+\rho)r}}{\beta} = \frac{be^{-(1+\rho)r}}{1-b+be^{-(1+\rho)r}} \leq 1,$$

from which it follows that

$$\frac{\rho\left[a^2q\alpha^{\frac{1}{1+\rho}-2}+b^2(1-q)\beta^{\frac{1}{1+\rho}-2}\right]e^{-(1+\rho)r}}{aq\alpha^{\frac{1}{1+\rho}-1}+b(1-q)\beta^{\frac{1}{1+\rho}-1}}\le\rho,$$

$$\frac{\left[aq\alpha^{\frac{1}{1+\rho}-1}+b(1-q)\beta^{\frac{1}{1+\rho}-1}\right]e^{-(1+\rho)r}}{q\alpha^{\frac{1}{1+\rho}}+(1-q)\beta^{\frac{1}{1+\rho}}}\le 1,$$

which, together with (A.29), yields

$$\left(\log h'(r)\right)'\le 0,$$

and therefore $h''(r)\le 0$, so that $g''(r)\le 0$. Thus, we see that $g(r)$ is concave with respect to r, and it is concluded that

$$\max_{r\ge 0} g(r)=g(0)=0.$$

Consequently,

$$\begin{aligned}
&\phi(\rho|W_B,q)\\
&\quad\stackrel{\Delta}{=}\ \max_{r\ge0}\phi(\rho|W_B,q,r)\\
&\quad=\ A_y\Delta\left[(1-q)(s_y+b)+q(s_y+a)\right]\\
&\qquad -A_y\Delta\left[(1-q)(s_y+b)^{\frac{1}{1+\rho}}+q(s_y+a)^{\frac{1}{1+\rho}}\right]^{1+\rho}\\
&\quad=\ A_y\Delta\left[s_y+b+q(a-b)\right]\\
&\qquad -A_y\Delta(s_y+b)\left[(1-q)+q\left(\frac{s_y+a}{s_y+b}\right)^{\frac{1}{1+\rho}}\right]^{1+\rho}\\
&\quad=\ A_y(a-b)\Delta\left[\frac{s_y+b}{a-b}+q\right]\\
&\qquad -A_y\Delta(s_y+b)\left[(1-q)+q\left(1+\frac{a-b}{s_y+b}\right)^{\frac{1}{1+\rho}}\right]^{1+\rho}\\
&\quad=\ A_y(a-b)\Delta\left[\frac{s_y+b}{a-b}+q\right]\\
&\qquad -A_y(a-b)\Delta\left(\frac{s_y+b}{a-b}\right)\left[(1-q)+q\left(1+\frac{a-b}{s_y+b}\right)^{\frac{1}{1+\rho}}\right]^{1+\rho}\\
&\quad=\ A_y^+\Delta\left[s_y^++q\right]-A_y^+\Delta s_y^+\left[1+\tau_y^+q\right]^{1+\rho},
\end{aligned}\tag{A.30}$$

where we have used the notation as in Theorem 10.3. Thus, the exponent per second $E_B^+(\rho,q) \stackrel{\Delta}{=} \phi(\rho|W_B,q)/\Delta$ is given by

$$E_B^+(\rho,q) = A_y^+\left[s_y^+ + q - s_y^+\left(1+\tau_y^+ q\right)^{1+\rho}\right], \tag{A.31}$$

which coincides with (10.24), from which equation (10.26) also follows, thus completing the another proof. □

Appendix III: *Another proof of Theorem 10.4*

In view of (5.29) and (5.31), we begin with

$$\phi(-\rho|W_E,q,r) = -\Gamma(1-\rho)r - \log\sum_{z=0}^{1} V_z^{1-\rho} \tag{A.32}$$

where

$$V_z = \sum_{v=0}^{1} q(v)\left[\sum_{x=0}^{1} W_E(z|x)P_{X|V}(x|v)e^{-(1-\rho)rx}\right]^{\frac{1}{1-\rho}} \quad (z=0,1). \tag{A.33}$$

We can compute them in the same way as in Appendix II to obtain

$$\begin{aligned}
&\phi(-\rho|W_B,q,r)\\
&\simeq \; -(1-\rho)\log\left(q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}\right) - r(1-\rho)\Gamma\\
&\quad +A_z\Delta\left(\frac{q\alpha_{0s}\alpha_0^{-1} + (1-q)\beta_{0s}\beta_0^{-1}}{q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}} - \left[\frac{q\alpha_{0s}^{\frac{1}{1-\rho}} + (1-q)\beta_{0s}^{\frac{1}{1-\rho}}}{q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}}\right]^{1-\rho}\right),
\end{aligned} \tag{A.34}$$

where we have set

$$\begin{aligned}
\alpha_0 &= 1-a+ae^{-(1-\rho)r},\\
\beta_0 &= 1-b+be^{-(1-\rho)r},\\
\alpha_{0s} &= (1-a)s_y + a(1+s_y)e^{-(1-\rho)r},\\
\beta_{0s} &= (1-b)s_y + b(1+s_y)e^{-(1-\rho)r}.
\end{aligned}$$

Set

$$g_0(r) = -(1-\rho)\log\left(q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}\right) - r(1-\rho)\Gamma,$$

then

$$g_0'(r) = -(1-\rho)\Gamma + (1-\rho)\left[\frac{aq\alpha_0^{\frac{1}{1-\rho}-1} + b(1-q)\beta_0^{\frac{1}{1-\rho}-1}}{q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}}\right]e^{-(1-\rho)r}.$$

Therefore it is obvious that $g_0(0) = 0$ and $g_0'(0) \leq 0$ if cost constraint is valid. We now want to show that $g_0(r)$ is concave with respect to r. To this end, it suffices to show that

$$h_0(r) = -(1-\rho)\log\left(q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}\right)$$

is concave with respect to r. First, we see that

$$\begin{aligned}
\left(\log h_0'(r)\right)' \quad = \quad & -\frac{\rho\left[a^2q\alpha_0^{\frac{1}{1-\rho}-2} + b^2(1-q)\beta_0^{\frac{1}{1-\rho}-2}\right]e^{-(1-\rho)r}}{aq\alpha_0^{\frac{1}{1-\rho}-1} + b(1-q)\beta_0^{\frac{1}{1-\rho}-1}} \\
& + \frac{\left[aq\alpha_0^{\frac{1}{1-\rho}-1} + b(1-q)\beta_0^{\frac{1}{1-\rho}-1}\right]e^{-(1-\rho)r}}{q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}} \\
& -(1-\rho).
\end{aligned} \tag{A.35}$$

with the same calculation as in Appendix II. From the definition of α_0 and β_0 we obtain

$$\begin{aligned}
ae^{-(1-\rho)r} &= \alpha_0 - (1-a), \\
be^{-(1-\rho)r} &= \beta_0 - (1-b),
\end{aligned}$$

and substituting them into (A.35) and rearranging it, we obtain

$$\begin{aligned}
&\left(\log h_0'(r)\right)' \\
&= -\frac{\rho\left[aq\alpha_0^{\frac{1}{1-\rho}-1} - a(1-a)q\alpha_0^{\frac{1}{1-\rho}-2} + b(1-q)\beta_0^{\frac{1}{1-\rho}-1} - b(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-2}\right]}{aq\alpha_0^{\frac{1}{1-\rho}-1} + b(1-q)\beta_0^{\frac{1}{1-\rho}-1}} \\
&\quad + \frac{\left[q\alpha_0^{\frac{1}{1-\rho}} - (1-a)q\alpha_0^{\frac{1}{1-\rho}-1} + (1-q)\beta_0^{\frac{1}{1-\rho}} - (1-b)(1-q)\beta_0^{\frac{1}{1-\rho}}\right]}{q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}}
\end{aligned}$$

$$
\begin{aligned}
& -\frac{q\alpha_0^{\frac{1}{1-\rho}}+(1-q)\beta_0^{\frac{1}{1-\rho}}}{q\alpha_0^{\frac{1}{1-\rho}}+(1-q)\beta_0^{\frac{1}{1-\rho}}}+\frac{\rho\left[aq\alpha_0^{\frac{1}{1-\rho}-1}+b(1-q)\beta_0^{\frac{1}{1-\rho}-1}\right]}{aq\alpha_0^{\frac{1}{1-\rho}-1}+b(1-q)\beta_0^{\frac{1}{1-\rho}-1}} \\
= & \ \frac{\rho\left[a(1-a)q(\alpha_0^{\frac{1}{1-\rho}-2}+b(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-2}\right]}{aq\alpha_0^{\frac{1}{1-\rho}-1}+b(1-q)\beta_0^{\frac{1}{1-\rho}-1}} \\
& -\frac{\left[(1-a)q\alpha_0^{\frac{1}{1-\rho}-1}+(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-1}\right]}{q\alpha_0^{\frac{1}{1-\rho}}+(1-q)\beta_0^{\frac{1}{1-\rho}}} \\
= & \ -(1-\rho)\frac{\left[a(1-a)q\alpha_0^{\frac{1}{1-\rho}-2}+b(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-2}\right]}{aq\alpha_0^{\frac{1}{1-\rho}-1}+b(1-q)\beta_0^{\frac{1}{1-\rho}-1}} \\
& +\frac{\left[a(1-a)q\alpha_0^{\frac{1}{1-\rho}-2}+b(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-2}\right]}{aq\alpha_0^{\frac{1}{1-\rho}-1}+b(1-q)\beta_0^{\frac{1}{1-\rho}-1}} \\
& -\frac{\left[(1-a)q\alpha_0^{\frac{1}{1-\rho}-1}+(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-1}\right]}{q\alpha_0^{\frac{1}{1-\rho}}+(1-q)\beta_0^{\frac{1}{1-\rho}}}. \qquad \text{(A.36)}
\end{aligned}
$$

We add the second term and the third term to obtain the fraction with a common denominator. The numerator is as follow:

$$
\begin{aligned}
& \left[a(1-a)q\alpha_0^{\frac{1}{1-\rho}-2}+b(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-2}\right]\cdot\left[q\alpha_0^{\frac{1}{1-\rho}}+(1-q)\beta_0^{\frac{1}{1-\rho}}\right] \\
& \quad -\left[(1-a)q\alpha_0^{\frac{1}{1-\rho}-1}+(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-1}\right]\cdot\left[aq\alpha_0^{\frac{1}{1-\rho}-1}+b(1-q)\beta_0^{\frac{1}{1-\rho}-1}\right] \\
= & \ q(1-q)\alpha_0^{\frac{1}{1-\rho}-2}\beta_0^{\frac{1}{1-\rho}-2} \\
& \times\left[b(1-b)\alpha_0^2+a(1-a)\beta_0^2-a(1-b)\alpha_0\beta_0-(1-a)b\alpha_0\beta_0\right] \\
= & \ q(1-q)\alpha_0^{\frac{1}{1-\rho}-2}\beta_0^{\frac{1}{1-\rho}-2}(a\beta_0-b\alpha_0)\left[(1-a)\beta_0-(1-b)\alpha_0\right] \\
= & \ -q(1-q)\alpha_0^{\frac{1}{1-\rho}-2}\beta_0^{\frac{1}{1-\rho}-2}(a-b)^2e^{-(1-\rho)r}, \qquad \text{(A.37)}
\end{aligned}
$$

where

$$a\beta_0-b\alpha_0$$

$$= a(1-b+be^{-(1-\rho)r}) - b(1-a+ae^{-(1-\rho)r})$$
$$= a-b$$

and

$$\begin{aligned}
&(1-a)\beta_0 - (1-b)\alpha_0 \\
&\quad = (1-a)(1-b+be^{-(1-\rho)r}) - (1-b)(1-a+ae^{-(1-\rho)r}) \\
&\quad = (1-a)be^{-(1-\rho)r} - a(1-b)e^{-(1-\rho)r} \\
&\quad = -(a-b)e^{-(1-\rho)r}.
\end{aligned}$$

Finally, through (A.36) and (A.37), we obtain

$$\begin{aligned}
&(\log(h_0'(r))' \\
&\quad = -(1-\rho)\frac{\left[a(1-a)q\alpha_0^{\frac{1}{1-\rho}-2} + b(1-b)(1-q)\beta_0^{\frac{1}{1-\rho}-2}\right]}{aq\alpha_0^{\frac{1}{1-\rho}-1} + b(1-q)\beta_0^{\frac{1}{1-\rho}-1}} \\
&\qquad - \frac{q(1-q)\alpha_0^{\frac{1}{1-\rho}-2}\beta_0^{\frac{1}{1-\rho}-2}(a-b)^2 e^{-(1-\rho)r}}{\left(aq\alpha_0^{\frac{1}{1-\rho}-1} + b(1-q)\beta_0^{\frac{1}{1-\rho}-1}\right)\left(q\alpha_0^{\frac{1}{1-\rho}} + (1-q)\beta_0^{\frac{1}{1-\rho}}\right)},
\end{aligned} \tag{A.38}$$

which yields

$$\left(\log(h_0'(r)\right)' \le 0,$$

and therefore $(h_0(r))'' \le 0$ so that $(g_0(r))'' \le 0$. Thus, we see that $g_0(r)$ is concave with respect to r, and it is concluded that

$$\max_{r\ge 0} g_0(r) = g_0(0) = 0.$$

Consequently

$$\begin{aligned}
&\phi(-\rho|W_E, q) \\
&\quad \stackrel{\Delta}{=} \max_{r\ge 0}\phi(-\rho|W_E, q, r) \\
&\quad = A_z^+\Delta\left[s_z^+ + q\right] - A_z^+\Delta s_z^+\left[1+\tau_z^+ q\right]^{1-\rho},
\end{aligned} \tag{A.39}$$

where we have used the notation as in Theorem 10.4. Thus, the exponent per second $E_E^+(\rho, q) \stackrel{\Delta}{=} \phi(-\rho|W_E, q)/\Delta$ is given by

$$E_E^+(\rho, q) = A_z^+\left[s_z^+ + q - s_z^+\left(1+\tau_z^+ q\right)^{1-\rho}\right], \tag{11.40}$$

which coincides with (10.28), from which equation (10.30) also follows, thus completing the another proof. □

Acknowledgments

The authors are very grateful to Vincent Tan for valuable discussions to improve the earlier manuscript.

References

[1] G. Vernam, "Secret signaling system," U.S. Patent, 1310719, 1919

[2] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no.4, pp. 656-715, 1949.

[3] C. H. Bennett and G. Brassard, "Quantum cryptography: public key distribution and coin tossing," *Proceedings of the IEEE International Conference on Computers Systems and Signal Processing,* Bangalore India, pp.175-179, 1984 ; For review, see N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography" *Rev. Mod. Phys.* Vol.74, no.1, pp.145-195, 2002

[4] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol.54, pp.1355-1387, 1975

[5] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol.24, no.3, pp.339-348, 1978

[6] A. Laourine and A. B. Wagner, "The degraded Poisson wiretap channel," *IEEE Transactions on Information Theory*, vol.IT-58, no.12, pp.7073-7085, 2012

[7] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Transactions on Information Theory*, vol.IT-57, no.6, pp. 3989-4001, 2011

[8] H. Koga and N. Sato, "On an upper bound of the secrecy capacity for a general wiretap channel," *IEEE International Symposium on Information Theory*, Adelaide, Australia, Sept. 4-9, 2005

[9] T.M. Cover and J.A. Thomas, *Elements of Information Theory,* 2nd ed., Wiley, New York, 2006

[10] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the

wiretap channel," *IEEE Transactions on Information Theory*, vol.IT-52, no.4, pp. 1562-1575, 2006

[11] M.R. Bloch and J.N.Laneman, "Secrecy from resolvability," *IEEE Transactions on Information Theory*, submitted.

[12] I. Csiszár, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol.32, no.1, pp. 40-47, 1996

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

[14] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed., Cambridge University Press, 2011

[15] M.S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day, San Francisco, 1964

[16] R. G. Gallager, *Information Theory and Reliable Communication,* Hoboken, NJ, Wiley, 1968

[17] S. Arimoto, "On the converse to the coding theorem for discrete memoryless channels," *IEEE Transactions on Information Theory*, vol.IT-19, pp. 357-359, 1973

[18] Y. Oohama, "Converse coding theorems for identification via channels," *IEEE Transactions on Information Theory*, vol.IT-59, no.2, pp. 744-759, 2013

[19] T. S. Han, *Information Spectrum Methods in Information Theory*, Springer, New York, 2003

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

[21] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Transactions on Information Theory*, vol.IT-399, no.3, pp. 752-772, 1993

[22] R.T. Rockafeller, *Convex Analysis*, Princeton University Press, Princeton, 1997

[23] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel–Part I," *IEEE Transactions on Information Theory*, vol.IT-34, no.6, pp.1449-1461, 1988

[24] A. D. Wyner, “Capacity and error exponent for the direct detection photon channel–Part II,” *IEEE Transactions on Information Theory*, vol.IT-34, no.6, pp.1462-1471, 1988

[25] A. Lapidoth, E. Telater and R. Urbanke, “On wide-band broadcast channels,” *IEEE Transactions on Information Theory*, vol.IT-49, no.12, pp. 3250-3258, 2003

[26] S.K.Lueng-Yan-Cheong and M.Hellamn, “The Gaussian wire-tap channel,” *IEEE Transactions on Information Theory*, vol.IT-24, no.4, pp.451-456, 1978

[27] Tie Liu and P. Vithwanath, “An extremal inequality motivated by multi terminal information-theoretic problems,” *IEEE Transactions on Information Theory*, vol.IT-53, no.5, pp.1839-1851, 2007