Achieving both positive secrecy rates of the users in two-way wiretap channel by individual secrecy

Chao Qi*, Bin Dai*, and Xiaohu Tang*
*Information Security and National Computing Grid Laboratory, Southwest Jiaotong University, Chengdu, China
E-mail: chaoqi@my.swjtu.edu.cn, daibin@home.swjtu.edu.cn, xhutang@home.swjtu.edu.cn

Abstract

In this paper, the individual secrecy of two-way wiretap channel is investigated, where two legitimate users’ messages are separately guaranteed secure against an external eavesdropper. For one thing, in some communication scenarios, the joint secrecy is impossible to achieve both positive secrecy rates of two users. For another, the individual secrecy satisfies the secrecy demand of many practical communication systems. Thus, firstly, an achievable secrecy rate region is derived for the general two-way wiretap channel with individual secrecy. In a deterministic channel, the region with individual secrecy is shown to be larger than that with joint secrecy. Secondly, outer bounds on the secrecy capacity region are obtained for the general two-way wiretap channel and for two classes of special two-way wiretap channels. The gap between inner and outer bounds on the secrecy capacity region is explored via the binary input two-way wiretap channels and the degraded Gaussian two-way wiretap. Most notably, the secrecy capacity regions are established for the XOR channel and the degraded Gaussian two-way wiretap channel. Furthermore, the secure sum-rate of the degraded Gaussian two-way wiretap channel under the individual secrecy constraint is demonstrated to be strictly larger than that under the joint secrecy constraint.

I. INTRODUCTION

With the wide usage of the wireless networks nowadays, the security of wireless communication has become a crucial issue. Due to the open nature of the wireless channel, the wireless links are more vulnerable to eavesdropping. However, in the dynamic wireless network, the traditional cryptography faces many challenges in handling the security problem, such as complex key distribution and management. By contrast, information theoretic secrecy guarantees secure communication against the eavesdropper even with unlimited computational power. In 1975, Wyner [1] introduced information theoretic secrecy to a noisy degraded broadcast channel and demonstrated that secure communication is possible without any shared key beforehand. Thereafter, information theoretic secrecy, a more powerful approach to wireless secure transmission, has attracted intensive attention [2]–[6].

As one of the classic multi-user channels, two-way channel models a large range of bidirectional communications, where two users exchange messages with each other through a common channel. For instance, two users talk with each other simultaneously via a full-duplex telephone networks; the power control centre (e.g. electricity company) interchanges information with the user via a smart grid network. The reliable communication of two-way channel was first studied by Shannon in [7], where inner and outer bounds on channel capacity region were presented. Later, Tekin and Yener [8] investigated the security along with reliability of the two-way channel in the presence of an external eavesdropper, which is referred to the two-way wiretap channel. Mainly, the two-way wiretap channel is explored in two secrecy criteria. One is the weak secrecy, requiring that the rate of information leakage to the eavesdropper vanishes. For the two-way wiretap channel with weak secrecy, both inner and outer bounds on the secrecy capacity were obtained. For the inner bound on the secrecy capacity, Tekin and Yener [8]–[10] and El Gamal et al. [11] respectively derived the achievable secrecy rate region for the Gaussian two-way wiretap channel and the general two-way wiretap channel. Specifically, reference [11] improves the results in [8]–[10] by a hybrid coding scheme combining the cooperative jamming and secret-key exchange mechanism. The outer bound on the secrecy capacity region of the
The degraded Gaussian two-way wiretap channel was studied in [12]. The other secrecy criterion is the strong secrecy, demanding that the information leakage to the eavesdropper, rather than the leakage rate, goes to zero. Regarding to the difficulty of studying strong secrecy, as we know, only Pierrot et al. [13] provided an achievable secrecy rate region with the strong secrecy of the general two-way wiretap channel.

So far, all the previous works on no matter weak secrecy or strong secrecy focus on the joint secrecy of two-way wiretap channel, assuring security of two legitimate users’ confidential messages together. However, if either of the legitimate users’ outputs is a degraded version of the eavesdropper’s output, achieving positive secrecy rates at both legitimate users is impossible with the joint secrecy (the details will be explained in the Lemma 1 in Section II). Such scenario is quite common, for instance the eavesdropper stays closer to the transmitter than the receiver does, as a result the legitimate receiver encounters more interferences and noises through the long distance transmission than the eavesdropper does. To achieve positive secrecy rates at both legitimate users, we introduce the individual secrecy of the two-way wiretap channel. Roughly speaking, individual secrecy requires that the rate of information leakage from each confidential message to the eavesdropper is made vanish. Comparatively, individual secrecy can be achieved by positive secrecy rates at both legitimate users. In fact, the individual secrecy constraint is also practical in other scenarios [14], [15]. For example, the secrecy criterion with the same definition is proposed in a multicast network [14], where a source node sends a set of message packets through the multicast network to the destination. The security in [14] requires that wiretapper gains no information about each packet, while still potentially obtains no meaningful information about the source. Under this secrecy constraint, the multicast capacity can be achieved [14]. Whereas, if the information leakage of all the packets goes to zero (the joint secrecy), it is impossible to achieve the multicast capacity [16]. Thus, the individual secrecy gains an advantage over the joint secrecy in [14].

Based on the analysis above, we investigate the individual secrecy of the two-way wiretap channel in this paper. Firstly, we derive an achievable secrecy rate region of the general two-way wiretap channel under the individual secrecy constraint. In order to illustrate the intuition of the result, a deterministic channel is provided to show that the achievable secrecy rate region under the individual secrecy constraint is strictly larger than that with the joint secrecy in [11]. Secondly, outer bounds on the secrecy capacity region are established for the general two-way wiretap channel and for two classes of special two-way wiretap channels. Further, the gap between the inner and outer bounds on the secrecy capacity region is explored via two cases: the binary input two-way wiretap channels and the degraded Gaussian channel. Most notably, we obtain the secrecy capacity region of the degraded Gaussian two-way wiretap channel under individual secrecy constraint. To the best of our knowledge, it is the first time to determine the secrecy capacity region for any kind of two-way wiretap channel in the literature.

The organization of this paper is as follows. In Section II we introduce the two-way wiretap channel with the individual secrecy. In Section III we present our results of the general two-way wiretap channel with the individual secrecy. A deterministic two-way wiretap channel is also given to illustrate the intuition behind the results. In Section IV we investigate binary-input two-way wiretap channels and the degraded Gaussian two-way wiretap channel with the individual secrecy. In the final section, we give the conclusions.

II. SYSTEM MODEL

Before discussing the system model, note that in this paper, we use capital letters, lower case letters and calligraphic letters to denote the random variables, sample values and alphabets, respectively. A similar convention is applied to the random vectors and their sample values. For example, $X^n$ denotes a random $n$-vector $(X_1, X_2, \cdots, X_n)$, and $x^n$ is a sample vector value in $X^n$.

In this paper, we study the two-way wiretap channel as shown in Fig. 1, where two legitimate users intend to exchange confidential messages with each other in the presence of an external eavesdropper. Particularly, we focus on the full-duplex scenario, where each of the legitimate users can send and receive messages simultaneously on the same degree of freedom.

Suppose $W_{1s}, W_{2s}$ are two message sets; $X_1, X_2$ are the finite channel input alphabets at user 1 and user 2; $Y_1, Y_2, Z$ are the channel output alphabets at user 1, user 2 and the eavesdropper, respectively. The discrete memoryless two-way wiretap channel is characterized by the transition probability distribution $p(y_1, y_2, z|x_1, x_2)$, where $x_1 \in X_1$, $x_2 \in X_2$ are the channel
inputs from user 1 and 2; \( y_1 \in \mathcal{Y}_1, y_2 \in \mathcal{Y}_2 \) and \( z \in \mathcal{Z} \) are channel outputs at user 1, user 2 and the eavesdropper. More specifically, the legitimate user \( i \) wants to transmit a confidential message \( W_i \in \mathcal{W}_i \) to the other user. The corresponding codeword \( X_i^n \in \mathcal{X}_i \) is sent at a transmission rate \( R_{is} = \frac{1}{n} H(W_{is}) \) for \( i = 1, 2 \). The channel output are \( Y_i^n \in \mathcal{Y}_i \) and \( Z^n \in \mathcal{Z} \) at at user \( i \) and the eavesdropper, respectively.

For such a two-way wiretap channel, a \( (2^{nR_{1s}}, 2^{nR_{2s}}, n) \) code consists of:

- Two independent message sets \( \mathcal{W}_{1s} = \{1, 2, \ldots, 2^{nR_{1s}}\} \), \( \mathcal{W}_{2s} = \{1, 2, \ldots, 2^{nR_{2s}}\} \).
- Two messages: \( W_{1s} \) and \( W_{2s} \) are independent and uniformly distributed over \( \mathcal{W}_{1s} \) and \( \mathcal{W}_{2s} \), respectively.
- Two encoders \( f_1 : \mathcal{W}_{1s} \rightarrow \mathcal{X}_{1s} \) and \( f_2 : \mathcal{W}_{2s} \rightarrow \mathcal{X}_{2s} \), which map each message \( w_{1s} \in \mathcal{W}_{1s} \) to a codeword \( x_{1s}^n \in \mathcal{X}_{1s} \); \( f_2 : \mathcal{W}_{2s} \rightarrow \mathcal{X}_{2s} \), which map each message \( w_{2s} \in \mathcal{W}_{2s} \) to a codeword \( x_{2s}^n \in \mathcal{X}_{2s} \).
- Two decoders \( g_1 : (\mathcal{Y}_{1s}, \mathcal{X}_{1s}) \rightarrow \mathcal{W}_{1s} \) and \( g_2 : (\mathcal{Y}_{2s}, \mathcal{X}_{2s}) \rightarrow \mathcal{W}_{2s} \), which map the received sequence \( y_{1s}^n \) and the sequence \( x_{1s}^n \) to a message \( \hat{w}_{1s} \); \( g_2 : (\mathcal{Y}_{2s}, \mathcal{X}_{2s}) \rightarrow \mathcal{W}_{2s} \), which map the received sequence \( y_{2s}^n \) and the sequence \( x_{2s}^n \) to a message \( \hat{w}_{2s} \).

For a given code, two metrics should be sufficed: reliability and security. The reliability is measured by the average error probabilities of decoding at legitimate user 1 and 2, defined as

\[
P_{e,1} = \frac{1}{2^{nR_{2s}}} \sum_{W_{2s}=1}^{2^{nR_{2s}}} Pr\{\hat{W}_{2s} \neq W_{2s}\};
\]

\[
P_{e,2} = \frac{1}{2^{nR_{1s}}} \sum_{W_{1s}=1}^{2^{nR_{1s}}} Pr\{\hat{W}_{1s} \neq W_{1s}\}.
\]

The individual security in this paper is defined by

\[
\frac{1}{n} I(W_{1s}; Z^n) \leq \tau_n, \quad \frac{1}{n} I(W_{2s}; Z^n) \leq \tau_n, \quad \lim_{n \to \infty} \tau_n = 0,
\]

**Definition 1.** The rate pair \((R_{1s}, R_{2s})\) is said to be achievable under the individual secrecy with \( R_{1s} = \frac{1}{n} H(W_{1s}) \), \( R_{2s} = \frac{1}{n} H(W_{2s}) \), if there exists a \( (2^{nR_{1s}}, 2^{nR_{2s}}, n) \) code such that

\[
P_{e,i} \leq \epsilon_n, \quad \text{for} \quad i = 1, 2
\]

\[
\frac{1}{n} I(W_{1s}; Z^n) \leq \tau_n, \quad \frac{1}{n} I(W_{2s}; Z^n) \leq \tau_n,
\]

\[
\lim_{n \to \infty} \epsilon_n = 0 \quad \text{and} \quad \lim_{n \to \infty} \tau_n = 0.
\]

Note that \( (3) \) indicates the reliability transmission constraint; \( (4) \) is the individual secrecy constraint.

**Remark 1.** For the joint weak secrecy (joint secrecy for short in this paper) in \([8], [11]\), the rate of information leakage rate of both the messages \( W_{1s} \) and \( W_{2s} \) is demanded vanishing, i.e.

\[
\frac{1}{n} I(W_{1s}, W_{2s}; Z^n) \leq \tau_n, \quad \lim_{n \to \infty} \tau_n = 0.
\]
If the coding schemes fulfill the \([3, 5]\) and the joint secrecy constraint \([6]\), then the rate pair \((R_{1s}, R_{2s})\) is said to be achievable under the joint secrecy constraint with \(R_{1s} = \frac{1}{n} H(W_{1s}), \ R_{2s} = \frac{1}{n} H(W_{2s})\).

However, the joint secrecy is not always affordable, such as the following lemma.

Lemma 1. Assume that in the two-way wiretap channels, the legitimate received symbol \(Y_1\) or \(Y_2\) is a degraded version of the received symbol \(Z\) at eavesdropper, i.e. \((X_1, X_2) \rightarrow Z \rightarrow Y_1\) or \((X_1, X_2) \rightarrow Z \rightarrow Y_2\) forms a Markov chain. Then, with the joint secrecy, the achievable secure transmission rates pair \((R_{1s}, R_{2s})\) with \(R_{1s} > 0, \ R_{2s} > 0\) is not available, while with the individual secrecy constraint, it is available with the \(R_{1s} > 0, \ R_{2s} > 0\).

Proof 1. The information leakage of two messages \(W_{1s}\) and \(W_{2s}\) are

\[
H(W_{1s}, W_{2s}|Z^n) = H(W_{1s}|Z^n) + H(W_{2s}|W_{1s}, Z^n)
\]

\[
\leq H(W_{1s}|Z^n) + H(W_{2s}|W_{1s}, Y^n_1)
\]

\[
\leq H(W_{1s}|Z^n) + n\epsilon_n
\]

\[
= nR_{1s} + n\epsilon_n
\]

(7)

where \((a)\) follows from the degraded assumption that \(Y_1\) is a degraded version of \(Z\); \((b)\) follows from the reliability transmission condition and Fano’s inequality with \(\lim_{n \to \infty} \epsilon_n = 0\).

- For the joint secrecy constraint

\[
\frac{1}{n} I(W_{1s}, W_{2s}; Z^n) \leq \tau_n, \quad \lim_{n \to \infty} \tau_n = 0,
\]

we can obtain that

\[
n(R_{1s} + R_{2s}) = H(W_{1s}, W_{2s})
\]

\[
= H(W_{1s}, W_{2s}|Z^n) + I(W_{1s}, W_{2s}; Z^n)
\]

\[
\leq H(W_{1s}, W_{2s}|Z^n) + n\tau_n
\]

\[
\leq nR_{1s} + n\epsilon_n + n\tau_n
\]

where \((d)\) follows from the joint secrecy constraint; \((d)\) follows from \((7)\). That is

\[
nR_{2s} \leq n\epsilon_n + n\tau_n
\]

(8)

As \(n\) goes to infinity (i.e. \(n \to \infty\)), according to \([5]\), we have \(\lim_{n \to \infty} \epsilon_n = 0, \ \lim_{n \to \infty} \tau_n = 0\). Therefore, by \((8)\) we have

\[
R_{2s} \leq 0.
\]

(9)

Similarly, if \(Y_2\) is a degraded version of \(Z\), then \(R_{1s} \leq 0\).

- For the individual secrecy constraint

\[
\frac{1}{n} I(W_{1s}; Z^n) \leq \tau_n, \quad \frac{1}{n} I(W_{2s}; Z^n) \leq \tau_n, \quad \lim_{n \to \infty} \tau_n = 0.
\]

we have

\[
n(R_{1s} + R_{2s}) = H(W_{1s}, W_{2s})
\]

\[
= H(W_{1s}, W_{2s}|Z^n) + I(W_{1s}, W_{2s}; Z^n)
\]

\[
\leq nR_{1s} + n\epsilon_n + I(W_{1s}; Z^n) + I(W_{2s}; Z^n|W_{1s})
\]

\[
\leq nR_{1s} + n\epsilon_n + n\tau_n + I(W_{2s}; Z^n|W_{1s})
\]

(10)
For the two-way wiretap channels with an external eavesdropper, an achievable secrecy rate region is given by Theorem 1.

As \( n \to \infty \), according to [5], that is

\[
R_{2s} \leq I(W_{2s}; Z^n|W_{1s}).
\]  

(11)

Similarly, \( R_{1s} \leq I(W_{1s}; Z^n|W_{2s}) \). Therefore, the individual secrecy can achieve both positive transmission rates pair \((R_{1s}, R_{2s})\). Moreover, if \( Y_1 \) is a degraded version of \( Z \), we have \( I(W_{2s}; Z^n|W_{1s}) \geq I(W_{2s}; Y^n_1|W_{1s}) \). This illustrates that \( R_{2s} \) could even achieve \( I(W_{2s}; Y^n_1|W_{1s}) \).

In summary, if \( Y_1 \) or \( Y_2 \) is a degraded version of \( Z \), the achievable secure transmission rates pair \((R_{1s} > 0, R_{2s} > 0)\) is not available under the joint secrecy constraint, while it is possible under the individual secrecy constraint. In the next section, we will give the exact achievable secrecy rate region of two-way wiretap channel with individual secrecy.

III. INDIVIDUAL SECRECY OF TWO-WAY WIERTAP CHANNEL

In this section, we present our main results of two-way wiretap channel with individual secrecy. Firstly, we derive an achievable secrecy rate region of the general two-way wiretap channel, further giving an intuitive interpretation of the result in a deterministic two-way wiretap channel. Secondly, we give an outer bound on the secrecy capacity.

A. An achievable secrecy rate region

**Theorem 1.** For the two-way wiretap channels with an external eavesdropper, an achievable secrecy rate region is given by

\[
\mathcal{R}^{Ind-I^n} \triangleq \text{convex closure of} \left\{ \bigcup_{p \in \mathcal{P}} \mathcal{R}^{Ind-I^n}(p) \right\}
\]

where \( \mathcal{P} \) denotes the set of all distribution of the random variables \( U_1, U_2, X_1, X_2 \) satisfying \( p(u_1u_2x_1x_2) = p(u_1)p(u_2)p(x_1|u_1)p(x_2|u_2) \);

\( \mathcal{R}^{Ind-I^n}(p) \) is the region of rate pairs \((R_{1s}, R_{2s})\) for \( p \in \mathcal{P} \), satisfying

\[
\begin{align*}
(R_{1s}, R_{2s}) : \\
R_{1s} \geq 0, R_{2s} \geq 0, \\
R_{1s} \leq I(U_1; Y_2|X_2) - I(U_1; Z) - I(U_2; Z|U_1) - I(U_2; Y_1|X_1)^+, \\
R_{2s} \leq I(U_2; Y_1|X_1) - I(U_2; Z) - I(U_1; Z|U_2) - I(U_1; Y_2|X_2)^+.
\end{align*}
\]

(12)

and \(|a|^+ = \max\{0, a\}\), \(|U_1| \leq |X_1| + 1\), \(|U_2| \leq |X_2| + 1\).

**Proof:** See the proof in Appendix [3].

Our achievable region is obtained by the stochastic encoding and channel prefixing, where the codeword \( U_1 \) and \( U_2 \) are drawn from two binning codebooks respectively, and then passed on to two virtual prefix channel respectively. Accordingly, the channel input \( X_1 \) and \( X_2 \) are generated according to \( p(x_1|u_1) \) and \( p(x_2|u_2) \), respectively. Indeed, the channel prefixing is an interpretation of cooperative jamming [17], which is a collaborative approach to improving the secrecy rate in a multi-user communication system.

Specially, since \( Z \) is related to \( X_1 \) and \( X_2 \) together, if the eavesdropper can decode part of message of user 2, it may help the eavesdropper to decode the confidential message \( W_{1s} \). Hence, when analyzing the individual secrecy of \( W_{1s} \), if \( R_2 \geq I(U_2; Z|U_1) \) then the codebook of user 2 is equally partitioned into \( 2^{R_{2s}} \) sub-codebooks with \( R_{2s} = R_2 - I(U_2; Z|U_1)_+ \), each part consisting of \( 2^{R_{2s}} \) codewords with \( R_{22} = I(U_2; Z|U_1) - \epsilon \). The secrecy analysis of \( W_{2s} \) works in a similar manner.

**Remark 2.** It is worth noting that \( R_{1s} \) and \( R_{2s} \) meet the conditions individually in (12). This phenomenon can be interpreted by considering the reliability and the individual secrecy of the system. Firstly, for the reliability, the rate pair \((R_{1s}, R_{2s})\) should satisfy the achievable rate region given by Shannon in [7], where \( R_{1s} \) and \( R_{2s} \) meet each condition separately with no trade-off between \( R_{1s} \) and \( R_{2s} \). Secondly, for the individual secrecy, \( R_{1s} \) and \( R_{2s} \) should meet \( \frac{1}{n} I(W_{1s}; Z) \leq \tau_n \) and \( \frac{1}{n} I(W_{2s}; Z) \leq \tau_n \), respectively. Therefore, as shown in Theorem [7] \( R_{1s} \) and \( R_{2s} \) are not directly interrelated with each other.
Unlike the individual secrecy, the results in \cite{our previous work} revealed a trade-off between $R_{1s}$ and $R_{2s}$ with the joint secrecy. This is because the joint secrecy constraint $\frac{1}{n}I(W_{1s}, W_{2s}; Z) \leq \tau_n$ is related to both $R_{1s}$ and $R_{2s}$ at the same time.

Applying Theorem 1 to a two-way wiretap channel where the eavesdropper receives as many messages as the legitimate users, we have the following corollary.

**Corollary 1.** Suppose that in the two-way wiretap channel the legitimate receivers and the eavesdropper receive the same amount of messages, i.e. $Y_1 = Y_2 = Z$, an achievable secrecy rate region with individual secrecy is given by

\[
R_1^{\text{Ind} - I_n} \triangleq \text{convex closure of } \{ \bigcup_{p \in \mathcal{P}} R_1^{\text{Ind} - I_n}(p) \},
\]

where $\mathcal{P}$ denotes the set of all distribution of the random variables $X_1, X_2$ satisfying $p(x_1, x_2) = p(x_1)p(x_2)$. $R_1^{\text{Ind} - I_n}(p)$ is the region of rate pairs $(R_{1s}, R_{2s})$ for $p \in \mathcal{P}$, satisfying

\[
\begin{align*}
(R_{1s}, R_{2s}) : \\
R_{1s} &\geq 0, R_{2s} \geq 0, \\
R_{1s} &\leq I(X_1; Y_2|X_2) - I(X_1; Z), \\
R_{2s} &\leq I(X_2; Y_1|X_1) - I(X_2; Z).
\end{align*}
\]

**Proof.** Setting $U_1 = X_1$ and $U_2 = X_2$ in \cite{previous work}, then

\[
R_{1s} \leq I(X_1; Y_2|X_2) - I(X_1; Z) - |I(X_2; Z|X_1) - I(X_2; Z|X_1)|^+
\]

\[
= I(X_1; Y_2|X_2) - I(X_1; Z)
\]

Similarly, $R_{2s} \leq I(X_2; Y_1|X_1) - I(X_2; Z)$. This completes the proof. \hfill \Box

**Remark 3.** With the joint secrecy, if $Y_1 = Y_2 = Z$ and $U_1 = X_1, U_2 = X_2$, the achievable secrecy rate region \cite{our previous work} is

\[
R_1^{J - I_n} \triangleq \text{convex closure of } \{ \bigcup_{p \in \mathcal{P}} R_1^{J - I_n}(p) \},
\]

where $\mathcal{P}$ denotes the set of all distribution of the random variables $X_1, X_2$ satisfying $p(x_1, x_2) = p(x_1)p(x_2)$. $R_1^{J - I_n}(p)$ is the region of rate pairs $(R_{1s}, R_{2s})$ for $p \in \mathcal{P}$, satisfying

\[
\begin{align*}
(R_{1s}, R_{2s}) : \\
R_{1s} &\geq 0, R_{2s} \geq 0, \\
R_{1s} &\leq I(X_2; Y_1|X_1), \\
R_{2s} &\leq I(X_1; Y_2|X_2), \\
R_{1s} + R_{2s} &\leq I(X_2; Y_1|X_1) + I(X_1; Y_2|X_2) - I(X_1, X_2; Z).
\end{align*}
\]

Note that, the last equation can be rewritten as

\[
R_{1s} + R_{2s} \leq I(X_2; Y_1|X_1) + I(X_1; Y_2|X_2) - I(X_1, X_2; Z)
\]

\[
= I(X_2; Y_1|X_1) - I(X_2; Z)
\]

or

\[
R_{1s} + R_{2s} \leq R_{1s} + R_{2s} \leq I(X_1; Z|X_2) - I(X_1; Z)
\]

By comparing \cite{corollary 1} with \cite{remark 3} and \cite{equation}, either $R_{1s}$ or $R_{2s}$ with individual secrecy is equal to the sum-rate $R_{1s} + R_{2s}$ with the joint secrecy, which indicates that the secrecy rate region $R_1^{J - I_n}$ with the joint secrecy is only half of the secrecy rate region $R_1^{\text{Ind} - I_n}$ with the individual secrecy.
B. An interpretation of Theorem [1] in a deterministic two-way wiretap channel

In this subsection, a deterministic two-way wiretap channel is studied to illustrate the intuition of the achievable secrecy rate region in Theorem [1]. Suppose that the deterministic two-way wiretap channel is described by

\[ Y_1 = X_1 \oplus X_2 \oplus N_1; \]
\[ Y_2 = X_1 \oplus X_2 \oplus N_2; \]
\[ Z = X_1 \oplus X_2 \oplus N_e; \]

where \( X_1, X_2, N_1, N_2, N_e \in \{0, 1\}; \) \( X_1 \) and \( X_2 \) are the binary channel inputs at user 1 and user 2, respectively; \( Y_1, Y_2 \) and \( Z \) are the binary channel outputs at the user 1, user 2 and the eavesdropper, respectively; \( N_1, N_2, N_e \) are the additive binary noise impairing user 1, user 2 and the eavesdropper, respectively. Then, the corresponding transition probabilities are given by

\[ p(N_1 = 1) = \varepsilon_1; \]
\[ p(N_2 = 1) = \varepsilon_2; \]
\[ p(N_e = 1) = \varepsilon_z. \]

Therefore, the transmission probabilities are

\[ p(y_1 \neq x_2|x_1) = \varepsilon_1; \]
\[ p(y_2 \neq x_1|x_2) = \varepsilon_2; \]
\[ p(z \neq x_1 \oplus x_2|x_1, x_2) = \varepsilon_z. \]

If the system does not have any eavesdropper, the model is reduced to a binary modulo-2 two-way channel that provides the reliability only.

**Lemma 2 ([7]).** For the full-duplex binary modulo-2 two-way channel, the achievable reliable transmit rate region \( \mathcal{R} \) is the union of non-negative rate pairs \((R_1, R_2)\) defined by

\[ R_1 \leq 1 - h(\varepsilon_2), \]
\[ R_2 \leq 1 - h(\varepsilon_1). \]

For the joint secrecy of the binary modulo-2 two-way wiretap channel, we have the following lemma.

**Lemma 3 ([11]).** For the full-duplex binary modulo-2 two-way wiretap channel with the joint secrecy, the achievable secrecy rate region \( \mathcal{R}^\text{Joint−In}_s \) is the union of non-negative rate pairs \((R_{1s}, R_{2s})\) defined by

\[ R_{1s} \leq 1 - h(\varepsilon_2), \]
\[ R_{2s} \leq 1 - h(\varepsilon_1), \]
\[ R_{1s} + R_{2s} \leq 1 + h(\varepsilon_z) - h(\varepsilon_1) - h(\varepsilon_2). \]

(17)

On the other hand, according to Theorem 1, the achievable secrecy rate region with the individual secrecy is given in the following corollary.

**Lemma 4.** For the full-duplex binary modulo-2 two-way wiretap channel with the individual secrecy, the achievable secrecy rate region \( \mathcal{R}^\text{Ind−In}_s \) is the union of non-negative rate pairs \((R_{1s}, R_{2s})\) defined by

\[ R_{1s} \leq 1 - h(\varepsilon_2), \]
\[ R_{2s} \leq 1 - h(\varepsilon_1). \]

**Proof:** See the proof of Corollary 2, 3, and 4 in Appendix A.

It is clear from Lemma 2, 3, and 4 that the achievable rate region \( \mathcal{R} \) is the same as \( \mathcal{R}^\text{Ind−In}_s \), while \( \mathcal{R}^\text{Joint−In}_s \) is smaller than \( \mathcal{R}^\text{Ind−In}_s \) with regard to the sum-rate constraint (17). We conclude the result in the following theorem.

**Theorem 2.** For the deterministic modulo-2 two-way wiretap channel, the achievable reliable transmit rate region \( \mathcal{R} \), the achievable secrecy rate region \( \mathcal{R}^\text{Joint−In}_s \) with the joint secrecy, and \( \mathcal{R}^\text{Ind−In}_s \) with the individual secrecy satisfy

\[ \mathcal{R}^\text{Joint−In}_s \subseteq \mathcal{R}^\text{Ind−In}_s = \mathcal{R}. \]
From the practical viewpoint, Theorem 2 reveals a great advantage of the individual secrecy that it can be achieved without any rate loss of the reliable transmission rate, yet there is a rate loss for the joint secrecy.

The geometric structures of $R_s$, $R_s^{Joint-In}$, and $R_s^{Ind-In}$ of binary modulo-2 two-way channel are depicted by four cases regarding to the value of $h(\varepsilon_1), h(\varepsilon_2)$ and $h(\varepsilon_z)$ in Fig. 2 where the boundary of $R_s$, $R_s^{Joint-In}$ and $R_s^{Ind-In}$ are plotted by the solid line, the dashed-dotted line and the dashed line, respectively. In Fig. 2, $R_s$ coincides with $R_s^{Ind-In}$ as a rectangle, and $R_s^{Joint-In}$ contains a missing corner due to the constraint $R_{1s} + R_{2s} \leq 1 + h(\varepsilon_z) - h(\varepsilon_1) - h(\varepsilon_2)$ in Lemma 3. Clearly, the individual secrecy provides a strictly larger secrecy rates region than the joint secrecy does, especially in the high rates region of $R_{1s}$ and $R_{2s}$.

C. An outer bound on the secrecy capacity of two-way wiretap channels with individual secrecy

Theorem 3. For the general two-way wiretapper channel, with the individual secrecy, an outer bound on the secrecy capacity is given by

$$R_s^{Ind-O} \triangleq \text{Conv}\{ \bigcup_{p \in P} R_s^{Ind-O}(p) \},$$

where $P$ denotes the set of all distribution of the random variables $U, V, X_1, X_2$ satisfying $p(uv_1v_2x_1x_2) = p(u)p(v_1|u)p(v_2|u)p(x_1x_2|uv_1v_2)$; $U, V_1$ and $V_2$ are the auxiliary random variables and $U \rightarrow (V_1, V_2) \rightarrow (X_1, X_2) \rightarrow (Y_1, Y_2, Z)$ forms a Markov
The outer bound (18) works for the general two-way wiretap channel with individual secrecy. Further, an outer bound derived for two classes of two-way channels in the following theorem.

**Theorem 4.** For the following two classes of two-way wiretapper channels,

1) the legitimate users and the eavesdropper receive the same amount of messages, i.e. \( Y_1 = Y_2 = Z = Y \);

2) the received message \( Z \) at the eavesdropper is a degraded version of both the messages at legitimate users, satisfying the Markov chain \( Y_1 \rightarrow Z \) and \( Y_2 \rightarrow Z \);

an outer bound on the secrecy capacity is given by

\[
\mathcal{R}_1^{\text{ind-O}} \triangleq \bigcup_{p \in \mathcal{P}} \mathcal{R}_1^{\text{ind-O}}(p),
\]

where \( \mathcal{P} \) denotes the set of all distribution of the random variables \( Q, X_1, X_2 \) with \( p(q|x_1 x_2) \), and \( \mathcal{R}_1^{\text{ind-O}}(p) \) is the region of rate pairs \( (R_{1s}, R_{2s}) \) for \( p \in \mathcal{P} \), satisfying

\[
\begin{align*}
(R_{1s}, R_{2s}) : & \\
R_{1s} & \geq 0, R_{2s} \geq 0, \\
R_{1s} & \leq I(X_1; Y_2|X_2, Q) - I(X_1; Z|Q), \\
R_{2s} & \leq I(X_2; Y_1|X_1, Q) - I(X_2; Z|Q).
\end{align*}
\]

the cardinality of the auxiliary random variables \( Q \) satisfies \( |Q| \leq |X_1||X_2| + 1. \)

**Proof:** See the proof in Appendix [D].

**Remark 4.** Later, Corollary [C] and Theorem [D] will be applied into the binary input two-way wiretap channels to show the gap between the inner and the outer bound on the secrecy capacity. Specially, for a degraded Gaussian channel we will proof that the inner bound in Theorem [C] and the outer bound in Theorem [D] coincide with each other, such that the secrecy capacity is fully established.

IV. **Binary input two-way wiretap channel and Degraded Gaussian two-way wiretap channel**

A. **Binary input two-way wiretap channel with individual secrecy**

In this subsection, we are interested in the binary input two-way wiretap channels when the legitimate users and the eavesdropper have the same channel output, i.e. \( Y_1 = Y_2 = Z \). As Shannon utilized the binary multiplying channel (BMC) to indicate the gap between the inner bound and the outer bound on the channel capacity of two-way channel, we also explore our main results in BMC to show the gap between the inner and the outer bound on the secrecy capacity. Considering the binary-input (i.e., \( x_1, x_2 \in \{0, 1\} \)) and binary-output (i.e., \( y_1 = y_2 = z \in \{0, 1\} \)) or ternary outputs (i.e., \( y_1 = y_2 = z \in \{0, 1, 2\} \)) transmission, the XOR channel and the Adder channel are also investigated.
For each channel, the achievable secrecy rate region $R_{Ind-I}^{1}$ and the outer bound on the secrecy capacity $R_{Ind-O}^{1}$ are derived from Corollary 1 and Theorem 4, respectively. For simplicity, we define the following operation

$$a \ast b := a(1-b) + (1-a)b,$$

for $0 \leq a, b \leq 1$

and the entropy function

$$h(a) := \begin{cases} -a \log a - (1-a) \log(1-a), & \text{if } 0 < a < 1 \\ 0, & \text{if } a = 0 \text{ or } 1. \end{cases}$$

Fig. 3: Transition diagrams of the binary-input two-way channels.

1) Binary Multiplying channel: The BMC is shown in Fig. 3(a), where the channel output is represented by $Y_1 = Y_2 = Z = X_1 \cdot X_2$. By Corollary 1, the achievable secrecy rate region $R_{BMC}^{Ind-I}$ for BMC with individual secrecy is the union of the following non-negative rate pair $(R_{1s}, R_{2s})$ over $X_1 \sim \text{Bern}(p_1), X_2 \sim \text{Bern}(p_2)$:

$$R_{1s} \leq p_2 h(p_1) + p_1 h(p_2) - h(p_1p_2),$$

$$R_{2s} \leq p_2 h(p_1) + p_1 h(p_2) - h(p_1p_2).$$

Fig. 4: Secrecy rate region of BMC channel.

The achievable secrecy rate region $R_{BMC}^{Ind-I}$ and the outer bound $R_{BMC}^{Ind-O}$ are shown in Fig. 4. Additionally, the achievable secrecy rate region with the joint secrecy $R_{BMC}^{J-I}$ is also plotted for comparison. The numerical results in Fig. 4 demonstrate that the region $R_{BMC}^{Ind-I}$ is twice as large as $R_{BMC}^{J-I}$, consistent with Remark 3. Moreover, it can be seen that the increase in $R_{2s}$ leads to the decrease in $R_{1s}$ on $R_{BMC}^{J-I}$, while $R_{1s}$ and $R_{2s}$ are greatly improved and achieve high secrecy rate simultaneously on $R_{BMC}^{Ind-I}$. However, the gap between $R_{BMC}^{Ind-I}$ and $R_{BMC}^{Ind-O}$ is still large.

2) Binary XOR channel: The XOR channel is shown in Fig. 3(b), where the channel output is represented by $Y_1 = Y_2 = Z = X_1 \oplus X_2$. By Corollary 1, the achievable secrecy rate region $R_{XOR}^{Ind-I}$ for XOR with individual secrecy is the union of the following non-negative rate pair $(R_{1s}, R_{2s})$ over $X_1 \sim \text{Bern}(p_1), X_2 \sim \text{Bern}(p_2)$:

$$R_{1s} \leq h(p_1) + h(p_2) - h(p_1 \ast p_2),$$

$$R_{2s} \leq h(p_1) + h(p_2) - h(p_1 \ast p_2).$$
Correspondingly, the achievable secrecy rate region with individual secrecy $R_{\text{Ind}}^{I} - R_{\text{In}}^{XOR}$, with the joint secrecy $R_{\text{Ind}}^{J} - R_{\text{In}}^{XOR}$ and the outer bound $R_{\text{XOR}}^{I}$ with individual secrecy are shown in Fig. 5. Clearly, $R_{\text{Ind}}^{J} - R_{\text{In}}^{XOR}$ is only half the size of $R_{\text{XOR}}^{I}$. Especially, the maximum achievable secrecy rate on $R_{\text{Ind}}^{I} - R_{\text{In}}^{XOR}$ is $(R_{1s}, R_{2s}) = (1, 1)$, which is also the maximum reliable rate without secrecy [7]. It indicates that the individual secrecy can be achieved with no rate loss of reliable transmission. Moreover, $R_{\text{Ind}}^{I} - R_{\text{In}}^{XOR}$ coincides with $R_{\text{XOR}}^{I}$, hence the individual secrecy capacity region of XOR channel is fully characterized with $(R_{1s} \leq 1, R_{2s} \leq 1)$.

3) Adder channel: The XOR channel is shown in Fig. 3 (c), where the channel output is represented by $Y_1 = Y_2 = Z = X_1 + X_2$. By Corollary 1, the achievable secrecy rate region $R_{\text{Ind}}^{I} - R_{\text{Ind}}^{XOR}$ for the binary Adder channel with individual secrecy is the union of the following non-negative rate pair $(R_{1s}, R_{2s})$ over $X_1 \sim \text{Bern}(p_1), X_2 \sim \text{Bern}(p_2)$:

$$R_{1s} \leq (p_1 \ast p_2) h \left( \frac{p_1(1 - p_2)}{p_1 \ast p_2} \right),$$

$$R_{2s} \leq (p_1 \ast p_2) h \left( \frac{p_2(1 - p_1)}{p_1 \ast p_2} \right).$$
Correspondingly, with individual secrecy, the achievable secrecy rate region $R_{Adder}^{Ind-In}$ and the outer bound $R_{Adder}^{Ind-O}$ are drawn in Fig. 8, where the achievable secrecy rate region with the joint secrecy $R_{Adder}^{J-In}$ is also plotted for comparison. As in BMC and XOR channel, in Adder channel $R_{Adder}^{Ind-In}$ is twice as large as $R_{Adder}^{J-In}$. Moreover, for the individual secrecy, the gap between $R_{Adder}^{Ind-I}$ and $R_{Adder}^{Ind-O}$ has narrowed considerably than that of the BMC.

**B. Degraded Gaussian two-way wiretap channel with individual secrecy**

In this subsection, we study a class of degraded Gaussian two-way wiretap channels with individual secrecy. We first define two classes of degraded channels.

In the two-way wiretap channel, suppose the channel inputs of the two users are $x_1$ and $x_2$, respectively; the channel output at the users and the eavesdropper are $y_1$, $y_2$ and $z$, respectively.

**Definition 2.** The two-way wiretap channel is physically degraded if the transition probability distribution satisfies

$$p(z, y_1 | x_1, x_2) = p(y_1 | x_1, x_2) p(z | y_1).$$

**Definition 3.** The two-way wiretap channel is physically degraded if the conditional marginal distribution is the same as that of a physically degraded two-way wiretap channel, i.e., there exists a distribution $p(z | y_1)$ such that

$$p(z | x_1, x_2) = \sum_{y_1} p(y_1 | x_1, x_2) p(z | y_1).$$

Assume that the channel is discrete and memoryless, and the channel outputs at the legitimate receivers and the eavesdropper are corrupted by additive Gaussian noise terms. Then, the channel outputs at each time $i$ are given by

$$Y_{1i} = X_{1i} + X_{2i} + Z_{1i};$$
$$Y_{2i} = X_{1i} + X_{2i} + Z_{2i};$$
$$Z_i = X_{1i} + X_{2i} + Z_{ei};$$

where $Z_{1i}$, $Z_{2i}$ and $Z_{ei}$ are independent zero-meaning additive Gaussian noises with $Z_{1i} \sim \mathcal{N}(0, N_1)$, $Z_{2i} \sim \mathcal{N}(0, N_2)$, $Z_{ei} \sim \mathcal{N}(0, N_e)$, and $N_e > N_1$, $N_e > N_2$. The average power constraints of the channel input sequences $X_1^n$ and $X_2^n$ are

$$\frac{1}{n} \sum_{i=1}^{n} E[X_{1i}^2] \leq P_1 \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^{n} E[X_{2i}^2] \leq P_2.$$ (24)

Under the assumption $N_e > N_1$, the output $Z$ is a stochastically degraded version of $Y_1$, since the marginal distribution $p(z | x_1, x_2)$ is the same as that of the following physical degraded Gaussian two-way channel:

$$Y_{1i} = X_{1i} + X_{2i} + Z_{1i};$$
$$Z_i = X_{1i} + X_{2i} + Z_{1i} + Z_{ei};$$

where $Z_{1i}$ is independent of $Z_{1i}$, being zero-meaning Gaussian noises with variance $Z_{ei} - Z_{1i}$. Similarly, under the assumption $N_e > N_2$, the output $Z$ is a stochastically degraded version of $Y_2$.

We have the fundamental limits of the degraded Gaussian two-way wiretap channel with individual secrecy in the following theorem.

**Theorem 5.** For the degraded Gaussian two-way wiretap channel with the individual secrecy, if the received symbol $Z$ at eavesdropper is a stochastically degraded version of the received symbol $Y_1$ and $Y_2$ at the legitimate users, the secrecy capacity is given by

$$R_{GTW}^{Ind-c} = \begin{cases} (R_{1s}, R_{2s}) : \\ R_{1s} \geq 0, R_{2s} \geq 0, \\ R_{1s} \leq \frac{1}{2} \log \left(\frac{P_1 + N_2}{N_1 P_1 + P_2 + N_e}\right), \\ R_{2s} \leq \frac{1}{2} \log \left(\frac{P_2 + N_1}{N_2 P_2 + P_1 + N_e}\right). \end{cases}$$ (27)
Proof: See Appendix E

Fig. 7: the secrecy capacity of the degraded Gaussian two-way wiretap channel with the individual secrecy and the achievable secrecy rate region with the joint secrecy, with $N_1 = 2$, $N_2 = 2$, $N_e = 3$, $P_1 = 300$, $P_2 = 300$.

In Fig. 7 we plot $\mathcal{R}^{\text{Ind-c}}_{\text{GTW}}$ and the achievable secrecy rate region $\mathcal{R}^{\text{Joint-}}_{\text{GTW}}$ with the joint secrecy for $N_1 = 2$, $N_2 = 2$, $N_e = 3$, $P_1 = 300$, $P_2 = 300$. Firstly, on $\mathcal{R}^{\text{Ind-c}}_{\text{GTW}}$, $R_{1s}$ and $R_{2s}$ achieve high rate region simultaneously, i.e. $R_{1s} = 3.1228$, $R_{1s} = 3.1228$. However, $R_{1s}$ and $R_{2s}$ can not achieve such high rate simultaneously as on $\mathcal{R}^{\text{Joint-}}_{\text{GTW}}$, where if $R_{1s}$ is as high as $R_{1s} = 3.1228$, then the $R_{2s}$ is only $R_{2s} = 0.2901$. Secondly, in lower secrecy rate region $R_{2s} < 0.2901$, the joint secrecy offers higher secrecy rate $R_{1s}$ than the individual secrecy does. This is because that two confidential messages are guaranteed secure under the joint secrecy constraint $\frac{1}{n}I(W_1, W_2; Z^n) \leq \tau_n$ which is referred to $R_{1s}$ and $R_{1s}$ together. By sacrificing the rate $R_{2s}$, the secrecy rate $R_{1s}$ can be improved, even achieving the $(R_{1s})_{\text{max}} = 3.4129$, $R_{2s} = 0$. Nevertheless, the individual secrecy provides higher secrecy sum-rate $R_{1s} + R_{2s}$ than the joint secrecy does. From Fig. 7 it can be seen $(R_{1s} + R_{2s})_{\text{Ind}} = 6.2456$ for the individual secrecy and $(R_{1s} + R_{2s})_{\text{Joint}} = 3.4129$ for the joint secrecy, which indicates that the sum-rate $R_{1s} + R_{2s}$ with individual secrecy is much larger than that with joint secrecy. Actually, the maximum sum-rate $R_{1s} + R_{2s}$ on $\mathcal{R}^{\text{Joint-}}_{\text{GTW}}$ is

$$
(R_{1s} + R_{2s})_{\text{Joint}} \leq \frac{1}{2} \log \frac{(P_1 + N_2)(P_2 + N_1)N_e}{N_2N_1(P_1 + P_2 + N_e)}.
$$

By (27), the sum-rate $R_{1s} + R_{2s}$ with the individual secrecy satisfies

$$
(R_{1s} + R_{2s})_{\text{Ind}} \leq \frac{1}{2} \log \frac{(P_1 + N_2)(P_2 + N_1)(P_1 + N_e)(P_2 + N_e)}{N_1N_2(P_1 + P_2 + N_e)(P_1 + P_2 + N_e)}.
$$

Such that

$$
(R_{1s} + R_{2s})_{\text{Joint}} - (R_{1s} + R_{2s})_{\text{Ind}} = \frac{1}{2} \log \frac{N_e(P_1 + P_2 + N_e)}{(P_1 + N_e)(P_2 + N_e)}
= \frac{1}{2} \log \frac{1}{1 + \frac{P_1}{N_e}} \times (1 + \frac{P_1}{P_2 + N_e})
< 0
$$

Theoretically, the sum-rate $R_{1s} + R_{2s}$ with the individual secrecy is strictly larger than that with the joint secrecy, hence it is consistent with the numerical results in Fig. 7.

V. CONCLUSION

In this paper, we investigated the fundamental limits of two-way wiretap channel with individual secrecy. Firstly, by channel prefixing approach and stochastic encoding, we derived an achievable secrecy rate region for the general two-way wiretap
channel. Secondly, we obtained outer bounds on the secrecy capacity region for the general two-way wiretap channel and for two classes of special two-way wiretap channels. The result showed that the individual secrecy creates an advantage over the joint secrecy for the achievable secrecy rate region in a binary modulo-2 two-way channel, where the region with individual secrecy was shown to be twice as large as that with joint secrecy. Particularly, the inner and the outer bound coincide with each other in XOR channel and degraded Gaussian two-way wiretap channel, hence the secrecy capacity regions were established. In addition, in the degraded Gaussian two-way wiretap channel, the individual secrecy gains larger secure sum-rate than the joint secrecy does.

**APPENDIX A**

**Proof of Lemma 2, 3, and 4**

In order to give the achievable reliable transmission rate region \( R \) and the secrecy rate region \( R_s^{Joint-In} \) and \( R_s^{Indi-In} \), we first calculate the following terms.

1. \( I(X_1; Y_2|X_2) \) and \( I(X_2; Y_1|X_1) \):
   
   \[
   I(X_1; Y_2|X_2) = H(Y_2|X_2) - H(Y_2|X_1, X_2) \\
   \leq 1 - \sum_{x_2} p(x_2) H(Y_2|X_1, x_2) \\
   = 1 - h(y_2 \neq x_1|x_2) \\
   = 1 - h(\epsilon_2)
   \]

   Similarly, we have

   \[
   I(X_2; Y_1|X_1) \leq 1 - h(\epsilon_1);
   \]

2. \( I(X_1; Y_2|X_2) + I(X_2; Y_1|X_1) - I(X_1, X_2; Z) \):
   
   \[
   I(X_1; Y_2|X_2) + I(X_2; Y_1|X_1) - I(X_1, X_2; Z) \\
   = H(Y_2|X_2) + H(Y_1|X_1) - H(Z) - H(Y_2|X_1X_2) - H(Y_1|X_1X_2) + H(Z|X_1X_2)
   \]

   By noting that,

   \[
   H(Y_2|X_2) + H(Y_1|X_1) - H(Z) \\
   = H(X_1 \oplus N_2) + H(X_2 \oplus N_1) - H(X_1 \oplus X_2 \oplus N) \\
   \overset{(a)}{=} H(X_1 \oplus N_2) + H(X_2 \oplus N_1) - H(X_1 \oplus N_2 \oplus X_2 \oplus N_1 \oplus \hat{N}_e) \\
   \overset{(b)}{\leq} H(X_1 \oplus N_2) + H(X_2 \oplus N_1) - H(X_1 \oplus N_2 \oplus X_2 \oplus N_1) \\
   = H(X_1 \oplus N_2) + H(X_1 \oplus N_2 \oplus X_2 \oplus N_1|X_1 \oplus N_2) - H(X_1 \oplus N_2 \oplus X_2 \oplus N_1) \\
   = H(X_1 \oplus N_2) - I(X_1 \oplus N_2 \oplus X_2 \oplus N_1|X_1 \oplus N_2) \\
   = H(X_1 \oplus N_2|X_1 \oplus N_2 \oplus X_2 \oplus N_1) \\
   \leq 1
   \]

   where \( (a) \) follows by setting \( \hat{N}_e = N_1 \oplus N_2 \oplus N_e \); \( (b) \) follows from the fact that conditioning does not increase entropy.

   Hence, we conclude that

   \[
   I(X_1; Y_2|X_2) + I(X_2; Y_1|X_1) - I(X_1, X_2; Z) \\
   \overset{(c)}{\leq} 1 - H(Y_2|X_1X_2) - H(Y_1|X_1X_2) + H(Z|X_1X_2) \\
   \overset{(d)}{=} 1 - h(\epsilon_1) - h(\epsilon_2) + h(\epsilon_2)
   \]
where (c) follows from \( \text{Lemma } 30 \); (d) follows from \( H(Y_2|X_1X_2) = h(\varepsilon_2) \), \( H(Y_1|X_1X_2) = h(\varepsilon_1) \), and \( H(Z|X_1X_2) = h(\varepsilon_z) \). 

- \( I(X_1;Y_2|X_2) - I(X_1;Z) \) and \( I(X_1;Y_2|X_2) - I(X_1;Z) \):

\[
I(X_1;Y_2|X_2) - I(X_1;Z) \\
= I(X_1;Y_2|X_2) + I(X_2;Z|X_1) - I(X_1, X_2; Z) \\
= H(Y_2|X_2) + H(Z|X_1) - H(Z) - H(Y_2|X_1X_2) - H(Z|X_1X_2) + H(Z|X_1X_2) \\
= H(Y_2|X_2) + H(Z|X_1) - H(Z) - H(Y_2|X_1X_2) \\
\leq 1 - h(\varepsilon_z)
\]

where (e) follows from the same process as \( \text{Lemma } 30 \), and \( H(Y_2|X_1X_2) = h(\varepsilon_2) \).

Similarly, we have

\[
I(X_2;Y_1|X_1) - I(X_2;Z) \leq 1 - h(\varepsilon_1).
\]

Based on these items above, the reliable transmission rate given by Shannon \( \text{[7]} \), the achievable reliable transmission rate pair \((R_1, R_2)\) satisfies

\[
R_{1s} \leq I(X_1;Y_2|X_2) \leq 1 - h(\varepsilon_2), \\
R_{2s} \leq I(X_2;Y_1|X_1) \leq 1 - h(\varepsilon_1).
\]

According to the results in previous work \( \text{[1]} \), we can obtain the achievable rate pair \((R_{1s}, R_{2s})\) for the deterministic modulo-2 two-way wiretap channel with the joint secrecy as

\[
R_{1s} \leq I(X_1;Y_2|X_2) - I(X_1;Z) \leq 1 - h(\varepsilon_2), \\
R_{2s} \leq I(X_2;Y_1|X_1) - I(X_2;Z) \leq 1 - h(\varepsilon_1), \\
R_{1s} + R_{2s} \leq I(X_1;Y_2|X_2) + I(X_2;Y_1|X_1) - I(X_1, X_2; Z) \leq 1 + h(\varepsilon_z) - h(\varepsilon_1) - h(\varepsilon_2).
\]

By Theorem \( \text{[1]} \) the secrecy rate pair \((R_{1s}, R_{2s})\) with the individual secrecy for the modulo-2 binary two-way wiretap channel satisfies

\[
R_{1s} \leq I(X_2;Y_1|X_1) - I(X_2;Z) \leq 1 - h(\varepsilon_2), \\
R_{2s} \leq I(X_1;Y_2|X_2) - I(X_1;Z) \leq 1 - h(\varepsilon_1).
\]

**APPENDIX B**

**PROOF OF THEOREM \( \text{[1]} \)**

With fixed probability density function \( p(u_1) \) and \( p(u_2) \), the codebooks are generated as follows.

1) **Codebook generation:** According to \( p(u_i) \), the user \( i \), \((i = 1, 2)\), randomly generates \( 2^{nR_i} \) independent and identically distributed (i.i.d) sequences \( u^n_i(w_{is}, w_{ir}) \), with \( (w_{is}, w_{ir}) \in [1 : 2^{nR_i}] \times [1 : 2^{nR_r}] \). Note that

\[
R_i = R_{is} + R_{ir}.
\]

When analysing the secrecy measurement \( \frac{1}{n}H(W_{1s}|Z^n) \), if \( R_2 \geq I(U_2;Z|U_1) \) then the codebook of user 2 is equally partitioned into \( 2^{R_{2s}} \) parts with \( R_{2s} = R_2 - I(U_2;Z|U_1) + \epsilon' \), each part consisting of \( 2^{R_{2s}} \) codewords with \( R_{2s} = I(U_2;Z|U_1) - \epsilon' \). Correspondingly, for the secrecy analysis of \( \frac{1}{n}H(W_{2s}|Z^n) \), if \( R_1 \geq I(U_1;Z|U_2) \) then the codebook of of user 1 is equally partitioned into \( 2^{R_{1s}} \) parts with \( R_{1s} = R_1 - I(U_1;Z|U_2) + \epsilon' \), each part consisting of \( 2^{R_{1s}} \) codewords with \( R_{1s} = I(U_1;Z|U_2) - \epsilon' \).

2) **Encoding:** To send message \( w_{1s} \), user 1 randomly chooses \( w_{1r} \in [1 : 2^{nR_{1r}}] \), finds \( u^n_{1r}(w_{1s}, w_{1r}) \), generates \( x^n_{1r} \) according to \( p(x_1|u_1) \) and sends \( x^n_{1r} \) to the channel. Similarly, to send message \( w_{2s} \), user 2 randomly chooses \( w_{2r} \in [1 : 2^{nR_{2r}}] \), finds \( u^n_{2r}(w_{2s}, w_{2r}) \), generates \( x^n_{2r} \) according to \( p(x_2|u_2) \) and sends \( x^n_{2r} \) to the channel.
3) Decoding: User 1 declares that \( \hat{w}_{2s} \) is sent by user 2 if \( u^n_2(\hat{w}_{2s}, \hat{w}_{2r}) \) is the unique sequence such that \( (u^n_2(\hat{w}_{2s}, \hat{w}_{2r}), x^n_1, y^n_1) \in T^n_e \). User 2 declares that \( \hat{w}_{1s} \) is sent by user 1 if \( u^n_1(\hat{w}_{1s}, \hat{w}_{1r}) \) is the unique sequence such that \( (u^n_1(\hat{w}_{1s}, \hat{w}_{1r}), x^n_2, y^n_2) \in T^n_e \).

4) Reliability Analysis: Based on the AEP and packing lemma [18], for sufficiently large \( n \), the average error probability of \( P_{e,1} \) and \( P_{e,2} \) goes to zero, if

\[
R_1 \leq I(U_1; Y_2 | X_2) - 4\epsilon, \quad R_2 \leq I(U_2; Y_1 | X_1) - 4\epsilon. \tag{32}
\]

5) Individual secrecy analysis: Firstly, we consider the equivocation of \( W_{1s} \) as follows.

\[
H(W_{1s} | Z^n)
= H(W_{1s}, W_{1r}, W_{2s}, U^n_1, U^n_2 | Z^n) - H(W_{1r}, W_{2s}, U^n_1, U^n_2 | W_{1s}, Z^n)
\]

\[
= n(I_{1s} + I_{1r} + I_{2s}) - nI(U_1, U_2; Z) - ne_n - H(W_{1r}, W_{2s}, U^n_1, U^n_2 | W_{1s}, Z^n) \tag{33}
\]

where (a) follows from the Markov chain \( (W_{1s}, W_{1r}, W_{2s}) \rightarrow (U^n_1, U^n_2) \rightarrow Z^n \), such that \( I(W_{1s}, W_{1r}, W_{2s}, U^n_1, U^n_2; Z^n) = I(U^n_1, U^n_2; Z^n) \); (b) follows from \( H(W_{1s}, W_{1r}, W_{2s}, U^n_1, U^n_2) = H(U^n_1, U^n_2) \), hence according to the codebook construction

\[
H(W_{1s}, W_{1r}, W_{2s}, U^n_1, U^n_2) = n(R_{1s} + R_{1r} + R_{2s}).
\]

And \( I(U^n_1, U^n_2; Z^n) \leq nI(U_1, U_2; Z) + ne_n \), which follows a similar proof of [19, Lemma 3].

Then the last term \( H(W_{1r}, W_{2s} | W_{1s}, Z^n) \) in (33) can be bounded in two different cases as follows.

1) If \( R_2 \leq I(U_2; Z | U_1) \), then

\[
H(W_{1r}, W_{2s}, U^n_1, U^n_2 | W_{1s}, Z^n) \tag{34}
\]

where (c) follows from the Fano’s inequality by taking

\[
R_{1r} + R_2 \leq I(U_1, U_2; Z) - \epsilon. \tag{35}
\]

Replacing the third terms in (33) by (34), we obtain

\[
H(W_{1s} | Z^n)
\geq n[R_{1s} + R_{1r} + R_{2s}] - nI(U_1, U_2; Z) - ne_n - ne' \tag{36}
\]

where (d) follows by taking

\[
R_{1r} + R_2 \geq I(U_1, U_2; Z) - 2\epsilon. \tag{37}
\]

2) If \( R_2 \geq I(U_2; Z | U_1) \), then

\[
H(W_{1r}, W_{2s}, U^n_1, U^n_2 | W_{1s}, Z^n)
= H(U^n_1, U^n_2 | W_{1s}, Z^n)
= H(U^n_1 | W_{1s}, Z^n) + H(U^n_2 | W_{1s}, Z^n, U^n_1)
\leq ne' + n[R_2 - I(U_2; Z | U_1)] + ne' \tag{38}
\]

where (e) follows from the Fano’s inequality by taking

\[
R_{1r} \leq I(U_1; Z) - \epsilon. \tag{39}
\]
The second term \( H(U_2^n|W_{1s}, Z^n, U_1^n) \) is bounded as follows. Since \( R_2 \geq I(U_2; Z|U_1) \), we consider the codebook by rate splitting as explained in 1) codebook generation. Therefore, we have

\[
H(U_2^n|W_{1s}, Z^n, U_1^n) = H(W_{21}, W_{22}|W_{1s}, U_1^n, Z^n) = H(W_{21}|W_{1s}, W_{1r}, Z^n) + H(W_{22}|W_{21}, W_{1s}, U_1^n, Z^n)
\]

\[
(f) \leq nR_{21} + ne'
\]

\[
= n[R_2 - I(U_2; Z|U_1)] + ne',
\]

where \((f)\) follows from \( H(W_{21}|W_{1s}, W_{1r}, Z^n) \leq H(W_{21}) = nR_{21}; \) \( H(W_{22}|W_{21}, W_{1s}, U_1^n, Z^n) \leq ne' \) holds by the Fano’s inequality by taking \( R_{22} \leq I(U_2; Z|U_1) \), which is due to the structure of the codebook.

Replacing the last term in \((33)\) by \((38)\), we obtain

\[
H(W_{1s}|Z^n)
\]

\[
\geq n(R_{1s} + R_{1r} + R_2) - nI(U_1, U_2; Z) - n\epsilon_n - [n(R_2 - I(U_2; Z|U_1)) + 2ne']
\]

\[
(g) \geq nR_{1s} - n(\epsilon_n + 2\epsilon'),
\]

where \((g)\) follows from

\[
R_{1r} \geq I(U_1; Z) - 2\epsilon.
\]

By \((39)\) and \((40)\), we have

\[
R_{1r} = I(U_1; Z).
\]

Combining \((37)\) and \((41)\) into a more compact form, \( R_{1r} \) can be rewritten as

\[
R_{1r} = I(U_1; Z) + |I(U_2; Z|U_1) - R_2| + .
\]

Following from a similar analysis of \( \frac{1}{n} H(W_{2s}|Z^n) \), we have

\[
R_{2r} = I(U_2; Z) + |I(U_1; Z|U_2) - R_1| + .
\]

6) secrecy rate analysis: Considering the reliability and the individual secrecy analysis, we obtained \((31), (32), (42), (43)\). After the Fourier-Motzkin elimination, the achievable secrecy rate region is the union of non-negative rate pair \((R_{1s}, R_{2s})\) satisfying

\[
\begin{align*}
R_{1s}, R_{2s} &\geq 0, \\
R_{1s} &\leq I(U_1; Y_2|X_2) - I(U_1; Z) - |I(U_2; Z|U_1) - I(U_2; Y_2|X_2)| + , \\
R_{2s} &\leq I(U_2; Y_1|X_1) - I(U_2; Z) - |I(U_1; Z|U_2) - I(U_1; Y_1|X_1)| + .
\end{align*}
\]

APPENDIX C

PROOF OF THEOREM

Proof:

First, we define the following auxiliary random variables to proceed to \( R_{1s} \).

\[
U_i = X_2^{i-1} Y_2^{i-1} Z_{i+1}^n, V_{1i} = (W_{1s}, U_i), V_{2i} = (W_{2s}, U_i)
\]
\[ nR_{1s} = H(W_{1s}) \leq H(W_{1s}|Z^n) + n\epsilon \]
\[ = H(W_{1s}|Z^n) - H(W_{1s}|Y_2^n, X_2^n) + H(W_{1s}|Y_2^n, X_2^n) + n\epsilon \]
\[ = H(W_{1s}) - I(W_{1s}; Z^n) - H(W_{1s}) + I(W_{1s}; Y_2^n, X_2^n) + H(W_{1s}|Y_2^n, X_2^n) + n\epsilon \]
\[ = -I(W_{1s}; Z^n) + I(W_{1s}; Y_2^n, X_2^n) + H(W_{1s}|Y_2^n, X_2^n) + n\epsilon \]
\[ \leq I(W_{1s}; X_2^n, Y_2^n) - I(W_{1s}; Z^n) + n\epsilon + n\delta_n \]
\[ = \sum_{i=1}^{n} [I(W_{1s}; X_2, Y_2|X_2^{i-1}, Y_2^{i-1}) - I(W_{1s}; Z_i|Z^n_{i+1})] + n\epsilon + n\delta_n \]
\[ = \sum_{i=1}^{n} [I(W_{1s}; Z^n_{i+1}; X_2, Y_2|X_2^{i-1}, Y_2^{i-1}) - I(Z^n_{i+1}; X_2, Y_2|X_2^{i-1}, Y_2^{i-1}, W_{1s}) + n\epsilon + n\delta_n \]
\[ = \sum_{i=1}^{n} [I(W_{1s}; X_2^{i-1}, Y_2^{i-1}, Z_i|Z^n_{i+1}) + I(X_2^{i-1}, Y_2^{i-1}; Z_i|Z^n_{i+1}, W_{1s})] + n\epsilon + n\delta_n \]
\[ \leq I(W_{1s}; Z^n_{i+1}; X_2, Y_2|X_2^{i-1}, Y_2^{i-1}) - I(W_{1s}; X_2^{i-1}, Y_2^{i-1}, Z_i|Z^n_{i+1})] + n\epsilon + n\delta_n \]
\[ = \sum_{i=1}^{n} [I(V_{1i}; X_2, Y_2|U_i) - I(V_{1i}; Z_i|U_i)] + n\epsilon + n\delta_n \]
\[ = n[I(V_1; X_2, Y_2|U) - I(V_1; Z|U)] + n\epsilon + n\delta_n \]

where (a) follows by the Fano’s inequality; (b) follows from the Csizsár sum identity [11]; (c) follows from the definition \( U_i = (X_2^{i-1}, Y_2^{i-1}, Z^n_{i+1}) \) and \( V_{1i} = (W_{1s}, U_i) \) in (45); and (d) follows from the standard procedure of introducing a time-sharing random variable.

Similarly, we can obtain
\[ R_{2s} \leq I(V_2; X_1, Y_1|U) - I(V_2; Z|U) + \epsilon + \delta_n \]

\[ \blacksquare \]

**APPENDIX D**

**PROOF OF THEOREM 4**

**Proof:**
1) For the two-way wiretap channels with $Y_1 = Y_2 = Z$, we first derive the outer bound of $R_{1s}$.

\[ nR_{1s} \leq H(W_{1s}|Z^n) + n\epsilon \]

\[ \overset{(a)}{=} H(W_{1s}|Y^n) - H((W_{1s}|X_1^n Y^n) + H(W_{1s}|X_2^n Y^n) + n\epsilon \]

\[ = I(W_{1s}; X_2^n|Y^n) + H(W_{1s}|X_2^n Y^n) + n\epsilon \]

\[ \overset{(b)}{\leq} I(W_{1s}; X_2^n|Y^n) + n\epsilon_1 + n\epsilon \]

\[ \leq I(W_{1s}; X_1^n; X_2^n|Y^n) + n\epsilon_1 + n\epsilon \]

\[ = I(X_1^n; X_2^n|Y^n) + I(W_{1s}; X_2^n|Y^n X_1^n) + n\epsilon_1 + n\epsilon \]

\[ \overset{(c)}{=} I(X_1^n; X_2^n|Y^n) + n\epsilon_1 + n\epsilon \]

\[ = I(X_1^n; X_2^n|Y^n) - I(X_1^n; Y^n) + n\epsilon_1 + n\epsilon \]

\[ \overset{(d)}{=} I(X_1^n; Y^n|X_2^n) - I(X_1^n; Y^n) + n\epsilon_1 + n\epsilon \]

\[ = \sum_{i=1}^{n} [I(X_1^n; Y_i|X_1^n Y_i^{i-1}) - I(X_1^n; Y_i|Y_i^{i-1})] + n\epsilon_1 + n\epsilon \]

\[ = \sum_{i=1}^{n} [H(Y_i|X_2^n Y_i^{i-1}) - H(Y_i|X_1^n X_2^n Y_i^{i-1}) - H(Y_i|Y_i^{i-1}) + H(Y_i|Y_i^{i-1} X_1^n)] + n\epsilon_1 + n\epsilon \]

\[ \overset{(e)}{=} \sum_{i=1}^{n} [I(X_1^n; Y_i|X_2^n Y_i^{i-1}) - H(Y_i|X_1^n X_2^n) - H(Y_i|Y_i^{i-1}) + H(Y_i|X_1^n Y_i^{i-1})] + n\epsilon_1 + n\epsilon \]

\[ = \sum_{i=1}^{n} [I(X_1^n; Y_i|X_2^n Y_i^{i-1}) - I(X_1^n; Y_i|Y_i^{i-1})] + n\epsilon_1 + n\epsilon \]

\[ \overset{(f)}{=} \sum_{i=1}^{n} [I(X_1^n; Y_i|X_2^n; Q_i, J = i) - I(X_1^n; Y_i|Q_i, J = i)] + n\epsilon_1 + n\epsilon \]

\[ \overset{(g)}{\leq} n[I(X_1^n; Y|X_2^n Q) - I(X_1^n; Y|Q)] + n\epsilon_1 + n\epsilon \]

where (a) follows from $Y_1 = Y_2 = Z = Y$; (b) follows from Fano’s inequality by taking $R_1 \leq I(X_1; Y_2 X_2)$; (c) follows from the coding scheme; (d) follows from the independence of $X_{1i}, X_{2i}$; the first and the last term of (e) follow that conditioning does not increase entropy, and the second term of (e) follows that $Y_i$ is independent of everything else given $X_{1i}, X_{2i}$ (Markov chain $Y_i^{i-1} \rightarrow (X_{1i}, X_{2i}) \rightarrow Y_i$); (f) follows from the definition of $Q_i = Y_i^{i-1}$ and $J = i$; (g) follows from that $J$ is uniformly distributed over $\{1, 2, \ldots, n\}$.

Similarly, we can obtain the outer bound $R_{2s} \leq I(X_1^n; Y|X_2^n) - I(X_1^n; Y|Q)$.

2) For the second class of channels, first, we define the following auxiliary random variables to proceed to $R_{1s}$.

\[ Q_i = Z^{i-1}. \] (45)
\[ nR_{1s} \leq H(W_{1s}|Z^n) + ne \]
\[ = H(W_{1s}|Z^n) - H((W_{1s}|X^n_2 Y^n_2 Z^n) + H(W_{1s}|X^n_2 Y^n_2 Z^n) + ne \]
\[ = I(W_{1s}; X^n_2 Y^n_2 Z^n) + H(W_{1s}|X^n_2 Y^n_2 Z^n) + ne \]
\[ \leq I(W_{1s}; X^n_2 Y^n_2 Z^n) + ne_1 + ne \]
\[ \leq I(W_{1s}X^n_1; X^n_2 Y^n_2 Z^n) + ne_1 + ne \]
\[ = I(X^n_1; X^n_2 Y^n_2 |Z^n) + I(W_{1s}; X^n_2 Y^n_2 |Z^n X^n_1) + ne_1 + ne \]
\[ \leq I(X^n_1; X^n_2 Y^n_2 |Z^n) + ne_1 + ne \]
\[ = H(X^n_1 |Z^n) - H(X^n_1 |X^n_2 Y^n_2 Z^n) + ne_1 + ne \]
\[ \leq I(X^n_1; X^n_2 Y^n_2) - I(X^n_1; Z^n) + ne_1 + ne \]
\[ = I(X^n_1; Y^n_2 |X^n_2) - I(X^n_1; Z^n) + ne_1 + ne \]
\[ = \sum_{i=1}^{n} [I(X^n_1; Y^n_{2i} |X^n_2) - I(X^n_1; Z^n)] + ne_1 + ne \]
\[ \leq \sum_{i=1}^{n} [I(X^n_1; Y^n_{2i} |X^n_2 Z^{i-1}) - I(X^n_1; Z^n)] + ne_1 + ne \]
\[ \leq \sum_{i=1}^{n} [I(X^n_1; Y^n_{2i} |X^n_2 Z^{i-1}) - I(X^n_1; X^n_{2i} Z^{i-1}) - I(X^n_1; Z^n)] + ne_1 + ne \]
\[ \leq \sum_{i=1}^{n} [I(X^n_1; Y^n_{2i} |X^n_2 Z^{i-1}) - I(X^n_1; X^n_{2i} Z^{i-1}) - I(X^n_1; Z^n)] + ne_1 + ne \]
\[ = \sum_{i=1}^{n} [I(X^n_1; Y^n_{2i} |X^n_2 Z^{i-1}) - I(X^n_1; Z^n)] + ne_1 + ne \]
\[ \leq \sum_{i=1}^{n} [I(X^n_1; Y^n_{2i} |X^n_2 Z^{i-1}) - I(X^n_1 ; Z^n)] + ne_1 + ne \]
\[ \leq n[I(X_1; Y_2 |X_2, Q, J = i) - I(X_1 ; Z^n)] + ne_1 + ne \]
\[ \leq n[I(X_1; Y_2 |X_2, Q) - I(X_1 ; Z^n)] + ne_1 + ne \]

where (a) follows from Fano’s inequality by taking \( R_1 \leq I(X_1; X_2 Y_2) \); (b) follows from the coding scheme; (c) follows from the degraded condition, i.e. \( Z^n \) is degraded of \( Y^n_2 \); (d) follows from the degraded condition, i.e. \( Z^n \) is degraded of \( Y^n_2 \); the first term of (e) follows that conditioning does not increase entropy, and the second term of (e) follows that \( Y_{2i} \) is independent of everything else given \( X_{1i} \) and \( X_{2i} \) (Markov chain \( Z^{i-1} \rightarrow (X_{1i} \ X_{2i}) \rightarrow Y_{2i} \)); (f) follows from that conditioning does not increase entropy; (g) follows from the definition of \( Q_i = Z^{i-1} \) and \( J = i \); (h) follows from that \( J \) is uniformly distributed over \( \{1, 2, \ldots, n\} \).

Similarly, we can obtain
\[ R_{2s} \leq H(Y_1|X_1 Z^{i-1}) - H(Y_1|X_1 X_2 Z^{i-1}) - H(Z_i|Z^{i-1}) + H(Z_i|Z^{i-1}X_2) \]
\[ \leq I(X_2; Y_1 |X_1, Q) - I(X_2; Z^n). \]
**APPENDIX E**

**Proof of Theorem 5**

**A. Proof of the Achievability**

Let $U_1 \sim N(0,(1-\alpha)P_1)$, $U_2 \sim N(0,(1-\beta)P_2)$, $X'_1 \sim N(0,\alpha P_1)$, $X'_2 \sim N(0,\beta P_2)$, and $U_1, U_2, X'_1, X'_2$ are independent with each other. $X_1 = U_1 + X'_1$, $X_2 = U_2 + X'_2$. The achievability proof follows by calculating the mutual information terms in Theorem 1 with the above definitions. Hence, the achievable secrecy rate region is

$$
\mathcal{R}_{GTW}^{\alpha} = \bigcup_{\alpha, \beta \in [0,1]} \begin{cases} 
(R_{1s}, R_{2s}) : \\
R_{1s} \geq 0, R_{2s} \geq 0, \\
R_{1s} \leq \frac{1}{2} \log \frac{(P_1 + N_2)(\alpha P_1 + P_2 + N_e)}{(\alpha P_1 + N_2)(P_1 + P_2 + N_e)}, \\
R_{2s} \leq \frac{1}{2} \log \frac{(P_2 + N_1)(P_1 + P_2 + N_e)}{(P_2 + N_1)(P_1 + P_2 + N_e)}. 
\end{cases}
$$

Further considering the convex hull operation, the maximum achievable secrecy rate region $\mathcal{R}_{GTW}^{\alpha}$ is achieved when $\alpha = 0$, $\beta = 0$, i.e.

$$
\mathcal{R}_{GTW}^{\alpha} = \bigcup_{\alpha, \beta \in [0,1]} \begin{cases} 
(R_{1s}, R_{2s}) : \\
R_{1s} \geq 0, R_{2s} \geq 0, \\
R_{1s} \leq \frac{1}{2} \log \frac{(P_1 + N_2)(P_2 + N_e)}{N_2(P_1 + P_2 + N_e)}, \\
R_{2s} \leq \frac{1}{2} \log \frac{(P_2 + N_1)(P_1 + N_e)}{N_1(P_1 + P_2 + N_e)}. 
\end{cases}
$$

**B. Proof of the Converse**

We further derive the outer bound on the secrecy region. From [47], we have

$$
nR_{1s} \leq \sum_{i=1}^{n} [H(Y_2|X_2) - H(Y_2|X_1, X_2) - H(Z_i|Z^{i-1}) + H(Z_i|X_1, Z^{i-1})] + n\epsilon_1
$$

$$
= \sum_{i=1}^{n} [H(Y_2|X_2) - H(Y_2|X_1, X_2) - H(Z_i|Q_i) + H(Z_i|X_1, Q_i)] - H(Z^n) + n\epsilon_1
$$

$$
= \sum_{i=1}^{n} [H(Y_2|X_2) - H(Y_2|X_1, X_2) + H(Z_i|X_1, Q_i)] - H(Z^n) + n\epsilon_1
$$

In order to obtain the outer bound on the secrecy rate region of Gaussian two-way wiretap channels, firstly we calculate the following series of entropy.

- We derive the entropy of $H(Z_i|X_1, Q_i)$, $H(Z_i|X_2, Q_i)$. Firstly,

$$
H(Z_i|X_1, Q_i)
$$

$$(a) \geq h(Z_i|X_1, X_2, Q_i)
$$

$$
= \frac{1}{2} \log 2\pi e N_e
$$

where $(a)$ follows from that conditioning does not increase entropy.

On the other hand,

$$
H(Z_i|X_1, Q_i)
$$

$$(b) \leq h(Z_i|X_1)
$$

$$
= h(X_1 + X_2, |X_1)
$$

$$
\leq h(X_2, |X_1)
$$

$$
= \frac{1}{2} \log 2\pi e (P_2 + N_e)
$$
where \((b)\) follows from that conditioning does not increase entropy.

Such that there exists some \(\beta \in [0, 1]\) such that
\[
H(Z_i | X_{1i}, Q_i) = \frac{1}{2} \log 2\pi e [N_e + \alpha (P_2 + N_e - N_e)]
\]
\[
= \frac{1}{2} \log 2\pi e (\beta P_2 + N_e)
\]
(51)

Similarly, we have some \(\beta \in [0, 1]\) such that
\[
H(Z_i | X_{2i}, Q_i) = \frac{1}{2} \log 2\pi e (\alpha P_1 + N_e)
\]
(52)

- We derive the entropy of \(H(Y_{2i} | X_{2i}, Q_i)\) and \(H(Y_{1i} | X_{1i}, Q_i)\).

By the entropy power inequality, we obtain
\[
2^{2h(Z_i | X_{1i} = x_{1i}, Q_i = q_i)} \geq 2^{2h(Y_{1i} | X_{1i} = x_{1i}, Q_i = q_i) + 2^{2h(Z_i | X_{1i} = x_{1i}, Q_i = q_i)}}
\]
\[
\geq 2^{2h(Y_{1i} | X_{1i} = x_{1i}, Q_i = q_i) + 2^{2h(Z_i | X_{1i} = x_{1i}, Q_i = q_i)}} + 2\pi e (N_e - N_1)
\]

That is
\[
h(Z_i | X_{1i} = x_{1i}, Q_i = q_i) \geq \frac{1}{2} \log [2^{2h(Y_{1i} | X_{1i} = x_{1i}, Q_i = q_i) + 2\pi e (N_e - N_1)]}
\]

Taking the expectation on both sides of the preceding equation, we have
\[
h(Z_i | X_{1i}, Q_i) = \mathbb{E} h(Z_i | X_{1i} = x_{1i}, Q_i = q_i)
\]
\[
\geq \frac{1}{2} \mathbb{E} \log [2^{2h(Y_{1i} | X_{1i} = x_{1i}, Q_i = q_i) + 2\pi e (N_e - N_1)]}
\]
\[
\geq \frac{1}{2} \log [2^{2h(Y_{1i} | X_{1i}, Q_i) + 2\pi e (N_e - N_1)]}
\]
(53)

where \((c)\) follows from Jensen’s inequality.

By (51),
\[
\frac{1}{2} \log 2\pi e (\beta P_2 + N_e) = h(Z_i | X_{1i}, Q_i) \geq \frac{n}{2} \log [2^{2h(Y_{1i} | X_{1i}, Q_i) + 2\pi e (N_e - N_1)]}
\]
i.e.,
\[
2\pi e (\beta P_2 + N_e) \geq 2^{2h(Y_{1i} | X_{1i}, Q_i) + 2\pi e (N_e - N_1)}
\]
\[
h(Y_{1i} | X_{1i}, Q_i) \leq \frac{1}{2} \log 2\pi e (\beta P_2 + N_1)
\]
(54)

Similarly, we have
\[
h(Y_{2i} | X_{2i}, Q_i) \leq \frac{1}{2} \log 2\pi e (\alpha P_1 + N_2)
\]
(54)

- We derive the entropy of \(H(Y_{1i} | X_{1i}, X_{2i}, Q_i)\) and \(H(Y_{2i} | X_{1i}, X_{2i}, Q_i)\).

\[
H(Y_{1i} | X_{1i}, X_{2i}, Q_i) = H(Y_{1i} | X_{1i}, X_{2i}) = \frac{1}{2} \log 2\pi e N_1
\]
(55)
\[
H(Y_{2i} | X_{1i}, X_{2i}, Q_i) = H(Y_{2i} | X_{1i}, X_{2i}) = \frac{1}{2} \log 2\pi e N_2
\]
(56)

- We derive the entropy of \(H(Z^n)\).

By [20] Lemma 1 or [3] Lemma 10, let \(g(P) = \frac{1}{2} \log (2\pi e P)\),
\[
H(X^n_1 + X^n_2) = H(X^n_1) + H(X^n_2) = \frac{n}{2} \log 2\pi e (P_1 + P_2) = n g(P_1 + P_2) = n v.
\]
Since $Z^n = X^n_1 + X^n_2 + Z^*_n$, then

$$H(Z^n) \geq n g(N_e + g^{-1}(v))$$

$$= n g(N_e + g^{-1}(g(P_1 + P_2)))$$

$$= n g(N_e + P_1 + P_2)$$

$$= \frac{n}{2} \log 2 \pi e (N_e + P_1 + P_2)$$

(57)

Hence,

$$n R_{1s} \leq \sum_{i=1}^{n} [H(Y_{2i}|X_1,Q_i) - H(Y_{2i}|X_1,X_2,Q_i) + H(Z_i|X_1,Q_i)] - H(Z^n) + n \epsilon_1$$

$$\leq \frac{d}{2} n \log 2 \pi e (\alpha P_1 + N_2) - \frac{n}{2} \log 2 \pi e N_2 + \frac{n}{2} \log 2 \pi e (\beta P_2 + N_e) - \frac{n}{2} \log 2 \pi e (N_e + P_1 + P_2)$$

$$\leq \frac{n}{2} \log (\alpha P_1 + N_2)(\beta P_2 + N_e)$$

$$+ \frac{n}{2} \log 2 \pi e (\beta P_2 + N_1)(\alpha P_1 + N_e) - \frac{n}{2} \log 2 \pi e (N_e + P_1 + P_2)$$

(58)

where $(d)$ follows by substituting (54), (56), (51) and (57).

Similarly,

$$n R_{2s} \leq \sum_{i=1}^{n} [H(Y_{1i}|X_2,Q_i) - H(Y_{1i}|X_1,X_2,Q_i) + H(Z_i|X_2,Q_i)] - H(Z^n) + n \epsilon_1$$

$$\leq a \frac{n}{2} \log 2 \pi e (\alpha P_1 + N_2) - \frac{n}{2} \log 2 \pi e N_2 + \frac{n}{2} \log 2 \pi e (\beta P_2 + N_e) - \frac{n}{2} \log 2 \pi e (N_e + P_1 + P_2)$$

$$\leq \frac{n}{2} \log (\alpha P_1 + N_2)(\alpha P_1 + N_e)$$

$$+ \frac{n}{2} \log 2 \pi e (\beta P_2 + N_1)(\alpha P_1 + N_e) - \frac{n}{2} \log 2 \pi e (N_e + P_1 + P_2)$$

Hence, the outer bound on secrecy capacity region is

$$\mathcal{R}_{GTW}^O \triangleq \bigcup_{\alpha, \beta \in [0,1]} \left\{ \begin{array}{l}
(R_{1s}, R_{2s}) : \\
R_{1s} \geq 0, R_{2s} \geq 0,
\end{array} \right. \left\{ \begin{array}{l}
R_{1s} \leq \frac{1}{2} \log \frac{(\alpha P_1 + N_2)(\beta P_2 + N_e)}{N_2(N_e + P_1 + P_2)}, \\
R_{2s} \leq \frac{1}{2} \log \frac{(\beta P_2 + N_1)(\alpha P_1 + N_e)}{N_1(N_e + P_1 + P_2)}.
\right. \right\}$$

Further considering the convex hull of $\alpha, \beta \in [0,1]$, the outer bound on the secrecy rate region is rewritten by $\alpha = \beta = 1$ as

$$\mathcal{R}_{GTW}^O \triangleq \left\{ \begin{array}{l}
(R_{1s}, R_{2s}) : \\
R_{1s} \geq 0, R_{2s} \geq 0,
\end{array} \right. \left\{ \begin{array}{l}
R_{1s} \leq \frac{1}{2} \log \frac{(P_1 + N_2)(P_2 + N_e)}{N_2(N_e + P_1 + P_2)}, \\
R_{2s} \leq \frac{1}{2} \log \frac{(P_2 + N_1)(P_1 + N_e)}{N_1(N_e + P_1 + P_2)}.
\right. \right\}$$

(59)

**REFERENCE**

[1] A. D. Wyner, “The wire-tap channel,” The bell system technical journal, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” IEEE transactions on information theory, vol. 24, no. 3, pp. 339–348, 1978.

[3] S. Leung-Yan-Cheong and M. E. Hellman, “The gaussian wire-tap channel,” Information Theory IEEE Transactions on, vol. 24, no. 4, pp. 451–456, 1978.

[4] Y. Liang and H. V. Poor, “Multiple-access channels with confidential messages,” IEEE Transactions on Information Theory, vol. 54, no. 3, pp. 976–1002, 2008.

[5] L. Lai and H. El Gamal, “The relay–eavesdropper channel: Cooperation for secrecy,” IEEE Transactions on Information Theory, vol. 54, no. 9, pp. 4005–4019, 2008.

[6] B. Dai and Z. Ma, “Multiple-access relay wiretap channel,” IEEE Transactions on Information Forensics and Security, vol. 10, no. 9, pp. 1835–1849, 2015.

[7] C. E. Shannon, “Two-way communication channels,” in Proc. 4th Berkeley Symp. Math. Stat. Prob, vol. 1. Citeseer, 1961, pp. 611–644.
[8] E. Tekin and A. Yener, “The gaussian multiple access wire-tap channel: wireless secrecy and cooperative jamming,” in 2007 Information Theory and Applications Workshop. IEEE, 2007, pp. 404–413.

[9] ———, “The general gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming,” IEEE Transactions on Information Theory, vol. 54, no. 6, pp. 2735–2751, 2008.

[10] ———, “Correction to: the gaussian multiple access wire-tap channel and the general gaussian multiple access and two-way wire-tap channels: achievable rates and cooperative jamming,” IEEE Transactions on Information Theory, vol. 56, no. 9, pp. 4762–4762, 2010.

[11] A. El Gamal, O. O. Koyluoglu, M. Youssef, and H. El Gamal, “Achievable secrecy rate regions for the two-way wiretap channel,” IEEE Transactions on Information Theory, vol. 59, no. 12, pp. 8099–8114, 2013.

[12] X. He and A. Yener, “The role of feedback in two-way secure communications,” IEEE Transactions on Information Theory, vol. 59, no. 12, pp. 8115–8130, 2013.

[13] A. J. Pierrot and M. R. Bloch, “Strongly secure communications over the two-way wiretap channel,” IEEE Transactions on Information Forensics and Security, vol. 6, no. 3, pp. 595–605, 2011.

[14] K. Bhattad and K. R. Narayanan, “Weakly secure network coding,” NetCod, Apr, vol. 104, 2005.

[15] S. Kadhe and A. Sprintson, “Weakly secure regenerating codes for distributed storage,” in International Symposium on Network Coding, 2014, pp. 1–6.

[16] N. Cai and R. W. Yeung, “Secure network coding,” in IEEE International Symposium on Information Theory, 2002. Proceedings, 2002, p. 323.

[17] S. Ulukus and E. Ekrem, “Cooperative secrecy in wireless communications,” in Securing Wireless Communications at the Physical Layer. Springer, 2009, pp. 143–172.

[18] A. El Gamal and Y. H. Kim, Network information theory. Cambridge university press, 2011.

[19] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, “Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions,” IEEE Transactions on Information Theory, vol. 54, no. 6, pp. 2493–2507, 2008.

[20] P. P. Bergmans, “A simple converse for broadcast channels with additive white gaussian noise (corresp.),” IEEE Trans. Information Theory, vol. 20, no. 2, pp. 279–280, 1974.