Two-way Channel In Information Theory

Yiwei Song  
University of Illinois at Chicago  
E-mail: sywad87@gmail.com

Abstract

Firstly, an introduction of two-way communications channel is given. Then Shannon’s inner and outer bounds are introduced. Upon Shannon’s bounds, tighter lower and upper bounds are found for general discrete memoryless two-way channels. After that, the capacity regions of some specific two-way channels are discussed. Finally, a summary is given.

I. INTRODUCTION

The two-way communications channel (TWC) was first introduced by Shannon in his original paper[1], in which he also derived the inner and outer bounds of TWC’s capacity region. Until now, some progress has been done on determining the capacity region of TWC upon Shannon’s inner and outer bounds. And some specific forms of TWCs are determined, while the general solution to capacity region of discrete memoryless two-way communications channel is still an open problem. Before the precise definition of TWCs is given, let’s firstly look at an intuitive example of two-way channel, which is taken from [14]:

Two persons want to send messages to each other through the two switches in their hands. The switches control a light bulb which both of them can observe (This is like a communication channel). The light is on when their switches are in the same state, while the light is off when their switches are in the different states. Obviously, the communication of the two persons can be coded in sequences of zeros and ones, corresponding to the switch states. As the first communications scheme, one person send one bit information by using his switch at each time interval, while the other person receive this message by observing the light. During this process, they can change their roles to communicate at both directions. As the second scheme, the two persons can send messages to each other simultaneously by using their switches. They can deduce the messages they received by the light and the state of their own switches. Of course, the second scheme is twice faster than the first one.

![Fig. 1. Two-way communications channel (taken from [1])](image)

The two-way communication channel is shown as Fig.1. $y_2$ is the communications output of input $x_1$ at a rate $R_1$ and $y_1$ is the output of $x_2$ at a rate $R_2$. The problem is to determine what pairs of rates $R_1$ and $R_2$ can be achieved with arbitrarily small error probabilities. And generally speaking, the difficulty is that the encoding of $x_1$ and decoding of $y_1$ can affect each other at terminal 1, and so do $x_2$ and $y_2$ at terminal 2.

II. CAPACITY REGION OF GENERAL DISCRETE MEMORYLESS TWC

A. Definition of discrete memoryless TWC

The discrete memoryless TWC (d.m. TWC) is defined as a set of transition probabilities $p(y_1, y_2|x_1, x_2)$, where $x_1 \in X_1$, $x_2 \in X_2$, $y_1 \in Y_1$, $y_2 \in Y_2$. The memoryless means that the current outputs only depend on current input. The two encoding functions of block code pair of length $n$ are defined as

\[
f_1(\Theta_1, Y_1^{i-1}) = X_1; i = 1, 2, \cdots, n
\]

\[
f_2(\Theta_2, Y_1^{i-1}) = X_2; i = 1, 2, \cdots, n
\]
And the decoding functions are defined as

\[ g_1(\Theta_1, Y_1^n) = \hat{\Theta}_2 \]
\[ g_2(\Theta_2, Y_2^n) = \hat{\Theta}_1 \]

Where \( W_1 \) and \( W_2 \) are the source messages, and \( i \) is the time interval. Here we can see the difficulty of determining the capacity region of d.m. TWC comes from the fact that the encoding not only depends on source messages but also on received signals at the terminal, and the decoding not only depends on received signals but also on sent messages. This makes the capacity region of TWC extremely difficult to determine.

B. Shannon’s inner and outer bound

In his original paper [1], Shannon states that for a memoryless discrete TWC there exists a convex region \( G \) of approachable rates. And any rates pair \( (R_1, R_2) \) in the region \( G \) is achievable, which means there exist codes that the two terminals can communicate at rates \( R_1 \) and \( R_2 \) in each direction with arbitrarily small error probability. Even the capacity region \( G \) is still not clear until now, Shannon gave an inner bound region \( G_I \) and outer bound region \( G_O \), which have the relationship with \( G \) as: \( G_I \subset G \subset G_O \) (Shown in Fig.3.)

The inner bound \( G_I \) is defined as the convex hull of points \( (R_1, R_2) \)

\[ R_1 = E_{p(x_1,x_2,y_2)} \left( \log \frac{p(x_1|x_2,y_2)}{p(x_1|x_2)} \right) \]
\[ = I(X_1;Y_2|X_2) \]  \[ (1) \]

\[ R_2 = E_{p(x_1,x_2,y_2)} \left( \log \frac{p(x_2|x_1,y_1)}{p(x_2|x_1)} \right) \]
\[ = I(X_2;Y_1|X_1) \]  \[ (2) \]
where $x_1$ and $x_2$ are independent which means that they are generated according to probability $p(x_1)p(x_2)$. Since $x_1$ and $x_2$ are independent, the expressions for $G_I$ above can also be written as

$$R_1 = E_{p(x_1,x_2,y_2)} \left( \log \frac{p(x_1|x_2,y_2)}{p(x_1)} \right)$$

$$= I(X_1; X_2, Y_2)$$  \hspace{1cm} (5)

$$R_2 = E_{p(x_1,x_2,y_2)} \left( \log \frac{p(x_2|x_1,y_1)}{p(x_2)} \right)$$

$$= I(X_2; X_1, Y_1)$$  \hspace{1cm} (6)

The outer bound $G_O$ is defined as the convex hull of points $(R_1, R_2)$, which is defined as in (1) (2) (3) (4). However, $x_1$ and $x_2$ are not necessarily independent. So they are generated according to arbitrary joint probability $p(x_1,x_2)$. Therefore, $G_O$ is defined as in (1) (2) (3) (4). The difference is that for $G_I$ $x_1$ and $x_2$ are generated independently, while for $G_O$ $x_1$ and $x_2$ are not necessarily independent. Arbitrary joint probability $p(x_1,x_2)$ has more flexibility (degree of freedom) than $p(x_1)p(x_2)$, and so it is easy to understand $G_O$ is a larger region than $G_I$. Also, we can consider the capacity region $G$ as a middle state between $G_O$ and $G_I$, and $x_1$ and $x_2$ are in a middle state between total independence and complete arbitrary.

There are a few things/properties we should know about Shannon's inner and outer bounds:

1. The three regions $G_I$, $G$ and $G_O$ are all convex and have the same intercepts on axes. [1]
2. Projectivity of $G_I$ and $G_O$: if $(R_1, R_2) \in G_I(G_O)$, $(R_1, 0), (0, R_2) \in G_I(G_O)$. This property is proved in [1].
3. It is noticed that both $G_I$ and $G_O$ are the convex hull of the points $(R_1, R_2)$ defined with the same expressions ((1)(2)(3)(4)). The difference is that for $G_I$ $x_1$ and $x_2$ are generated independently, while for $G_O$ $x_1$ and $x_2$ are not necessarily independent. Arbitrary joint probability $p(x_1,x_2)$ has more flexibility (degree of freedom) than $p(x_1)p(x_2)$, and so it is easy to understand $G_O$ is a larger region than $G_I$. Also, we can consider the capacity region $G$ as a middle state between $G_O$ and $G_I$, and $x_1$ and $x_2$ are in a middle state between total independence and complete arbitrary.
4. Note that the expressions of $(R_1, R_2)$ defined for inner bound is not necessarily a convex region (For example: ‘push-to-talk’ channel introduced in [1]). But inner bound, as the convex hull of these points, is convex region. While the expressions of $(R_1, R_2)$ defined for outer bound is a convex region, so the region itself is the outer bound. These properties are proved in [1][14]
5. For some specific channels (such as binary erasure-noiseless TWC, modulo-2 adding channel, restricted channel...), the inner bound coincides with the outer bound (Which means probability product $p(x_1)p(x_2)$ is sufficient to maximize $R_1 + \lambda R_2$), so the capacity region can be determined. For some other channels (such as binary multiplying channel), they have different inner and outer bound, the capacity region of these channels are still not known.

C. Tighter lower and upper bounds than Shannon’s bounds

Even though we still do not know the capacity region of general d.m. TWC, researchers have made some progress to derive a tighter lower and upper bound than Shannon’s inner and outer bound respectively. In this section, a tighter lower bound found by Han [6] and a tighter upper bound found by Zhang et al. [7] are introduced.

Firstly, we look at Han’s lower bound, which is obtained by means of test channel and auxiliary random variables. (Shown in Fig.4.)

$$R_1 = E_{p(x_1,x_2,y_2)} \left( \log \frac{p(x_1|x_2,y_2)}{p(x_1)} \right)$$

$$= I(X_1; X_2, Y_2)$$  \hspace{1cm} (6)

$$R_2 = E_{p(x_1,x_2,y_2)} \left( \log \frac{p(x_2|x_1,y_1)}{p(x_2)} \right)$$

$$= I(X_2; X_1, Y_1)$$  \hspace{1cm} (8)

Where $U_1, U_2$ represent the new message information; $\tilde{U}_1, \tilde{U}_2$ represent previous message information; and $\tilde{W}_1, \tilde{W}_2$ represent feedback information from output terminals. And $\tilde{W}_1 = (X_1, Y_1), \tilde{W}_2 = (X_1, Y_1)$. So now
the transition probability of the test channel is:
\[ p^*(y_1, y_2 | (u_1, \tilde{u}_1, \tilde{w}_1), (u_2, \tilde{u}_2, \tilde{w}_2)) = \sum_{x_1, x_2} p(y_1, y_2 | x_1, x_2) q_1(x_1 | u_1, \tilde{u}_1) q_2(x_2 | u_2, \tilde{u}_2, \tilde{w}_2) \]

Han stated that the capacity region \( R \) of d.m. TWC is the convex hull of rates pair \( (R_1, R_2) \) satisfying:
\[
R_1 \leq I(\tilde{U}_1; X_2, Y_2, \tilde{U}_2, \tilde{W}_2)
\]
\[
R_2 \leq I(\tilde{U}_2; X_1, Y_1, \tilde{U}_1, \tilde{W}_1)
\]
for all possible \( p(u_1, u_2, \tilde{u}_1, \tilde{u}_2, \tilde{w}_1, \tilde{w}_2, x_1, x_2, y_1, y_2) \). He also proved:
\[
G_I \subset R \subset G_O
\]
and there exists TWC’s for which
\[
G_I \not= R
\]
Moreover, Han also gave a general coding scheme which can achieve all the rate pairs his capacity region \( R \). These show that Han’s region \( R \) is truly a tighter lower bound than Shannon’s inner bound for general d.m. TWC.
Zhang et al. made an improvement upon Shannon’s outer bound. Their main idea is to impose conditions on \((X_1, X_2)\) intermediate between independence and complete generality \([\cdot]\). This idea is abstractly mentioned above. Zhang et al. use random variables \((Z_1, Z_2)\) to specify the intermediate.
Zhang et al. gave three sets of rates pairs \((R_1, R_2)\) and proved the capacity region \( R \) of TWC is a subset of them respectively. The first set includes all rates pairs \((R_1, R_2)\) satisfying
\[
R_1 \leq \min(H(X_1|Z_1), I(X_1; Y_2 | X_2, Z_2))
\]
\[
R_2 \leq \min(H(X_2|Z_2), I(X_2; Y_1 | X_1, Z_1))
\]
for all possible joint distributions \( p(x_1, x_2, y_1, y_2, z_1, z_2) = p(z_1, z_2) p(x_1 | z_1) p(x_2 | z_2) p(y_1, y_2 | x_1, x_2) \)
The second set includes all rates pairs \((R_1, R_2)\) satisfying
\[
R_1 \leq \min(H(X_1|X), I(X_1; Y_2 | X_2))
\]
\[
R_2 \leq \min(H(X_2|X), I(X_2; Y_1 | X_1))
\]
for all possible joint distributions \( p(x_1, x_2, y_1, y_2, x) = p(x)p(x_1 | x) p(x_2 | x) p(y_1, y_2 | x, x_2) \)
The third set is the intersection of two sets \( R' \) and \( R'' \). \( R' \) include all rates pairs \((R_1, R_2)\) satisfying
\[
R_1 \leq \min(H(X_1|X), I(X_1; Y_2 | X_2))
\]
\[
R_2 \leq \min(H(X_2|X), I(X_2; Y_1 | X_1))
\]
for all possible joint distributions given in the second set.
\( R'' \) include all rates pairs \((R_1, R_2)\) satisfying
\[
R_1 \leq \min(H(X_1|X), I(X_1; Y_2 | X_2, X))
\]
\[
R_2 \leq \min(H(X_2|X), I(X_2; Y_1 | X_1))
\]
for all possible joint distributions given in the second set.
In the three sets defined above, \( Z_1, Z_2, X \) are auxiliary random variables, which satisfy the joint distributions form given in each set. And the second and third sets can be seen as a simpler but looser version of the first set.[\cdot]. All the three sets given by Zhang et al. are the tighter upper bounds than Shannon’s outer bound.
Finally, it is interesting to notice that both Han and Zhang et al. get their results by means of auxiliary random variables. And these tighter bounds can apply to a special form of TWC, Binary Multiplying Channel, which will be discussed later.

III. Capacity region of some specific TWCs
Even though there is no closed expression or practical solution to determine the capacity region of general d.m. TWC, some specific TWCs’ capacity region is equal to Shannon’s inner bound and some ones have even tighter bounds than that for general TWC.
A. Semi-symmetric TWC

TWC is said to have a semi-symmetric structure if for all possible interchanges of inputs alphabets of $X_1$, the output alphabets $Y_1$ or $Y_2$ or both $Y_1$ and $Y_2$ can be interchanged to leave the transition probability matrix $p(y_1, y_2|x_1, x_2)$ the same.

The semi-symmetric TWC is shown to have a coincided inner and outer bound, which means the maximum for a dependent assignment of $p(x_1, x_2)$ is actually obtained with $x_1$ and $x_2$ independent. This property is proved in [1] and [14].

An example of semi-symmetric TWC is called binary erasure noiseless TWC with transition probabilities as: $y_1 = x_2$ (this is a perfect noiseless binary channel). And if $x_2 = 0$, $y_2 = x_1$. If $x_2 = 1$, $p(y_2 = 0|x_2 = 1) = p(y_2 = 1|x_2 = 1) = 0.5$. The transition probability matrix is shown in Fig. 5. Notice that for all possible interchanges of inputs alphabets of $x_1$ (here $1 \leftrightarrow 0$), if output alphabets of $y_2$ is interchanged simultaneously, the transition probability is unaltered. The capacity region $(R_1, R_2)$ of binary erasure-noiseless TWC is given in [1] by

$$R_1 = (1 - \beta)H(\alpha)$$

$$R_2 = H(\beta)$$

where $0 \leq \alpha, \beta \leq 1$. (shown in Fig. 6.)

\[
\begin{array}{c|cccc}
  x_1 x_2 & 00 & 01 & 10 & 11 \\
  \hline
  y_1 y_2 & 00 & 01 & 10 & 11 \\
  \hline
  00 & 1 & 0 & 0 & 0 \\
  01 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
  10 & 0 & 1 & 0 & 0 \\
  11 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
\end{array}
\]

Fig. 5. Transition probability of binary erasure-noiseless TWC (taken from [14])

Fig. 6. The capacity region of binary erasure-noiseless TWC (taken from [14])
B. Symmetric TWC

TWC is said have a symmetric structure if for all possible interchanges of inputs alphabets of $X_1$ and $X_2$, the output alphabets $Y_1$ or $Y_2$ or both $Y_1$ and $Y_2$ can be interchanged to leave the transition probability matrix $p(y_1, y_2|x_1, x_2)$ the same. Since the symmetric TWC is also a semi-symmetric TWC, capacity region of symmetric TWC is shown to equal the inner bound. And more specifically, the capacity region of symmetric TWC is a rectangle formed by the point $(0,0)$ and the point corresponding to the input distribution $p(x_1, x_2) = \frac{1}{|X_1||X_2|}$, which is proved in [1] [14].

An example of symmetric TWC is modulo-2 adding channel: The inputs and outputs are all binary, and $y_1 = y_2 = x_1 \oplus x_2$. The receiver can simply deduce the message from the received signal and message sent by itself. So it is easy to see that the communication rates of two direction do not affect each other. According to the conclusion above, the capacity region of modulo-2 adding channel is a rectangle formed by the point $(0,0)$ and the point corresponding to the input distribution $p(x_1, x_2) = \frac{1}{4}$, which is the point $(1,1)$. (Shown in Fig.7.).

![Fig. 7. The capacity region of modulo-2 adding channel (taken from [1])](image)

C. Restricted TWC

Compared with the general TWC, restricted TWC is easier to deal with. In restricted TWC encoding only depend on the source messages to be transmitted, while in the general TWC encoding also depends the signal received at the terminal. The encoding function in restricted TWC can be expressed as

\[
\begin{align*}
    f_1(W_1, Y_1^{i-1}) &= X_1 \\
    f_2(W_2, Y_2^{i-1}) &= X_2
\end{align*}
\]

which is simpler than the encoding function in general TWC. For decoding process, the restricted TWC is the same as general TWC.

Obviously, the inputs $X_1$ and $X_2$ do not depend on outputs $Y_1$ and $Y_2$ any more. And $X_1$ and $X_2$ in restricted TWC are independent to each other. Thus, for restricted d.m. TWC, the capacity region is equal to the inner bound [\].

D. Single-output TWC (T-Channel)

As the name suggested, Single-output TWC has only one output $Y = Y_1 = Y_2$, and the the transition probability is shown as $p(y|x_1, x_2)$. The capacity region of single-output TWC is also not known yet, but besides the upper and lower bounds discussed in the first section, single-output channel has even more tighter bounds. Moreover, for some classes of single-output TWC, the capacity region is equal to the Shannon’s inner bound. These results are given and proved by Hekstra and Willems in [8]:

Up to now, the tightest upper bound of single-output TWC is given by means of an adaptive parallel channel. And this bound applies well to BMC, which will be shown in the next subsection.
Moreover, there are two classes of single-output TWC, of which the capacity region is equal to inner bounds. First class: A single-output TWC for which there exists a mapping \( f \) from \( X_2 \times Y \) into \( X_1 \) such that \( p(y|x_1, x_2) = 0 \) if \( x_1 \neq f(y, x_2) \). Second class: A single-output TWC for which there exists a finite alphabet \( U \) and two mappings \( f_1 : Y \times X_1 \rightarrow U \) and \( f_2 : Y \times X_2 \rightarrow U \) such that \( p(y|x_1, x_2) = 0 \) if \( f_1(y, x_1) \neq f_2(y, x_2) \), and, for which \( I(X_1; X_2) = 0 \) implies that \( I(X_1; X_2|Y, U) = 0 \) where \( U \) is a random variable and \( U = f_1(Y, X_1) = f_2(Y, X_2) \).

**E. Binary Multiplying Channel (BMC)**

No other TWC attract more attention than Binary Multiplying Channel (BMC). Since BMC was introduced by Blackwell, researchers had been trying to determine the capacity region of BMC for decades. Even though it is still not precisely determined, some progress have been made to constrain the capacity region to a comparatively small scope.

As one of the single-output TWC, BMC inherit all the natures and bounds of single-output TWC. Specifically for BMC, all inputs and outputs are binary and \( Y_1 = Y_2 = X_1X_2 \). Different from many TWCs discussed before, BMC is a typical TWC which has differing inner and outer bounds. According to the definitions, the inner bound of BMC is calculated as convex hull of points \((R_1, R_2)\) with expressions:

\[
R_1 = p_2 H(p_1) \\
R_2 = p_1 H(p_2)
\]

where \( p_1 = P(x_1 = 1) \) and \( p_2 = P(x_2 = 1) \) The outer bound of BMC is calculated as:

\[
R_1 = (1 - p_1) H\left(\frac{p_2}{1 - p_1}\right) \\
R_2 = (1 - p_2) H\left(\frac{p_1}{1 - p_2}\right)
\]

where \( p_1 = P(1, 0), p_2 = P(0, 1), 1 - p_1 - p_2 = P(1, 1). \) Note that \( P(0, 0) = 0 \). This is because that any assignment in which \( P(0, 0) \) is positive can be improved by transferring this probability to one of the other possible pairs \([\ldots]\). The inner and outer bound of BMC is shown in Fig.8.

![Fig. 8. The inner and outer bounds of BMC](image-url)
bound of BMC are (0.6169, 0.6169) and (0.6946, 0.6946) respectively. A code devised by D.W. Hagelbarger can achieve the rates pair (0.571, 0.571). And a improved Hagelbarger code suggested by Shannon can yields rates pair (0.593, 0.593). Note that both of them are inside the inner bound.

In the fifteen years after Shannon’s original paper, no encoding technique was found to achieve rates exceeding the inner bound for any TWC. So Meulen in [3] began to doubt that whether there exists such an encoding technique. Then Schalkwijk in [4] gave a coding scheme of BMC, which achieves the rates pair (0.61914, 0.61914) and outperforms the inner bound. So, Schalkwijk gave a positive answer to Meulen’s doubt and proved that the inner bound is not the capacity region for BMC, and of course for general d.m. TWC. Then in [5], Schalkwijk uses a technique called bootstrapping to further extend the capacity region of BMC, which can achieve rates pair (0.63056, 0.63056). For the upper bound of BMC, Zhang et al. proved that the capacity region of BMC does not include (0.64891, 0.64891) in [7]. And then Hekstra and Willems [8] showed that an upper bound for maximizing \( R_1 + R_2 \) is (0.64628, 0.64628). Thus, until now \((R'_1, R'_2)\) for real capacity region is between (0.63056,0.63056) and (0.64628,0.64628).

F. Gaussian TWC

In the previous sections, several kinds of discrete memoryless TWCs are discussed. Now we talk about a continuous memoryless two-way channel, Gaussian two-way channel, the capacity of which is found by Han [6]. [6] defined memoryless Gaussian two-way channel as:

\[
Y_1 = aX_1 + bX_2 + N_1 \\
Y_2 = cX_1 + dX_2 + N_2
\]

where \(X_1 \in X_1, X_2 \in X_2, Y_1 \in Y_1, Y_2 \in Y_2, X_1, X_2, Y_1\) and \(Y_2\) are the same set of real numbers. \(N_1\) and \(N_2\) are dependent Gaussian additive noises with zero mean and variances \(\sigma_1^2, \sigma_2^2\) respectively. Also, the inputs symbol has a power constraints defined as usual way:

\[
E(\sum_{t=1}^{n} x^2_{1t}) \leq nP_1 \\
E(\sum_{t=1}^{n} x^2_{2t}) \leq nP_2
\]

[] proves the capacity region of gaussian TWC is given by a set \((R_1, R_2)\) with the expressions:

\[
R_1 \leq \frac{1}{2} \log(1 + \frac{c^2P_1}{\sigma_2^2}) \\
R_2 \leq \frac{1}{2} \log(1 + \frac{b^2P_2}{\sigma_1^2})
\]

Notice that the capacity region does not depend on the parameters a and d. It is easy to understand because both terminals know the symbols sent by themselves. And when they decode the received signals, they can just simply subtract the symbols sent by themselves. It is like the modulo-2 adding channel, where the rates of each direction do not affect each other. Generally speaking, Gaussian TWC shows some symmetric natures.

IV. SUMMARY

For general d.m. TWC, the capacity region still not known. Shannon’s inner and outer bounds restrict the capacity region to a certain level. And a tighter lower and upper bound were proposed to further specify the capacity region. For specific TWC forms, besides the results stated above, they have some special characteristics, which may make their capacity region have more tighter bounds, or even to determine. Specifically, semi-symmetric TWC, symmetric TWC, restricted TWC and some classes of single-output TWC are proved to have their capacity regions equaling their inner bounds. While the capacity region of other sing-output channel, including BMC, is still not precisely known, but the tightest bounds up to now are discussed in this survey. For a continuous memoryless TWC, Gaussian TWC, the capacity region is proved to be the same as that of two independent Gaussian channels. What makes the capacity region of TWC so difficult to determine is the fact that the encoding at each terminal not only depends on the messages to send but also on the symbols received,
and the decoding depends on not only the symbols received but also the messages sent. BMC is regarded as an representative showing the difficulty of TWC.

Finally, it is also interesting to notice that in the recent years, some derivatives or new forms of TWC emerged, such as two-way relay channel [18], two-way wiretap channel [19,20] and multi-way channel [21]. These provide more research topics of two-way channels.

REFERENCES