A SCHEME FOR ENLARGING THE CAPACITY REGION
OF MULTIPLE-ACCESS CHANNELS USING FEEDBACK

BY

THOMAS M. COVER
S.K. LEUNG-YAN-CHEONG

TECHNICAL REPORT NO. 17
MARCH 1, 1976

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT ENG 76-03684

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
A SCHEME FOR ENLARGING THE CAPACITY REGION
OF MULTIPLE-ACCESS CHANNELS USING FEEDBACK

by

THOMAS M. COVER
S.K. LEUNG-YAN-CHEONG

TECHNICAL REPORT NO. 17
MARCH 1, 1976

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT ENG 76-03684
(formerly NSF Grant GK-34363)

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
A SCHEME FOR ENLARGING THE CAPACITY REGION
OF MULTIPLE-ACCESS CHANNELS USING FEEDBACK

Thomas M. Cover†
S.K. Leung-Yan-Cheong ‡‡

Abstract

Gaarder and Wolf [1] have shown by an example that it is possible to increase the capacity region of a discrete memoryless multiple-access channel through the use of feedback. In this note, a scheme for achieving this on general multiple-access channels is proposed and analyzed. The additive white Gaussian noise multiple-access channel is first considered and an achievable rate region using this scheme is calculated. Then the scheme is applied to discrete memoryless multiple-access channels. Sufficient conditions under which it is possible to increase the capacity region are fairly weak, indicating that feedback can almost always increase the capacity region.

†Stanford University. The work of this author was partially supported by NSF Grant GK 34363, and SRI ARPA Contract DAHC-15-73-C-0187.

‡‡Stanford University. The work of this author was supported under NSF Grants GK-33250 and ENG-10173.
A SCHEME FOR ENLARGING THE CAPACITY REGION
OF MULTIPLE-ACCESS CHANNELS USING FEEDBACK

Thomas M. Cover†
S.K. Leung-Yan-Cheong ††

Abstract

Gaarder and Wolf [1] have shown by an example that it is possible to
increase the capacity region of a discrete memoryless multiple-access chan-
nel through the use of feedback. In this note, a scheme for achieving this
on general multiple-access channels is proposed and analyzed. The additive
white Gaussian noise multiple-access channel is first considered and an
achievable rate region using this scheme is calculated. Then the scheme is
applied to discrete memoryless multiple-access channels. Sufficient condi-
tions under which it is possible to increase the capacity region are fairly
weak, indicating that feedback can almost always increase the capacity re-
gion.

†Stanford University. The work of this author was partially supported by
NSF Grant GK 34363, and SRI ARPA Contract DAHC-15-73-C-0187.

††Stanford University. The work of this author was supported under
NSF Grants GK-33250 and ENG-10173.
I Introduction

In a recent paper [1], Gaarder and Wolf showed by an example that it is possible to increase the capacity region of a discrete memoryless multiple-access channel through the use of feedback. The example which they used was a noiseless binary erasure multiple-access channel. In this note, we propose a feedback scheme for enlarging the capacity region of general multiple-access channels. The scheme consists of two stages. During the first stage (stage 1), the two transmitters send information reliably to each other at the maximum possible rate. Stage 1 ends when each transmitter has complete knowledge of the other's message. The rate of transmission in stage 1 will be too high for reliable transmission to the receiver. However, the stage 1 transmissions will enable the receiver to narrow down the set of possible transmitted messages to a considerably smaller set of "typical" messages. With probability arbitrarily close to 1, the set of "typical" messages will contain the actual transmitted message. The receiver and the two transmitters then arrange the messages in the "typical" set in some lexicographic ordering. This sets up stage 2 during which the two transmitters totally cooperate to send the index of the actual transmitted message.

In retrospect, it is interesting to note that the scheme used by Gaarder and Wolf is a special case of the scheme proposed here.

In the next section, we introduce some notation and summarize known results concerning the capacity region of memoryless multiple-access channels. In section III the model of the memoryless multiple-access channel with feedback is introduced. In section IV, we show that the proposed scheme increases the achievable rate region for the additive white Gaussian
noise (AWGN) multiple-access channel. In section V, we make a digression on typical sequences and recall some basic results which will be needed in section VI, where we consider discrete memoryless multiple-access channels.

II Some General Remarks and Notations

In a multiple-access channel several transmitters communicate with a single receiver. The channel is characterized by its input terminals with alphabets $x_1, x_2, \ldots, x_m$, its output terminal with alphabet $y$, and a set of conditional probability measures on the output signal $Y$ given the input signals $X_1, X_2, \ldots, X_m$ (see Fig. 1).

![Diagram of a multiple-access channel](image)

Figure 1. Multiple-Access Channel.
Throughout this note we shall assume that the channels are memoryless, i.e.

\[ P(y|x_1, x_2) = \prod_{i=1}^{n} P(y_i|x_{1i}, x_{2i}) \quad (1) \]

where

\[ y = (y_1, \ldots, y_n) \]
\[ x_1 = (x_{11}, \ldots, x_{1n}) \]
\[ x_2 = (x_{21}, \ldots, x_{2n}) \]

For simplicity, we shall deal with only two senders. The sources will be considered to be independent. (See [2] for an investigation of the savings that can be achieved when the sources are dependent.) Recently Ahlswede [3], Liao [4], and Slepian and Wolf [2] have determined the capacity region \( C \) of a two-input single output discrete memoryless channel: \( C = \text{closure of the convex hull of the union, over all input distributions } p_{X_1, X_2}(\cdot, \cdot) \) with independent \( X_1, X_2 \), of the sets of rate pairs \( R = (R_1, R_2) \) satisfying

\[ 0 \leq R_1 \leq I(X_1; Y | X_2) \quad (2a) \]
\[ 0 \leq R_2 \leq I(X_2; Y | X_1) \quad (2b) \]
\[ 0 \leq R_1 + R_2 \leq I(X_1, X_2; Y) \quad (2c) \]

The AWGN multiple-access channel is the most commonly encountered continuous alphabet channel \( (X_1 = X_2 = Y = R) \). The output signal \( Y \) is the sum \( X_1 + X_2 + Z \) where \( X_1 \) and \( X_2 \) are the input signals and \( Z \) is a zero-mean Gaussian (noise) random variable independent of \( X_1 \) and \( X_2 \) with variance \( E Z^2 = N \). There are average power constraints \( P_1 \) and \( P_2 \) on the inputs which require that the encoded messages \( X_1 \) and \( X_2 \) (which have \( n \) components) satisfy
\[
\frac{1}{n} \sum_{t=1}^{n} (x_i(t))^2 \leq P_i, \quad i = 1, 2
\]  

(3)

The capacity region \( \mathcal{C} \) for the AWGN multiple-access channel has been determined by Wyner [5] and Cover [6]:

\( \mathcal{C} = \) set of all rate pairs \( R = (R_1, R_2) \) satisfying

\[
R_1 \leq \frac{1}{2} \log \left(1 + \frac{P_1}{N}\right)
\]

(4a)

\[
R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_2}{N}\right)
\]

(4b)

\[
R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_1+P_2}{N}\right)
\]

(4c)

III Feedback Channel

The memoryless multiple-access channel with feedback is depicted in figure 2 below.

![Diagram](image)

**Figure 2.** Multiple-Access Channel with Feedback.
The $i^{th}$ symbol $x_{1i}$ sent by transmitter 1 depends upon the message index $k \in \{1,2,\ldots,2^nR\}$ that transmitter 1 wishes to send and upon the previous receiver symbols $y_1, y_2, \ldots, y_{i-1}$. A similar statement holds for the $i^{th}$ symbol $x_{2i}$ sent by the second transmitter.

IV The AWGN Multiple-Access Channel

For illustration, we will analyze in detail the symmetric case in which the average power constraints on both inputs are the same, namely $P_1 = P_2 = P$. We will then show how to generalize the scheme to the asymmetric case.

![AWGN Multiple-Access Channel with Feedback](image)

Figure 3. AWGN Multiple-Access Channel with Feedback.

The basic idea is for each transmitter to send at full power to the receiver, but at rates corresponding to the capacity of the channel when the transmitter's own signal is subtracted out. Thus each transmitter will know very quickly what the other transmitter is sending. However,
the receiver \( y \) is still confused because the total rate has exceeded his capacity. In stage 2, the transmitters now use coherent transmission to send the missing bits in the receiver's knowledge of the intended messages.

The model for the AWGN multiple-access channel with feedback is shown in figure 3. Our proof that feedback can enlarge the achievable rate region makes use of the usual random coding argument. Consider the ensemble of randomly generated codes of blocklength \( n \) and rates \( R_{12} = R_{21} = R^* \) obtained as follows. Each codeword \( \tilde{x}_1 \) is an \( n \)-sequence of independent outcomes of a zero-mean Gaussian random variable with variance \( P' = P - \eta \) where \( \eta > 0 \) is a quantity which will be made to tend to zero. We generate \( 2^{nR_{12}} \) independent such codewords. This will be the code used by the first transmitter during stage 1. Similarly the code used by the second transmitter during stage 1 is obtained by generating \( 2^{nR_{21}} \) independent codewords \( \tilde{x}_2 \) as above.

Let us suppose that we wish to send one of \( 2^{nR^*} \) equiprobable messages to \( Y \) from each source. Because of the symmetry induced by the random coding, we can assume the messages actually transmitted to be \((1,1)\). From the results on coding for single-input single-output AWGN channels, we know that at the end of stage 1, \( \tilde{x}_1 \) can be guessed at the second transmitter with arbitrarily small probability of error, say \( P_{e_{1,2}} < \varepsilon / 5 \), \( \varepsilon > 0 \) if

\[
R_{12} \leq \frac{1}{2} \log \left( 1 + \frac{P_1}{N} \right) \text{ bits/transmission} \tag{5}
\]

and the blocklength \( n \) is sufficiently large. Similarly, \( \tilde{x}_2 \) can be estimated at the first transmitter with arbitrarily small probability of error, say \( P_{e_{1,1}} < \varepsilon / 5 \), \( \varepsilon > 0 \) if

-6-
\[ R_{21} \leq \frac{1}{2} \log \left( 1 + \frac{P'}{N} \right) \text{ bits/transmission} \] (6)

and \( n \) is sufficiently large. In particular, we can set \( R_{12} = R_{21} = R^* = \frac{1}{2} \log \left( 1 + \frac{P'}{N} \right) \). We note that stage 1 requires \( n \) transmissions.

Let us denote the \( n \)-sequence received at \( Y \) during stage 1 by \( \tilde{y} \). In the following, we shall be concerned with the set of \( 2^{nR^*} \) codewords obtained by taking all possible sums \( \tilde{x} = x_1 + x_2 \). By the independence of \( x_1 \) and \( x_2 \), each component of \( \tilde{x} \) is a zero-mean Gaussian r.v. with variance \( 2P' \). We shall say that a codeword \( \tilde{x} \) is linked to \( \tilde{y} \) if \( \tilde{x} \) lies within the \( n \)-sphere of radius \( \sqrt{n(N+\epsilon)} \) centered at \( \tilde{y} \). By the law of large numbers, we know that with probability \( P_c = 1 - \frac{\epsilon}{5} \), \( \epsilon > 0 \) the correct codeword will be linked to \( \tilde{y} \).

We now proceed to consider the set \( S_y \) of codewords which, at the end of stage 1, are linked to \( \tilde{y} \). Let \( |S_y| \) denote the cardinality of \( S_y \). Then the average \( |S_y| \), taken over the ensemble of random codes and possible input messages is \( E |S_y| = L_1 + L_2 + 1 \), (see Eq. (40) for a full explanation in a general context), where

\[ L_1 = \text{expected number of codewords linked to the received sequence when a code with } 2(2^nR^*-1) \text{ independent codewords whose } n \text{ components are generated independently according to a zero-mean Gaussian random variable with variance } P' \text{ is used over an AWGN channel with noise variance } N. \]

\[ L_2 = \text{expected number of codewords linked to the received sequence when a code with } (2^nR^*-1)^2 \text{ independent codewords whose } n \text{ components are independent zero-mean Gaussian r.v. with variance } 2P' \text{ is used over the same channel.} \]

From Shannon [7] we have that for sufficiently large \( n \), \( \forall \epsilon > 0 \),

-7-
\[ L_1 \leq 2(2^n R^* - 1)^2 - n \left( \frac{1}{2} \log \left( 1 + \frac{p'_i}{N} \right) - \varepsilon \right) \]  

\[ L_2 \leq (2^n R^* - 1)^2 - n \left( \frac{1}{2} \log \left( 1 + \frac{2p'_i}{N} \right) - \varepsilon \right) \]  

So,

\[ L_1 + L_2 < 2^{n R^*} \left( \frac{-n}{2} \log \left( 1 + \frac{p'_i}{N} \right) + 2^{n R^*} \frac{n}{2} \log \left( 1 + \frac{2p'_i}{N} \right) \right) 2^{n \varepsilon} \]

\[ \forall \varepsilon > 0 \]

\[ = \left( \frac{n}{2} \log \left( 1 + \frac{p'_i}{N} \right) - \frac{1}{2} \log \left( 1 + \frac{2p'_i}{N} \right) \right) 2^{n \varepsilon} \]

\[ \leq 2 \left( \frac{n}{2} \log \left( 1 + \frac{p'_i}{N} \right) - \frac{1}{2} \log \left( 1 + \frac{2p'_i}{N} \right) + \varepsilon' \right) \]

\[ \Delta \]

\[ = K, \varepsilon' > \varepsilon \]  

Using Markov's inequality, we obtain

\[ P_b \leq \Pr \left\{ \left| S_Y \right| > K2^n \varepsilon_1 \right\} \leq 2^{-n \varepsilon_1}, \varepsilon_1 > 0 \]

\[ < \varepsilon/5, \forall \varepsilon > 0 \]  

Note by inspection of Eq. (11) that in stage 1 the uncertainty of the receiver is reduced by \( \frac{n}{2} \log \left( 1 + \frac{2p'_i}{N} \right) \) bits, precisely that which could be obtained (without feedback) if 1 and 2 were actually trying to communicate with \( y \) rather than with each other.

Let \( P_F = \Pr \left\{ \| x_1(1) \|^2 > n\varepsilon \text{ or } \| x_2(1) \|^2 > n\varepsilon \right\} \)

where \( \| x_i(1) \|^2 = \frac{n}{2} \sum_{j=1}^{i} x_{ij}^2(1), i = 1, 2. \)
Now, \( \frac{1}{n} \sum_{j=1}^{n} x_{ij}^2(1) \) is the arithmetic average of \( n \) independent identically distributed random variables with expected value \( \mathbb{E}(x_{ij}^2(1)) = P'< P \).

By the law of large numbers, we know that \( \Pr \left\{ \|x_1(1)\|^2 > nP, \ i = 1,2 \right\} \) can be made arbitrarily small. By the union bound, we can let \( P_F < \frac{\epsilon}{5} \).

Define \( E_1 \) as the event that stage 1 is successful, i.e.,

1. \( \|x_1(1)\|^2 \leq nP \) and \( \|x_2(1)\|^2 \leq nP \).

2. \( x_1 \) and \( x_2 \) are correctly decoded at transmitters 2 and 1 respectively.

3. the correct codeword \( x(1) \) is linked to the received sequence \( y \).

4. \( |S_y| < K2^{n\epsilon_1} \).

Let \( E_{1}^c \) denote the complement of this event. Then, by the union bound,

\[
\Pr(E_{1}^c) \leq P_{e_{1,2}} + P_{e_{1,1}} + (1-P_c) + P_b + P_F < \epsilon
\]

(15)

where \( \Pr(\cdot) \) denotes expectation over the choice of codebooks and possible input messages.

Therefore, there exists at least one code which satisfies the power constraints and has an average probability of failure in stage 1 less than \( \epsilon \). Since we can choose \( n \) to be arbitrarily small, the limiting value of \( R^* \) is \( \frac{1}{2} \log (1 + \frac{P}{N}) \).

Assuming the first stage is successful, in stage 2 we have to send at most \( n[\log (1 + \frac{P}{N}) - \frac{1}{2} \log (1 + \frac{2P}{N}) + \epsilon] \) bits to the receiver in order to specify completely the correct codeword. But in stage 2, the two transmitters can cooperate totally since they both know the correct
They both send the same signal. Thus, the signals add coherently at the receiver for stage 2. Because of the additive nature of the channel, total cooperation between the two senders allows reliable transmission (i.e., with a probability of error $P_{e_2} < \varepsilon$) up to a rate of 
\[
\frac{1}{2} \log (1 + \frac{4P}{N}) \text{ bits/transmission i.e., the effective power is now} 
\]
\[
(\sqrt{P} + \sqrt{P})^2 = 4P \text{ instead of } P + P = 2P.
\]

We define the probability of failure $P_f$ as the probability that at the end of stage 2, the receiver does not correctly estimate the messages sent from the 2 transmitters. Thus $P_f \leq \Pr(E_1^c) + P_{e_2} < 2\varepsilon$, which means the proposed scheme allows reliable transmission.

We now calculate the effective achievable rate. The number of transmissions required for stage 2 is

\[
\frac{n \left[ \log (1 + \frac{P}{N}) - \frac{1}{2} \log (1 + \frac{2P}{N}) + \varepsilon \right]}{\frac{1}{2} \log (1 + \frac{4P}{N})} \quad (16)
\]

We recall that stage 1 requires $n$ transmissions and that the total amount of information conveyed from the two senders to the receiver is $n \log (1 + \frac{P}{N})$ bits. So the overall effective rate $R_1 + R_2$ is

\[
\frac{\log (1 + \gamma) \log (1 + 4\gamma)}{\log (1 + 4\gamma) + 2 \log (1 + \gamma) - \log (1 + 2\gamma)} \quad \text{bits/transmission} \quad (17)
\]

where $\gamma \triangleq \frac{P}{N}$.
Figure 3 shows the point, *, which can be achieved by the proposed scheme for $\gamma = 5$.

![Graph showing capacity region with no feedback and total cooperation line.](image)

Figure 4. Capacity Region of AWGN Multiple-Access Channel with $\gamma = 5$. Rates are in nats/transmission.

In the asymmetric case, suppose $P_1 > P_2$. We would like the two transmitters to learn each other's message after an equal number of transmissions so that they are ready to cooperate at the same time. So we choose to transmit $nR_1$ and $nR_2$ bits to the receiver from transmitters 1 and 2 respectively where $R_1 = \frac{1}{2} \log \left(1 + \frac{P_1}{N}\right)$ and $R_2 = \frac{1}{2} \log \left(1 + \frac{P_2}{N}\right)$. It is not difficult to see that this scheme will yield a point outside the capacity region with no feedback.

V Jointly Typical Sequences

In this section we recall some basic results concerning typical
sequences which will be used to find out when feedback can increase the capacity region of discrete memoryless multiple-access channels.

Let \( \{X^{(1)}, X^{(2)}, \ldots, X^{(k)}\} \) denote a finite collection of discrete random variables with some fixed joint distribution \( p(x^{(1)}, x^{(2)}, \ldots, x^{(k)}) \). Let \( S \) denote an ordered subset of these random variables and consider \( n \) independent copies of \( S \). Thus,

\[
\Pr[S = s] = \prod_{i=1}^{n} \Pr[S_i = s_i] \quad (18)
\]

For example, if \( S = (X^{(j)}, X^{(k)}) \), then

\[
\Pr[S = s] = \Pr\left(\left(\tilde{X}^{(j)}, \tilde{X}^{(k)}\right) = \left(X^{(j)}, X^{(k)}\right)\right) = \prod_{i=1}^{n} p\left(X^{(j)}, X^{(k)}\right) \quad (19)
\]

**Definition:** The set \( A_\varepsilon \) of jointly \( \varepsilon \)-typical \( n \)-sequences \( \left(\tilde{X}^{(1)}, \tilde{X}^{(2)}, \ldots, \tilde{X}^{(k)}\right) \) is defined by

\[
A_\varepsilon\left(X^{(1)}, X^{(2)}, \ldots, X^{(k)}\right) = \left\{\left(\tilde{X}^{(1)}, \tilde{X}^{(2)}, \ldots, \tilde{X}^{(k)}\right) \in \left(\tilde{X}(1)\right)^n \times \left(\tilde{X}(2)\right)^n \times \ldots \times \left(\tilde{X}(k)\right)^n : -\frac{1}{n}\log p\left(S\right) - H(S) \leq \varepsilon, \quad \forall S \subseteq \left\{X^{(1)}, X^{(2)}, \ldots, X^{(k)}\right\}\right\} \quad (20)
\]

where \( \tilde{S} \) denotes the ordered set of sequences in \( \left\{\tilde{X}^{(1)}, \ldots, \tilde{X}^{(k)}\right\} \) corresponding to \( S \). Let \( A_\varepsilon(S) \) denote the restriction of \( A_\varepsilon \) to the
coordinates corresponding to $S$. Thus, for example, for $S = (x^{(1)}, x^{(3)})$,

$$A_{\varepsilon}(x^{(1)}, x^{(3)}) = \left\{ (x^{(1)}, x^{(3)}) : \left| -\frac{1}{n} \log p(x^{(1)}, x^{(3)}) - H(x^{(1)}, x^{(3)}) \right| < \varepsilon, \right. \left. \left| -\frac{1}{n} \log p(x^{(1)}) - H(x^{(1)}) \right| < \varepsilon, \right. \left. \left| -\frac{1}{n} \log p(x^{(3)}) - H(x^{(3)}) \right| < \varepsilon \right\} \tag{21}$$

We now recall three basic lemmas. For a proof of these lemmas, see Forney [8] and Cover [9].

**Lemma 1:** For any $\varepsilon > 0$, there exists an integer $n$ such that $A_{\varepsilon}(S)$ satisfies, for all $S \subseteq \{x^{(1)}, \ldots, x^{(k)}\}$,

(i) $Pr(A_{\varepsilon}(S)) \geq 1 - \varepsilon,$

(ii) $s \in A_{\varepsilon}(S) \Rightarrow \left| -\frac{1}{n} \log p(s) - H(S) \right| < \varepsilon \tag{22}$

(iii) $(1-\varepsilon)2^{n(H(S)-\varepsilon)} \leq |A_{\varepsilon}(S)| \leq 2^{n(H(S)+\varepsilon)}$

**Lemma 2:** Let the discrete random variables $X, Y$ have joint distribution $p(x, y)$. Let $X'$ and $Y'$ be independent with the marginals

$$p(x) = \sum_{y} p(x, y)$$

$$p(y) = \sum_{x} p(x, y)$$

Let $(X, Y) \sim \prod_{i=1}^{n} p(x_{i}, y_{i})$ and $(X', Y') \sim \prod_{i=1}^{n} p(x_{i})p(y_{i})$ where $X_{i}$ and $Y_{i}$, $1 \leq i \leq n$ are independent and identically distributed as $X$ and $Y$.
respectively. Then

$$\Pr\{(X', Y') \in A_{\epsilon} (X, Y)\} \leq 2^{-n[I(X; Y) - \epsilon]} \quad (23)$$

Lemma 3: Let the discrete random variables $X, Y, Z$ have joint distribution $p(x, y, z)$. Let $X', Y'$ be conditionally independent given $Z$, with the marginals

$$p(x|z) = \sum_y \frac{p(x, y, z)}{p(z)} \quad (24)$$

$$p(y|z) = \sum_x \frac{p(x, y, z)}{p(z)}$$

Let $(X, Y, Z) \sim \prod_{i=1}^{n} p(x_i, y_i, z_i)$ and

$$(X', Y', Z) \sim \prod_{i=1}^{n} p(x_i|z_i)p(y_i|z_i)p(z_i) \quad .$$

Then

$$\Pr\{(X', Y', Z) \in A_{\epsilon} (X, Y, Z)\} \leq 2^{-n[I(X; Y|Z) - \epsilon]} \quad (25)$$

VI The Discrete Memoryless Multiple-Access Channel

In this section, we shall find conditions under which the proposed scheme will increase the capacity region of discrete memoryless multiple-access channels. In particular we shall prove the following.
Theorem 1:

Let $Q_1$ and $Q_2$ denote probability distributions on $X_1$ and $X_2$ respectively. Suppose that $Q^* = (Q_1^*, Q_2^*)$ achieves the maximum of $I(X_1, X_2; Y)$ over all product distributions $Q_1(x_1)Q_2(x_2)$. Then the proposed feedback scheme allows reliable transmission at an overall effective rate of

$$R_1 + R_2 = \frac{AB}{A + B - C}$$

(26)

where

$$A = I_{Q^*}(X_1; Y|X_2) + I_{Q^*}(X_2; Y|X_1)$$

(27.a)

$$B = \max_{p(x_1, x_2)} I(X_1, X_2; Y)$$

(27.b)

and

$$C = I_{Q^*}(X_1, X_2; Y)$$

(27.c)

is the maximum non-feedback sum rate.

Remark: The maximum in the expression for $B$ is taken over all joint distributions of $X_1$ and $X_2$.

Corollary: Sufficient conditions under which feedback will increase the capacity region are $A, B > C$, i.e.,

(i) $I_{Q^*}(X_1; Y|X_2) + I_{Q^*}(X_2; Y|X_1) > I_{Q^*}(X_1, X_2; Y)$

(28)

(ii) $\max_{p(x_1, x_2)} I(X_1, X_2; Y) > I_{Q^*}(X_1, X_2; Y)$

(29)

Remark: Notice that (ii) is a necessary condition for feedback to increase the capacity region.
Proof of Corollary: We want to show that

\[
\frac{AB}{A+B-C} > C \quad (30)
\]

From equations (28) and (29), we have \( A > C \) and \( B > C \).

So let \( A = C + \alpha \) and \( B = C + \beta \) where \( \alpha, \beta > 0 \).

Now \( \alpha \beta > 0 \)

Therefore

\[
c^2 + \alpha c + c\beta + \alpha \beta > c^2 + \alpha c + \beta c
\]

i.e.

\[
\frac{(C + \alpha)(C + \beta)}{C + \alpha + \beta} > C \quad (34)
\]

i.e.

\[
\frac{AB}{A+B-C} > C \quad \text{Q.E.D.} \quad (35)
\]

We now proceed to prove Theorem 1.

Consider a randomly generated code of blocklength \( n \) and rates \( R_{12}, R_{21} \) obtained as follows. Each codeword \( X_i \), \( i = 1, 2 \) is an \( n \)-sequence of independent outcomes of a random variable distributed according to \( Q_i \). We generate \( 2^{nR_{12}} \) independent \( X_1 \)'s and \( 2^{nR_{21}} \) independent \( X_2 \)'s. Let these be indexed as \( X_1(k) \), \( k \in [1, 2^{nR_{12}}] \) and \( X_2(\ell) \), \( \ell \in [1, 2^{nR_{21}}] \).

Let \( K, L \) be independent random variables drawn according to uniform distributions on \([1, 2^{nR_{12}}]\) and \([1, 2^{nR_{21}}]\) respectively. Let the code be chosen randomly from the above ensemble. As in section IV, we first proceed to upper bound the average probability \( \overline{Pr}[E_1^c] \) that stage 1 is
unsuccessful. By the symmetry induced by the random coding, we see that each transmitted message \((k, \ell)\) yields the same probability of error. So hence forth we shall assume that the actual transmitted message is \((1, 1)\).

The decoding rule for estimating the message of the first transmitter at the second transmitter (at the end of stage 1) is as follows. If \(\tilde{y}\) is received, declare \(\hat{k} = k\) was sent if and only if there is only one \(k \in \left\lfloor 1, 2^{nR_{12}} \right\rfloor\) such that \((x_1(k), x_2(1), \tilde{y})\) are jointly typical. Using Lemmas 1 and 3, it can be shown (for details see [9]) that \(k\) can be decoded with arbitrarily small probability of error, say,

\[ P_{e_{1,2}} < \frac{\varepsilon}{4} \text{ if } R_{12} < I_{Q^*}(X_1; Y | X_2) \tag{36} \]

and \(n\) is sufficiently large. Similarly \(\ell\) can be estimated at the first transmitter with arbitrarily small probability of error, say,

\[ P_{e_{1,1}} < \frac{\varepsilon}{4} \text{ if } R_{21} < I_{Q^*}(X_2; Y | X_1) \tag{37} \]

We now consider the set \(S_y\) of codewords which, at the end of stage 1, are jointly typical with \(\tilde{y}\). From Lemma 1, we know that \((x_1(1), x_2(1))\) will be jointly typical with \(\tilde{y}\) with high probability, say,

\[ P_c > 1 - \frac{\varepsilon}{4} \cdot \]
Let \( \psi_{k\ell}(\sim) = \begin{cases} 1, & (x_1(k), x_2(\ell), y) \text{ are typical} \\ 0, & \text{otherwise} \end{cases} \) (38)

Then \( |S_y| = \sum_{k,\ell} \psi_{k\ell}(\sim) \) (39)

\[
E|S_y| = \sum_{k=1,\ell=1}^{2} E\psi_{k\ell}(\sim) + \sum_{k\neq 1,\ell=1}^{2} E\psi_{k\ell}(\sim) + \sum_{k=1,\ell\neq 1}^{2} E\psi_{k\ell}(\sim) + \sum_{k\neq 1,\ell\neq 1}^{2} E\psi_{k\ell}(\sim)
\] (40)

Using Lemmas 2 and 3, it can be shown that

\[
E\psi_{k\ell}(\sim) \leq 2^{-n} \left[ I_{Q^*}(X_1; X_2; Y) - \varepsilon \right], \quad k\neq 1, \ell\neq 1
\] (41)

\[
E\psi_{1\ell}(\sim) \leq 2^{-n} \left[ I_{Q^*}(X_2; Y|X_1) - \varepsilon \right], \quad k=1, \ell\neq 1
\] (42)

\[
E\psi_{k1}(\sim) \leq 2^{-n} \left[ I_{Q^*}(X_1; Y|X_2) - \varepsilon \right], \quad k\neq 1, \ell=1
\] (43)

Therefore,

\[
E|S_y| \leq 1 + \binom{nR_{12}-1}{2} \binom{nR_{21}-1}{2} 2^{-n} \left[ I_{Q^*}(X_1; X_2; Y) - \varepsilon \right]
\]

\[
+ \binom{nR_{12}-1}{2} 2^{-n} \left[ I_{Q^*}(X_1; Y|X_2) - \varepsilon \right]
\]

\[
+ \binom{nR_{21}-1}{2} 2^{-n} \left[ I_{Q^*}(X_2; Y|X_1) - \varepsilon \right]
\] (44)
Taking the limit in Equations (36) and (37) we can set \( R_{12} = I_{Q^*}(X_1; Y|X_2) \) and \( R_{21} = I_{Q^*}(X_2; Y|X_1) \). Then (44) becomes

\[
\mathbb{E}[S_{Y^n}] < 1 + 2.2^n \epsilon + 2^n \left[ I_{Q^*}(X_1; Y|X_2) + I_{Q^*}(X_2; Y|X_1) - I_{Q^*}(X_1, X_2; Y) + \epsilon \right]
\]

\[
< 2^{n(D + \epsilon_1)}
\]  \hspace{1cm} (45)

where \( D = I_{Q^*}(X_1; Y|X_2) + I_{Q^*}(X_2; Y|X_1) - I_{Q^*}(X_1, X_2; Y) \)  \hspace{1cm} (47)

Using Markov's inequality we obtain

\[
P_b \overset{\triangle}{=} \text{Pr}\left\{ |S_{Y^n}| > 2^{n(D + \epsilon_1)} \right\} \leq 2^{-n\epsilon_2}
\]

\[
< \frac{\epsilon}{4} \quad \text{say}
\]  \hspace{1cm} (48)

From Equation (15)

\[
\text{Pr}\left\{ E^c_1 \right\} \leq P_{e_1,2} + P_{e_1,1} + (1 - P_c) + P_b \quad \text{say}
\]

\[
< \epsilon
\]  \hspace{1cm} (49)

Since \( \text{Pr}\left\{ E^c_1 \right\} < \epsilon \), there must exist at least one code with an average probability of failure (taken over \( K_L \)) in stage 1 less than \( \epsilon \).

Assuming the first stage is successful, in stage 2 we have to send at most \( n(D + \epsilon_1) \) bits to the receiver in order to completely specify the correct code word. In stage 2, total cooperation between the two transmitters allows transmission up to a rate of \( \max_{p(x_1, x_2)} I(X_1, X_2; Y) \) with arbitrary small probability of error, say \( P_{e_2} < \epsilon \). Thus the probability of failure \( P_f \leq \text{Pr}\left\{ E^c_1 \right\} + P_{e_2} < 2\epsilon \), which means that the scheme
allows reliable transmission. A straightforward calculation shows that
the overall effective sum rate using this scheme is

\[
R_1 + R_2 = \frac{\left[ I_{Q^*}(X_1; Y|X_2) + I_{Q^*}(X_2; Y|X_1) \right]}{\max_{p(x_1, x_2)} I(X_1, X_2; Y)}
\]

(51)

This completes the proof of Theorem 1.

VII Concluding Remarks

The capacity region of multiple-access channels with feedback can
generally be increased by the scheme of transmitting first at high rates
until both transmitters know each other's messages, then cooperatively at
a lower rate to resolve the remaining receiver ambiguity. An open problem
would be to determine the full capacity region with feedback and conditions,
if any, under which the above scheme is optimal.
References


