An Achievable Rate Region for the Multiple-Access Channel with Feedback

THOMAS M. COVER, FELLOW, IEEE, AND CYRIL S. K. LEUNG, MEMBER, IEEE

Abstract—An achievable rate region $R_1 \leq I(X_1; Y|X_2, U)$, $R_2 \leq I(X_2; Y|X_1, U)$, $R_1 + R_2 \leq I(X_1, X_2; Y)$, where $p(u, x_1, x_2, y) = p(u)p(x_1|u)p(x_2|u)p(y|x_1, x_2)$, **is established for the multiple-access channel with feedback. Time sharing of these achievable rates yields the rate region of this paper. This region generally exceeds the achievable rate region without feedback and exceeds the rate point found by Gaarder and Wolf for the binary erasure multiple-access channel with feedback. The presence of feedback allows the independent transmitters to understand each other's intended transmissions before the receiver has sufficient information to achieve the desired decoding. This allows the transmitters to cooperate in the transmission of information that resolves the residual uncertainty of the receiver. At the same time, independent information from the transmitters is superimposed on the cooperative correction information. The proof involves list codes and block Markov encoding.**

I. INTRODUCTION

GAARDER AND WOLF [1] have demonstrated that it is possible to increase the capacity region of a discrete memoryless multiple-access channel (Fig. 1) through the use of feedback.

The basic idea behind the scheme in [1] and the extension in [2] can be described in two stages. During 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 one, 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 lexicographic order. This sets up stage 2 during which the two transmitters cooperate to send the index of the actual transmitted message. It is shown in [2] that such a scheme generally enlarges the capacity region.

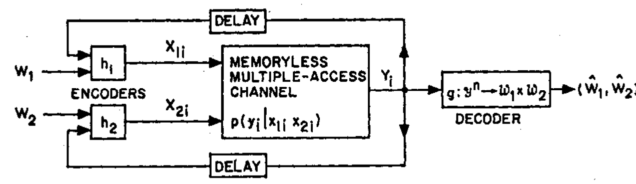


Fig. 1. Multiple-access channel with feedback.

In this paper we superimpose stage 1 and stage 2 as is done with degraded broadcast channels [3]. Before describing how this scheme works, we compare its performance with that of [1]. The noiseless binary erasure multiple-access channel shown in Fig. 2 will be used as an example. For this channel, Gaarder and Wolf show that a rate pair $(R_1, R_2) = (0.76, 0.76)$ is achievable where R_i, $i = 1,2$, refers to the transmission rate from the ith sender to the receiver in bits per transmission. Throughout this paper, information and entropy will be measured in bits. Their method is as follows. For the first m transmissions, send m independent bits from each transmitter. By the structure of the channel it can be seen that approximately $m/2$ transmissions will be perceived perfectly by the receiver Y, and the remaining $m/2$ will be ambiguous. We use the next n transmissions to transmit the correct X_1 values for the ambiguous transmissions at the cooperative channel capacity of $\log_2 3$ bits/transmission. Thus $n = (m/2)/\log_2 3$. Hence the total average rate sum is given by $R_1 = R_2$, $R_1 + R_2 = (m + m)/(m + m/2 \log_2 3)) = 2/(1 + 1/\log 9) = 1.52037$.

We shall show that for this same channel, the new scheme can achieve $R_1 = R_2 = 0.79113 \cdots$. An upper bound is $R_1 = R_2 = (1/2)\log 3 = 0.79248 \cdots$, which is achieved by total cooperation. The achievable points and upper bound are shown in Fig. 3. It can be seen that the performance of the new scheme is very close to the upperbound. This improvement over the two-stage scheme is perhaps surprising since the noiseless binary erasure multiple-access channel seems well suited to the scheme in [1].

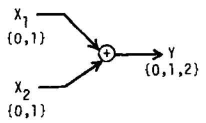
In the next section we review the capacity region of memoryless multiple-access channels, introduce the channel with feedback, and state our main result. Section III contains a digression on typical sequences. We recall some

Manuscript received April 13, 1977; revised August 1, 1980. This work was supported in part by the National Science Foundation under Grant ECS-78-23334, in part by Stanford Research Institute ARPA under Contract DAHC-15-C-0187, and in part by the Natural Sciences and Engineering Research Council of Canada under Grant A1052. This paper was presented at the IEEE International Symposium on Information Theory, Cornell University, Ithaca, NY, September 1977. (A revision of this paper was allowed to lapse and the paper was resubmitted to this TRANSACTIONS in October 1979.)

T. M. Cover is with the Departments of Electrical Engineering and Statistics, Stanford University, Stanford, CA 94305.

C. S. K. Leung is with the Department of Electrical Engineering, University of British Columbia, Vancouver, BC, Canada.

0018-9448/81/0500-0292$00.75



(a)

X_1	X_2	Y
0	0	0
0	1	1
1	0	1
1	1	2

} "erasures"

(b)

Fig. 2. (a) Noiseless binary erasure multiple-access channel. (b) Channel input output relationship.

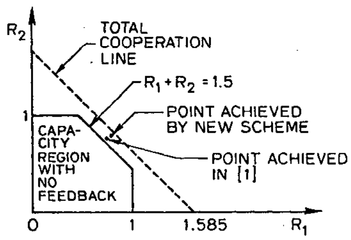


Fig. 3. Achievable points and upperbound for the noiseless binary erasure multiple-access channel.

basic results that will be needed to establish an achievable rate region in Section IV. Section V discusses the cardinality bound on an auxiliary random variable U used in defining the rate region. Finally the additive white Gaussian noise (AWGN) multiple access channel is treated in Section VI.

II. Preliminaries and Main Result

In a multiple-access channel, several transmitters communicate with a single receiver. The channel is characterized by its input terminals with alphabets $\mathcal{X}_1, \mathcal{X}_2, \cdots, \mathcal{X}_M$, its output terminal with alphabet $\mathcal{Y}$, and a set of conditional probability measures on the output signal Y given the input signals $X_1, X_2, \cdots, X_M$ (see Fig. 4). We shall deal with only two senders.

Definitions: A (two-user) memoryless *multiple-access channel* is denoted by $\{(\mathcal{X}_1,\mathcal{X}_2),\mathcal{Y}, p(y|x_1,x_2)\}$, where $\mathcal{X}_1$ and $\mathcal{X}_2$ are the input alphabets, $\mathcal{Y}$ is the output alphabet, and $p(\cdot|x_1,x_2)$ is a collection of probability mass functions on $\mathcal{Y}$ indexed by the input symbols $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$. The channel is said to be *discrete* if $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}$ are finite sets.

The channel is *memoryless* if

$$P(y_k|x_{11},x_{12},\cdots,x_{1k},x_{21},x_{22},\cdots,x_{2k},y_1,y_2,\cdots,y_{k-1}) = p(y_k|x_{1k},x_{2k}),$$

where x_{1j}, x_{2j} and y_j denote the inputs and outputs of the channel at time j.

The feedback does not enter into the definition until a code is defined. An $((M_1,M_2),n)$ *code* for the *multiple-access channel with feedback* is given by the following:

i) a collection of encoding functions x_s: $\{1,2,\cdots,M_s\} \times \mathcal{Y}^n \to \mathcal{X}_s^n$, $s=1,2$, where the kth coordinate of $\boldsymbol{x}_s \in \mathcal{X}_s^n$ is given by the function

$$x_{sk}(w_s,(y_1,y_2,\cdots,y_{k-1})), \quad k=1,2,\cdots,n,\ w_s=1,2,\cdots,M_s.$$

ii) a decoding function g: $\mathcal{Y}^n \to \{1,2,\cdots,M_1\} \times \{1,2,\cdots,M_2\}$. We shall write $g(y_1,\cdots,y_n)=(\hat{w}_1,\hat{w}_2)$.

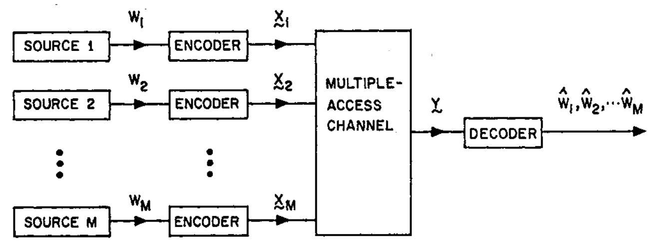


Fig. 4. Multiple-access channel.

That is, each sender s sends a sequence of symbols depending on his desired message w_s and the *past* symbols $y_1, y_2, \cdots, y_{k-1}$ fed back from the receiver.

We shall use the average probability of error criterion defined as the probability of error under the assumption that the messages are random variables W_1 and W_2 drawn according to a uniform distribution over $\{1,2,\cdots,M_1\} \times \{1,2,\cdots,M_2\}$. Thus we define the *error probability* for the code to be

$$P_n = P\{(W_1,W_2) \neq g(Y_1,\cdots,Y_n)\} = \frac{1}{M_1M_2}\sum_{w_1=1,w_2=1}^{M_1,M_2} P\{g(\boldsymbol{Y}) \neq (w_1,w_2)|W_1=w_1,W_2=w_2\}. \tag{1}$$

We are interested in the rates at which asymptotically error free transmission can take place. Hence we shall say that the rate pair (R_1,R_2) is *achievable* for the memoryless multiple-access channel with feedback if there exists a sequence of $((M_1,M_2),n)$ codes with

$$R_1 \le \frac{1}{n}\log M_1, \quad R_2 \le \frac{1}{n}\log M_2,$$

such that $P_n \to 0$ as $n \to \infty$. The *capacity region* is the closure of the set of all achievable rates.

Ahlswede [5], and Liao [6], (see also Slepian and Wolf [4]) have determined the capacity region R^* of a two-input single output discrete memoryless channel with no feedback: R^* is the 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 $\boldsymbol{R}=(R_1,R_2)$ satisfying

$$R_1 < I(X_1;Y|X_2), \tag{2a}$$

$$R_2 < I(X_2;Y|X_1), \tag{2b}$$

$$R_1+R_2 < I(X_1,X_2;Y). \tag{2c}$$

The following theorem is the primary result of this paper.

Theorem 1: Let U be a discrete random variable which takes on values in the set $\mathcal{U}=\{1,2,\cdots,m\}$, where $m=\min\{\|\mathcal{X}_1\|\cdot\|\mathcal{X}_2\|,\|\mathcal{Y}\|\}$ and $\|\mathcal{X}_i\|$ denotes the alphabet size of user i. Consider the set $\mathcal{P}$ of all joint distributions of the form

$$P_{UX_1X_2Y}(u,x_1,x_2,y) = P_U(u)P_{X_1|U}(x_1|u) \cdot P_{X_2|U}(x_2|u)P_{Y|X_1X_2}(y|x_1,x_2), \tag{3}$$

where $P_{Y|X_1X_2}(y|x_1,x_2)$ is fixed by the channel. For each

$P \in \mathcal{P}$, denote by $\mathcal{R}(P)$ the set of all rate pairs (R_1, R_2) satisfying

$$R_1 < I(X_1; Y | X_2, U), \tag{4a}$$

$$R_2 < I(X_2; Y | X_1, U), \tag{4b}$$

$$R_1 + R_2 < I(X_1, X_2; Y), \tag{4c}$$

where the mutual informations are computed for P. Then the achievable rate region $\mathcal{R}^*$ is given by the closure of the convex hull of $\cup \mathcal{R}(P)$, where the union is over all $P \in \mathcal{P}$.

Using Theorem 1, we observe[1] that for the noiseless binary erasure multiple-access channel of Fig. 2, the rate pair $(R_1, R_2) = (0.79113, 0.79113)$ can be achieved by letting $\Pr\{U=0\} = \Pr\{U=1\} = 1/2$, and letting X_1 and X_2 be conditionally independent given $U=u$ with $\Pr\{X_i \neq u\} = 0.237663$, $i = 1,2$.

III. Jointly Typical Sequences

In this section we recall some basic results concerning typical sequences which will be used to establish Theorem 1. See Wolfowitz [11] for the first systematic exploitation of typical sequences in formal capacity proofs.

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

$$\Pr\{\mathbf{S} = \mathbf{s}\} = \prod_{i=1}^{n} \Pr\{S_i = s_i\}. \tag{5}$$

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

$$\Pr\{\mathbf{S} = \mathbf{s}\} = \Pr\left\{ (\mathbf{X}^{(j)}, \mathbf{X}^{(k)}) = (\mathbf{x}^{(j)}, \mathbf{x}^{(k)}) \right\} \tag{6}$$

$$= \prod_{i=1}^{n} p\left(x_i^{(j)}, x_i^{(k)}\right).$$

Definition: The set A_ϵ of *jointly ϵ-typical* n-sequences $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)})$ is defined by

$$A_\epsilon(X^{(1)}, X^{(2)}, \cdots, X^{(k)}) = \Big\{ (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}) \in (\mathcal{X}^{(1)})^n \times (\mathcal{X}^{(2)})^n \times \cdots \times (\mathcal{X}^{(k)})^n : \left| -\frac{1}{n} \log p(\mathbf{s}) - H(S) \right| < \epsilon, \ \forall S \subseteq \{X^{(1)}, X^{(2)}, \cdots, X^{(k)}\} \Big\}, \tag{7}$$

where $\mathbf{s}$ denotes the ordered set of sequences in $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(k)}$ corresponding to S. Let $A_\epsilon(S)$ denote the restriction of A_ϵ to the coordinates corresponding to S. Thus, for example, for $S = \{X^{(1)}, X^{(3)}\}$,

$$A_\epsilon(X^{(1)}, X^{(3)}) = \Big\{ (\mathbf{x}^{(1)}, \mathbf{x}^{(3)}) : \left| -\frac{1}{n} \log p(\mathbf{x}^{(1)}, \mathbf{x}^{(3)}) - H(X^{(1)}, X^{(3)}) \right| < \epsilon, \left| -\frac{1}{n} \log p(\mathbf{x}^{(1)}) - H(X^{(1)}) \right| < \epsilon, \left| -\frac{1}{n} \log p(\mathbf{x}^{(3)}) - H(X^{(3)}) \right| < \epsilon \Big\}. \tag{8}$$

We now recall three basic lemmas. For a proof of these lemmas, see Forney [9] and Cover [10].

Lemma 1: For any $\epsilon > 0$, there exists an integer n_0, such that for all $n \geq n_0$, $A_\epsilon(S)$ satisfies, for all $S \subseteq \{X^{(1)}, \cdots, X^{(k)}\}$:

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

ii) $\mathbf{s} \in A_\epsilon(S) \Rightarrow \left| -\frac{1}{n} \log p(\mathbf{s}) - H(S) \right| < \epsilon$,

iii) $(1-\epsilon) 2^{n(H(S)-\epsilon)} \leq \| A_\epsilon(S) \| \leq 2^{n(H(S)+\epsilon)}$. (9)

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 $(\mathbf{X}, \mathbf{Y}) \sim \prod_{i=1}^n p(x_i, y_i)$ and $(\mathbf{X}', \mathbf{Y}') \sim \prod_{i=1}^n p(x_i) p(y_i)$ where X_i and Y_i, $1 \leq i \leq n$ are independent and identically distributed as X and Y, respectively. Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}') \in A_\epsilon(X, Y)\} \leq 2^{-n[I(X;Y) - \epsilon]}. \tag{10}$$

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

$$p(x|z) = \sum_y p(x, y, z)/p(z),$$

$$p(y|z) = \sum_x p(x, y, z)/p(z). \tag{11}$$

Let

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

and

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

Then

$$\Pr\{(\mathbf{X}', \mathbf{Y}', \mathbf{Z}) \in A_\epsilon(X, Y, Z)\} \leq 2^{-n[I(X;Y|Z) - 7\epsilon]}. \tag{12}$$

[1]These calculations were contributed by E. C. van der Meulen.

IV. Proof of an Achievable Rate Region for the Two-Sender Multiple-Access Channel with Feedback

In this section, we describe a coding scheme for the multiple-access channel with feedback and show that it achieves the rate region defined in Theorem 1. This scheme is related to that in [1], but incorporates two additional features, namely the idea of superposition of information and a block Markov stationary coding structure. The scheme uses a large number B of blocks, each of length n. In block b, $1<b<B$, the transmitters send enough information to the receiver to enable him to resolve any uncertainty left over from block $b-1$. Superimposed on this information is some new independent information which each transmitter wishes to convey to the receiver. The rate of this new information is small enough so that each transmitter can reliably determine the other's message through the feedback links. Although the transmission rate of the new information will be too high for reliable transmission to the receiver, the receiver will be able to restrict the set of possible transmitted messages to a considerably smaller list of "typical" messages that with high probability will contain the correct message W. All the receiver needs in order to learn W is the index of W in some prearranged ordering of the typical set.

At the end of block b, the receiver learns block $b-1$ reliably and has some uncertainty (to be resolved in block $b+1$) about the new superimposed information. Also at the end of block b, each transmitter has complete knowledge of the other's message, thus enabling the transmitters to cooperate in resolving the receiver's residual uncertainty about block b in block $b+1$. This scheme results in a "steady-state" noncooperative infusion of new information and a cooperative resolution of the residual uncertainty.

We now use a random coding technique to prove Theorem 1.

Random Coding: First fix a choice of $P_U(u)$, $P_{X_1|U}(x_1|u)$, $P_{X_2|U}(x_2|u)$. Generate a sequence of 2^{nR_0} independent identically distributed random vectors $\boldsymbol{u}=(u_1,u_2,\cdots,u_n)$ according to $P_U(u_1,u_2,\cdots,u_n)=\prod_{i=1}^n P_U(u_i)$. Index these vectors as $\boldsymbol{u}(j)$, $j\in[1,2^{nR_0}]$. For each $\boldsymbol{u}(j)$, we generate 2^{nR_1} conditionally independent n-sequences $\boldsymbol{x}_1^j(k)$, $k\in[1,2^{nR_1}]$ and 2^{nR_2} conditionally independent n-sequences $\boldsymbol{x}_2^j(l)$, $l\in[1,2^{nR_2}]$. The components of $\boldsymbol{x}_s^j(\cdot)$, $s=1,2$, are generated independently according to $P_{X_s|U}(x_i|u_i(j))$. Thus for example, the n-sequence $\boldsymbol{x}_1^j(k)$ has probability

$$\Pr\{\boldsymbol{X}_1(k)=\boldsymbol{x}_1|\boldsymbol{U}(j)=\boldsymbol{u}\}=\prod_{i=1}^{n}P(x_{1i}|u_i(j)), \quad (13)$$

where $u_i(j)$ is the ith component of $\boldsymbol{u}(j)$.

It is intended that the "cloud center" $\boldsymbol{u}(j)$ will be correctly received by Y during the block in which it is sent. The "satellite" indices k and l will be decoded correctly by senders 2 and 1, but only partially understood by the receiver.

In the first block no cooperative information is sent: the transmitters and receiver use a predetermined index j' and encode $k\in[1,2^{nR_1}]$ and $l\in[1,2^{nR_2}]$ into $\boldsymbol{x}_1^{j'}(k)$ and $\boldsymbol{x}_2^{j'}(l)$, respectively. In the last block the transmitters send no new information and the decoder receives enough information to resolve the residual uncertainty about the penultimate block. We note that if the number of blocks B used in the scheme is large, the effective transmission rates over B blocks will be only negligibly affected by the rates in blocks 1 and B.

Encoding: Suppose that j^* is the index which is to be sent to the receiver (in block b) in order to resolve his residual uncertainty about the new messages (sent by the two transmitters) in block $b-1$. We will shortly derive an upperbound on the range of the index j^*. Also, let us denote the two new indices (or messages) to be sent in block b by (k^*,l^*) where $k^*\in[1,2^{nR_1}]$ and $l^*\in[1,2^{nR_2}]$. Then in block b, the first encoder sends $\boldsymbol{x}_1^{j^*}(k^*)$ and the second encoder sends $\boldsymbol{x}_2^{j^*}(l^*)$. Let the n-sequence obtained by the receiver be $\boldsymbol{y}$.

Decoding: The decoding procedures at the end of block b at the receiver and the two transmitters are as follows.

1) The receiver declares $\hat{j}=j$ was sent if and only if there is a unique j such that $(\boldsymbol{u}(j),\boldsymbol{y})$ are jointly typical. From Lemmas 1 and 2, we know that j^* can be decoded by receiver Y with arbitrarily small probability of error $P_{e_1}<\epsilon/5B$ if j^* takes on less than $2^{n(I(U;Y)-\epsilon)}$ values and n is sufficiently large.

2) Transmitter 1 estimates the message of the second transmitter by declaring $\hat{l}=l$ if and only if there is a unique l such that $(x_1^{j^*}(k^*),x_2^{j^*}(l),\boldsymbol{y})$ are jointly typical. Using Lemmas 1 and 3, it can be shown (for details see [10]) that $\hat{l}=l^*$ with arbitrarily small probability of error $P_{e_{2,1}}<\epsilon/5B$ if

$$R_2<I(X_2;Y|X_1,U), \quad (14)$$

and n is sufficiently large.

3) Similarly, the second transmitter can guess k^* with arbitrarily small probability of error $P_{e_{1,2}}<\epsilon/5B$ if

$$R_1<I(X_1;Y|X_2,U). \quad (15)$$

Cardinality of S_y: At the end of block b, the two transmitters know each other's message (with small probability of error). We now consider the set S_y of codewords which at the end of block b are jointly typical with $\boldsymbol{y}$. From Lemma 1, we know that $(x_1^{j^*}(k^*),x_2^{j^*}(l^*))$ will be jointly typical with $\boldsymbol{y}$ with high probability, say $P_c>1-\epsilon/5B$. Let

$$\Psi_{kl}(y)=\begin{cases}1, & \left(\boldsymbol{x}_1^{j^*}(k),x_2^{j^*}(l),\boldsymbol{y}\right) \text{ are typical,}\\ 0, & \text{otherwise.}\end{cases} \quad (16)$$

Then the cardinality of S_y is the random variable

$$\|S_y\|=\sum_{k,l}\Psi_{k,l}(y), \quad (17)$$

and

$$E\|S_y\| = E\Psi_{k^*,l^*}(y) + \sum_{k\neq k^*, l=l^*} E\Psi_{k,l}(y) + \sum_{k=k^*, l\neq l^*} E\Psi_{k,l}(y) + \sum_{k\neq k^*, l\neq l^*} E\Psi_{k,l}(y). \quad (18)$$

We now bound the random variable $\|S_Y\|$.

Lemma 4: For any $\epsilon>0$, $R_1>0$, $R_2>0$, there exists an n_0 such that for $n\geqslant n_0$,

$$P\{\|S_Y\| > 2^{n(R_1+R_2-I(X_1,X_2;Y|U)+\epsilon)}\} < \epsilon.$$

Proof: Using Lemmas 2 and 3, it follows that

$$E\Psi_{k,l}(Y) \leqslant 2^{-n[I(X_1,X_2;Y|U)-\epsilon]}, \qquad k\neq k^*,\ l\neq l^*, \quad (19)$$

$$E\Psi_{k,l}(Y) \leqslant 2^{-n[I(X_2;Y|X_1,U)-\epsilon]}, \qquad k=k^*,\ l\neq l^*, \quad (20)$$

$$E\Psi_{k,l}(Y) \leqslant 2^{-n[I(X_1;Y|X_2,U)-\epsilon]}, \qquad k\neq k^*,\ l=l^*. \quad (21)$$

Therefore,

$$E\|S_Y\| \leqslant 1 + (2^{nR_1}-1)(2^{nR_2}-1)2^{-n[I(X_1,X_2;Y|U)-\epsilon]} + (2^{nR_1}-1)2^{-n[I(X_1;Y|X_2,U)-\epsilon]} + (2^{nR_2}-1)2^{-n[I(X_2;Y|X_1,U)-\epsilon]}. \quad (22)$$

If

$$R_1 < I(X_1;Y|X_2,U)-\epsilon, \quad (23)$$

$$R_2 < I(X_2;Y|X_1,U)-\epsilon, \quad (24)$$

then n can be chosen so that the last two terms in (22) are less than ϵ. Thus,

$$E\|S_Y\| < (1+\epsilon) + 2^{n[R_1+R_2-I(X_1,X_2;Y|U)-\epsilon]} \quad (25)$$

$$= (1+\epsilon) + 2^{n(A-\epsilon)} \quad (26)$$

$$\leqslant 2^{nA}, \qquad \text{for } n \text{ sufficiently large, and } A>\epsilon, \quad (27)$$

where

$$A = R_1+R_2-I(X_1,X_2;Y|U). \quad (28)$$

Using Markov's inequality we obtain

$$P_b \triangleq \Pr\{\|S_Y\| > 2^{n\epsilon_1}2^{nA}\} \leqslant 2^{-n\epsilon_1}, \qquad \text{for any } \epsilon_1>0 \quad (29)$$

$$\leqslant \epsilon/5B, \qquad \text{for sufficiently large } n. \quad (30)$$

This completes the proof of Lemma 4.

Thus with high probability $1-P_b$, the index required (in block $b+1$) in order to resolve the receiver's residual uncertainty about block b takes on no more than $2^{n(A+\epsilon_1)}$ values. However, recall that in the first step of the decoding procedure, we require that the index have fewer than $n(I(U;Y)-\epsilon)$ bits. Thus we need

$$A+\epsilon_1 = R_1+R_2-I(X_1,X_2;Y|U)+\epsilon_1 < I(U;Y)-\epsilon, \quad (31)$$

or equivalently

$$R_1+R_2 < I(X_1,X_2;Y|U)+I(U;Y)-(\epsilon+\epsilon_1) \quad (32)$$

$$= I(X_1,X_2;Y)-(\epsilon+\epsilon_1). \quad (33)$$

By choosing n sufficiently large and ϵ, ϵ_1 sufficiently small, we can make R_1+R_2 as close to $I(X_1,X_2;Y)$ as desired in (33).

Bounding the Probability of Error: For the above scheme, we will declare an error in block b if one or more of the following events happen.

- E_1 Decoding step 1 fails (i.e., j^* incorrectly decoded).
- E_2 Decoding step 2 fails (i.e., transmitter 1 incorrectly estimates the message of transmitter 2).
- E_3 Decoding step 3 fails (i.e., message of first transmitter incorrectly decoded at second transmitter).
- E_4 $(x_1^{j^*}(k^*), x_2^{j^*}(l^*))$ not jointly typical with y.
- E_5 $\|S_y\| > 2^{n(A+\epsilon_1)}$.

Using the union bound, we can upperbound the probability of error $\bar{P}_{e,b}$ (averaged over the choice of codebooks and possible input messages) in block b by

$$\bar{P}_{e,b} \leqslant \Pr\{E_1\} + \Pr\{E_2\} + \cdots + \Pr\{E_5\}$$
$$= P_{e_1} + P_{e_{2,1}} + P_{e_{1,2}} + (1-P_c) + P_b \quad (34)$$
$$< \epsilon/B.$$

This proves that there exists at least one code with an average probability of error in block b of less than ϵ/B. The union bound then states that the average probability of error in B blocks is less than ϵ. Q.E.D.

V. The Cardinality of U

A rate region is considered to be presented in satisfactory form only when it is computable from its characterization. Thus Shannon's original theorem

$$C = \sup_{p(x)} I(X;Y) \quad (35)$$

is satisfactory, while the equally correct

$$C = \sup_n \sup_{p(x_1,\cdots,x_n)} (1/n) I(X_1,\cdots,X_n;Y_1,\cdots,Y_n) \quad (36)$$

is not, simply because the first involves a maximization of a convex function over a compact set and the second does not. The first expression leads naturally to an algorithm of finite length such that for any n a computation of finite length will yield an approximation C_n, $|C_n-C|\leqslant 1/n$. The second expression has no such natural interpretation. Equation (35) is said to be a "single letter" characterization of capacity.

Yet another computability consideration enters the rate region characterization in Theorem 1 of this paper. It appears that

$$R_1 \leqslant I(X_1;Y|X_2,U),$$
$$R_2 \leqslant I(X_2;Y|X_1,U), \quad (37)$$
$$R_1+R_2 \leqslant I(X_1,X_2;Y),$$

is a single letter characterization. However this characterization becomes useful for computation only when the cardinality of the auxiliary random variable U is bounded. This bound limits the computation to that of a well behaved function over a compact set. Salehi [15] has investigated the region in (37) and has proved that $\|U\| \leq \min\{\|X_1\| \cdot \|X_2\|, \|Y\|\}$ is sufficient to generate the entire region.

VI. The AWGN Multiple-Access Channel

We first recall the capacity region of the AWGN multiple-access channel. An achievable rate region using the feedback scheme proposed in Section I will then be given.

The AWGN multiple-access channel is the most commonly studied continuous alphabet channel $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \mathcal{R}$. 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 $EZ^2 = N$. There are average power constraints P_1 and P_2 on the inputs which require that the encoded messages $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ (which have n components x_{it}, $t = 1,2,\cdots,n$) satisfy

$$\frac{1}{n}\sum_{t=1}^{n} x_{it}^2 \leq P_i, \qquad i = 1,2. \tag{38}$$

The capacity region $\mathcal{C}$ for the AWGN multiple-access channel has been determined by Wyner [7] and Cover [8] based on the work of [5, 6]: $\mathcal{C}$ = set of all rate pairs $\boldsymbol{R} = (R_1, R_2)$ satisfying

$$R_1 \leq \frac{1}{2}\log\left(1 + \frac{P_1}{N}\right), \tag{39a}$$

$$R_2 \leq \frac{1}{2}\log\left(1 + \frac{P_2}{N}\right), \tag{39b}$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(1 + \frac{P_1 + P_2}{N}\right). \tag{39c}$$

The model for the AWGN multiple-access channel with feedback is shown in Fig. 5.

We now proceed to determine an achievable rate region using the feedback scheme of Section I. The basic idea in block b, $1 < b < B$, is for transmitter 1 to devote a fraction α_1 of its power P_1 to the transmission of new information and the remaining fraction $\bar{\alpha}_1 = (1 - \alpha_1)$ to the resolution of the receiver's residual uncertainty about the messages in block $b - 1$. Similarly, transmitter 2 uses α_2 of its power P_2 to send new information and $\bar{\alpha}_2$ to help resolve the receiver's residual uncertainty. The code book for each block is chosen as follows. We first generate a sequence of n independent identically distributed (i.i.d.) normal random variables with zero mean and unit variance. Then we independently generate 2^{nR_0} such sequences. At the first transmitter we scale the components of each of these sequences by $\sqrt{\bar{\alpha}_1 P_1}$. At the second transmitter these same sequences are scaled by $\sqrt{\bar{\alpha}_2 P_2}$. Furthermore, we independently generate 2^{nR_i} n-sequences each of which consists of n i.i.d. Gaussian random variables with mean zero and variance $\alpha_i P_i$ for use at transmitter i. At each transmitter i,

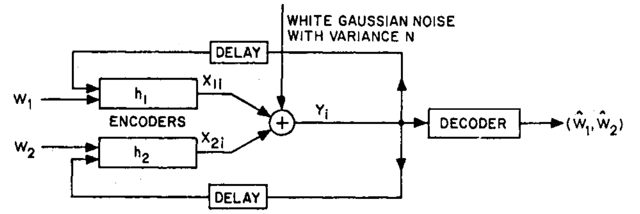


Fig. 5. AWGN multiple-access channel with feedback.

the two sequences of power $\alpha_i P_i$ and $\bar{\alpha}_i P_i$ are added to form the channel input.

We note that the receiver's residual uncertainty is resolved at a rate corresponding to coherent transmission from the two encoders, i.e., at an effective power

$$\left(\sqrt{\bar{\alpha}_1 P_1} + \sqrt{\bar{\alpha}_2 P_2}\right)^2 = \bar{\alpha}_1 P_1 + \bar{\alpha}_2 P_2 + 2\sqrt{\bar{\alpha}_1 P_1 \bar{\alpha}_2 P_2}.$$

The receiver first decodes the information needed for resolving its residual uncertainty and then subtracts off the corresponding signal before estimating the new information. Using standard superposition arguments, it is possible to show that any rate pair (R_1, R_2) satisfying

$$R_1 < \frac{1}{2}\log\left(1 + \frac{\alpha_1 P_1}{N}\right), \tag{40a}$$

$$R_2 < \frac{1}{2}\log\left(1 + \frac{\alpha_2 P_2}{N}\right), \tag{40b}$$

$$R_1 + R_2 < \frac{1}{2}\log\left(1 + \frac{P_1 + P_2 + 2\sqrt{\bar{\alpha}_1 \bar{\alpha}_2 P_1 P_2}}{N}\right), \tag{40c}$$

for $0 \leq \alpha_i \leq 1$, is achievable.[2]

It is fairly easy to see that a necessary condition on α_1 and α_2 to achieve a point on the boundary of the rate region given by (40) is

$$\left(1 + \frac{\alpha_1 P_1}{N}\right)\left(1 + \frac{\alpha_2 P_2}{N}\right) = \left(1 + \frac{P_1 + P_2 + 2\sqrt{\bar{\alpha}_1 \bar{\alpha}_2 P_1 P_2}}{N}\right). \tag{41}$$

Using (40) and (41), an achievable rate region for the AWGN multiple-access channel with feedback with $P_1/N = P_2/N = 10$ has been computed. The results are shown in Fig. 6. We might note that if $P_1/N = P_2/N = \alpha$, then the pair

$$(R_1^*, R_2^*) = \left(\frac{1}{2}\log\left(2\sqrt{1+\alpha} - 1\right), \frac{1}{2}\log\left(2\sqrt{1+\alpha} - 1\right)\right) \tag{42}$$

is achievable.

Acknowledgment

The improvement in the achievable rate region over the authors' previous effort in [2] was stimulated by extensive

[2]Recently Leung and Ozarow [13] have examined an optimal deterministic feedback scheme which shows that the region defined by (40) is smaller than the capacity region.

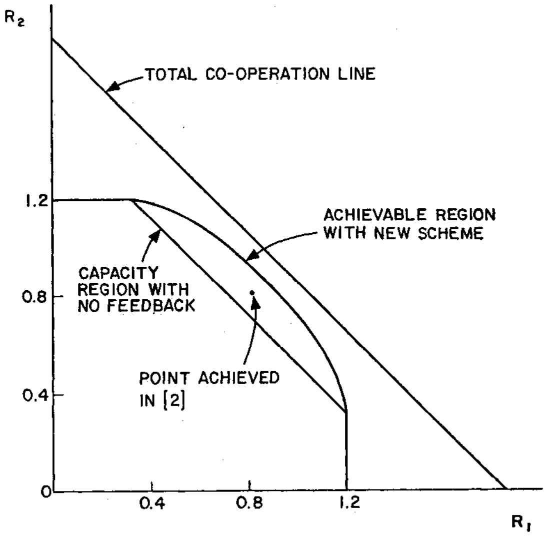


Fig. 6. Achievable rate region for AWGN multiple-access channel with $P_1/N = P_2/N = 10$.

discussion with E. C. van der Meulen. Throughout these discussions he graciously contributed comments and calculations. The authors also wish to thank P. Bergmans for detailed discussions and calculations on the Gaussian version of the multiple-access channel (MAC) feedback channel and A. El Gamal for efforts on the as yet unproved converse. All of the above participated in a cooperative workshop atmosphere during July and August of 1976. Also during this work, A. Carleial [14] independently generated the region in the Gaussian case.

References

[1] N. T. Gaarder and J. K. Wolf, "The capacity region of a multiple-access discrete memoryless channel can increase with feedback," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 100–102, Jan. 1975.

[2] T. M. Cover and S. K. Leung-Yan-Cheong, "A scheme for enlarging the capacity region of multiple-access channels using feedback," Dept. of Stat., Stanford Univ., Stanford, CA, Tech. Rep. 17, Mar. 1976.

[3] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 2–14, Jan. 1972.

[4] D. Slepian and J. K. Wolf, "A coding theorem for multiple-access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, pp. 1037–1076, Sept. 1973.

[5] R. Ahlswede, "Multi-way communication channels," in *Proc. Second Int. Symp. Inform. Transmiss.*, Tsahkadsor, Armenia, USSR, Hungarian Press, 1971.

[6] H. Liao, "Multiple-access channels," Ph.D. dissertation, Dept. Elec. Eng., Univ. Hawaii, Honolulu, 1972.

[7] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 2–10, Jan. 1974.

[8] T. M. Cover, "Some advances in broadcast channels," in *Advances in Communication Systems*, vol. 4, *Theory and Applications*, ed. by A. Viterbi, Ed. San Francisco: Academic, 1975.

[9] G. D. Forney, "Information theory," course notes for EE-376 Elec. Eng. Dept., Stanford University, Stanford, CA, 1972.

[10] T. M. Cover, "An achievable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 399–404, July 1975.

[11] J. Wolfowitz, *Coding Theorems of Information Theory*, 2nd ed. Berlin: Springer, 1964.

[12] R. F. Ahlswede and J. Korner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 629–637, Nov. 1975.

[13] L. Ozarow and S. K. Leung, "Deterministic feedback coding schemes for the additive white Gaussian noise broadcast and multiple-access channels," submitted to *IEEE Trans. Inform. Theory*.

[14] A. B. Carleial, "On the capacity of the Gaussian multiple-access channel with feedback," *Rev. Brasileira Tecnol.*, vol. 8, 1977.

[15] M. Salehi, "Cardinality bounds on auxiliary variables in multiple-user theory via the method of Ahlswede and Korner," Tech. Rep. 33, Stanford Univ., Stanford, CA, Aug. 1978.