AN ACHIEVABLE RATE REGION FOR THE BROADCAST CHANNEL

BY

THOMAS M. COVER

TECHNICAL REPORT NO. 10
SEPTEMBER 18, 1974

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT GK-34363

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
AN ACHIEVABLE RATE REGION FOR THE BROADCAST CHANNEL

by

THOMAS M. COVER

TECHNICAL REPORT NO. 10
September 18, 1974

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCES FOUNDATION GRANT GK-34363

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
AN ACHIEVABLE RATE REGION FOR THE BROADCAST CHANNEL

Thomas M. Cover
Stanford University

Abstract

Let $p(y_1,y_2|x)$ denote a discrete memoryless channel with single source $X$ and two independent receivers $Y_1$ and $Y_2$. We exhibit an achievable region of rates $(R_{11},R_{12},R_{22})$ at which independent information can be sent respectively to receiver 1, to both receivers 1 and 2, and to receiver 2. The achievability of the region is shown by using a version of the asymptotic equipartition property involving many simultaneous "typicality" constraints. These results immediately generalize to yield an achievable rate region for the m-sender n-receiver channel in terms of standard mutual information quantities.
1. Introduction

We shall define a two-receiver memoryless broadcast channel, denoted
by \((X, p(y_1, y_2 | x), Y_1 \times Y_2)\) or by \(p(y_1, y_2 | x)\), to consist of three
finite sets \(X, Y_1, Y_2\) and a collection of probability distributions
\(p(\cdot, \cdot | x)\) on \(Y_1 \times Y_2\), one for each \(x \in X\). The interpretation is that
\(x\) is an input to the channel and \(y_1\) and \(y_2\) are the respective out-
puts at receiver terminals 1 and 2 as shown in Fig. 1. The problem
is to communicate simultaneously with receivers 1 and 2 as efficiently
as possible.

This problem was posed in [1], and a sequence of contributions to the
solution has appeared in [2,..,9]. Our results are closely related to
those of [8,9]. The as yet unpublished [8] contains an independent state-
ment of the achievability of the region \(\mathcal{R}\) treated in this paper and out-
lines a proof based on results in [9].
The $n^{th}$ extension for a broadcast channel is the broadcast channel

$$
\left( X^n, p(y_1^n, y_2^n | x), y_1^n \times y_2^n \right),
$$

(1)

where $p(y_1^n, y_2^n | x) = \prod_{j=1}^{n} p(y_{1j}, y_{2j} | x_j)$, for $x \in X^n$, $y_1 \in Y_1^n$, $y_2 \in Y_2^n$.

An $((M_{11}, M_{22}, M_{12}), n)$ code for a broadcast channel consists of three sets of integers

$$
\begin{align*}
M_{\sim 11} &= \{1, 2, \ldots, M_{11}\} \\
M_{\sim 12} &= \{1, 2, \ldots, M_{12}\} \\
M_{\sim 22} &= \{1, 2, \ldots, M_{22}\}
\end{align*}
$$

(2)

an encoding function

$$
x: M_{\sim 11} \times M_{\sim 12} \times M_{\sim 22} \rightarrow X^n,
$$

(3)

and two decoding functions

$$
\begin{align*}
g_1: y_1^n &\rightarrow M_{\sim 11} \times M_{\sim 12}; g_1(y_1^n) = (\hat{\jmath}, \hat{k}) \\
g_2: y_2^n &\rightarrow M_{\sim 12} \times M_{\sim 22}; g_2(y_2^n) = (\hat{k}, \hat{\ell})
\end{align*}
$$

(4)

The set $\{x(j, k, \ell) | (j, k, \ell) \in M_{\sim 11} \times M_{\sim 12} \times M_{\sim 22}\}$ is called the set of codewords. As illustrated in Fig. 1, we think of integers $j$ and $\ell$ as being arbitrarily chosen by the transmitter to be sent to receivers 1 and 2, respectively. The integer $k$ is also chosen by the transmitter and is intended to be received by both receivers. Thus, $k$ is the "common" part of the message, and $j$ and $\ell$ are the "independent" parts.
of the message.

If the message \((j,k,\ell)\) is sent, let

\[
\lambda(j,k,\ell) = \Pr\{g_1(Y_1) \neq (j,k) \text{ or } g_2(Y_2) \neq (k,\ell)\}
\]

(5)
denote the probability of error, where we note that \(Y_1, Y_2\) are the only chance variables in the above expression. We define the average probability of error of the code as

\[
P(e) = P_n(e) = \frac{1}{M} \sum_{j,k,\ell} \lambda(j,k,\ell)
\]

(6)

where

\[
M = M_{11} M_{22} M_{12}
\]

(7)

This probability of error is calculated under a special distribution; namely, the uniform distribution over the codewords.

We shall define the rate \((R_{11}, R_{12}, R_{22})\) of an \(((M_{11}, M_{12}, M_{22}), n)\) code by

\[
R_{11} = \frac{1}{n} \log M_{11}
\]

\[
R_{12} = \frac{1}{n} \log M_{12}
\]

(8)

\[
R_{22} = \frac{1}{n} \log M_{22}
\]

all defined in bits/transmission. Thus, \(R_{11}\) is the rate of transmission of independent information to receiver 1, for \(i = 1, 2\), and \(R_{12}\) is the portion of the information common to both receivers.
Definition. The rate \((R_{11}, R_{12}, R_{22})\) is said to be achievable by a broadcast channel if, for any \(\epsilon > 0\) and for all \(n\) sufficiently large, there exists an \((M_{11}, M_{12}, M_{22}, n)\) code with

\[
\begin{align*}
M_{11} &\geq 2^{nR_{11}} \\
M_{12} &\geq 2^{nR_{12}} \\
M_{22} &\geq 2^{nR_{22}}
\end{align*}
\] (9)

such that \(P_n(\epsilon) < \epsilon\).

Comment. Note that the total number \(M = M_{11}M_{12}M_{22}\) of codewords for a code satisfying (9) must exceed \(2^{n(R_{11}+R_{12}+R_{22})}\).

Definition. The capacity region \(\mathcal{R}^*\) for a broadcast channel is the set of all achievable rates \((R_{11}, R_{12}, R_{22})\).

The goal of this paper is to determine an achievable region \(\mathcal{R}\) and thus bound \(\mathcal{R}^*\) below by \(\mathcal{R} \subseteq \mathcal{R}^*\).

Comment. The extension of the definition of the broadcast channel from two receivers to \(k\) receivers is notationally cumbersome but straightforward, given the following comment. The index sets \(M_{11}, M_{12}, M_{22}\) are replaced by \(2^k - 1\) index sets \(I(\theta), \theta \in \{0,1\}^k, \theta \neq 0\), with the interpretation that the integer \(i(\theta)\) selected in index set \(I(\theta) = \{1,2,\ldots,M(\theta)\}\) is intended (by the proper code selection) to be received correctly by every receiver \(j\) for which \(\theta_j = 1\) in \(\theta = (\theta_1, \theta_2, \ldots, \theta_k)\). Then, for example, the rate of transmission over the \(n^{th}\) extension of a broadcast channel to the \(i^{th}\) receiver will be given by
\[ R_1 = \frac{1}{n} \log \prod_{\theta \in \{0,1\}^k} M(\theta) = \frac{1}{n} \sum_{\theta_{i} = 1} \log M(\theta). \] (10)

In the two-receiver broadcast channel, the corresponding sets in the new notation are \( M_{12} = I(1,1), M_{11} = I(1,0), M_{22} = I(0,1). \)

2. **An Achievable Region \( \mathcal{R} \)**

In this section we shall exhibit an achievable \( \mathcal{R} \). We shall prove the achievability of \( \mathcal{R} \) in section 4.

First we shall define three auxiliary random variables \( U, R, V \) taking values in finite sets \( \mathcal{U}, \mathcal{R}, \mathcal{V} \). Let \( x : \mathcal{U} \times \mathcal{R} \times \mathcal{V} \rightarrow \mathcal{X} \) denote an arbitrary mapping of the auxiliary random variables into the input alphabet. The picture we have in mind is given in Fig. 2.

FIG. 2 : Auxilliary Random Variables
For each assignment of probability distributions \( p(u), p(r), p(v) \) and mapping function \( x(\cdot) \), we associate the joint probability distribution function

\[
p(u, r, v, y_1, y_2) = p(u)p(r)p(v)p(y_1, y_2|x(u, r, v))
\]

are defined in the usual way.

Define

\[
I = (I_1, I_2, \ldots, I_6) = \left( I(U; R, Y_1), I(V; R, Y_2), I(R; U, Y_1), I(R; V, Y_2), I(U, R; Y_1), I(V, R; Y_2) \right)
\]

Let \( \mathcal{J} \) denote the set of all \( I \in \mathbb{R}^6 \) generated by all assignments of \( p(u), p(r), p(v), x(\cdot) \). Let \( C_\mathcal{J} \) denote the convex hull of \( \mathcal{J} \). Let \( R(I) \) denote the set of all \( (R_{11}, R_{12}, R_{22}) \in \mathbb{R}^3 \) satisfying the six inequalities

\[
\begin{align*}
R(I): & & R_{11} < I_1 \\
& & R_{22} < I_2 \\
& & R_{12} < I_3 \\
& & R_{12} < I_4 \\
& & R_{11} + R_{12} < I_5 \\
& & R_{12} + R_{22} < I_6
\end{align*}
\]
Theorem 1. The region

$$\mathcal{R} = \bigcup_{I \in C_{o}(\mathcal{J})} R(I)$$

(15)

is achievable.

We can express the capacity region in another form. Observe that an arbitrary point \( I \) on the boundary of \( C_{o}(\mathcal{J}) \) can always be expressed as the convex combination of no more than 6 (extreme) points of \( \mathcal{J} \). Let

$$I = \sum_{i=1}^{6} p(q^{(i)}) \frac{I}{q^{(i)}}, \quad I \in q^{(i)} \in \mathcal{J}$$

(16)

be the desired convex combination, where \( q^{(i)} = (p(u|q^{(i)}), p(r|q^{(i)}), p(v|q^{(i)}), x_{q^{(i)}}(\cdot)) \) is an element in the set of all assignments \( (p, x(\cdot)) \) and \( \frac{I}{q^{(i)}} \) is the vector of mutual informations induced by this assignment.

Let \( Q \) denote a random variable with \( \Pr(Q = q^{(i)}) = p(q^{(i)}), p(q^{(i)}) \geq 0, i = 1, 2, \ldots, 6, \sum p(q^{(i)}) = 1 \). It then follows from inspection of the definition of mutual information that, for example,

$$\sum_{i=1}^{6} p(q^{(i)}) I_{q^{(i)}}(U; R, Y_1) = I(U; R, Y_1 | Q).$$

(17)

Thus, \( \mathcal{R} \) can be expressed as follows.

Theorem 1'. \( \mathcal{R} \) is the union of all \( (R_{11}, R_{12}, R_{22}) \in \mathbb{R}^3 \) satisfying the inequalities
\[ R_{11} < I(U;R,Y_1|Q) \]
\[ R_{22} < I(V;R,Y_2|Q) \]
\[ R_{12} < I(R;U,Y_1|Q) \]
\[ R_{12} < I(R;V,Y_2|Q) \]
\[ R_{11} + R_{12} < I(U,R,Y_1|Q) \]
\[ R_{12} + R_{22} < I(V,R,Y_2|Q) \]

where the union is over all r.v.'s Q such that Q takes on no more than 6 values in the set of all assignments \((p,x(\cdot))\). This is the same region as in Theorem 1.

The following conjecture is more a result of wishful thinking than of soundly based intuition. If true, this conjecture would substantially reduce the computation necessary to compute \( \bar{R} \).

**Conjecture.** The set of achievable rates in Theorem 1 is unchanged if the alphabet sizes of the auxiliary random variables are restricted to \( U = \bar{U} = \bar{V} = \{0,1,2,\ldots,m-1\} \), \( m = |X| \) and if the mapping function \( x(\cdot) \) is given by
\[ x = u + r + v \pmod{|X|} \]

3. **Jointly Typical Sequences**

Before proceeding with the proof, we shall define a simultaneous asymptotic equipartition property. This will allow us to decode messages at the receiver by simply checking to see which of the possible
input messages is jointly "typical" with the received output. If there is one and only one such candidate, we shall declare this to be the message.

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 r.v.'s, and consider \( n \) independent copies of \( S \). Thus,

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

(20)

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

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

(21)

We know by the law of large numbers that for a given subset \( S \) of r.v.'s

\[- \frac{1}{n} \log p(S_1, S_2, \ldots, S_n) = - \sum_{i=1}^{n} \log p(S_i) \to H(S) \]

(22)

with probability one. This convergence takes place simultaneously with probability one for all \( 2^k \) subsets

\[
S \subseteq \{X^{(1)}, X^{(2)}, \ldots, X^{(k)}\}
\]

(23)

Definition. The set \( A_\epsilon \) of jointly \( \epsilon \)-typical \( n \)-sequences \( (x^{(1)}, x^{(2)}, \ldots, x^{(k)}) \) is defined by
\[ A_\varepsilon(x^{(1)}, x^{(2)}, \ldots, x^{(k)}) = \]

\[ A_\varepsilon = \left\{ \left( x^{(1)}_1, x^{(2)}_1, \ldots, x^{(k)}_1 \right) \in \left( x^{(1)}_2 \right)^n \times \left( x^{(2)}_2 \right)^n \times \ldots \times \left( x^{(k)}_2 \right)^n : \right. \]

\[ \left| - \frac{1}{n} \log p(s) - H(S) \right| \leq \varepsilon, \quad \forall S \subseteq \left\{ x^{(1)}, x^{(2)}, \ldots, x^{(k)} \right\}, \quad (24) \]

where \( s \) denotes the ordered set of sequences in \( \left\{ x^{(1)}, \ldots, x^{(k)} \right\} \) corresponding to \( S \). Let \( A_\varepsilon(S) \) denote the restriction of \( A_\varepsilon \) to the coordinates corresponding to \( S \).

Thus, for example,

\[ A_\varepsilon(x^{(1)}, x^{(2)}) = \left\{ \left( x^{(1)}_1, x^{(2)}_1 \right) : \right. \]

\[ \left| - \frac{1}{n} \log p(x^{(1)}_1, x^{(2)}_1) - H(x^{(1)}_1, x^{(2)}_2) \right| < \varepsilon, \]

\[ \left| - \frac{1}{n} \log p(x^{(1)}_1) - H(x^{(1)}_1) \right| < \varepsilon, \]

\[ \left| - \frac{1}{n} \log p(x^{(2)}_1) - H(x^{(2)}_1) \right| < \varepsilon \} \quad (25) \]

The following is a version of the asymptotic equipartition property involving simultaneous constraints. (Compare with Forney [10].)

**Lemma 1.** For any \( \varepsilon > 0 \), there exists an integer \( n \) such that \( A_\varepsilon(S) \) satisfies

1. \[ \Pr\left\{ A_\varepsilon(S) \right\} \geq 1 - \varepsilon, \quad \forall S \subseteq \left\{ x^{(1)}, \ldots, x^{(k)} \right\} \]

\[ (26) \]

2. \[ s \in A_\varepsilon(S) \Rightarrow \left| - \frac{1}{n} \log p(s) - H(S) \right| < \varepsilon \]
(iii) \( (1 - \epsilon) 2^{n(H(S) - \epsilon)} \leq |A_\epsilon(S)| \leq 2^{n(H(S) + \epsilon)} \). \hspace{1cm} (26)

Proof. (i) follows from the convergence of the random variables in the definition of \( A_\epsilon(S) \). (ii) follows immediately from the definition of \( A_\epsilon(S) \) in Eq. (24). (iii) follows from

\[
1 \geq \sum_{s \in A_\epsilon(S)} p(s) \geq \sum_{A_\epsilon(S)} 2^{-n(H(S) + \epsilon)} = |A_\epsilon(S)| 2^{-n(H(S) + \epsilon)}, \hspace{1cm} (27)
\]

and

\[
(1 - \epsilon) \leq \sum_{s \in A_\epsilon(S)} p(s) \leq |A_\epsilon(S)| 2^{-n(H(S) - \epsilon)}. \hspace{1cm} (28)
\]

Corollary. If \((w, z) \in A_\epsilon(W, Z)\), then

\[
2^{-n(H(W | Z) + 2\epsilon)} \leq p(w | z) \leq 2^{-n(H(W | Z) - 2\epsilon)}. \hspace{1cm} (29)
\]

Proof. \((w, z) \in A_\epsilon(W, Z)\) implies

\[
2^{-n(H(W, Z) + \epsilon)} \leq p(w, z) \leq 2^{-n(H(W, Z) - \epsilon)}, \hspace{1cm} (30)
\]

and

\[
2^{-n(H(Z) + \epsilon)} \leq p(z) \leq 2^{-n(H(Z) - \epsilon)}. \hspace{1cm} (31)
\]

The corollary follows from \( p(w | z) = p(w, z) / p(z) \).

We shall need to know the probability that conditionally independent sequences are jointly typical. Let the discrete r.v.'s \( W, Z, Q \) have
joint distribution \( p(w,z,q) \). Let \( W', Z' \) be conditionally independent given \( Q \), with the marginals

\[
p(w | q) = \sum_z p(w,z,q) / p(q)
\]

\[
p(z | q) = \sum_w p(w,z,q) / p(q)
\]  \( (32) \)

The unconditional version of the following lemma has been observed and proved by Forney [10] as crucial in giving the natural proof of Shannon's second theorem. This lemma has also been independently used by the author on source compression for dependent ergodic sources [11].

**Lemma 2.** Let \( (W,Z,Q) \sim \prod_{i=1}^n p(w_i,z_i,q_i) \) and \( (W',Z',Q) \sim \prod_{i=1}^n p(w_i | q_i) \cdot p(z_i | q_i) p(q_i) \) as above. Then, for \( n \) such that \( \Pr[A_\varepsilon] \geq 1 - \varepsilon \),

\[
(1 - \varepsilon) 2^{-n(I(W;Z|Q) + \varepsilon)} \leq \Pr\left\{ (W',Z',Q) \in A_\varepsilon (W,Z,Q) \right\}
\]

\[
\leq 2^{-n(I(W;Z|Q) - 7\varepsilon)}
\]  \( (33) \)

**Proof.** From the corollary, \( (w,z,q) \in A_\varepsilon (W,Z,Q) \) implies

\[
2^{-n(H(W|Q) + H(Z|Q) + H(Q) + 6\varepsilon)} \leq p(w | q) p(z | q) p(q)
\]

\[
\leq 2^{-n(H(W|Q) + H(Z|Q) + H(Q) - 6\varepsilon)}
\]  \( (34) \)

and

\[
2^{-n(H(W,Z,Q) + \varepsilon)} \leq p(w,z,q) \leq 2^{-n(H(W,Z,Q) - \varepsilon)}
\]  \( (35) \)
Observe that

\[ H(W|Q) + H(Z|Q) + H(Q) - H(W,Z,Q) = H(W|Q) + H(Z|Q) - H(W,Z|Q) = I(W;Z|Q) \]  

(36)

Consequently,

\[
p(w,z,q) \ 2^{-n(I(W;Z|Q)+7\varepsilon)} \leq p(w|q) \ p(z|q) \ p(q) \]

\[
\leq p(w,z,q) \ 2^{-n(I(W;Z|Q)-7\varepsilon)}. \quad (37)
\]

Now, summing the terms in the above equation over \( A_\varepsilon \),

\[
(1 - \varepsilon) \ 2^{-n(I(W;Z|Q)+7\varepsilon)} \leq \sum_{A_\varepsilon} 2^{-n(I(W;Z|Q)+7\varepsilon)} p(w,z,q)
\]

\[
\leq \sum_{A_\varepsilon} p(w|q) \ p(z|q) \ p(q)
\]

\[
= \text{Pr}\{ (W',Z',Q) \in A_\varepsilon \ (W,Z,Q) \}
\]

\[
\leq 2^{-n(I(W;Z|Q)-7\varepsilon)} \sum_{A_\varepsilon} p(w,z,q)
\]

\[
\leq 2^{-n(I(W;Z|Q)-7\varepsilon)}. \quad (38)
\]

Q.E.D.
4. The Achievability of $\mathcal{R}$

For any $I \in C_0(j)$, we shall show how to achieve any rate $(R_{11}, R_{12}, R_{22})$ satisfying the 6 inequalities in Eq. (18). Consider a given assignment $p(u|q), p(r|q), p(v|q), x(\cdot|q), p(q), q \in \{q^{(1)}, q^{(2)}, \ldots, q^{(6)}\}$ and the associated $I \in C_0(j)$ given in Eq. (17).

Random Encoding for Theorem 1'

First, generate a sequence of $n$ i.i.d. r.v.'s $Q = (Q_1, Q_2, \ldots, Q_n)$. Here, $Q_k = q_k$ plays the role of a time-sharing parameter that at each time $k$ informs the transmitter and both receivers that the mode of operation is $Q_k = q_k$, where $q_k$ is one of the 6 modes in $\{q^{(1)}, q^{(2)}, \ldots, q^{(6)}\}$. Conditioned on $Q = q$, generate $2^{nR_{11}}$ random n-sequences of random variables drawn according to $p(u|q)$; $2^{nR_{12}}$ random n-sequences of r.v.'s drawn according to $p(r|q)$; $2^{nR_{22}}$ random n-sequences of r.v.'s drawn according to $p(v|q)$. Index the strings by $j = 1, 2, \ldots, 2^{nR_{11}}$; $k = 1, 2, \ldots, 2^{nR_{12}}$; $\ell = 1, 2, \ldots, 2^{nR_{22}}$; respectively. Thus, for example, the $j^{th}$ n-sequence (word) $u(j)$ has probability

$$\Pr[u(j) = u(j)|Q = q] = \prod_{i=1}^{n} p(u_i(j)|q_i) . \quad (39)$$

Also, $u(j), r(k), v(\ell), \forall j, k, \ell$, are conditionally independent given $Q$.

To each $(j,k,\ell)$ there corresponds a triple of n-sequences $(u(j), r(k), v(\ell))$ and the codeword

$$x(j,k,\ell) = (x_1(j,k,\ell), x_2(j,k,\ell), \ldots, x_n(j,k,\ell)) , \quad (40)$$
where
\[ x_m(j, k, \ell) = x(u_m(j), r_m(k), v_m(\ell) | q_m) \]  \hspace{1cm} (41)

The codebook consists of the \( M \) \( n \)-sequences \( x(i, j, k) \), \((j, k, \ell) \in M_{11} \times M_{12} \times M_{22}\).

**Decoding Rule**

Both receivers know \( \mathcal{Q} \). If \( y_1 \) is received, declare \((\hat{j}, \hat{k}) = (j, k) \) was sent if there is one and only one pair \((j, k) \in M_{11} \times M_{12}\) such that \((u(j), r(k), y_1, q) \in A_\varepsilon(U, R, Y_1, Q)\), i.e., if there is only one input pair \((j, k)\) that is jointly typical with the output.

If \( y_2 \) is received, declare \((\hat{k}, \hat{\ell}) = (k, \ell) \) was sent if there is one and only one pair \((k, \ell) \) such that \((r(k), v(\ell), y_2, q) \in A_\varepsilon(R, V, Y_2, Q)\).

**Proof of the Achievability of \( \mathcal{R} \).** Let \( J, K, L \) be independent r.v.'s drawn according to uniform distributions on \( M_{11}, M_{12}, M_{22} \), respectively. Let the code be chosen randomly according to the encoding description. Then the probability of error (over \( J, K, L \) and the random code) is given by
\[ P(e) = \Pr[(\hat{J}, \hat{K}) \neq (J, K) \text{ or } (\hat{K}, \hat{L}) \neq (K, L)] \]  \hspace{1cm} (42)

By the symmetry induced by the random coding, we see that each transmitted message \((j, k, \ell)\) yields the same probability of error. Thus, setting \((j, k, \ell) = (1, 1, 1)\), we have precisely
\[ P(e) = \Pr[(\hat{J}, \hat{K}) \neq (1, 1) \text{ or } (\hat{K}, \hat{L}) \neq (1, 1)] \]  \hspace{1cm} (43)

Let \( E(j, k, 1) \) denote the event \((u(j), r(k), y_1, Q) \in A_\varepsilon(U, R, Y_1, Q)\). This is the event that \((u(j), r(k), y_1, Q)\) is jointly \( \varepsilon \)-typical,
thus implying that \((j,k)\) are candidates for the decoded message \(g(Y_1)\).

Similarly, let \(E(k,\ell,2)\) denote the event \((R(k),V(\ell),Y_2) \in A_\epsilon(R,V,Y_2,Q)\).

Now

\[
\overline{P}(\epsilon) = \Pr\left\{E^c(1,1,1) \cup E^c(1,1,2) \bigcup_{(j,k) \neq (1,1)} E(j,k,1) \bigcup_{(k,\ell) \neq (1,1)} E(k,\ell,2)\right\}
\]

\[
\leq \Pr[E^c(1,1,1)] + \Pr[E^c(1,1,2)] + \sum_{k \neq 1} \Pr[E(1,k,1)]
\]

\[
+ \sum_{j \neq 1} \Pr[E(j,1,1)] + \sum_{j \neq 1} \Pr[E(j,k,1)] + \sum_{(k,\ell) \neq (1,1)} \Pr[E(k,\ell,2)].
\]

The first two terms correspond to the event that the correct codeword does not fall in the decoding set. The last terms correspond to the event that some incorrect codeword falls in the decoding set.

Choosing \(n\) sufficiently large that

\[
\Pr\left\{A_\epsilon(U,R,V,Y_1,Y_2)\right\} \geq 1 - \epsilon,
\]

we see from Lemma 1 that

\[
\Pr\left\{E^c(1,1,1)\right\} < \epsilon, \Pr\left\{E^c(1,1,2)\right\} < \epsilon.
\]

Consider the event \(E(j,1,1)\). We observe, for \(j \neq 1, k = 1\), that \(U(j)\) is conditionally independent of \((R(1),Y_1)\) given \(Q\). Also, the distribution of \(U(j)\) given \(Q\) is the same as that of \(U(1)\) given \(Q\).

Thus, Lemma 2 applies with the substitution \(U(j) = \overline{w}'\) and \((R(1),\overline{Y}_1) = \overline{z}'\) in Eq. (33). That is, for \(j \neq 1,

\[
\Pr\left\{E(j,1,1)\right\} = \Pr\left\{\left(\overline{U}(j),R(1),\overline{Y}_1,Q\right) \in A_\epsilon(U,R,Y_1,Q)\right\} \leq 2^{-n(I(U;R,Y_1|Q) - 7\epsilon)}
\]
Consequently,

$$
\sum_{j=1}^{M_{11}} \text{Pr}[E(j,1,1)] \leq 2^{-n(I(U;R,Y_1|Q)-R_{11}-7\epsilon)}
$$

(48)

Therefore, this term can be made less than $\varepsilon$ for

$$R_{11} \leq I(U;R,Y_1|Q) - 7\epsilon - \frac{1}{n} \log \frac{1}{\varepsilon}.
$$

(49)

Thus, this term $\to 0$ as $n \to \infty$ if

$$R_{11} < I(U;R,Y_1|Q).
$$

(50)

This is the first condition in Theorem 1'.

Similarly, applying Lemma 2 for the terms

$$
E(1,k,1) \quad \text{with} \quad (W',Z',Q) = (R(k),(U(1),Y_1),Q)
$$

$$
E(j,k,1) \quad \text{with} \quad (W',Z',Q) = ((U(j),R(k)),Y_1,Q)
$$

$$
E(k,1,2) \quad \text{with} \quad (W',Z',Q) = (R(k),(V(1),Y_2),Q)
$$

$$
E(1,\ell,2) \quad \text{with} \quad (W',Z',Q) = (V(\ell),(R(1),Y_2),Q)
$$

$$
E(k,\ell,2) \quad \text{with} \quad (W',Z',Q) = ((R(k),V(\ell)),Y_2,Q)
$$

(51)

for $j \neq 1$, $k \neq 1$, $\ell \neq 1$, we find (first letting $n \to \infty$, then $\varepsilon \to 0$) that $\bar{P}_n(e) \to 0$ whenever the conditions of Theorem 1' are satisfied.

Finally, if $\bar{P}_n(e) < \varepsilon$, there must exist at least one $((M_{11},M_{22}',M_{12}'),n)$ code with $P_n(e) < \varepsilon$. Thus, $\bar{P}_n(e) \to 0$ implies that there exists a sequence of $((M_{11}',M_{22}',M_{12}'),n)$ codes with $P_n(e) \to 0$ for any $(R_{11}',R_{12}',R_{22}') \in \mathbb{R}$. Q.E.D.
Acknowledgment

I would like to thank Dirk Hughes-Hartogs for his frequent interactions on this problem. I have also benefited from discussions with Carroll Keilers, Marty Hellman, Aydano Carleial, and Aaron Wyner.
References


