Two-User Interference Channels with Correlated Information Sources*

HARRY H. TAN

School of Engineering, University of California, Irvine, Irvine, California 92717

Inner and outer bounds on the capacity region of two-user discrete memoryless interference channels (IFC) are obtained in the communication situation where private messages are sent by each sender as well as a common message by both senders to its corresponding receiver. A limiting expression for the capacity region is also derived. Special cases of the IFC are considered.

I. INTRODUCTION

A two-user interference channel (IFC) consists of two senders situated at separate input terminals simultaneously transmitting messages to two corresponding receivers situated at separate output terminals. Furthermore, there is no other direct cross-communication between any of the four terminal points of the channel. The motivation for studying the IFC is to better understand the crosstalk problem in practical communication systems. In the IFC each sender generally has to contend with the disturbance from the other sender’s usage of the channel in addition to external channel noise disturbances.

Information— theoretic studies of the IFC when the message sources at the two channel input terminals are statistically independent have previously been carried out by Ahlswede (1971, 1974), Carleial (1975a, 1975b, 1978), Bergmans (1976) and Sato (1977). An excellent survey of this work as well as work on other multiple-terminal channels can be found in Van der Meulen (1977). The purpose of this paper is to study the IFC in the situation when the message sources at the two channel input terminals are correlated in a special way. Specifically we assume three statistically independent information sources, two of which are called private sources and the third a common source. Then each sender has available and transmits the outputs of the common source and one private source to its corresponding receiver. This type of correlated information sources has previously been considered by Slepian and Wolf (1973) in connection with a different multiple-terminal channel, the multiple-access channel.

* This work was supported by the National Science Foundation under Grant ENG 76-21827.
This paper is organized as follows. A precise definition of the communication situation involving the two-user discrete memoryless IFC that is considered in this paper is given in Section II. Inner and outer bounds on the capacity region as well as a limiting expression for the capacity region are given in Section III. Section IV considers special cases of the IFC. The random coding coding proof used to establish the inner bound to the capacity region is discussed in Section V. A summary of the results in the paper is given in Section VI. Details of many of the proofs and derivations are relegated to the appendices.

II. DEFINITIONS

A two-user discrete memoryless interference channel (IFC) \( \mathcal{H} = (\mathcal{X}_1 \times \mathcal{X}_2, P(y_1, y_2 | x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2) \) consists of finite sender (1) and sender (2) channel input alphabets \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) respectively, finite receiver (1) and receiver (2) channel output alphabets \( \mathcal{Y}_1 \) and \( \mathcal{Y}_2 \) respectively and joint channel transition probability function \( P(y_1, y_2 | x_1, x_2) \). Let

\[
\begin{align*}
P_1(y_1 | x_1, x_2) &= \sum_{y_2} P(y_1, y_2 | x_1, x_2), \\
P_2(y_2 | x_1, x_2) &= \sum_{y_1} P(y_1, y_2 | x_1, x_2),
\end{align*}
\]

be the receiver (1) and receiver (2) marginal channel transition probability functions. Also, let \( \mathcal{X}_1^n, \mathcal{X}_2^n, \mathcal{Y}_1^n \) and \( \mathcal{Y}_2^n \) denote the sets of \( n \)-sequences of elements from \( \mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1 \) and \( \mathcal{Y}_2 \) respectively and let \( P_n, P_1.n \) and \( P_2.n \) denote the \( n \)th memoryless extensions of \( P, P_1 \) and \( P_2 \) respectively.

The communication situation that we are mainly concerned with in this paper, as shown in Figure 1, has three message sources \( S_0, S_1 \) and \( S_2 \) which output three independent messages \( k, m, i \) for transmission. Here we assume that \( 1 \leq k \leq M_0, 1 \leq m \leq M_1 \) and \( 1 \leq i \leq M_2 \) and that each message triplet is equiprobable. The message pair \( (k, m) \) is encoded by sender (1) into
a codeword $x_{1,km} \in \mathcal{X}_1^N$ and the message pair $(k, i)$ is encoded by sender (2) into a codeword $x_{2,ki} \in \mathcal{X}_2^N$. Receiver (1) must estimate $(k, m)$ and receiver (2) $(k, i)$. The senders and receivers are not allowed to collaborate.

In this communication situation a $(N, M_0, M_1, M_2, \lambda)$-code consists of two sets $\{x_{1,km} \in \mathcal{X}_1^N; 1 \leq k \leq M_0, 1 \leq m \leq M_1\}$, $\{x_{2,ki} \in \mathcal{X}_2^N; 1 \leq k \leq M_0, 1 \leq i \leq M_2\}$ of codewords; and two sets of pairwise disjoint decoding subsets $\{A_{km} \subset \mathcal{Y}_1^N\}$, $\{B_{ki} \subset \mathcal{Y}_2^N\}$ such that

$$\frac{1}{M_0 M_1 M_2} \sum_{k, m, i} P_{1,N}(A_{km} \mid x_{1,km}, x_{2,ki}) \geq 1 - \lambda,$$

$$\frac{1}{M_0 M_1 M_2} \sum_{k, m, i} P_{2,N}(B_{ki} \mid x_{1,km}, x_{2,ki}) \geq 1 - \lambda.$$ 

Thus a $(N, M_0, M_1, M_2, \lambda)$-code is of block-length $N$ and arithmetic average error probability at each receiver less than $\lambda$. The transmission rate of this code is defined by a rate triple $(R_0, R_1, R_2)$ where

$$R_i = \frac{(\log M_i)}{N}, \quad i = 1, 2, 3.$$ 

A rate triple $(R_0, R_1, R_2)$ is said to be achievable if for every $\varepsilon > 0$ and $\lambda \in (0, 1)$, there exist $(N, M_0, M_1, M_2, \lambda)$-codes for all $N$ sufficiently large such that $M_i = \lfloor e^{N(R_i - \varepsilon)} \rfloor$ for $i = 1, 2, 3$ ($\lfloor x \rfloor$ denotes the smallest integer greater than or equal to $x$). The capacity region $\mathcal{C}(\mathcal{X})$ of the IFC $\mathcal{X}$ in this communication situation is defined to be the set of all achievable rate triples.

This paper is concerned with the problem of specifying $\mathcal{C}(\mathcal{X})$ for general IFC $\mathcal{X}$ in this special communication situation. This problem was suggested by Van der Meulen in his survey paper as Problem XVII (Van der Meulen (1977)). We shall obtain some results that go towards solving this particular problem for general IFC’s and which solves it in some special cases. Previous work (Ahlswede (1971, 1974)), Carleial (1975b, 1978) and Sato (1977) on the IFC $\mathcal{X}$ has been concerned with the communication situation where there are private messages only and no common message for transmission. That communication situation corresponds to the situation of statistically independent message sources at the two sender terminals whereas the communication situation defined here assumes a special type of correlation between these message sources. In terms of the notation here, the information theoretic problem considered in these previous works is concerned with the specification of the private messages only capacity region; that is, the set of all achievable rate pairs $(R_1, R_2)$ such that $(0, R_1, R_2) \in \mathcal{C}(\mathcal{X})$. The results that we will derive in this paper include some of these previous works as special cases, in particular the achievable rate region of Carleial (1975b, 1978), the outer bound of Sato (1977), and the limiting expression of Ahlswede (1971).
III. The Capacity Region $\mathcal{C}(\mathcal{X})$

In this section we give inner and outer bounds on $\mathcal{C}(\mathcal{X})$ as well as a less useful limiting expression for $\mathcal{C}(\mathcal{X})$ for a general IFC $\mathcal{X}$.

Consider an IFC $\mathcal{X} = (\mathcal{X}_1 \times \mathcal{X}_2, P(y_1, y_2 | x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ and a set $\mathcal{R}_i(\mathcal{X})$ of rate triples defined as follows. Define a test channel for $\mathcal{X}$ to be any finite-alphabet random vector $(Z, U_1, U_2, X_1, X_2)$ having joint probability mass function (pmf) of the form

$$\Pr[Z = z, U_1 = u_1, U_2 = u_2, X_1 = x_1, X_2 = x_2] = Q_0(z)Q_{10}(u_1 | z)Q_{20}(u_2 | z)Q_{1}(x_1 | u_1)Q_{2}(x_2 | u_2),$$

where $X_1$ and $X_2$ have alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$ respectively and where $Q_0$ is a pmf and $Q_{10}, Q_{20}, Q_1, Q_2$ are conditional pmfs. Note that (7) implies that $(Z, (U_1, U_2), (X_1, X_2))$ is a Markov Chain. Let $\mathcal{D}_i$ be the set of all such test channels $(Z, U, X)$, where $U = (U_1, U_2)$, $X = (X_1, X_2)$. Furthermore let $\mathcal{D}_i(\mathcal{X})$ be the set of all random vectors $(Z, U, X, Y)$ such that (i) $(Z, U, X) \in \mathcal{D}_i$, (ii) $Y_1$ and $Y_2$ have alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ respectively, where $Y = (Y_1, Y_2)$, (iii) $(Z, U, X, Y)$ is a Markov Chain such that $\Pr[Y_1 = y_1, Y_2 = y_2 | X_1 = x_1, X_2 = x_2] = P(y_1, y_2 | x_1, x_2)$. For a given $(Z, U, X, Y) \in \mathcal{D}_i(\mathcal{X})$, denote

$$a_1 = \min[I(U_1; Y_1 | Z), I(U_1; Y_2 | Z)],$$

$$b_1 = \min[I(U_2; Y_1 | U_1Z), I(U_2; Y_2 | U_2Z)],$$

$$a_2 = \min[I(U_1; Y_1 | U_2Z), I(U_1; Y_2 | U_2Z)],$$

$$b_2 = \min[I(U_2; Y_1 | Z), I(U_2; Y_2 | Z)],$$

$$a_3 = \min[I(U_1; Y_1 | Z), I(U_1; Y_2 | U_2Z)],$$

$$b_3 = \min[I(U_2; Y_1 | U_1Z), I(U_2; Y_2 | Z)],$$

$$a_4 = \min[I(U_1; Y_1 | U_2Z), I(U_1; Y_2 | Z)],$$

$$b_4 = \min[I(U_2; Y_1 | Z), I(U_2; Y_2 | U_1Z)],$$

where the average mutual informations $I(U_i; Y_i | Z)$, etc. are given as in Gallager (1968). Next define for $i = 1, 2, 3, 4, \mathcal{R}_i(\mathcal{X})$ to be the set of all rate triples $(R_0, R_1, R_2)$ satisfying the following inequality constraints (9) for some $(Z, U, X, Y) \in \mathcal{D}_i(\mathcal{X})$.

$$R_1 \leq I(X_1; Y_1 | U_1 U_2) + a_i,$$

$$R_2 \leq I(X_2; Y_2 | U_1 U_2) + b_i,$$

$$R_0 + R_1 + \frac{b_i}{I(X_2; Y_2 | U_1 U_2) + b_i} R_2 \leq I(U_2X_1; Y_1),$$

$$R_0 + R_2 + \frac{a_i}{I(X_1; Y_1 | U_1 U_2) + a_i} R_1 \leq I(U_1X_2; Y_2).$$
Finally define
\[ \mathcal{R}_I(\mathcal{K}) = \overline{\text{co}} \left( \bigcup_{i=1}^{4} \mathcal{R}_i(\mathcal{K}) \right), \]  
(10)
where \( \overline{\text{co}}(A) \) denotes the closed convex hull of \( A \). The following theorem, which states that \( \mathcal{R}_I(\mathcal{K}) \) is an inner bound to \( \mathcal{C}(\mathcal{K}) \) for general IFC \( \mathcal{K} \), is a main result of this paper. We shall defer until Section V the random coding proof of this theorem.

**Theorem 1.** Every rate triple in \( \mathcal{R}_I(\mathcal{K}) \) is achievable for general IFC \( \mathcal{K} \). Hence \( \mathcal{R}_I(\mathcal{K}) \subseteq \mathcal{C}(\mathcal{K}) \).

In the communication situation where there are private messages only and no common message present for transmission, Carleial (1975b, 1978) obtained a set of achievable rate pairs \( (R_1, R_2) \) such that \( (0, R_1, R_2) \in \mathcal{C}(\mathcal{K}) \) defined as follows. For \( i = 1, 2, 3, 4 \), let \( \mathcal{R}'_I(\mathcal{K}) \) be the set of all rate triples in \( \mathcal{R}_I(\mathcal{K}) \) satisfying the inequality constraints (9) for some \( (Z, U, X, Y) \in \mathcal{D}_I(\mathcal{K}) \) such that \( Z = \text{constant almost surely} \). Let \( \mathcal{R}'_I(\mathcal{K}) \) be defined as \( \mathcal{R}_I(\mathcal{K}) \) in (10) with the \( \mathcal{R}_I(\mathcal{K}) \) replaced by \( \mathcal{R}_I(\mathcal{K}) \). Then Carleial’s set of achievable rate pairs is precisely the set of all \( (R_1, R_2) \) such that \( (0, R_1, R_2) \in \mathcal{R}'_I(\mathcal{K}) \). Hence Theorem 1 contains this result of Carleial as a special case. Moreover an examination of the extreme points in each \( \mathcal{R}_I(\mathcal{K}) \) indicates that under the restriction \( R_0 = 0 \), the inner bound \( \mathcal{R}_I(\mathcal{K}) \) is exactly Carleial’s achievable rate region: \( \mathcal{R}_I(\mathcal{K}) \) with the restriction \( R_0 = 0 \).

In the private messages only communication situation, Sato (1977) has recently obtained an outer bound on the set of all \( (R_1, R_2) \) such that \( (0, R_1, R_2) \in \mathcal{C}(\mathcal{K}) \). Let us now consider the generalization of Sato’s result to obtain an outer bound on \( \mathcal{C}(\mathcal{K}) \). For an IFC \( \mathcal{K} \), let \( \mathcal{D}_0(\mathcal{K}) \) be the set of all \( (Z, X, Y) \) such that \( (Z, U, X, Y) \in \mathcal{D}_I(\mathcal{K}) \) where \( U_1 = U_2 = Z \) almost surely. Furthermore let \( \mathcal{R}_0(\mathcal{K}) \) be the set of all rate triples \( (R_0, R_1, R_2) \) satisfying the following inequality constraints (11) for some \( (Z, X, Y) \in \mathcal{D}_0(\mathcal{K}) \).

\[ R_1 \leq I(X_1; Y_1 | X_2 Z), \]
\[ R_2 \leq I(X_2; Y_2 | X_1 Z), \]
\[ R_1 + R_2 \leq I(X_1 X_2; Y_1 Y_2 | Z), \]
\[ R_0 + R_1 + R_2 \leq I(X_1 X_2; Y_1 Y_2). \]

Then the following proposition gives an outer bound to \( \mathcal{C}(\mathcal{K}) \). The proof of this proposition, which is similar to Sato’s proof, is given in Appendix A.

**Proposition 1.** \( \mathcal{C}(\mathcal{K}) \subseteq \overline{\text{co}}(\mathcal{R}_0(\mathcal{K})) \) for general IFC \( \mathcal{K} \).
Note from (3) and (4) that the error probabilities depend only on the marginal transition probabilities of the channel and not on the joint channel transition probabilities. Hence the notion of achievable rate triples and the capacity region \( C(\mathcal{H}) \) also depend only on the marginal channel transition probabilities. Thus all IFC \( \mathcal{H} \) with the same marginal channel transition probability functions \( P_1 \) and \( P_2 \) given by (1) and (2) have the same capacity region \( C(\mathcal{H}) \). Now it is clear from (11) that the outer bound \( \overline{c}(R_0(\mathcal{H})) \) depends on the joint channel transition probabilities of \( \mathcal{H} \). Hence using the technique of Sato (1977), a tighter outer bound on \( C(\mathcal{H}) \) for a given IFC \( \mathcal{H} \) can be obtained by using the intersection of \( \overline{c}(R_0(\mathcal{H}')) \) over all IFC \( \mathcal{H}' \) with the same marginal transition probabilities as \( \mathcal{H} \). Thus we have the following theorem.

**Theorem 2.** For general IFC \( \mathcal{H} \),

\[
C(\mathcal{H}) \subseteq \bigcap_{\mathcal{H}'} \overline{c}(R_0(\mathcal{H}')), \tag{12}
\]

where the intersection in (12) is over all IFC \( \mathcal{H}' \) with the same marginal transition probability functions \( P_1 \) and \( P_2 \) as \( \mathcal{H} \).

Finally we note that Sato’s (1977) outer bound on the private messages only capacity region is obtained by restricting \( R_0 = 0 \) and \( Z = \text{constant} \) almost surely in the definition (11) for \( R_0(\mathcal{H}) \).

The inner and outer bounds on \( C(\mathcal{H}) \) given in Theorems 1 and 2 involve optimization of single-letter average mutual informations. We have not been able to give a similar characterization of \( C(\mathcal{H}) \) in terms of single-letter average mutual information quantities. It appears that even for the private messages only case, a characterization of the capacity region in terms of single-letter average mutual information is still not known. However, Ahlswede (1971) has obtained a limiting expression for the private messages only capacity region. We next show how Ahlswede’s limiting expression can be modified to characterize \( C(\mathcal{H}) \). Although such limiting expressions appear to be of little computational use, we shall give this result for the sake of completeness.

In order to give this limiting expression for \( C(\mathcal{H}) \), let us consider for a given IFC \( \mathcal{H} \) its \( n \)th extension \( \mathcal{H}^n = \{X_1^n \times X_2^n, \ P_n(y_1, y_2 | x_1, x_2), \ Y_1^n \times Y_2^n \} \). So \( \mathcal{H}^n \) is an IFC with sender (1) and (2) channel input alphabets \( X_1^n \) and \( X_2^n \) respectively, receiver (1) and (2) channel output alphabets \( Y_1^n \) and \( Y_2^n \) respectively and joint channel transition probability function \( P_n \). Hence we may define for \( \mathcal{H}^n \) the set \( D_0(H)^n \) of all random vectors \( (Z, X, Y) \) where \( X = (X_1, X_2) \) and \( Y = (Y_1, Y_2) \) analogous to the specification of \( D_0(H) \) used in (11). Here of course each \( (Z, X, Y) \in D_0(\mathcal{H}) \) is such that \( X_1 \) and \( X_2 \) have alphabets \( X_1^n \) and \( X_2^n \) respectively, \( Y_1 \) and \( Y_2 \) have alphabets \( Y_1^n \) and \( Y_2^n \) respectively, and \( (Z, X, Y) \) has joint pmf of the form \( Q_0(z) \ P_1(x_1 | z) \ Q_2(x_2 | z) \ P_n(y_1, y_2 | x_1, x_2) \). Define \( \mathcal{H}^n(\mathcal{H}^n) \) to be the set of all rate triples
(R_0, R_1, R_2) satisfying the following inequality constraints (13) for some (Z, X, Y) \in D_0 \mathcal{R}^n.

\begin{align*}
R_1 & \leq I(X_1; Y_1 | Z), \\
R_0 + R_1 & \leq I(ZX_1; Y_1), \\
R_2 & \leq I(X_2; Y_2 | Z), \\
R_0 + R_2 & \leq I(ZX_2; Y_2).
\end{align*}

(13)

Now it is easy to see that Theorem 1 gives \( R^n(\mathcal{H}) \subset C(\mathcal{H}) \) which then implies that \( (1/n) R^n(\mathcal{H}) \subset C(\mathcal{H}) \). For example when \( n = 1 \), \( R(\mathcal{H}) \subset R_1(\mathcal{H}) \) since (13) is a special case of (9) when \( U_1 = U_2 = Z \) almost surely. Since \( R_2(\mathcal{H}) \subset C(\mathcal{H}) \) by Theorem 1, so \( R(\mathcal{H}) \subset C(\mathcal{H}) \). The case when \( n > 1 \) is similarly proved. This shows that \( \bigcup_{n=1}^{\infty} \text{co}(1/n) R^n(\mathcal{H}) \subset C(\mathcal{H}) \). In Appendix B the converse statement is proved which then establishes the following theorem.

**Theorem 3.** For general IFC \( \mathcal{H} \),

\[ C(\mathcal{H}) = \bigcup_{n=1}^{\infty} \text{co}\left(\frac{1}{n} R^n(\mathcal{H})\right). \]  

(14)

**IV. Special Cases**

In this section we consider two special cases of the IFC: the twin channel IFC and the separate channels IFC (terminology due to Sato (1977)). In the first case the inner and outer bounds on \( C(\mathcal{H}) \) described in the last section coincide. In the second case the inner bound is equal to \( C(\mathcal{H}) \), but the outer bound is generally not tight.

Let us first consider the twin channel IFC. An IFC \( \mathcal{H} = (X_1 \times X_2, P(y_1, y_2 | x_1, x_2), Y_1 \times Y_2) \) is called a twin channel if the two receiver outputs are equal almost surely; that is, if

\[ P(y_1, y_2 | x_1, x_2) = \delta(y_1, y_2) P_1(y_1 | x_1, x_2), \]  

(15)

where \( \delta \) is the Kronecker delta function. A twin channel IFC \( \mathcal{H} \) can be viewed as being composed of two identical multiple access channels (Ahlswede (1974)) and hence its capacity region \( C(\mathcal{H}) \) is known from the previous work of Slepian and Wolf (1973). Let us first examine the outer bound \( \text{co}(R_0(\mathcal{H})) \) on \( C(\mathcal{H}) \). In view of (15) we must have \( Y_1 = Y_2 = Y \) in (11) and hence \( R_0(\mathcal{H}) \) is the set of all rate triples satisfying the following inequalities (16) for some \( (Z, X, Y) \in D_0 \mathcal{H} \).
\displaymath
\begin{align*}
R_1 & \leq I(X_1; Y \mid X_2Z), \\
R_2 & \leq I(X_2; Y \mid X_1Z), \\
R_1 + R_2 & \leq I(X_1X_2; Y \mid Z), \\
R_0 + R_1 + R_2 & \leq I(X_1X_2Y).
\end{align*}

(16)

Let us now examine the inner bound \( \mathcal{R}_1(\mathcal{H}) \) to \( \mathcal{C}(\mathcal{H}) \). Let \( \mathcal{T}_1(\mathcal{H})[\mathcal{T}_2(\mathcal{H})] \) be the set of all rate triples satisfying the following inequalities (17) [(18)] for some \( (Z, X, Y) \in \mathcal{D}_0 \mathcal{H} \) (here \( Y = (Y_1, Y_2) \) and \( Y_1 = Y_2 = Y \) a.s., because of (15)).

\displaymath
\begin{align*}
R_1 & \leq I(X_1; Y \mid Z), \\
R_2 & \leq I(X_2; Y \mid X_1Z), \\
R_0 + R_1 + R_2 & \leq I(X_1X_2Y). \\
R_1 & \leq I(X_1; Y \mid X_2Z), \\
R_2 & \leq I(X_2; Y \mid Z), \\
R_0 + R_1 + R_2 & \leq I(X_1X_2Y). \\
\end{align*}

(17)

(18)

Then the inequalities (17) [(18)] are a special case of (9) when \( i = 1 \) \( [i = 2] \) with \( U_1 = X_1 \) and \( U_2 = X_2 \) almost surely. Hence \( \mathcal{T}_1(\mathcal{H}) \subset \mathcal{R}_1(\mathcal{H}) \) and \( \mathcal{T}_2(\mathcal{H}) \subset \mathcal{R}_2(\mathcal{H}) \), and by Theorem 1 \( \overline{\text{co}}(\mathcal{T}_1(\mathcal{H}) \cup \mathcal{T}_2(\mathcal{H})) \subset \mathcal{R}_1(\mathcal{H}) \subset \mathcal{C}(\mathcal{H}) \subset \overline{\text{co}}(\mathcal{R}_0(\mathcal{H})). \) However, it can be easily shown by using a simple convexity argument that \( \overline{\text{co}}(\mathcal{T}_1(\mathcal{H}) \cup \mathcal{T}_2(\mathcal{H})) = \overline{\text{co}}(\mathcal{R}_0(\mathcal{H})). \) Thus we have proved the following theorem, which agrees with the work of Slepian and Wolf (1973).

**Theorem 4.** For a twin channel IFC \( \mathcal{H} \), \( \mathcal{R}_1(\mathcal{H}) = \overline{\text{co}}(\mathcal{R}_0(\mathcal{H})) = \mathcal{C}(\mathcal{H}). \)

An IFC in which there is no mutual interference between the two users is said to have separate channels. That is, a separate channels IFC \( \mathcal{H} = (\mathcal{X}_1 \times \mathcal{X}_2, P(y_1, y_2 \mid x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2) \) has the property that

\displaymath
\begin{align*}
P_1(y_1 \mid x_1, x_2) & = P_1(y_1 \mid x_1), \\
P_2(y_2 \mid x_1, x_2) & = P_2(y_2 \mid x_2).
\end{align*}

(19)

Let us examine the inner bound \( \mathcal{R}_1(\mathcal{H}) \) to the capacity region of a separate channels IFC \( \mathcal{H} \). Let \( \mathcal{T}_1(\mathcal{H}) \) be the set of all rate triples satisfying the following inequalities (20) for some \( (Z, X, Y) \in \mathcal{D}_0 \mathcal{H} \).

\displaymath
\begin{align*}
R_1 & \leq I(X_1; Y_1 \mid Z), \\
R_2 & \leq I(X_2; Y_2 \mid Z), \\
R_0 + R_1 & \leq I(X_1; Y_1), \\
R_0 + R_2 & \leq I(X_2; Y_2).
\end{align*}

(20)
Now \( I(ZX_1; Y_i) = I(X_1; Y_i) \) for \((Z, X, Y) \in \mathcal{D}_0 \mathcal{H}\) because of (19). Hence (20) is a special case of (9) with \( U_1 = U_2 = Z \) almost surely. Hence by Theorem 1, \( \overline{\operatorname{co}}(\mathcal{F}_i(\mathcal{H})) \subset \mathcal{R}(\mathcal{H}) \subset \mathcal{C}(\mathcal{H}) \). The following theorem which is proved in Appendix B states that \( \overline{\operatorname{co}}(\mathcal{F}_i(\mathcal{H})) \) is actually equal to \( \mathcal{C}(\mathcal{H}) \).

**THEOREM 5.** For a separate channels IFC \( \mathcal{H} \), \( \mathcal{C}(\mathcal{H}) = \mathcal{R}(\mathcal{H}) = \overline{\operatorname{co}}(\mathcal{F}_i(\mathcal{H})) \).

Note that restricting \( R_0 = 0 \) and \( Z = \) constant almost surely in the definition (20) for \( \mathcal{F}_i(\mathcal{H}) \) results in the private messages only capacity region of the separate channels IFC derived by Sato (1977). It turns out that in this case the intersection outer bound \( \bigcap_{\mathcal{H}'} \overline{\operatorname{co}}(\mathcal{R}_0(\mathcal{H}')) \) on \( \mathcal{C}(\mathcal{H}) \) given in Theorem 2 is not generally tight. To see this let \( C_1 \) and \( C_2 \) be the respective capacities of discrete memoryless channels with transition probabilities \( P_1(y_1 | x_1) \) and \( P_2(y_2 | x_2) \) given by (19). Now, it follows from (11) that the maximum \( R_0 \) rate in the intersection bound of Theorem 2 is

\[
\min_{\mathcal{H}'} \max_{(Z, X, Y) \in \mathcal{D}_0 \mathcal{H}'} I(X_1X_2 ; Y_1Y_2) \leq I(X_1; Y_1) + I(X_2; Y_2) \text{ for } (Z, X, Y) \in \mathcal{D}_0 \mathcal{H}'.
\]

Hence \( \max_{(Z, X, Y) \in \mathcal{D}_0 \mathcal{H}'} I(X_1X_2 ; Y_1Y_2) \geq C_1 + C_2 \) for all \( \mathcal{H}' \). Hence the maximum \( R_0 \) rate in the intersection outer bound of Theorem 2 is at least \( C_1 + C_2 \). However it is clear from (20) and Theorem 5 that the maximum possible \( R_0 \) rate in \( \mathcal{C}(\mathcal{H}) \) is \( \min(C_1, C_2) \). Thus the outer bound is not tight in general for separate channels IFC.

V. PROOF OF THEOREM 1

In order to prove Theorem 1, let us consider the following somewhat redundant communication situation involving the IFC \( \mathcal{H} \). This situation, as shown in Figure 2, has five message sources \( S_0 \), \( S_{11} \), \( S_{10} \), \( S_{22} \), \( S_{20} \) which output independent messages \( k, m_1, m_2, i_1, i_2 \) respectively where \( 1 \leq k \leq M_0 \),

![Fig. 2. Communication Situation for capacity region \( \mathcal{C}^*(\mathcal{H}) \).](image)
1 \leq m_1 \leq M_{11}, 1 \leq m_2 \leq M_{10}, 1 \leq i_1 \leq M_{22}, 1 \leq i_2 \leq M_{20}, \text{ and all message quintuplets are equiprobable. The message triplet } (k, m_1, m_2) \text{ is encoded by sender (1) into a codeword } x_{1,km_1m_2} \in \mathcal{X}_1^N \text{ and the message triplet } (k, i_1, i_2) \text{ is encoded by sender (2) into a codeword } x_{2,ki_1i_2} \in \mathcal{X}_2^N. \text{ Receiver (1) must estimate } (h, m_1, m_2, i_2) \text{ and receiver (2) must estimate } (k, i_1, i_2, m_2). \text{ Hence here the } S_0 \text{ source emits a common message that is transmitted by both senders to its corresponding receiver. However the } S_{10} \text{ source emits a message that is transmitted by sender (1) only but to both receivers. The } S_{11} \text{ source emits a private message that is transmitted by sender (i) to receiver (i). So the difference between this situation and the communication situation described in Section II is the presence of the } S_{10} \text{ and } S_{11} \text{ message sources.}

In this communication situation, a \((N, M_0, M_{11}, M_{10}, M_{22}, M_{20}, \lambda)\)-code consists of two sets \(\{x_{1,km_1m_2} \in \mathcal{X}_1^N: 1 \leq k \leq M_0, 1 \leq m_1 \leq M_{11}, 1 \leq m_2 \leq M_{10}\}\), \(\{x_{2,ki_1i_2} \in \mathcal{X}_2^N: 1 \leq k \leq M_0, 1 \leq i_1 \leq M_{22}, 1 \leq i_2 \leq M_{20}\}\) of codewords; and two sets of pairwise disjoint decoding subsets \(\{A_{km_1m_2} \subset \mathcal{Y}_1^N\}, \{B_{ki_1i_2} \subset \mathcal{Y}_2^N\}\) so that

\[
\frac{1}{M_0M_{11}M_{10}M_{22}M_{20}} \sum_{k, m_1, m_2, i_1, i_2} P_{1,N}(A_{km_1m_2} | x_{1,km_1m_2}, x_{2,ki_1i_2}) \geq 1 - \lambda, \tag{21}
\]

\[
\frac{1}{M_0M_{11}M_{10}M_{22}M_{20}} \sum_{k, m_1, m_2, i_1, i_2} P_{2,N}(B_{ki_1i_2} | x_{1,km_1m_2}, x_{2,ki_1i_2}) \geq 1 - \lambda.
\]

A rate quintuple \((R_0, R_{11}, R_{10}, R_{22}, R_{20})\) is then said to be achievable if for every \(\epsilon > 0\) and \(\lambda \in (0, 1)\), there exist \((N, M_0, M_{11}, M_{10}, M_{22}, M_{20}, \lambda)\)-codes for all \(N\) sufficiently large such that \(M_0 = \lfloor e^{N(R_0 - \epsilon)} \rfloor\) and \(M_i = \lfloor e^{N(R_i - \epsilon)} \rfloor\) for \((i,j) \in \{11, 10, 22, 20\}\). The capacity region \(\mathcal{C}^*(\mathcal{K})\) of the IFC \(\mathcal{K}\) in this communication situation is then defined to be the set of all achievable rate quintuples.

Referring to the definitions in Section II, it is clear that \(\mathcal{C}(\mathcal{K})\) is precisely the intersection of \(\mathcal{C}^*(\mathcal{K})\) with the \(R_{10} = 0\) and \(R_{20} = 0\) rate planes. The following proposition gives another characterization of \(\mathcal{C}(\mathcal{K})\) in terms of \(\mathcal{C}^*(\mathcal{K})\) which is useful for proving Theorem 1.

**Proposition 2.** Let \(\mathcal{C}(\mathcal{K})\) be the set of all rate triples \((R_0, R_1, R_2)\) that satisfy the following constraints (22) for some \(\alpha_1, \alpha_2 \in [0, 1]\) and some \((R_0^*, R_{11}^*, R_{10}^*, R_{22}^*, R_{20}^*) \in \mathcal{C}^*(\mathcal{K})\).

\[
R_0 = R_0^*, \quad \alpha_1 R_1 = R_{10}^*, \quad \alpha_2 R_2 = R_{20}^*, \quad (1 - \alpha_1)R_1 = R_{11}^*, \quad (1 - \alpha_2)R_2 = R_{22}^*. \tag{22}
\]

Then \(\mathcal{C}(\mathcal{K}) = \mathcal{C}(\mathcal{K})\).
Proof. It is clear that $\mathcal{C}(\mathcal{X}) \subset \mathcal{C}(\mathcal{X})$ since the set of all rate triples satisfying (22) for some rate quintuple in $\mathcal{C}(\mathcal{X})$ so that $\alpha_1 = \alpha_3 = 0$ is precisely the intersection of $\mathcal{C}(\mathcal{X})$ with the $R_{10} = 0$ and $R_{20} = 0$ rate planes. In Appendix D it is proved that $\mathcal{C}(\mathcal{X}) \subset \mathcal{C}(\mathcal{X})$. Q.E.D.

Our strategy now to prove Theorem 1 will be to first derive an inner bound to $\mathcal{C}(\mathcal{X})$ and then to apply the implication $\mathcal{C}(\mathcal{X}) \subset \mathcal{C}(\mathcal{X})$ of Proposition 2 on this inner bound to obtain $\mathcal{R}_1(\mathcal{X})$. Define for $i = 1, 2, 3, 4$, $\mathcal{R}_i(\mathcal{X})$ to be the set of all rate quintuples $(R_0, R_{11}, R_{10}, R_{22}, R_{20})$ satisfying the following inequality constraints (23) for some $(Z, U, X, Y) \in \mathcal{D}_1\mathcal{X}$:

\begin{align*}
R_{10} &\leq a_i \\
R_{20} &\leq b_i \\
R_{12} &\leq I(X_1; Y_1 | U_1 U_2), \\
R_{22} &\leq I(X_2; Y_2 | U_1 U_2),
\end{align*}

(23) where $a_i$ and $b_i$ are given in (8). Next define

$$\mathcal{R}_i^*(\mathcal{X}) = \text{co} \left( \bigcup_{i=1}^{4} \mathcal{R}_i(\mathcal{X}) \right).$$

(24)

The following theorem which states that every rate quintuple in $\mathcal{R}_i^*(\mathcal{X})$ is achievable is proved in Appendix C.

**Theorem 6.** Every rate quintuple in $\mathcal{R}_i^*(\mathcal{X})$ is achievable for a general IFC $\mathcal{X}$. Hence $\mathcal{R}_i^*(\mathcal{X}) \subset \mathcal{C}(\mathcal{X})$.

Now note from (9), (22) and (23) that every rate triple which satisfies (9) can be written in the form (22) for

$$\begin{align*}
\alpha_1 &= \frac{a_i}{I(X_1; Y_1 | U_1 U_2)} + a_i, \\
\alpha_2 &= \frac{b_i}{I(X_2; Y_2 | U_1 U_2)} + b_i,
\end{align*}$$

(25)

and for some rate quintuple satisfying (23). Hence Theorem 1 is a direct consequence of the implication $\mathcal{C}(\mathcal{X}) \subset \mathcal{C}(\mathcal{X})$ in Proposition 2 and Theorem 6. This then concludes its proof. This proof implicitly gives the rationale for the inequality constraints (9) that specify the rate region $\mathcal{R}_i(\mathcal{X})$. Operationally the rate triples in $\mathcal{R}_i(\mathcal{X})$ are attained by the additional transmission of a portion.
\( \alpha_1 \) of the private \( S_1 \) message intended for receiver (1) to receiver (2) and a portion \( \alpha_2 \) of the private \( S_2 \) message intended for receiver (2) to receiver (1), where \( \alpha_1 \) and \( \alpha_2 \) are given in (25).

VI. SUMMARY

Summarizing, we have obtained inner and outer bounds on the capacity region of two-user discrete memoryless interference channels (IFC) in the communication situation where private messages are sent by each sender as well as a common message by both senders to its corresponding receiver. A limiting expression for the capacity region was derived. In the special case of the twin user IFC the inner and outer bounds coincide and give the capacity region. However, the inner bound is tight and the outer bound is generally not tight for the special case of separate channels IFC. An interesting speculation is whether the inner bound is tight in general. We are not optimistic that this is so. This is because the inner bound under the restriction \( R_0 = 0 \) is the inner bound obtained by Carleial (1975b, 1978) for the private messages only communication situation. Although it does not appear to be known whether Carleial’s inner bound is tight for discrete memoryless IFC’s, Carleial (1975b, 1978) has shown that his inner bound is not tight for the Gaussian IFC. This fact lends to our pessimism regarding the tightness of our inner bound.

One question that we have not yet resolved is whether the size of the \( ZU_1U_2 \) alphabets may be bounded in the specification (9) of \( \mathcal{R}(\mathcal{X}) \) and whether the size of the \( Z \) alphabet may be bounded in the specification (11) of \( \mathcal{C}(\mathcal{R}_0(\mathcal{X})) \). In each case computation of the inner and outer bound would be simpler if these alphabets can be constrained in size. Hence this is an open problem of considerable importance.

APPENDIX A

Proof of Proposition 1

We shall omit many of the detailed steps of the proof that are virtually identical to techniques used by Slepian and Wolf (1973). Consider for the moment any \((N, M_0, M_1, M_2, \lambda)\)-code with codewords \( \{x_{1,km}\}_{k=1,m=1}^{M_1,M_2} \) and \( \{x_{2,klh}\}_{h=1,m=1}^{M_0,M_2} \). This code attains a transmission rate triple \((R_0, R_1, R_2)\) given by (5). Let \( S_0, S_1, S_2 \) be random variables denoting the outputs of the three message sources, \( X_i = (X_{i1},...,X_{iN}) \) be a random vector representing the output of the sender \((i)\) encoder and \( Y_i = (Y_{i1},...,Y_{iN}) \) denote the corresponding received vector at receiver \((i)\). Hence the random vector \((S_0, S_1, S_2, X_1, X_2, Y_1, Y_2)\) has joint pmf
\[
\begin{align*}
\Pr[S_0 = k, S_1 = m, S_2 = i, X_1 = x_1, X_2 = x_2, Y_1 = y_1, Y_2 = y_2] &= P_{S_0}(k) P_{S_1}(m) P_{S_2}(i) P_{X_1|S_0,S_1}(x_1 | k, m) P_{X_2|S_0,S_2}(x_2 | k, i) \\
&\cdot P_{Y_1,Y_2}(y_1, y_2 | x_1, x_2),
\end{align*}
\]

where \( P_n \) is the \( n \)th memoryless extension of the joint channel transition probability function \( P \) and where \( P_{S_0}, P_{S_1}, P_{S_2} \) are uniform probability mass functions and

\[
P_{X_1|S_0,S_1}(x_1 | k, m) = \begin{cases} 1 & \text{if } x_1 = x_{1,km} \\ 0 & \text{otherwise} \end{cases},
\]

\[
P_{X_2|S_0,S_2}(x_2 | k, i) = \begin{cases} 1 & \text{if } x_2 = x_{2,ki} \\ 0 & \text{otherwise} \end{cases}.
\]

Furthermore let \((S_0^*, S_1^*)\) be the receiver (1) estimate of the transmitted \((S_0, S_1)\) message pair and \((S_0', S_2')\) be the receiver (2) estimate of the transmitted \((S_0, S_2)\) message pair. Hence \(S_0^*\) and \(S_1^*\) are random variables that are functions of \(Y_1\) and \(S_0\) and \(S_2^*\) are random variables that are functions of \(Y_2\). Next define

\[
\begin{align*}
P_{e,1} &= \Pr[S_1^* \neq S_1], \\
P_{e,2} &= \Pr[S_2^* \neq S_2], \\
P_{e,3} &= \Pr[S_1^* \neq S_1 \text{ or } S_2^* \neq S_2], \\
P_{e,4} &= \Pr[S_0^* \neq S_0 \text{ or } S_1^* \neq S_1 \text{ or } S_2^* \neq S_2].
\end{align*}
\]

Thus, we must have \(P_{e,1} \leq \Pr[S_0^* \neq S_0 \text{ or } S_1^* \neq S_1] \leq \lambda\) and also \(P_{e,2} \leq \lambda\) and \(P_{e,3} \leq P_{e,4} \leq 2\lambda\) by a similar argument. Finally, let \(h(\alpha) = -\alpha \log \alpha - (1 - \alpha) \log(1 - \alpha)\) be the binary entropy function.

First, by using Fano inequality arguments similar to that establishing Lemmas 1–3 of Slepian and Wolf (1973), it is easy to show that

\[
(1 - P_{e,1}) \log M_1 \leq h(P_{e,1}) + I(X_1; Y_1 | X_2 S_0),
\]

\[
(1 - P_{e,2}) \log M_2 \leq h(P_{e,2}) + I(X_2; Y_2 | X_1 S_0),
\]

\[
(1 - P_{e,3}) \log M_1 M_2 \leq h(P_{e,3}) + I(X_1X_2; Y_1Y_2 | S_0),
\]

\[
(1 - P_{e,4}) \log M_0 M_1 M_2 \leq h(P_{e,4}) + I(X_1X_2Y_1Y_2).
\]

Next using arguments similar to that establishing Lemma 4 of Slepian and Wolf (1973), it is easy to show that

\[
I(X_1; Y_1 | X_2 S_0) \leq \sum_{n=1}^{N} I(X_{1n}; Y_{1n} | X_{2n} S_0),
\]
\[ I(\mathbf{X}_2; \mathbf{Y}_2 | \mathbf{X}_1; S_0) \leq \sum_{n=1}^{N} I(X_{2n}; Y_{2n} | X_{1n}; S_0), \]

\[ I(\mathbf{X}_1; \mathbf{Y}_2 | S_0) \leq \sum_{n=1}^{N} I(X_{1n}; X_{2n}; Y_{1n} Y_{2n} | S_0), \]

\[ I(\mathbf{X}_1; \mathbf{Y}_2) \leq \sum_{n=1}^{N} I(X_{1n}; X_{2n}; Y_{1n} Y_{2n}). \quad (A.3) \]

So using (5), (A.2) and (A.3) we have

\[ P_{e,1} \geq 1 - 1/N R_1 - (1/N) \sum_{n=1}^{N} I(X_{1n}; \mathbf{Y}_1 | X_{2n}; S_0)/R_1, \]

\[ P_{e,2} \geq 1 - 1/N R_2 - (1/N) \sum_{n=1}^{N} I(X_{2n}; \mathbf{Y}_2 | X_{1n}; S_0)/R_2; \]

\[ P_{e,3} \geq 1 - 1/N (R_1 + R_2) - (1/N) \sum_{n=1}^{N} I(X_{1n}; X_{2n}; Y_{1n} Y_{2n} | S_0)/(R_1 + R_2), \]

\[ P_{e,d} \geq 1 - 1/N (R_0 + R_1 + R_2)
- (1/N) \sum_{n=1}^{N} I(X_{1n}; X_{2n}; Y_{1n} Y_{2n})/(R_0 + R_1 + R_2). \quad (A.4) \]

Next consider a random vector \( T \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_2 \mathbf{X}_2 \mathbf{Y}_1 \mathbf{Y}_2 \) satisfying the following conditions:

1. \( T \) and \( Z \) are independent.
2. \( T \) takes on equiprobable values in \( \{1, \ldots, N\} \) and \( Z \) is distributed as \( S_0 \).
3. \( \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_1, \mathbf{Y}_2 \) take on values in \( \mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2 \) respectively.
4. \( ((T \mathbf{Z}_1), (\mathbf{X}_1), (\mathbf{Y}_1)) \) is Markov so that \( \Pr[\mathbf{X}_1 = x_1, \mathbf{X}_2 = x_2 | \mathbf{Z}_1 = z_1, \mathbf{Z}_2 = z_2] = P(y_1, y_2 | x_1, x_2), \Pr[\mathbf{X}_1 = x_1, \mathbf{X}_2 = x_2 | T = t, Z = k] = \Pr[\mathbf{X}_1 = x_1 | T = t, Z = k] \cdot \Pr[\mathbf{X}_2 = x_2 | T = t, Z = k] \) where \( \Pr[\mathbf{X}_i = x_i | T = t, Z = k] = \Pr[X_{in} = x_i | S_0 = k] \).

Now it follows from these conditions that \( ((T \mathbf{Z}_1), \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_1, \mathbf{Y}_2) \in \mathcal{D}_0 \mathcal{X} \). Also it is easy to show that

\[ I(\mathbf{X}_1^c; \mathbf{Y}_1 | \mathbf{X}_2; (T \mathbf{Z}_1)) = (1/N) \sum_{n=1}^{N} I(X_{1n}; Y_{1n} | X_{2n}; S_0), \]

\[ I(\mathbf{X}_2^c; \mathbf{Y}_2 | \mathbf{X}_1; (T \mathbf{Z}_1)) = (1/N) \sum_{n=1}^{N} I(X_{2n}; Y_{2n} | X_{1n}; S_0), \]

\[ I(\mathbf{X}_1^c \mathbf{X}_2^c; \mathbf{Y}_1 \mathbf{Y}_2 | (T \mathbf{Z}_1)) = (1/N) \sum_{n=1}^{N} I(X_{1n} X_{2n}; Y_{1n} Y_{2n} | S_0). \quad (A.5) \]
Moreover

\[ I(\tilde{X}_1 \tilde{X}_2; \tilde{Y}_1 \tilde{Y}_2) = I(TZ \tilde{X}_1 \tilde{X}_2; \tilde{Y}_1 \tilde{Y}_2) \]

\[ \geq I(Z \tilde{X}_1 \tilde{X}_2; \tilde{Y}_1 \tilde{Y}_2 | T) \]

\[ = (1/N) \sum_{n=1}^{N} I(X_{1n}X_{2n}; Y_{1n}Y_{2n}). \quad (A.6) \]

Hence since \(((TZ), \tilde{X}_1, \tilde{X}_2, \tilde{Y}_1, \tilde{Y}_2) \in \mathcal{D}_0 \mathcal{H}\), it follows from (11), (A.5) and (A.6) that if \((R_0, R_1, R_2)\) satisfies

\[ R_1 \leq (1/N) \sum_{n=1}^{N} I(X_{1n}; Y_{1n} | X_{2n}S_0), \]

\[ R_2 \leq (1/N) \sum_{n=1}^{N} I(X_{2n}; Y_{2n} | X_{1n}S_0), \]

\[ R_1 + R_2 \leq (1/N) \sum_{n=1}^{N} I(X_{1n}X_{2n}; Y_{1n}Y_{2n} | S_0), \]

\[ R_0 + R_1 + R_2 \leq (1/N) \sum_{n=1}^{N} I(X_{1n}X_{2n}; Y_{1n}Y_{2n}), \quad (A.7) \]

then \((R_0, R_1, R_2) \in \overline{co}(\mathcal{R}_0(\mathcal{H}))\). Now to complete the proof we need only use the line of argument of Cover (1972). Namely, if \((R_0, R_1, R_2) \notin \overline{co}(\mathcal{R}_0(\mathcal{H}))\), then a geometric and convexity argument similar to the argument used to establish Equation (46) of Cover (1972), establishes the existence of a \(\delta > 0\) so that at least one of the following inequalities holds:

\[ (1/N) \sum_{n=1}^{N} I(X_{1n}; Y_{1n} | X_{2n}S_0) < (1 - \delta)R_1, \]

\[ (1/N) \sum_{n=1}^{N} I(X_{2n}; Y_{2n} | X_{1n}S_0) < (1 - \delta)R_2, \]

\[ (1/N) \sum_{n=1}^{N} I(X_{1n}X_{2n}; Y_{1n}Y_{2n} | S_0) < (1 - \delta)(R_1 + R_2), \]

\[ (1/N) \sum_{n=1}^{N} I(X_{1n}X_{2n}; Y_{1n}Y_{2n}) < (1 - \delta)(R_0 + R_1 + R_2). \quad (A.8) \]

Now if one of the inequalities in (A.8) holds then the RHS of the corresponding inequality in (A.4) is bounded below by \(\delta\) minus a positive term that decreases to zero with \(1/N\) as \(N \to \infty\). Hence one of the \(P_{e_i} \to 0\) as \(N \to \infty\) if
(R_0, R_1, R_2) \notin \overline{co}(\mathcal{R}_0(\mathcal{X})). Since \lambda \geq P_{e,1}, P_{e,2} and 2\lambda \geq P_{e,0}, P_{e,4}, \lambda \to 0 as N \to \infty if (R_0, R_1, R_2) \notin \overline{co}(\mathcal{R}_0(\mathcal{X})). This proves that no rate triple outside \overline{co}(\mathcal{R}_0(\mathcal{X})) is achievable.

APPENDIX B

1. Proof of the Converse Statement in Theorem 3

We need to show that \mathcal{C}(\mathcal{X}) \subset \bigcup_{n=1}^{\infty} \overline{co}(\{1/n, \mathcal{R}_n(\mathcal{X})\}). To do this we shall use a proof similar in some respects to that in Appendix A. So consider for the moment any \((N, M_0, M_1, M_2, \lambda)\)-code with codewords \{x_{1,k_1}, x_{2,k_2}\} and that attains a rate triple \((R_0, R_1, R_2)\) given by (5). Let \(S_0, S_1, S_2, S_0^*, S_1^*, S_2^*, X_i, Y_i\) be as before in Appendix A. Also define \(P_{e,1}\) and \(P_{e,2}\) as in (A.1c) and define

\[
\tilde{P}_{e,1} = \Pr[S_0^* \neq S_0 \text{ or } S_1^* \neq S_1],
\]

\[
\tilde{P}_{e,2} = \Pr[S_0^* \neq S_0 \text{ or } S_2^* \neq S_2].
\]

So for \(i = 1, 2\), \(P_{e,i} \leq \tilde{P}_{e,i} \leq \lambda\). Now since \(S_0^*\) and \(S_1^*\) are functions of \(Y_1\) we can write \(S_0^* = g_1(Y_1)\) and \(S_1^* = g_2(Y_1)\). Next, using the notation given in (A.1), we can write

\[
P_{e,1} = \sum_{k,m} P_{S_0}(k) P_{S_1}(m) P_{x_1|S_0,S_1}(x_1 | k, m)
\]

\[
\cdot \sum_{k,m,i} P_{x_1|S_0,S_1}(x_1 | k, m) P_{x_2|S_0,S_1}(x_2 | k, i) P_{1n}(y_1 | x_1, x_2)
\]

\[
= \sum_{k,m,i} P_{S_0}(k) P_{S_1}(m) P_{S_2}(i) P_{x_1|S_0,S_1}(x_1 | k, m) P_{x_2|S_0,S_1}(x_2 | k, i) P_{1n}(y_1 | x_1, x_2)
\]

\[
= \sum_{k,m,i} P_{S_0}(k) P_{S_1}(m) P_{S_2}(i) \Pr[Y_1 = y_1 | S_0 = k, S_1 = m, S_2 = i]
\]

\[
= \sum_{k,m,i} \Pr[Y_1 = y_1, S_0 = k, S_1 = m]
\]

where \(P_{1n}\) is the \(n\)th memoryless extension of \(P_1\) given in (1). Here (1) in (B.2) is because of (A.1b), (2) because \(P_{1n}(y_1 | x_{1,k}, x_{2,i}) = \Pr[Y_1 = y_1 | S_0 = k, S_1 = m, S_2 = i]\) and (3) is obtained by summing out the variable \(i\). Similarly we can also show that

\[
\tilde{P}_{e,1} = \sum_{k,m} \sum_{y_1: (g_1(y_1), g_2(y_1)) \neq (k,m)} \Pr[Y_1 = y_1, S_0 = k, S_1 = m].
\]
Next standard Fano inequality derivation (Theorem 4.3.2 of Gallager (1968)) can be used with (B.2) and (B.3) to yield
\[ H(S_1 | Y_1 S_0) \leq P_{e,1} \log M_1 + h(P_{e,1}), \]  
\[ H(S_0 S_1 | Y_1) \leq \tilde{P}_{e,1} \log M_0 M_1 + h(\tilde{P}_{e,1}). \]  
(B.4)

Similarly, it can be shown that
\[ H(S_2 | Y_2 S_0) \leq P_{e,2} \log M_2 + h(P_{e,2}), \]  
\[ H(S_0 S_2 | Y_2) \leq \tilde{P}_{e,2} \log M_0 M_2 + h(\tilde{P}_{e,2}). \]  
(B.5)

Moreover it is easy to show that for \( i = 1, 2, \)
\[ I(S_i Y_i | S_0) \leq I(S_i X_i; Y_i | S_0) = I(X_i; Y_i | S_0), \]  
\[ I(S_0 S_i Y_i) \leq I(S_0 S_i X_i; Y_i) = I(S_0 X_i; Y_i). \]  
(B.6)

So (5), (B.4)–(B.6) gives for \( i = 1, 2, \)
\[ P_{e,i} \geq 1 - \frac{1}{NR_i} - (1/|NR_i|) I(X_i; Y_i | S_0), \]  
\[ \tilde{P}_{e,i} \geq 1 - \frac{1}{N(R_0 + R_i)} - (1/|N(R_0 + R_i)|) I(S_0 X_i; Y_i). \]  
(B.7)

Moreover in view of (13), if
\[ R_0 + R_i \leq (1/N) I(S_0 X_i; Y_i), \]  
\[ R_i \leq (1/N) I(X_i; Y_i | S_0), \]  
(B.8)

for \( i = 1, 2, \) then \( (R_0, R_1, R_2) \in (1/N) \mathcal{R}^N(\mathcal{X}^N). \)

So suppose that
\[ (R_0, R_1, R_2) \notin \bigcup_{n=1}^{\infty} \mathcal{C}_n \mathcal{R}^n(\mathcal{X}^n). \]  
(B.9)

Then \( (R_0, R_1, R_2) \notin \mathcal{C}_n(1/N) \mathcal{R}^n(\mathcal{X}^n) \) for every \( N \geq 1. \) Then a repetition of the line of argument following (A.7) in Appendix A used along with (B.7) and (B.8) shows that any sequence of \( (N, M_0, M_1, M_2, \lambda^{(N)}) \)-codes attaining a rate triple that satisfies (B.9) is such that \( \lambda^{(N)} \rightarrow 0 \) as \( N \rightarrow \infty. \) This proves that \( \mathcal{C}(\mathcal{X}) \subset \bigcup_{n=1}^{\infty} \mathcal{C}_n(1/N) \mathcal{R}^n(\mathcal{X}^n) \).

2. Proof of Theorem 5

We need only show that \( \mathcal{C}(\mathcal{X}) \subset \overline{\mathcal{C}(\mathcal{X}^N)}. \) Fortunately much of this proof proceeds as the above proof of the converse part of Theorem 3. So again consider for the moment any \( (N, M_0, M_1, M_2, \lambda) \)-code and let \( S_0, S_1, S_2, S_0^x, S_1^x, S_0', S_2', X_i, Y_i, P_{e,i} \) and \( \tilde{P}_{e,i} \) be as before in the above proof of the
converse part of Theorem 3. Since that proof above applies to general IFC's $\mathcal{X}$, we can conclude that the inequalities (B.7) are still valid here. Next, using the separate channels assumption (19),

$$\Pr[Y_i = y_i \mid X_1 = x_1, X_2 = x_2] = \Pr[Y_i = y_i \mid X_i = x_i],$$

(B.10)

for $i = 1, 2$. A direct consequence of (B.10) and (A.1a) then is that, for $i = 1, 2$,

$$I(S_i; X_i) = I(X_i; Y_i).$$

(B.11)

Now it is easy to show that for $i = 1, 2$,

$$I(X_i; Y_i \mid S_0) \leq N \sum_{n=1}^{N} I(X_{in}; Y_{in} \mid S_0),$$

$$I(X_i; Y_i) \leq N \sum_{n=1}^{N} I(X_{in}; Y_{in}).$$

(B.12)

So (5), (B.7), (B.11) and (B.12) yield that for $i = 1, 2$,

$$P_{e,i} \geq 1 - 1/NR_i - (1/N) \sum_{n=1}^{N} I(X_{in}; Y_{in} \mid S_0)/R_i,$$

$$P_e \geq 1 - 1/N(R_0 + R_i) - (1/N) \sum_{n=1}^{N} I(X_{in}, Y_{in})/(R_0 + R_i).$$

(B.13)

Now an argument similar to that establishing (A.7) yields that if

$$R_i \leq (1/N) \sum_{n=1}^{N} I(X_{in}; Y_{in} \mid S_0),$$

$$R_0 + R_i \leq (1/N) \sum_{n=1}^{N} I(X_{in}; Y_{in}),$$

(B.14)

for $i = 1, 2$, then $(R_0, R_1, R_2) \in \mathcal{F}_i(\mathcal{X})$. Finally, a repetition of the argument following (A.7) in Appendix A can now be used with (B.7) and (B.8) to conclude that no rate triple outside $\Omega(\mathcal{F}_i(\mathcal{X}))$ is achievable. This concludes the proof.

**APPENDIX C**

*Proof of Theorem 6*

Let $\mathcal{X}^*(\mathcal{X})$ be the set of all rate quintuples $(R_0, R_1, R_1, R_2, R_2, R_3)$ satisfying the following inequality constraints (C.1a) and (C.1b) for some $(Z, U, X, Y) \in \mathcal{D}_i \mathcal{X}$. 
\[ R_0 + R_{10} + R_{20} + R_{11} \leq I(U_2X_1; Y_1), \]
\[ R_{10} + R_{20} + R_{11} \leq I(U_2X_1; Y_1 | Z), \]
\[ R_{10} + R_{11} \leq I(X_1; Y_1 | U_2Z), \]
\[ R_{20} \leq I(U_2; Y_1 | U_1Z), \]
\[ R_{11} \leq I(X_1; Y_1 | U_1U_2), \]
\[ R_{20} + R_{11} \leq I(U_2X_1; Y_1 | U_1Z). \]
\[ (C.1a) \]

\[ R_0 + R_{10} + R_{20} + R_{22} \leq I(U_1X_2; Y_2), \]
\[ R_{10} + R_{20} + R_{22} \leq I(U_1X_2; Y_2 | Z), \]
\[ R_{20} + R_{22} \leq I(X_2; Y_2 | U_1Z), \]
\[ R_{10} \leq I(U_1; Y_2 | U_2Z), \]
\[ R_{22} \leq I(X_2; Y_2 | U_1U_2), \]
\[ R_{10} + R_{22} \leq I(U_1X_2; Y_2 | U_2Z). \]
\[ (C.1b) \]

Now since for \( i, j, k = 1, 2 \) such that \( i \neq k \) we have
\[ I(U_i; Y_i | U_iZ) \geq I(U_i; Y_i | Z), \]
and
\[ I(X_i; Y_i | U_iU_2) = I(X_i; Y_i | U_iU_2Z), \]
it follows that
\[ \overline{co}(U_{i=1}^4 \mathcal{B}_i^*(\mathcal{H})) \subset \overline{co}(\mathcal{H}^*(\mathcal{H})). \]
Hence Theorem 6 is a corollary to the following theorem which we will prove here. (Although we do not do so here, it can be shown that actually \( \overline{co}(U_{i=1}^4 \mathcal{B}_i^*(\mathcal{H})) = \overline{co}(\mathcal{H}^*(\mathcal{H})) \) so that Theorem 6 is equivalent to Theorem C below.)

**THEOREM C.** Every rate quintuple in \( \overline{co}(\mathcal{H}^*(\mathcal{H})) \) is achievable for general IFC \( \mathcal{H} \).

**Proof.** A random coding proof will be used to establish that any rate quintuple satisfying the inequality constraints (C.1a) and (C.1b) for a given \( (Z, U, X, Y) \in \mathcal{D}_i\mathcal{H} \) is achievable. This then would establish the theorem. So fix an arbitrary \( (Z, U, X, Y) \in \mathcal{D}_i\mathcal{H} \) for the remainder of the proof. Let \( Q_0, Q_{10}, Q_{20}, Q_1 \) and \( Q_2 \) be given as in (7) and \( Q_{0,N}, Q_{10,N}, Q_{20,N}, Q_{1,N}, \) and \( Q_{2,N} \) denote their respective Nth extensions. For a given rate quintuple \( (R_0, R_{11}, R_{10}, R_{22}, R_{20}) \) let \( M_0 = [e^{NR_0}] \) and \( M_{ij} = [e^{NR_{ij}}] \) and consider the following random code ensemble:

1. First choose \( M_0 \) subcluster centers \( z_k, 1 \leq k \leq M_0 \), independently each according to the pmf \( Q_{0,N} \).

2. Next choose \( M_0M_{10} \) 1-cluster centers \( u_{1,km_0}, 1 \leq k \leq M_0, 1 \leq m_0 \leq M_{10} \) by passing each \( z_k \) \( M_{10} \) times independently through a discrete memoryless channel with transition probabilities \( Q_{10} \).

643/44/1-7
(3) Now choose $M_0 M_{10} M_{11}$ codewords $x_{1,km,m_2}$, $1 \leq k \leq M_0$, $1 \leq m_1 \leq M_{11}$, $1 \leq m_2 \leq M_{10}$ by passing each $u_{1,km}$ through a discrete memoryless channel with transition probabilities $Q_1$.

(4) Finally choose $M_0 M_{20}$ 2-cluster centers $u_{2,ki_2}$, $1 \leq k \leq M_0$, $1 \leq i_2 \leq M_{20}$ and $M_0 M_{20} M_{22}$ codewords $x_{2,ki_2,i_2}$, $1 \leq k \leq M_0$, $1 \leq i_2 \leq M_{20}$, $1 \leq i_2 \leq M_{22}$ similarly using $Q_{20}$ and $Q_{21}$ in place of $Q_{10}$ and $Q_1$ respectively.

Hence this code ensemble implies that the set $\{z_k\}_{k=1}^{M_0}$, $\{u_{1,km}\}_{k=1,m=1}^{M_0 M_{10}}$, $\{u_{2,ki_2}\}_{k=1,i_2=1}^{M_0 M_{20}}$, $\{x_{1,km,m_2}\}_{k=1,m_1=1,m_2=1}^{M_0 M_{11} M_{10}}$, $\{x_{2,ki_2,i_2}\}_{k=1,i_2=1,i_2=1}^{M_0 M_{20} M_{22}}$ of subcluster centers, cluster centers and codewords have the following joint pmf

$$Q_{1,N}(x_{1,km,m_2} | u_{1,km}) Q_{10,N}(u_{1,km} | z_k) Q_{2,N}(x_{2,ki_2,i_2} | u_{2,ki_2}) Q_{20,N}(u_{2,ki_2} | z_k) Q_{0,N}(z_k).$$

(C.2)

The subcluster centers $z_k$ and cluster centers $u_{1,km}$, $u_{2,ki_2}$ are used solely for the purpose of decoding at the two receivers. The indices $k$, $m_1$, $m_2$, $i_1$, and $i_2$ represent the $S_0$, $S_{11}$, $S_{10}$, $S_{22}$ and $S_{20}$ source messages respectively. Hence receiver (1) seeks to correctly decode the transmitted $(k, m_1, m_2, i_2)$ and receiver (2) seeks to correctly decode the transmitted $(k, i_1, i_2, m_2)$. In order to specify the decoding rules, define

$$P_3(y_1 | x_1, u_2) = \sum_{x_2} P_1(y_1 | x_1, x_2) Q_2(x_2 | u_2)$$

$$P_4(y_1 | u_1, u_2) = \sum_{x_2} P_2(y_1 | x_1, u_2) Q_1(x_1 | u_1)$$

$$P_5(y_1 | u_2, z) = \sum_{u_1} P_4(y_1 | u_1, u_2) Q_{10}(u_1 | z)$$

$$P_6(y_1 | u_1, z) = \sum_{u_1} P_4(y_1 | u_1, u_2) Q_{20}(u_2 | z)$$

$$P_7(y_1 | z) = \sum_{u_1} P_4(y_1 | u_1, z) Q_{10}(u_1 | z)$$

$$P_8(y_1) = \sum_{z} P_7(y_1 | z) Q_0(z).$$

(C.3)

Similarly define $P_{8}^\delta(y_2 | x_2, u_1) = \sum_{x_2} P_6(y_2 | x_1, x_2) Q_1(x_1 | u_1)$ and $P_{8}^\delta(y_2 | u_1, u_2)$, $P_8^\delta(y_2 | u_1, z)$, $P_8^\delta(y_2 | u_2, z)$, $P_8^\delta(y_2 | z)$ and $P_8^\delta(y_2)$. Moreover let $P_{i,N}$ and $P_{i,N}^\delta$ denote the $N$th memoryless extensions of $P_i$ and $P_i^\delta$ respectively.

Now fix an arbitrary $\epsilon > 0$ and define the following sets of jointly typical sequences.

$$A_1(k, m_1, m_2, i_2) = \left\{ y_1 : \frac{1}{N} \log \frac{P_{8,N}(y_1 | x_{1,km,m_2}, u_{2,ki_2})}{P_{8,N}(y_1)} - I(U_kX_1, Y_1) \leq \epsilon \right\},$$
Next obtain the sets $A_j^*(k, i_1, i_2, m_2)$ of jointly typical sequences from the corresponding $A_j(k, m_1, m_2, i_2)$ by interchanging the (1) and (2) indices in the average mutual information terms and by making the following substitutions in (C.4). Finally define

$$A(k, m_1, m_2, i_2) = \bigcap_{j=1}^{6} A_j(k, m_1, m_2, i_2),$$

$$A^*(k, i_1, i_2, m_2) = \bigcap_{j=1}^{6} A_j^*(k, i_1, i_2, m_2).$$

Then the decoding rule at receiver (1) is to decode the received $y_1$ as $(k, m_1, m_2, i_2)$ iff $y_1 \in A(k, m_1, m_2, i_2)$ for one and only one $(k, m_1, m_2, i_2)$. Similarly
the decoding rule at receiver (2) is to decode the received $y_2$ as $(k, i_1, i_2, m_2)$ if $y_2 \in A^*(k, i_1, i_2, m_2)$ for one and only one $(k, i_1, i_2, m_2)$.

Intuitively the decoding rule at receiver (1) is to decode the received $y_1$ as $(k, m_1, m_2, i_2)$ if there is one and only one $(x_{k, u_1, k_2, x_1, k_3, m_2, u_2, k_4})$ that is jointly typical with $y_1$. For a more lengthy discussion of joint typicality, see Forney (1972) and Cover (1975). A similar comment applies to the receiver (2) decoding rule. For $j = 1, 2$ let

$$P_{e,j}(k, m_1, m_2, i_1, i_2) = \Pr[\text{decoding error at receiver (j)} \mid x_{1, k_1, m_2}, x_{2, k_1, i_2} \text{ sent}]$$

(C.6)

Now if $x_{1, k_1, m_2}$ and $x_{2, k_1, i_2}$ are sent, a decoding error at receiver (1) can occur if and only if one of the following two error events occur: (i) the event that $y_1 \notin A(k, m_1, m_2, i_2)$ (ii) the event that $y_1 \in A(k', m'_1, m'_2, i'_2)$ for some $(k', m'_1, m'_2, i'_2) \neq (k, m_1, m_2, i_2)$. The probability of the first event is overbounded by the sum of the probabilities of $y_1 \notin A(k, m_1, m_2, i_2)$ over $j = 1, \ldots, 6$. This total sum is given by $P_{e,1}^{(0)}$ in (C.8) below. The probability of the second event is overbounded by the sum of the probabilities of $y_1 \in A(k', m'_1, m'_2, i'_2)$ over all indices $(k', m'_1, m'_2, i'_2) \neq (k, m_1, m_2, i_2)$. Each of the probabilities that $y_1 \in A(k', m'_1, m'_2, i'_2)$ can be further overbounded by the probability that $y_1 \notin A_j(k', m'_1, m'_2, i'_2)$ for $j = 1, \ldots, 6$. Hence (C.9) below reflects all the various parts of this sum where in the expression for $P_{e,1}^{(j)}$ the event $y_1 \in A(k', m'_1, m'_2, i'_2)$ is overbounded by the event $y_1 \notin A_j(k', m'_1, m'_2, i'_2)$. Hence we have established the following inequality (C.7) for each fixed $(k, m_1, m_2, i_2)$:

$$P_{e,1}(k, m_1, m_2, i_1, i_2) \leq \sum_{j=0}^{6} P_{e,1}^{(j)}$$

(C.7)

where

$$P_{e,1}^{(0)} = \sum_{j=1}^{6} \sum_{y_1 \notin A_j(k, m_1, m_2, i_2)} P_{1,N}(y_1 \mid x_{1, k_1, m_2}, x_{2, k_1, i_2})$$

(C.8)

and where

$$P_{e,1}^{(1)} = \sum_{k' \neq k} \sum_{m'_1 \neq m_1} \sum_{m'_2 \neq m_2} \sum_{i'_2} P_{1,N}(y_1 \mid x_{1, k_1, m_2}, x_{2, k_1, i_2})$$

$$P_{e,1}^{(2)} = \sum_{m'_1 \neq m_1} \sum_{m'_2 \neq m_2} \sum_{i'_2} P_{1,N}(y_1 \mid x_{1, k_1, m_2}, x_{2, k_1, i_2})$$

$$P_{e,1}^{(3)} = \sum_{m'_1 \neq m_1} \sum_{m'_2 \neq m_2} \sum_{i'_2} P_{1,N}(y_1 \mid x_{1, k_1, m_2}, x_{2, k_1, i_2})$$

$$P_{e,1}^{(4)} = \sum_{i'_2} \sum_{y_1 \notin A(k, m_1, m_2, i_2)} P_{1,N}(y_1 \mid x_{1, k_1, m_2}, x_{2, k_1, i_2})$$
Let us now consider the random code ensemble average of the probabilities $P_{e,1}'$ above. So for the remainder of this appendix let an overbar denote averaging over the random code ensemble, that is, averaging with respect to the joint pmf (C.2). First it follows from the weak law of large numbers as in Cover (1975) and Forney (1972) applied to jointly typical sequences that for $1 \leq j \leq 6$,

$$\lim_{N \to \infty} \sum_{y_1 \in A_j(k',m_1',m_2',i_2')} P_{1,N}(y_1 \mid x_1,k'm_1',x_2,k'i_2') = 0 \quad (C.10)$$

Next, in order to overbound $P_{e,1}'$, let us consider

$$\sum_{y_1 \in A_j(k',m_1',m_2',i_2')} P_{1,N}(y_1 \mid x_1,k'm_1',x_2,k'i_2')$$

for a fixed $k' \neq k$ and any $m_1', m_2', i_2'$. Note from (C.4) that $A_j(k', m_1', m_2', i_2')$ depends only on $x_1,k'm_1',x_2,k'i_2'$ in the random code ensemble. Hence this ensemble average is over the set \{ $x_1,k'm_1',x_2,k'i_2'$, $u_1,km_2$, $u_2,k'i_2'$, $z_k$, $x_1,k'm_1',u_1,k'm_2',u_2,k'i_2',z_k$ \} of random sub-cluster centers, cluster centers and codewords. From the specification of the random code ensemble (C.2) note that since $k' \neq k$,

(i) $z_k$ and $z_{k'}$ are independent

(ii) $u_1,km_2$ and $u_1,k'm_2'$ are independent for any $m_2'$.

(iii) $u_2,k'i_2'$ and $u_2,k'i_2'$ are independent for any $i_2'$.

(iv) $x_1,k'm_2$ and $x_1,k'm_1'm_2'$ are independent for any $m_1'$ and $m_2'$.

In taking the random code ensemble average of $P_{e,1}'$, these properties (i)–(iv) allow us to consider the sum over all indices $m_1', m_2'$ and $i_2'$ in the expression for $P_{e,1}'$ given in (C.9). So formally from (C.2) we can write

$$\sum_{y_1 \in A_j(k',m_1',m_2',i_2')} P_{1,N}(y_1 \mid x_1,k'm_1',x_2,k'i_2')$$

$$= \sum_{j' \in A_j(k',m_1',m_2',i_2')} \sum_{u_1,km_2} \sum_{u_1,k'm_2'} \sum_{u_2,k'i_2'} \sum_{x_1,k'm_1'} \sum_{x_1,k'm_2'} \sum_{x_2,k'i_2'}$$

$$\cdot \left[ Q_1,N(x_1,k'm_2 \mid u_1,km_2) Q_{01,N}(u_1,km_2 \mid z_k) Q_{02N}(x_2,k'i_2' \mid u_2,k'i_2') \right.$$
\[ Q_{01,N}(u_{1,k,m_1}, z_k) \leq e^{-N[I(U;X_1;Y_1) - R_0 - R_1 - R_0 - R_1 - \varepsilon]}, \]

where (1) in (C.11a) is obtained by summing out the appropriate variables and using (C.3) and (2) is because \[ P_{8,N}(y_1) \leq P_{8,N}(y_1 | x_{1,k,m_1,m_2}, u_{2,k'}, e^{-N[I(U;X_1;Y_1) - \varepsilon]} \]

for \( y_1 \in A_k(k', m_1', m_2', i_2') \).

So combining (C.9) and (C.11a) we obtain

\[ P_{e,1} = \sum_{k \neq k'} \sum_{m_1', m_2', i_2'} \sum_{y_1 \in A_k(k', m_1', m_2', i_2')} P_{1,N}(y_1 | x_{1,k,m_1,m_2}, x_{2,k,i_2}) \]

\[ \leq (M_0 - 1) M_0 M_1 e^{-N[I(U;X_1;Y_1) - \varepsilon]}, \]

where (1) in (C.11b) is because \( M_0 - 1 \leq e^{NR_{10}} \) and \( M_1 - 1 \leq e^{NR_{11}} \) implies that \( M_0 e^{-N_{10}} + 1 \leq 2e^{NR_{10}} \) and similarly \( M_1 e^{-N_{11}} \leq 2e^{NR_{11}} \).

Similarly in overbounding \( P_{e,1} \) we note that \( A_k(k, m_1', m_2', i_2) \) depends only on \( x_{1,k,m_1,m_2}, u_{2,k,i_2} \) and \( z_k \) in the code ensemble. Hence the ensemble average of \( \sum_{y_1 \in A_k(k, m_1', m_2', i_2')} P_{1,N}(y_1 | x_{1,k,m_1,m_2}, x_{2,k,i_2}) \) is over the set \{ \( x_{1,k,m_1,m_2}, u_{2,k,i_2}, u_{1,k,m_2}, x_{1,k,m_2}, z_k \) \}. Since for \( m_1 \neq m_2 \) and any \( m_1', m_2' \) and \( u_{1,k,m_2} \) are conditionally independent conditioned on \( z_k \) and \( x_{1,k,m_2} \) and \( x_{1,k,m_2} \) are also conditionally independent conditioned on \( z_k \), this allows us to proceed as in (C.11a) and (C.11b) to establish that

\[ P_{e,1} = \sum_{m_1,m_2} \sum_{u_{2,k,i_2}} P_{5,N}(y_1 | u_{2,k,i_2}, z_k) \cdot Q_{01,N}(u_{1,k,m_1}, z_k) \leq 2e^{-N[I(X_1;Y_1;U_2) - R_0 - R_{11} - \varepsilon]}, \]
Next in overbounding \( P_{e,1}^{(3)} \) we note that the ensemble average of 
\[ \sum_{i_1 \in A_1} \sum_{i_2 \in A_2} \sum_{m_1 \in C_1} \sum_{m_2 \in C_2} \sum_{k \in \mathcal{K}} \sum_{z \in \mathcal{Z}} P_1, N(y_1 \mid x_1, k, m_1, m_2, z; \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{z}) \] is over the set \{ \mathbf{x}_1, \mathbf{x}_2, \mathbf{z} \} in the code ensemble. Now for \( m_2 \neq m_1, i_2 \neq i_1 \) and any \( m_1 \), conditioned on \( z \), \( \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}_1, \mathbf{x}_2 \) and \( z \) are conditionally independent and so are also \( \mathbf{u}_a, k, i \) and \( \mathbf{u}_b, k, i \) and also \( \mathbf{x}_1, k, m_1, m_2 \) and \( \mathbf{x}_1, k, m_1, m_2 \). Hence similar to (C.12) we can establish that 
\[ P_{e,1}^{(3)} = \sum_{i_1 \in A_1} \sum_{i_2 \in A_2} \sum_{m_1 \in C_1} \sum_{m_2 \in C_2} \sum_{k \in \mathcal{K}} \sum_{z \in \mathcal{Z}} P_1, N(y_1 \mid x_1, k, m_1, m_2, z; \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{z}) \]
\[ = 2 e^{-N \left[ I(U_2 X_1; Y_1 I Z) - R_{10} - R_{20} - R_{11} \right]} \] (C.13)

Continuing in a similar fashion we have 
\[ P_{e,1}^{(4)} = \sum_{i_1 \in A_1} \sum_{i_2 \in A_2} \sum_{m_1 \in C_1} \sum_{m_2 \in C_2} \sum_{k \in \mathcal{K}} \sum_{z \in \mathcal{Z}} P_1, N(y_2 \mid x_2, k, m_1, m_2, z; \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{z}) \]
\[ = e^{-N \left[ I(U_2 X_1; Y_1 I Z) - R_{20} - R_{11} \right]} \] (C.14)

and
\[ P_{e,1}^{(5)} = \sum_{i_1 \in A_1} \sum_{i_2 \in A_2} \sum_{m_1 \in C_1} \sum_{m_2 \in C_2} \sum_{k \in \mathcal{K}} \sum_{z \in \mathcal{Z}} P_1, N(y_1 \mid x_1, k, m_1, m_2, z; \mathbf{u}_1, \mathbf{u}_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{z}) \]
\[ = e^{-N \left[ I(U_2 X_1; Y_1 I Z) - R_{20} - R_{11} \right]} \] (C.16)

It then follows from (C.7) to (C.16) that 
\[ \lim_{N \to \infty} P_{e,1}(k, m_1, m_2, i_1, i_2) = 0 \]
if the rate quintuple \( (R_0 - 2\epsilon, R_{11} - 2\epsilon, R_{10} - 2\epsilon, R_{20} - 2\epsilon) \) satisfies the inequality constraints (C.1a). By a similar argument, in view of the symmetry between receivers (1) and (2), it can be established that 
\[ \lim_{N \to \infty} P_{e,2}(k, m_1, m_2, i_1, i_2) = 0 \]
if the rate quintuple \( (R_0 - 2\epsilon, R_{11} - 2\epsilon, \)
$R_{10} - 2\epsilon, R_{22} - 2\epsilon, R_{20} - 2\epsilon$) satisfies the inequality constraints (C.1b). This then establishes the theorem. Q.E.D.

**APPENDIX D**

*Proof that $\mathcal{C}(\mathcal{X}) \subseteq \mathcal{C}(\mathcal{X})$*

Fix an achievable rate quintuple $(R_0^*, R_1^*, R_2^*, R_{10}^*, R_{22}^*)$ in $\mathcal{C}(\mathcal{X})$, $\alpha_1 \alpha_2 \in [0, 1]$ and let $(R_0, R_1, R_2)$ be a rate triple satisfying (22). We need to show that this rate triple is achievable. We shall do this by using a coding scheme of the second communication situation. The basic idea is to split each of the two private message sources in this first communication situation that we are concerned with into two separate message sources resulting in a total of five message sources as in the second communication situation. The message rates involved in this splitting up of the private message sources is given in (22).

The formal proof is as follows. Fix an arbitrary $\delta > 0$ and for integer $N \geq 1$

$$M_l = \lfloor e^{N(R_l - \delta)} \rfloor,$$

for $l = 0, 1, 2$. We are concerned with the first communication situation in which the $(S_0, S_1, S_2)$ message sources output message triplets $(k, m, i)$ where $1 \leq k \leq M_0, 1 \leq m \leq M_1, 1 \leq i \leq M_2$ as depicted in Figure 1. Suppose we can find a five-message set of sources $(S_0^*, S_1^*, S_2^*, S_{10}^*, S_{22}^*)$ as in the second communication situation of Figure 2, which outputs message quintuplets $(k^*, m_1^*, m_2^*, i_1^*, i_2^*)$ where $1 \leq k^* \leq M_0^*, 1 \leq m_1 \leq M_1^*, 1 \leq m_2 \leq M_{10}^*, 1 \leq i_1 \leq M_{21}^*, 1 \leq i_2 \leq M_{22}^*$, so that the following two conditions are satisfied:

(i) there is a one-to-one correspondence between the message triplets $(k, m, i)$ and the message quintuplets $(k^*, m_1^*, m_2^*, i_1^*, i_2^*)$

(ii) for each $\lambda \in (0, 1)$ there exist $(N, M_0^*, M_1^*, M_{10}^*, M_{21}^*, M_{22}^*, M_0^*, \lambda)$-codes for all $N$ sufficiently large.

Now use a coding scheme in which the output of the $(S_0, S_1, S_2)$ message sources is transformed to a five-message source set using the one-to-one transformation given by (i) and subsequently coded with the codes given by (ii). Then the receivers can reliably decode the output of the five message source set and use the one-to-one transformation given by (i) to reliably recover the output of the original message sources. This then establishes the achievability of $(R_0, R_1, R_2)$.

In order to show the existence of such a five-message source set, define for $l = 1, 2$

$$M_l = e^{N(1-\alpha_l)R_l-\delta_l}, \quad (D.2)$$
Express each \( m, 1 \leq m \leq M_1 \), as a binary vector of length \( \lceil \log_2 M_1 \rceil \) bits. Let \( m_1^* \) be the integer representation of the first \( \lceil \log_2 M_1 \rceil \) bits of this binary vector and let \( m_2^* \) be the integer representation of the remaining bits of this binary vector representing \( m \). Similarly express each \( i, 1 \leq i \leq M_2 \), as a binary vector of length \( \lceil \log_2 M_2 \rceil \) bits and let \( i_1^* \) be the integer representation of the first \( \lceil \log_2 M_2 \rceil \) bits of this vector and \( i_2^* \) the integer representation of the remaining bits. Finally let \( k^* = k \). It is clear that this defines a one-to-one correspondence between \((k, m, i)\) and \((k^*, m_1^*, m_2^*, i_1^*, i_2^*)\) thus satisfying condition (i). Note that since the \((k, m, i)\) message sources are statistically independent and equiprobable, it can be easily shown that the \((k^*, m_1^*, m_2^*, i_1^*, i_2^*)\) message sources are also independent and equiprobable. Moreover

\[
M_0^* = M_0 ,
\]
\[
M_1^* = 2^{\lceil \log_2 M_1 \rceil} ,
\]
\[
M_2^* = 2^{\lceil \log_2 M_1 - \lceil \log_2 M_1 \rceil \rceil} ,
\]
for \( l = 1, 2 \). Now from (D.2), for \( l = 1, 2 \), \( \lceil \log_2 M_1 \rceil < (\log_2 e) N[(1 - \alpha_l) R_l - \delta] + 1 \). Thus from (D.3) and (22), for \( l = 1, 2 \),

\[
M_1^* < e^{N[(1 - \alpha_l) R_l - \delta - (\log_2 e) / N]}
\]
\[
e^{N[R_1^* - \delta - (\log_2 e) / N]} \quad (D.4)
\]

Similarly \( \lceil \log_2 M_1 \rceil - \lceil \log_2 M_1 \rceil < (\log_2 e) N[\alpha_l R_l - 2\delta + (1/N)] + 1 \). Thus from (D.3) and (22),

\[
M_2^* < e^{N[\alpha_l R_l - 2\delta + (1 + \log_2 e) / N]}
\]
\[
e^{N[R_2^* - 2\delta + (1 + \log_2 e) / N]} \quad (D.5)
\]

Finally from (D.1), (D.3) and (22),

\[
M_0^* = [e^{N(R_0^* - \delta)}]. \quad (D.6)
\]

So from (D.4)-(D.6) it follows that condition (ii) is satisfied since by assumption the rate quintuple \((R_0^*, R_1^*, R_2^*, R_{10}^*, R_{20}^* ) \in \mathcal{C}^0(\mathcal{X})\).

ACKNOWLEDGMENT

The author wishes to thank the reviewer for many helpful suggestions.

RECEIVED: September 9, 1977; REVISED: July 24, 1979
REFERENCES


FORNEY, G. D. (1972), Information Theory, course notes for EE 376, Stanford University Electrical Engineering Department, Stanford, California.


Printed in Belgium