INFORMATION AND CONTROL 29, 185–203 (1975)

The Capacity Region of a Channel with s Senders and r Receivers

MICHAEL L. ULREY*

Wichita State University, Wichita, Kansas

The study of multi-way channels was initiated by Shannon in his basic paper "Two-way communication channels" (Shannon, 1961). Ahlswede (1971b) has defined and classified multi-way channels of various kinds and proved simple characterizations for the capacity regions of channels with (a) two senders and one receiver, and (b) three senders and one receiver.

Subsequently, Ahlswede (1974) found a new approach to the coding problem for a channel with two senders and one receiver which led to an alternative characterization of the capacity region of this channel. This approach seems to be more canonical than the earlier one, and was used successfully in determining the capacity region of a channel with two senders and two receivers in case both senders send messages simultaneously to both receivers (Ahlswede, 1974). In the earlier paper, Ahlswede conjectured that the results of that paper would hold for any channel with $s \geqslant 2$ senders and one receiver. A conjecture of the later paper was that its results would hold for any channel with $s \geqslant 2$ senders and $r \geqslant 1$ receivers in case all senders send independent messages simultaneously to all receivers.

In this paper, we have proved the latter conjecture to be true. The characterization we get for the special case $s = 3$ and $r = 1$ is different from that of Ahlswede's (1971b) earlier paper. All of our results are obtained under the assumption of independent sources.

1. THE CHANNEL MODEL AND STATEMENT OF THE CODING PROBLEM

In this paper, a noisy, discrete, stationary, memoryless channel with $s \geqslant 2$ senders and $r \geqslant 1$ receivers is studied.

Let $X_1, X_2, \ldots, X_s$ and $Y_1, Y_2, \ldots, Y_r$ be finite sets; $X_1, \ldots, X_s$ denote the input alphabets and $Y_1, \ldots, Y_r$ the output alphabets of the channel to be

* Present address. Department of Mathematics, Wichita State University, Wichita, Kansas 67208.

described. For every $t = 1, 2, \ldots$, let $X_k^t = X_k$ and $Y_j^t = Y_j$ for all $k = 1, \ldots, s$ and $j = 1, \ldots, r$. Let n be a positive integer and define

$$X_k(n) = \prod_{t=1}^{n} X_k^t \quad \text{and} \quad Y_j(n) = \prod_{t=1}^{n} Y_j^t \qquad \text{for all} \quad k = 1, \ldots, s$$
$$\text{and} \quad j = 1, \ldots, r.$$

For each $k = 1, \ldots, s$, $X_k(n)$ is the set of words of length n with letters from the alphabet X_k which can be sent over the channel; similarly, for each $j = 1, \ldots, r$, $Y_j(n)$ is the set of words of length n with letters from the alphabet Y_j which can be received over the channel. Further define

$$\hat{X}^t = \hat{X} = \prod_{k=1}^{s} X_k^t \quad \text{and} \quad \hat{Y}^t = \hat{Y} = \prod_{j=1}^{r} Y_j^t \qquad \text{for all} \quad t = 1, \ldots, n.$$

If M is an $n \times s$ matrix, let M_k^t be the element in the t-th row and k-th column of M, M^t be the t-th row of M, and M_k the k-th column of M for all $t = 1, \ldots, n$ and $k = 1, \ldots, s$. Similarly define $\mathbf{M}_j^t$, $\mathbf{M}^t$ and $\mathbf{M}_j$ for an $n \times r$ matrix $\mathbf{M}$ for all $t = 1, \ldots, n$ and $j = 1, \ldots, r$. Then let

$$\mathscr{M} = \{M\colon M \text{ is an } n \times s \text{ matrix and } M_k \in X_k(n) \text{ for all } k = 1, \ldots, s\}$$

and

$$\boldsymbol{\mathscr{M}} = \{\mathbf{M}\colon \mathbf{M} \text{ is an } n \times r \text{ matrix and } \mathbf{M}_j \in Y_j(n) \text{ for all } j = 1, \ldots, r\}.$$

If $x_k(n) \in X_k(n)$ for all $k = 1, \ldots, s$, by $M = (x_1(n), \ldots, x_s(n))$ we shall mean the matrix $M \in \mathscr{M}$ with $M_k = x_k(n)$ for $k = 1, \ldots, s$.

The column M_k ($k = 1, \ldots, s$) of an $M \in \mathscr{M}$ represents a word of length n sent across the channel by the k-th sender. The row M^t ($t = 1, \ldots, n$) of an $M \in \mathscr{M}$ represents an s-tuple of letters, one letter from each sender, sent across the channel at instant t. Similarly, the column $\mathbf{M}_j$ ($j = 1, \ldots, r$) of an $\mathbf{M} \in \boldsymbol{\mathscr{M}}$ represents a word of length n received by the j-th receiver, while the row $\mathbf{M}^t$ ($t = 1, \ldots, n$) represents an r-tuple of letters, one letter intended for each receiver, received over the channel at instant t.

Let $\omega(\cdot \mid \cdot)$ be a non-negative function defined on $\hat{X} \times \hat{Y}$ such that $\sum_{\hat{y} \in \hat{Y}} \omega(\hat{y} \mid \hat{x}) = 1$ for all $\hat{x} \in \hat{X}$. Then the channel transmission probabilities are given by

$$P_n(\mathbf{M} \mid M) = \prod_{t=1}^{n} \omega(\mathbf{M}^t \mid M^t) \qquad \text{for all} \quad M \in \mathscr{M} \quad \text{and} \quad \mathbf{M} \in \boldsymbol{\mathscr{M}}. \tag{1.1}$$

The probability that the r words $\mathbf{M}_1, \ldots, \mathbf{M}_r$ are received, given that the s words $M_1, \ldots, M_s$ are sent, is given by $P_n(\mathbf{M} \mid M)$. The channel with s senders and r receivers is then completely described by the input alphabets $X_1, \ldots, X_s$, the output alphabets $Y_1, \ldots, Y_r$, and the channel probability function $\omega(\cdot \mid \cdot)$.

As to how this channel is actually used, we assume throughout this paper that all of the s senders send independent messages simultaneously to all of the r receivers. In keeping with the notation of Ahlswede (1971b), this communication situation is denoted by (P, T_{sr}), where the P refers to the transmission probabilities defined in (1.1).

A code concept appropriate to the communication situation (P, T_{sr}) is now introduced. Let $N_1, \ldots, N_s$ be positive integers and define $N = \prod_{k=1}^{s} N_k$, $\vec{N} = (N_1, \ldots, N_s)$ and $\bar{I} = \{(i_1, \ldots, i_s) \mid i_k$ is an integer and $1 \leqslant i_k \leqslant N_k$ for $k = 1, \ldots, s\}$.

A code $-(n, \vec{N})$ for (P, T_{sr}) is a system $\{(M(\bar{\imath}), A_j(\bar{\imath})): \bar{\imath} \in \bar{I}, j = 1, \ldots, r\}$ such that

(i) $M(\bar{\imath}) \in \mathscr{M}$ for all $\bar{\imath} \in \bar{I}$

(ii) There exists a collection

$$C = \{M_k(i_k): M_k(i_k) \in X_k(n) \text{ for all } i_k = 1, \ldots, N_k \text{ and } k = 1, \ldots, s\} \text{ such that } M_k(\bar{\imath}) = M_k(i_k) \text{ for all } \bar{\imath} = (i_1, \ldots, i_s) \in \bar{I} \text{ and } k = 1, \ldots, s \tag{1.2}$$

(iii) $A_j(\bar{\imath}) \subsetneq Y_j(n)$ for all $\bar{\imath} \in \bar{I}, j = 1, \ldots, r$

(iv) $A_j(\bar{\imath}) \cap A_j(\bar{\imath}') = \phi$ whenever $\bar{\imath} \neq \bar{\imath}'$, for all $j = 1, \ldots, r$.

For each $j = 1, \ldots, r$, define $P_n^j(\cdot \mid \cdot)$ on $\mathscr{M} \times Y_j(n)$ by $P_n^j(y_j(n) \mid M) = \sum_{\mathscr{M}(y_j(n))} P_n(\mathbf{M} \mid M)$, where $\mathscr{M}(y_j(n)) = \{\mathbf{M}: \mathbf{M} \in \mathscr{M}$ and $\mathbf{M}_j = y_j(n)\}$. Then if λ is a real number with $0 < \lambda < 1$, a code $-(n, \vec{N}, \lambda)$ is a code $-(n, \vec{N})$ such that

$$\frac{1}{N} \sum_{\bar{\imath} \in \bar{I}} \sum_{j=1}^{r} P_n^j(A_j(\bar{\imath})^c \mid M(\bar{\imath})) \leqslant \lambda, \tag{1.3}$$

where $A_j(\bar{\imath})^c$ denotes the complement of $A_j(\bar{\imath})$.

An s-tuple $(R_1, \ldots, R_s)$ of real numbers is called a *s-tuple of achievable rates* for (P, T_{sr}) if for all $\epsilon > 0$ and $0 < \lambda < 1$, and for all n sufficiently large, there is a code $-(n, \vec{N}, \lambda)$ for (P, T_{sr}) such that $(1/n) \log N_k \geqslant R_k - \epsilon$ for all $k = 1, \ldots, s$.

The set of all s-tuples of achievable rates is denoted by $G(P, T_{sr})$. Following the terminology of Shannon (1961), $G(P, T_{sr})$ is called the *capacity region*.

The problem then is to find a simple ("single letter") characterization for $G(P, T_{sr})$ in case $s \geqslant 2, r \geqslant 1$. In Section 3 we obtain such a characterization for $s \geqslant 2$ and $r = 1$ and in Section 4 generalize to the case $s \geqslant 2, r \geqslant 1$.

2. A General Fano-Type Estimate

In this section, a Fano-type lemma (Fano, 1952, 1954; Gallager, 1968; Wolfowitz, 1964) is proved, which in Section 3 enables us to obtain an outer bound on the capacity region $G(P, T_{s1})$. We assume that $r = 1$ throughout Sections 2 and 3. Thus, since there is only one output alphabet, we denote it by Y, and by $Y(n)$ the n-th Cartesian product of Y with itself.

Now the so-called rate functions are defined, which are useful in the formulation and proofs of Lemma I, Theorem I, and Theorem 2. Let $I = \{1, 2,..., \alpha\}$ be a finite indexing set and $J = \{i_1, i_2,..., i_j\} \subseteq I$ where $i_1 < i_2 < \cdots < i_j$. Let $A_1, A_2,..., A_\alpha$ and B be finite sets and define $A(J) = A_{i_1} \times \cdots \times A_{i_j}$ for all $J \subseteq I$, $J \neq \phi$. Let $q(\cdot)$ be a probability distribution (p.d.) on $A(I)$, and for each non-empty $J \subseteq I$, denote by $q_J(\cdot)$ the marginal distribution of $q(\cdot)$ on $A(J)$. (Note $q_I(\cdot) = q(\cdot)$). Finally let $Q(\cdot \mid \cdot)$ be a non-negative function defined on $A(I) \times B$ such that

$$\sum_{b \in B} Q(b \mid \hat{a}) = 1 \qquad \text{for all} \quad \hat{a} \in A(I).$$

Then for all non-empty $J \subseteq I$, define the rate function

$$R_J(q, Q, A(I), B) = \sum_{b \in B} \sum_{\hat{a} \in A(I)} q(\hat{a}) Q(b \mid \hat{a})$$

$$\times \log \frac{Q(b \mid \hat{a})}{\sum_{\hat{u} \in A(J, \hat{a})} q_J((u_{i_1},..., u_{i_j})) Q(b \mid \hat{u})} \tag{2.1}$$

where, if $\hat{a} = (a_1,..., a_\alpha) \in A(I)$, then $A(J, \hat{a}) = \{\hat{u}: \hat{u} = (u_1,..., u_\alpha) \in A(I)$ and $u_k = a_k$ for all $k \notin J\}$. When the input and output alphabets and the transmission probabilities are understood, we will write $R_J(q)$ for $R_J(q, Q, A(I), B)$. Also, if $I = \{1\}$ (that is, there is only one "input alphabet"), we will write $R(q, Q, A_1, B)$ for $R_I(q, Q, A_1, B)$.

Suppose that for each $t = 1,..., n$ and $k = 1,..., s$ a p.d. $p_k{}^t(\cdot)$ on X_k is given. Then define p.d.'s $p^t(\cdot)$, $p_k(\cdot)$ and $p(\cdot)$ on X, $X_k(n)$ and $\mathscr{M}$, respectively, by

$$p^t(\hat{x}^t) = \prod_{k=1}^{s} p_k{}^t(x_k{}^t) \qquad \text{for all} \quad t = 1,\ldots, n$$

$$p_k(x_k(n)) = \prod_{t=1}^{n} p_k{}^t(x_k{}^t) \qquad \text{for all} \quad k = 1,\ldots, s \tag{2.2}$$

$$p(M) = \prod_{k=1}^{s} p_k(x_k(n))$$

where $\hat{x}^t = (x_1{}^t,\ldots, x_s{}^t) \in \hat{X}$, $x_k(n) = (x_k{}^1,\ldots, x_k{}^n) \in X_k(n)$ and

$$M = (x_1(n),\ldots, x_s(n)) \in \mathscr{M}.$$

Now specify $p_k{}^t(\cdot)$ (for $t = 1,\ldots, n$ and $k = 1,\ldots, s$) as follows. Assume a code — $(n, \vec{N})$ for (P, T_{s1}) is given and let C be as defined in (ii) of (1.2). Then if $M_k{}^t(i_k)$ denotes the t-th component of $M_k(i_k)$,

$$p_k{}^t(x) = \frac{|\{i_k : M_k{}^t(i_k) = x,\ 1 \leqslant i_k \leqslant N_k\}|}{N_k} \tag{2.3}$$

for all $t = 1,\ldots, n$, $k = 1,\ldots, s$ and $x \in X_k$.

The following is a generalized Fano-type lemma. It was first stated and proved in (Ahlswede, 1971b) for the case $s = 3$, $r = 1$. Van der Meulen (1974) has given an improved statement of and much improved proof of the following lemma.

LEMMA 1. *Given a code* $-(n, \vec{N}, \lambda)$ *for* (P, T_{s1}). *Let* $p_k{}^t(\cdot)$ *be defined as in* (2.3) *and* $p^t(\cdot)$ *as in* (2.2). *Then for all nonempty* $D \subseteq \{1,\ldots, s\}$, *there is a number* $k_D(\lambda, n)$ *such that*

$$\log\left(\prod_{k\in D} N_k\right) \leqslant \sum_{t=1}^{n} R_D(p^t, \omega, \hat{X}, Y) + k_D(\lambda, n)$$

where $(1/n)\, k_D(\lambda, n) \to 0$ *as* $n \to \infty$, $\lambda \to 0$.

Proof. The argument is a generalization of the one in (Ahlswede, 1971b). For ease of notation it is assumed that $D = \{1,\ldots, d\}$ for some integer d, $1 \leqslant d \leqslant s$. The extension to arbitrary D presents only notational difficulties, and will be omitted. We do the case $1 \leqslant d < s$ first; the case $d = s$ requires a different argument.

Let the given code be denoted $\{(M(\bar{\imath}), A(\bar{\imath})): \bar{\imath} \in \bar{I}\}$ where $A(i) \subseteq Y(n)$ for all $\bar{\imath} \in \bar{I}$. Let the given set of codewords C be as denoted in (ii) of (1.2). Then consider the probability space $(\Omega, \bar{\mu})$ where

$$\Omega = \{\overline{M}: \overline{M} = (M_{d+1}(i_{d+1}), \ldots, M_s(i_s))$$

for some

$$i_k, 1 \leqslant i_k \leqslant N_k, \text{ for all } k = d+1, \ldots, s\}$$

and $\bar{\mu}$ is the equidistribution on Ω.

For each $\overline{M} = (M_{d+1}(j_{d+1}), \ldots, M_s(j_s)) \in \Omega$ we define a non-stationary discrete memoryless channel (depending on $\overline{M}$) as follows. The input alphabet is $\tilde{X} = \prod_{k=1}^{d} X_k$ and the output alphabet is Y. For each $t = 1, \ldots, n$, a function $\omega_{\overline{M}}{}^t(\cdot \mid \cdot)$ is defined on $\tilde{X} \times Y$ by

$$\omega_{\overline{M}}{}^t(y \mid \tilde{x}) = \omega(y \mid (x_1, \ldots, x_d, M_{d+1}^t(j_{d+1}), \ldots, M_s{}^t(j_s)))$$

for all $\tilde{x} = (x_1, \ldots, x_d) \in \tilde{X}$ and $y \in Y$. If we define

$$\tilde{\mathscr{M}} = \{\tilde{M} : \tilde{M} \text{ is an } n \times d \text{ matrix and } \tilde{M}_k \in X_k(n) \text{ for all } k = 1, \ldots, d\},$$

then the transmission probabilities are given by

$$\tilde{P}_{\overline{M}}(y(n) \mid \tilde{M}) = \prod_{t=1}^{n} \omega_{\overline{M}}{}^t(y^t \mid \tilde{M}^t) \qquad \text{for all} \quad y(n) = (y^1, \ldots, y^n) \in Y(n)$$

and $\tilde{M} \in \tilde{\mathscr{M}}$.

Given this non-stationary d.m.c. $(\tilde{X}, Y, \{w_{\overline{M}}^t : t = 1, \ldots, n\})$, we construct a code for it as follows. Let $\tilde{N} = \prod_{k=1}^{d} N_k$, $\overline{N} = \prod_{k=d+1}^{s} N_k$ and

$$\bar{I}_{\overline{M}} = \{\bar{\imath} : \bar{\imath} = (i_1, \ldots, i_s) \in \bar{I} \text{ and } i_k = j_k \text{ for } k = d+1, \ldots, s\}.$$

Then for each $\overline{M} = (M_{d+1}(j_{d+1}), \ldots, M_s(j_s)) \in \Omega$, consider the system

$$\{(\tilde{M}, A(\bar{\imath})): \tilde{M} = (M_1(i_1), \ldots, M_d(i_d)) \text{ for some } \bar{\imath} \in \bar{I}_{\overline{M}}\}. \tag{2.4}$$

This code, although originally meant for the channel with d senders and one receiver, can be regarded as a code for the one-way channel described above, by letting the d senders coalesce.

Furthermore, if $\lambda' = \lambda - \lambda \log \lambda$, then there is a set $B \subseteq \Omega$ with $|B| \geqslant [(\lambda' - \lambda)/(\lambda')]\overline{N}$ and such that for all $\overline{M} \in B$, the code in (2.4) is a code $-(n, N)$ with average probability of error λ' (see (Wolfowitz, 1964) for notation) for the corresponding non-stationary channel.

To prove this, let a r.v. L^* be defined on Ω by

$$L^*(\overline{M}) = \frac{1}{\tilde{N}} \sum_{\bar{\imath} \in \bar{I}_{\overline{M}}} P_n(A(\bar{\imath})^c \mid M(\bar{\imath}))$$

for all $\bar{M} = (M_{d+1}(j_{d+1}),..., M_s(j_s)) \in \Omega$. By (1.3) we have $EL^* \leqslant \lambda$, where the expectation is taken with respect to $\bar{\mu}(\cdot)$. Hence, by Markov's inequality, the set $B = \{L^* < \lambda'\}$ satisfies $| B | \geqslant [(\lambda' - \lambda)/(\lambda')] \bar{N}$ as asserted.

Therefore, by Fano's Lemma, for all $\bar{M} \in B$,

$$\log \tilde{N} \leqslant \frac{R(\tilde{\mu}, \tilde{P}_{\bar{M}}, \tilde{\mathscr{M}}, Y(n)) + 1}{1 - \lambda'} \tag{2.5}$$

where $\tilde{\mu}(\cdot)$ is the p.d. defined on $\tilde{\mathscr{M}}$ by

$$\tilde{\mu}(\tilde{M}) = \begin{cases} \dfrac{1}{\tilde{N}}, & \text{if } \tilde{M} \text{ is in the code in (2.4)} \\ 0 & \text{otherwise.} \end{cases}$$

Now for each $t = 1,..., n$, let $\tilde{\mu}^t(\cdot)$ be a p.d. on $\tilde{X}$ defined by

$$\tilde{\mu}^t(\tilde{x}) = \sum_{\{\tilde{M} | \tilde{M}^t = \tilde{x}\}} \tilde{\mu}(\tilde{M}) \qquad \text{for all} \quad \tilde{x} \in \tilde{X}.$$

From an argument similar to that of the proof of Theorem 4.2.1 in (Gallager, 1968), it can be concluded that

$$R(\tilde{\mu}, \tilde{P}_{\bar{M}}, \tilde{\mathscr{M}}, Y(n)) \leqslant \sum_{t=1}^{n} R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y). \tag{2.6}$$

Then (2.5) and (2.6) yield

$$\log \tilde{N} \leqslant \frac{\sum_{t=1}^{n} R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y) + 1}{1 - \lambda'} \tag{2.7}$$

for all $\bar{M} \in B$.

Averaging (2.7) over all $\bar{M} \in B$ gives

$$\log \tilde{N} \leqslant (1 - \lambda')^{-1} \left[\sum_{t=1}^{n} \frac{1}{\bar{N}} \sum_{\bar{M} \in \Omega} R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y) \right]$$

$$+ \left\{ (1 - \lambda')^{-1} \left[\sum_{t=1}^{n} \sum_{\bar{M} \in B} \frac{1}{| B |} R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y) \right. \right.$$

$$\left. \left. - \sum_{t=1}^{n} \sum_{\bar{M} \in \Omega} \frac{1}{\bar{N}} R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y) \right] \right\}. \tag{2.8}$$

Since $0 \leqslant R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y) \leqslant \log|\tilde{X}|$ for all $t = 1,..., n$ and $|B| \geqslant [(\lambda' - \lambda)/(\lambda')]\bar{N}$, the term in curly brackets in (2.8) is smaller than

$$E(n, \lambda) = n(\log|\tilde{X}|)(1 - \lambda')^{-1}\left(\frac{\lambda}{\lambda' - \lambda}\right)$$
$$= n(1 - \lambda')^{-1}(\log\lambda)(\log|\tilde{X}|)$$

Hence

$$\log\tilde{N} \leqslant (1 - \lambda')^{-1}\left[\sum_{t=1}^{n}\frac{1}{\bar{N}}\sum_{\bar{M}\in\Omega} R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y)\right] + E(n, \lambda). \tag{2.9}$$

Note that $\tilde{\mu}^t(\tilde{x}) = p_1^t(x_1) \cdots p_d^t(x_d)$ for all $\tilde{x} = (x_1,..., x_d) \in \tilde{X}$. This fact, together with the definitions of $R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y)$ and $R_D(p^t, \omega, \hat{X}, Y)$ yield

$$\frac{1}{\bar{N}}\sum_{\bar{M}\in\Omega} R(\tilde{\mu}^t, \omega_{\bar{M}}^t, \tilde{X}, Y) = R_D(p^t, \omega, \hat{X}, Y). \tag{2.10}$$

Putting together (2.9) and (2.10) gives

$$\log\tilde{N} \leqslant \sum_{t=1}^{n} R_D(p^t, \omega, \hat{X}, Y)$$
$$+ \left\{\lambda'(1 - \lambda')^{-1}\left[\sum_{t=1}^{n} R_D(p^t, \omega, \hat{X}, Y)\right]\right.$$
$$\left. + (1 - \lambda')^{-1} + E(n, \lambda)\right\}. \tag{2.11}$$

The term in curly brackets in (2.11) is smaller than

$$k_D(\lambda, n) = \lambda'(1 - \lambda')^{-1}(n\log|\hat{X}|) + (1 - \lambda')^{-1} + E(n, \lambda)$$

since $0 \leqslant R_D(p^t, \omega, \hat{X}, Y) \leqslant \log|\hat{X}|$ for all nonempty $D \subseteq \{1,..., s\}$. This fact and (2.11) yield the conclusion of the theorem, in case $1 \leqslant d < s$.

Now assume $d = s$. Again we construct an auxiliary channel, this time a d.m.c. It has input alphabet X, output alphabet Y, and channel probability function $\omega(\cdot\mid\cdot)$. Then the code $\{(M(\bar{\imath}), A(\bar{\imath})) \mid \bar{\imath} \in \bar{I}\}$, although originally intended for the multi-way channel, can be regarded as a code $-(n, N)$ (again, see (Wolfowitz, 1964) for notation) for the d.m.c., by letting the s senders coalesce. Furthermore, by (1.3) this code has average probability of error λ. Thus by Fano's Lemma,

$$\log N \leqslant \frac{R(\hat{\mu}, P_n, \mathcal{M}, Y(n)) + 1}{1 - \lambda} \tag{2.12}$$

where

$$\hat{\mu}(M) = \begin{cases} \dfrac{1}{N}, & \text{if } M = M(\bar{\imath}) \quad \text{for some } \bar{\imath} \in \bar{I} \\ 0, & \text{otherwise} \end{cases}$$

for all $M \in \mathscr{M}$.

By Theorem 4.2.1 in (Gallager, 1968), we have

$$R(\hat{\mu}, P_n, \mathscr{M}, Y(n)) \leqslant \sum_{t=1}^{n} R(\hat{\mu}^t, \omega, \hat{X}, Y) \tag{2.13}$$

where $\hat{\mu}^t(\hat{x}) = \sum_{\{M: M^t = \hat{x}\}} \hat{\mu}(M)$ for all $\hat{x} \in \hat{X}$. Noting that $\hat{\mu}^t(\hat{x}) = p^t(\hat{x})$ for all $\hat{x} \in \hat{X}$ and $t = 1,..., n$, (2.12) and (2.13) yield

$$\log N \leqslant \frac{\sum_{t=1}^{n} R(p^t, \omega, \hat{X}, Y) + 1}{1 - \lambda}.$$

Since $0 \leqslant R(p^t, \omega, \hat{X}, Y) \leqslant \log |\hat{X}|$ for all $t = 1,..., n$,

$$\log N \leqslant \sum_{t=1}^{n} R(p^t, \omega, \hat{X}, Y) + k(\lambda, n)$$

where $k(\lambda, n) = n\lambda(1 - \lambda)^{-1} \log |\hat{X}| + (1 - \lambda)^{-1}$, and the proof is complete.

3. Capacity Region of a Channel with s Senders and One Receiver

Order the $l = 2^s - 1$ non-empty subsets of $\{1,..., s\}$ and call them $D(1),..., D(l)$. To make the notation less complicated, we denote $R_D(q, \omega, \hat{X}, Y)$ simply by $R_D(q)$. Then define

$$F(Y) = \{(R_{D(1)}(q),..., R_{D(l)}(q)): q = q_1 \times \cdots \times q_s \text{ for some } q_1(\cdot),..., q_s(\cdot), \text{ where } q_k(\cdot) \text{ is a p.d. on } X_k \text{ for } k = 1,..., s\}.$$

Also define

$$\bar{F}(Y) = \left\{\vec{R}: \vec{R} = \frac{1}{n} \sum_{t=1}^{n} (R_{D(1)}(p^t),..., R_{D(l)}(p^t)) \text{ where } p^t = p_1{}^t \times \cdots \times p_s{}^t \text{ and } p_k{}^t \text{ is a p.d. on } X_k \text{ for } t = 1,..., n \text{ and } k = 1,..., s\right\}$$

Let $F^*(Y)$ denote the convex hull of $F(Y)$ and $\vec{R} = (R_1{}^*,..., R_l{}^*)$ denote an arbitrary member of $F^*(Y)$. (Note that $F(Y) \subseteq \bar{F}(Y) \subseteq F^*(Y)$).

Then let

$$G(\vec{R}, Y) = \left\{(R_1, ..., R_s): \sum_{k \in D(m)} R_k \leqslant R_m^* \text{ for all } m = 1, ..., l\right\}$$

$$\text{and } G(Y) = \bigcup_{\vec{R} \in F^*(Y)} G(\vec{R}, Y).$$

LEMMA 2. *$G(Y)$ is convex, closed under projections, and compact in the usual topology of Euclidean s-space.*

Proof. The facts that $G(Y)$ is convex, closed under projections and bounded are immediate from the definition of $G(Y)$. It only remains to show it is closed.

A p.d. $q(\cdot)$ on a finite set A with $|A| = a$ can be viewed as a "probability vector" $q = (q, ..., q_a)$ where q_k, for all k, $1 \leqslant k \leqslant a$, is the probability attached to the k-th element of A in some ordering. Thus $q_k \geqslant 0$ for all k, $1 \leqslant k \leqslant a$, and $\sum_{k=1}^{a} q_k = 1$. Viewed in this sense, the set of all product p.d.'s on $\hat{X}$ becomes a compact subset of Euclidean $|\hat{X}|$-space. Then by the continuity of the rate functions, $F(Y)$ is a compact subset of Euclidean l-space. Since the convex hull of a compact set in a Euclidean space is also compact, $F^*(Y)$ is compact.

Let $\bar{R}(1)$, $\bar{R}(2)$, $\bar{R}(3)$,... be a sequence of elements of $G(Y)$ where $\lim_{n\to\infty} \bar{R}(n)$ exists and equals $\bar{R}$, say. We will be done if we show that $\bar{R} \in G(Y)$.

For all $n = 1, 2, 3, ...$ there exist $\vec{R}(n) \in F^*(Y)$ such that $\bar{R}(n) \in G(\vec{R}(n), Y)$. Let $\bar{R} = (R_1, ..., R_s)$, $\bar{R}(n) = (R_1(n), ..., R_s(n))$ and $\vec{R}(n) = (R_1^*(n), ..., R_l^*(n))$ for all $n = 1, 2, 3, ...$. By the boundedness of $F^*(Y)$ there is a $\beta < 0$ such that $R_m^*(n) \leqslant \beta$ for all $n = 1, 2, 3, ...$, and $m = 1, ..., l$.

Let $\epsilon > 0$. Then there is a positive integer $n(\epsilon)$ such that $n \geqslant n(\epsilon)$ implies $\sum_{k\in D(m)} R_k - \epsilon \leqslant R_m^*(n) \leqslant \beta$ for all $m = 1, ..., l$. Hence there is a subsequence $\{n_t\}_{t=1}^{\infty}$ of $\{n\}_{n=1}^{\infty}$ such that for all $m = 1, ..., l$, $\lim_{t\to\infty} R_m^*(n_t)$ exists and equals R_m^*, say. Since $F^*(Y)$ is closed, $\vec{R} = (R_1^*, ..., R_l^*) \in F^*(Y)$.

Furthermore, $\sum_{k\in D(m)} R_k - \epsilon \leqslant R_m^*$ for all $m = 1, ..., l$. Since ϵ was arbitrary, $\sum_{k\in D(m)} R_k \leqslant R_m^*$ for all $m = 1, ..., l$. Hence $\bar{R} \in G(\vec{R}, Y)$, which implies $\bar{R} \in G(Y)$.

THEOREM 1. *The capacity region $G(P, T_{s_1}) = G(Y)$.*

Proof. First we show $G(P, T_{s1}) \subseteq G(Y)$. Let $(R_1, ..., R_s) \in G(P, T_{s1})$. Let $\epsilon > 0$ and $0 < \lambda < 1$. Then for all n sufficiently large, there is a code $-(n, \vec{N}, \lambda)$ such that $(1/n) \log N_k \geqslant R_k - \epsilon$ for all $k = 1, ..., s$. Using this fact and Lemma 1, it can be concluded that, for any $\delta > 0$, if ϵ and λ are chosen sufficiently small and n sufficiently large,

$$\sum_{k\in D} R_k \leqslant \frac{1}{n}\sum_{t=1}^{n} R_D(p^t) + \delta$$

for all non-empty $D \subseteq \{1,..., s\}$. Since

$$\frac{1}{n}\sum_{t=1}^{n} (R_{D(1)}(p^t),..., R_{D(l)}(p^t)) \in F^*(Y),$$

$(R_1 - \delta,..., R_s - \delta) \in G(Y)$. Because δ was arbitrary, $(R_1 ,..., R_s)$ belongs to the closure of $G(Y)$, and hence $G(Y)$, since it is closed.

Now for the direct half—we use Shannon's random coding method. Suppose that the following items, to be specified later, are given: positive integers n, $N_1 ,..., N_s$ and a collection $\{p_k{}^t: 1 \leqslant t \leqslant n,\ 1 \leqslant k \leqslant s\}$ of p.d.'s, where $p_k{}^t(\cdot)$ is a p.d. on X_k for all $t = 1,..., n$ and $k = 1,..., s$. Let p.d.'s $p^t(\cdot)$, $p_k(\cdot)$ and $p(\cdot)$ be defined as in (2.2) in terms of the $p_k{}^t$'s.

Let $\mathscr{C}$ denote the collection of all sets of codewords C as defined in (ii) of (1.2). Define a p.d. $p^*(\cdot)$ on $\mathscr{C}$ by

$$p^*(C) = \prod_{k=1}^{s} \prod_{i_k=1}^{N_k} p_k(M_k(i_k)) \qquad \text{for all} \quad C \in \mathscr{C}. \tag{3.1}$$

Then choose a set of codewords C at random according to the p.d. $p^*(\cdot)$. For each $\bar{\imath} \in \bar{I}$, then, $M(\bar{\imath})$ will be the matrix satisfying $M_k(\bar{\imath}) = M_k(i_k)$ for all $k = 1,..., s$.

Once the codewords have been chosen, define maximum likelihood decoding sets (depending on C) by

$$A(\bar{\imath}) = \{y(n): y(n) \in Y(n) \text{ and } P_n(y(n) \mid M(\bar{\imath})) > P_n(y(n) \mid M(\bar{\jmath})) \text{ for all } \bar{\jmath} \neq \bar{\imath}\}.$$

The average error for the code $\{(M(\bar{\imath}), A(\bar{\imath})) \mid \bar{\imath} \in \bar{I}\}$ is

$$\lambda(C) = \frac{1}{N} \sum_{\bar{\imath}\in\bar{I}} P_n(A(\bar{\imath})^c \mid M(\bar{\imath})).$$

If $\lambda^*(\mathscr{C})$ is a r.v. which takes the value x with probability $p^*\{C: \lambda(C) = x\}$, for all real x, then the random coding method requires a suitable upper bound on $E\lambda^*(\mathscr{C}) = \sum_{C\in\mathscr{C}} p^*(C)\, \lambda(C)$. We proceed now to derive such a bound.

Note that

$$\sum_{C\in\mathscr{C}} p^*(C)\, P_n(A(\bar{\imath})^c \mid M(\bar{\imath})) = \sum_{C\in\mathscr{C}} p^*(C)\, P_n(A(\bar{\jmath})^c \mid M(\bar{\jmath}))$$

for all $\bar{\imath}, \bar{\jmath} \in \bar{I}$. Thus if $\lambda_1^*(\mathscr{C})$ is a r.v. taking the value x with probability $p^*\{C: P_n(A(\bar{1})^c \mid M(\bar{1})) = x\}$ for all real x, where $\bar{I} = (1, 1,..., 1)$, then

$$E\lambda_1^*(\mathscr{C}) = E\lambda^*(\mathscr{C}). \tag{3.2}$$

Now

$$\begin{aligned} E\lambda_1^*(\mathscr{C}) &= \sum_{C\in\mathscr{C}} p^*(C)\, P_n(A(\bar{1})^c \mid M(\bar{1})) \\ &= \sum_{M(\bar{1})\in\mathscr{M}} p(M(\bar{1}))\, P_n(A(\bar{1})^c \mid M(\bar{1})) \\ &= \sum_{M\in\mathscr{M}} \sum_{y(n)\in Y(n)} p(M)\, P_n(y(n) \mid M) \\ &\quad \times p^*\{C: M(\bar{1}) = M \text{ and } P_n(y(n) \mid M(\bar{1})) \\ &\quad \leqslant P_n(y(n) \mid M(\bar{\imath})) \text{ for some } \bar{\imath} \neq \bar{1}\} \\ &\leqslant \sum_{\bar{\imath}\neq\bar{1}} \sum_{M\in\mathscr{M}} \sum_{y(n)\in Y(n)} p(M)\, P_n(y(n) \mid M) \\ &\quad \times p^*\{C: M(\bar{1}) = M \text{ and } P_n(y(n) \mid M(\bar{1})) \\ &\quad \leqslant P_n(y(n) \mid M(\bar{\imath}))\}. \end{aligned}$$

The object then is to bound from above, for each $\bar{\imath} \neq \bar{1}$, the corresponding term in the sum of the last expression in (3.3).

Let $\bar{\imath} \neq \bar{1}$ be fixed and let $D = \{k: i_k \neq 1,\ 1 \leqslant k \leqslant s\}$. Let $d = |D|$. Then if $M \in \mathscr{M}$, $\tilde{M}$ shall denote the $n \times d$ matrix obtained from M by deleting the $(s - d)$ columns with indices $k \notin D$. Likewise $\overline{M}$ shall denote the $n \times (s - d)$ matrix obtained from M by deleting the d columns with indices $k \in D$. $\tilde{\mathscr{M}}$ and $\overline{\mathscr{M}}$ denote the collections of all matrices $\tilde{M}$ and $\overline{M}$, respectively, as M ranges over $\mathscr{M}$. Also, if $U \in \tilde{\mathscr{M}}$ and $V \in \overline{\mathscr{M}}$, then UV denotes the matrix $M \in \mathscr{M}$ with $\tilde{M} = U$ and $\overline{M} = V$.

Define p.d.'s $\tilde{p}(\cdot)$ and $\bar{p}(\cdot)$ on $\tilde{\mathscr{M}}$ and $\overline{\mathscr{M}}$, respectively, by

$$\tilde{p}(U) = \sum_{V\in\overline{\mathscr{M}}} p(UV) \qquad \text{for all} \quad U \in \tilde{\mathscr{M}}$$

and $\bar{p}(V) = \sum_{U\in\tilde{\mathscr{M}}} p(UV)$ for all $V \in \overline{\mathscr{M}}$. Note that $p(M) = \tilde{p}(\tilde{M})\,\bar{p}(\overline{M})$ for all $M \in \mathscr{M}$.

For all non-empty $D \subseteq \{1,..., s\}$, define the "information function" $I_D(\cdot, \cdot)$ by

$$I_D(y(n), M) = \log \frac{P_n(y(n) \mid M)}{\sum_{U \in \tilde{\mathscr{M}}} \tilde{p}(U) P_n(y(n) \mid U\bar{M})}$$

for all $y(n) \in Y(n)$ and $M \in \mathscr{M}$. Let $\tilde{N} = \prod_{k \in D} N_k$. Let α be a constant to be specified later, and define $B(V) = \{(y(n), U): y(n) \in Y(n), U \in \tilde{\mathscr{M}}$, and $I_D(y(n), UV) \geqslant \log \alpha\tilde{N}\}$ for all $V \in \bar{\mathscr{M}}$.

Then we have that

$$\sum_{M \in \mathscr{M}} \sum_{y(n) \in Y(n)} p(M) P_n(y(n) \mid M) p^*\{C: M(\bar{1}) = M \text{ and } P_n(y(n) \mid M(\bar{1})) \leqslant P_n(y(n) \mid M(\bar{\imath}))\}$$
$$\leqslant \sum_{V \in \bar{\mathscr{M}}} \bar{p}(V) \Big[\sum_{B(V)} \tilde{p}(U) P_n(y(n) \mid UV) p^*\{C: M(\bar{1}) = UV \text{ and } P_n(y(n) \mid M(\bar{1})) \leqslant P_n(y(n) \mid M(\bar{\imath}))\} + \sum_{B(V)^c} \tilde{p}(U) P_n(y(n) \mid UV) \Big]. \quad (3.4)$$

Using the definition of $B(V)$, a little calculation yields

$$\sum_{B(V)} \tilde{p}(U) P_n(y(n) \mid UV) p^*\{C: M(\bar{1}) = UV \text{ and } P_n(y(n) \mid M(\bar{1})) \leqslant P_n(y(n) \mid M(\bar{\imath}))\} \leqslant \frac{1}{\alpha \tilde{N}} \text{ for all } V \in \bar{\mathscr{M}}. \quad (3.5)$$

As for the second term in square brackets on the *RHS* of the inequality in (3.4), let $I_D{}^*$ be a r.v. which takes the value x with probability

$$\sum_{B(x)} p(M) P_n(y(n) \mid M)$$

for all real x, where $B(x) = \{(y(n), M): I_D(y(n), M) = x\}$. Then

$$EI_D{}^* = \sum_{t=1}^{n} R_D(p^t) \qquad \text{for all} \quad D \subseteq \{1,..., s\}, \quad D \neq \varnothing. \quad (3.6)$$

Now choose α so that

$$\alpha\tilde{N} < \exp\left\{\sum_{t=1}^{n} R_D(p^t) - k\sqrt{n}\right\} \quad (3.7)$$

for some positive constant k. From (3.6) and (3.7) it follows that

$$\sum_{V \in \bar{\mathscr{M}}} \bar{p}(V) \sum_{B(V)^c} \tilde{p}(U) P_n(y(n) \mid UV) \leqslant \sum_{T} p(M) P_n(y(n) \mid M) \quad (3.8)$$

where $T = \{(y(n), M): I_D(y(n), M) < EI_D^* - k\sqrt{n}\}$. By Chebyshev's inequality, the RHS of the inequality (3.8) is bounded above by $\mathrm{Var}(I_D^*)/k^2 n$. It is known (for example see Wolfowitz, 1964, Chapter 8) that there is a constant k_0, independent of n, such that $\mathrm{Var}(I_D^*) \leqslant k_0 n$, for $n = 1, 2, \ldots$.

Combining these facts together with (3.2)–(3.5) and (3.8), we have

$$E\lambda^*(\mathscr{C}) \leqslant \frac{l}{\alpha N} + \frac{k_0 l}{k^2}. \tag{3.9}$$

Now we are ready to show that $G(Y) \subseteq G(P, T_{s1})$. Let $(R_1, \ldots, R_s) \in G(Y)$. Let $\epsilon > 0$, $0 < \delta < \epsilon$ and $0 < \lambda < 1$. Then there is an l-tuple

$$\vec{R} = (R_1^*, \ldots, R_l^*) \in F^*(Y)$$

such that

$$\sum_{k \in D(m)} R_k \leqslant R_m^* \qquad \text{for all} \quad m = 1, \ldots, l. \tag{3.10}$$

Also it is possible to find a $\vec{R}' = (R_1', \ldots, R_l') \in \bar{F}(Y)$ such that

$$| R_m^* - R_m' | < \delta/2 \qquad \text{for} \quad m = 1, \ldots, l. \tag{3.11}$$

Let $n(\delta)$ be a positive integer and $\{q_k^t \mid 1 \leqslant t \leqslant n(\delta),\ 1 \leqslant k \leqslant s\}$ be a collection of p.d.'s such that

$$R_m' = \frac{1}{n(\delta)} \sum_{t=1}^{n(\delta)} R_{D(m)}(q^t) \qquad \text{for} \quad m = 1, \ldots, l.$$

Find a positive integer n_0 such that if $n \geqslant n_0$,

$$\frac{n(\delta)}{n} R_m' + \frac{n(\delta) \log | \hat{X} |}{n} < \frac{\delta}{2}. \tag{3.12}$$

Choose α and k sufficiently large so that

$$\frac{l}{\alpha N} < \frac{\lambda}{2} \qquad \text{and} \qquad \frac{k_0 l}{k^2} < \frac{\lambda}{2}. \tag{3.13}$$

Then find a positive integer n_1 such that, whenever $n \geqslant n_1$,

$$k\sqrt{n} + \log \alpha + s \log 2 - n(\epsilon - \delta) < 0. \tag{3.14}$$

Let n be an integer satisfying

$$n \geqslant \max(n(\delta), n_0, n_1). \tag{3.15}$$

Define a collection $\{p_k{}^t \mid t = 1,..., n; k = 1,..., s\}$ of p.d.'s by

$$p_k{}^t = q_k^{[t]} \qquad \text{for all} \quad t = 1,..., n, \quad k = 1,..., s, \tag{3.16}$$

where $[t] \equiv t(\text{mod } n(\delta))$ and $1 \leqslant [t] \leqslant n(\delta)$.

Now define

$$R_m'' = \frac{1}{n} \sum_{t=1}^{n} R_{D(m)}(p^t) \qquad \text{for} \quad m = 1,..., l.$$

A little calculation shows that

$$| R_m'' - R_m{}' | \leqslant \frac{n(\delta)}{n} R_m{}' + \frac{n(\delta) \log | \hat{X} |}{n} \qquad \text{for all} \quad m = 1,..., l. \tag{3.17}$$

Thus (3.11), (3.12) and (3.17) yield

$$| R_m'' - R_m{}^* | < \delta \qquad \text{for all} \quad m = 1,..., l. \tag{3.18}$$

Now if

$$N_k = \langle e^{n(R_k - \epsilon)} \rangle \qquad \text{for all} \quad k = 1,..., s, \tag{3.19}$$

where $\langle x \rangle$ denotes the smallest integer $\geqslant x$, choose a code $-(n, \vec{N})$ at random as described earlier in the proof using (3.15), (3.16) and (3.19) to specify n, $\{p_k{}^t: 1 \leqslant t \leqslant n, 1 \leqslant k \leqslant s\}$ and $\{N_k: k = 1,..., s\}$.

Now (3.9) will hold if α satisfies (3.7). But if m is the index such that $D = D(m)$, then by (3.10), (3.14), (3.18) and (3.19), we have

$$\begin{aligned}
\alpha \tilde{N} &\leqslant \alpha \prod_{k \in D} \{\exp[n(R_k - \epsilon)] + 1\} \\
&\leqslant \alpha 2^d \prod_{k \in D} \{\exp[n(R_k - \epsilon)]\} \\
&\leqslant \alpha 2^s \exp\left\{\sum_{k \in D} n(R_k - \epsilon)\right\} \\
&\leqslant \alpha 2^s \exp\{nR_m{}^* - n\epsilon\} \\
&\leqslant \alpha 2^s \exp\{nR_m'' - n\delta - n\epsilon\}
\end{aligned}$$

$$= \alpha 2^s \exp \left\{ \sum_{t=1}^{n} R_D(p^t) + n\delta - n\epsilon \right\}$$

$$= \exp \left\{ \sum_{t=1}^{n} R_D(p^t) - k\sqrt{n} + (k\sqrt{n} + \log \alpha + s \log 2 - n(\epsilon - \delta)) \right\}$$

$$\leqslant \exp \left\{ \sum_{t=1}^{n} R_D(p^t) - k\sqrt{n} \right\}.$$

Thus (3.9) holds, and by (3.13), $E\lambda^*(\mathscr{C}) \leqslant \lambda$. Since the expected error averaged over all randomly chosen codes does not exceed λ, there must exist a code $-(n, \vec{N}, \lambda)$. Thus for all $\epsilon > 0$ and $0 < \lambda < 1$ and $n \geqslant \max(n(\delta), n_0, n_1)$, there is a code $-(n, \vec{N}, \lambda)$ for (P, T_{s_1}). Therefore $(R_1, ..., R_s) \in G(P, T_{s_1})$, that is, $(R_1, ..., R_s)$ is a set of achievable rates, and the theorem is proved.

4. Capacity Region of a Channel with s Senders and r Receivers

We now characterize $G(P, T_{sr})$ for $s \geqslant 2$ and $r \geqslant 2$.

For all $j = 1, ..., r$, define $\omega_j(\cdot \mid \cdot)$ on $\hat{X} \times Y_j$ by

$$\omega_j(y \mid \hat{x}) = \sum_{\hat{Y}(y,j)} \omega(\hat{y} \mid \hat{x}) \qquad \text{for all } \hat{x} \in \hat{X} \text{ and all } y \in Y_j,$$

where $\hat{Y}(y, j) = \{\hat{y} : \hat{y} = (y_1, ..., y_r) \in \hat{Y} \text{ and } y_j = y\}$. Then if $q(\cdot)$ is a p.d. on $\hat{X}$, define for each $j = 1, ..., r$ and non-empty $D \subseteq \{1, ..., s\}$, $R_D{}^j(q) = R_D(q, \omega_j, \hat{X}, Y_j)$.

Let ρ denote a finite set of s-tuples $(q_1, ..., q_s)$ where $q_k(\cdot)$ is a p.d. on X_k for $k = 1, ..., s$. Let $q = q_1 \times \cdots \times q_s$, and let $\mu(\cdot)$ be a p.d. on ρ.

To each pair (ρ, μ) is assigned a vector $\vec{R}(\rho, \mu)$ as follows. Let $\bar{R}_j(q) = [R^j_{D(1)}(q), ..., R^j_{D(l)}(q)]$ for all $j = 1, ..., r$ and define $\vec{R}(\rho, \mu) = (\tilde{R}_1, ..., \tilde{R}_l)$ where $\tilde{R}_m = \min\{\tilde{R}_m{}^1, ..., \tilde{R}_m{}^s\}$ for all $m = 1, ..., l$, and $\tilde{R}_m{}^k$ is the mth component of $\sum_\rho \mu(q_1, ..., q_s) \bar{R}_k(q)$ for $k = 1, ..., s$ and $m = 1, ..., l$.

Then denote $F(\hat{Y}) = \{\vec{R} : \vec{R} = \vec{R}(\rho, \mu) \text{ for some } (\rho, \mu)\}$. Define $G(\vec{R}, \hat{Y}) = \{(R_1, ..., R_s) : \sum_{k \in D(m)} R_k \leqslant \tilde{R}_m \text{ for } m = 1, ..., l\}$, and $G(\hat{Y}) = \bigcup_{\vec{R} \in F(\hat{Y})} G(\vec{R}, \hat{Y})$.

We remark that $G(\hat{Y})$ is convex, closed under projections and compact in the usual topology of Euclidean s-space.

Theorem 2. *The capacity region* $G(P, T_{sr}) = G(\hat{Y})$.

Proof. First we show $G(P, T_{sr}) \subseteq G(\hat{Y})$. Let $(R_1, ..., R_s) \in G(P, T_{sr})$. Then for all $\epsilon > 0$ and $0 < \lambda < 1$ and all n sufficiently large, there exists a code $-(n, \vec{N}, \lambda)$ for (P, T_{sr}) such that $(1/n) \log N_k \geqslant R_k - \epsilon$ for all $k = 1,..., s$. Let $D \subseteq \{1,..., s\}$, $D \neq \phi$, and let $p_k^t(\cdot)$ be defined as in (2.3) for $t = 1,..., n$ and $k = 1,..., s$. By Lemma 1, we can find a number $k_D(\lambda, n)$ such that

$$\log\left(\prod_{k \in D} N_k\right) \leqslant \sum_{t=1}^{n} R_D{}^j(p^t) + k_D(\lambda, n)$$

for all $j = 1,..., r$, where $(1/n)\, k_D(\lambda, n) \to 0$ as $\lambda \to 0$ and $n \to \infty$.

Then let $\rho = \{(p_1{}^t,..., p_s{}^t): t = 1,..., n\}$ and $\mu(\cdot)$ be the equidistribution on ρ. By arguing as in Theorem 1, it follows that for all $\delta > 0$, if ϵ and λ are sufficiently small, and n sufficiently large, $(R_1 - \delta,..., R_s - \delta) \in G(\vec{R}, \hat{Y})$ where $\vec{R} = \vec{R}(\rho, \mu)$. Thus $(R_1 ,..., R_s)$ belongs to the closure of $G(\hat{Y})$, and hence to $G(\hat{Y})$, since it is closed.

Finally we show $G(\hat{Y}) \subseteq G(P, T_{sr})$. Let $(R_1 ,..., R_s) \in G(\hat{Y})$. Then there exists a $\vec{R} = \vec{R}(\rho, \mu) = (\tilde{R}_1 ,..., \tilde{R}_l) \in F(\hat{Y})$ such that $\sum_{k \in D(m)} R_k \leqslant \tilde{R}_m$ for $m = 1,..., l$. Let $\epsilon > 0$, $0 < \delta < \epsilon$ and $0 < \lambda < 1$. Find a positive integer $n(\delta)$ and a collection $\{q_k{}^t: 1 \leqslant t \leqslant n(\delta),\ 1 \leqslant k \leqslant s\}$ of p.d.'s, where $q_k{}^t(\cdot)$ is a p.d. on X_k for $t = 1,..., n(\delta)$ and $k = 1,..., s$, such that the following holds: if $\rho' = \{(q_1{}^t,..., q_s{}^t): 1 \leqslant t \leqslant n(\delta)\}$, $\mu'(\cdot)$ is the equidistribution on ρ', and $\vec{R}' = (\tilde{R}_1',..., \tilde{R}_l') = \vec{R}(\rho', \mu')$, then $|\, \tilde{R}_m - \tilde{R}_m' \,| < \delta/2$ for $m = 1,..., l$.

For all $n \geqslant n(\delta)$, define the collection $\{p_k{}^t: 1 \leqslant t \leqslant n,\ 1 \leqslant k \leqslant s\}$ of p.d.'s by (3.16), in terms of the $q_k{}^t$'s above. Also define $N_k = \langle e^{n(R_k - \epsilon)} \rangle$ for $k = 1,..., s$.

Now select a set $C \in \mathscr{C}$ of codewords at random according to the p.d. $p^*(\cdot)$ (defined in terms of the $p_k{}^t$'s above) as in the proof of Theorem 1. Define decoding sets for all $\bar{\imath} \in \bar{I}$ and $j = 1,..., r$ by

$$\begin{aligned} A_j(\bar{\imath}) = \{y_j(n)\colon y_j(n) \in Y_j(n) \text{ and } P_n{}^j(y_j(n) \mid M(\bar{\imath})) \\ > P_n{}^j(y_j(n) \mid M(\bar{\imath}')) \text{ for all } \bar{\imath}' \neq \bar{\imath}\}. \end{aligned}$$

Now for all $j = 1,..., r$ and $C \in \mathscr{C}$, define

$$f_j(C) = \frac{1}{N} \sum_{\bar{\imath} \in \bar{I}} P_n{}^j(A_j(\bar{\imath})^c \mid M(\bar{\imath})).$$

Also let $f_j{}^*(\mathscr{C})$ be a r.v. taking the value x with probability $p^*\{C\colon f_j(C) = x\}$ for all real x.

Then for each $j = 1,..., r$ argue as in the proof of Theorem 1 with $A(\bar{\imath})$ replaced by $A_j(\bar{\imath})$, $P_n(\cdot \mid \cdot)$ replaced by $P_n{}^j(\cdot \mid \cdot)$, and λ replaced by λ/r^2, to conclude that for all sufficiently large n, $Ef_j{}^*(\mathscr{C}) \leqslant \lambda/r^2$.

By a Lemma of Shannon (1961), there exists a $C \in \mathscr{C}$ such that $f_j(C) \leqslant rEf_j^*(\mathscr{C}) \leqslant \lambda/r$ for all $j = 1,\ldots, r$. Thus $\sum_{j=1}^{r} f_j(C) \leqslant \lambda$ and (1.3) holds. Therefore $(R_1, \ldots, R_s) \in G(P, T_{sr})$ and the proof is complete.

Remark. At the time the original manuscript of this paper was submitted, the author had just received a copy of "A Coding Theorem for Multiple Access Channels with Correlated Sources", by D. Slepian and J. K. Wolf, in which the authors obtained results for a channel with two senders and one receiver for certain correlated sources.

Acknowledgments

The author would like to express his gratitude to Professor Rudolf Ahlswede for his invaluable advice in the preparation of the author's doctoral dissertation, upon which this paper is based. Also, the comments and suggestions of the referee were a great help in making some of the proofs more rigorous and the paper as a whole more readable.

Received: June 5, 1975

References

Ahlswede, R. (1971a), On two-way communication channels and a problem by Zarankiewicz, presented at the Sixth Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes.

Ahlswede, R. (1971b), Multi-way communication channels, *in* Second International Symposium on Information Theory, Publishing House of the Hungarian Academy of Sciences, 23–52.

Ahlswede, R. (1974), The capacity region of a channel with two senders and two receivers, *The Annals of Probability* 2, 805–814.

Fano, R. M. (1952, 1954), Statistical theory of communication, Notes on a course given at MIT.

Gallager, R. G. (1968), "Information Theory and Reliable Communication," Wiley, New York.

Shannon, C. E. (1957), Certain results in coding theory for noisy channels, *Inform. Contr.* 1, 6–25.

Shannon, C. E. (1961), Two-way communication channels, *Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 611–694, University of California Press, Berkeley.

Slepian, D. and Wolf, J. K. (1973), A coding theorem for multiple access channels with correlated sources, *Bell. System Tech. J.* **52**.

Wolfowitz, J. (1964), "Coding Theorems of Information Theory," Second Edition, Springer-Verlag, Berlin-Heidelberg-New York.

Van der Meulen, E. C. (1972), "Multiterminal communication channels," presented at the Second Conference on Stochastic Processes and Applications, Katholieke Universiteit te Leuven.

Van Der Meulen, E. C. (1974), On a problem by Ahlswede regarding the capacity region of certain multiway channels, *Inform. Contr.* **25**.