stichting
mathematisch
centrum

AFDELING MATHEMATISCHE STATISTIEK SW 24/73 DECEMBER

Edward C. van der MEULEN

RANDOM CODING THEOREMS FOR THE GENERAL DISCRETE MEMORYLESS BROADCAST CHANNEL

Prepublication

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS (MOS) subject classification scheme (1970): 94A15

Random coding theorems for the general discrete memoryless broadcast channel *)

by

Edward C. van der Meulen **)

Abstract

Three different communication situations are considered for the general, non-degraded, discrete memoryless broadcast channel with two components. One of these situations includes the case of sending common, but also separate, information to both receivers. For each communication situation a random coding inner bound on the capacity region is derived. An example is given showing that in one situation the inner bound contains pairs of rate points dominating the time-sharing line. Each capacity region is also described by a limiting expression.

The relationship with the results of Cover and Bergmans on degraded broadcast channels is brought out, and the connection with other multi-way channels, in particular the channel with two senders and two receivers, is also shown.

---

*) This paper is not for review; it is meant for publication in a journal.

**) From January 1, 1974: Department of Statistics, University of Rochester, Rochester, New York 14627, U.S.A.

CONTENTS

I. INTRODUCTION AND SUMMARY OF RESULTS

In a basic paper Cover [5] analyzed the so-called broadcast channel, the problem being how to send information from a single source simultaneously to several receivers. As one of his results, Cover obtained a random coding inner bound on the capacity region of the broadcast channel for the special case when this channel factors out into two binary symmetric channels, one of which is noiseless.

Subsequently, Bergmans [4] generalized Cover's results to the case of the discrete memoryless broadcast channel with degraded components, and stated and proved in a rigorous way a random coding theorem for this class of channels.

Recently, Gallager [7] proved a weak converse, showing that the random coding inner bound obtained by Bergmans is indeed the capacity region of the discrete memoryless degraded broadcast channel.

In the present paper we extend the results of Cover and Bergmans to the general case of a discrete memoryless broadcast channel with two components. In particular, the restriction that the broadcast channel is of degraded type is removed. In Section II the definitions and concepts of the paper are developed.

In Section III a random coding theorem is proved for the situation in which different messages are sent to both receivers. Our proof is based on existing random coding theorems, which were obtained earlier by Ahlswede [1] and the author [15] for the channel with two senders and two receivers. An example, originally due to Blackwell, is presented which shows that it is possible to transmit in this situation at pairs of rates well above the time-sharing line.

Our approach involves the consideration of cascades of multi-way channels. A theorem is proved regarding the use of pure pre-multiplying channels. This theorem can be regarded as a first step towards a more general theory of partial ordering of multi-way channels. It is shown that there is a close connection between the broadcast channel with two components, and the channel with two senders and two receivers introduced by Shannon [10].

Also, a comparison is made with the Cover-Bergmans random coding scheme. It is shown that there are similarities with, but also differences between our random coding scheme for the present communication situation, and their scheme for the degraded broadcast channel. Finally, a limiting expression is derived for the capacity region in this situation.

In Section IV a random coding theorem is proved for the situation in which one message is sent to both receivers, and another message is sent to only one of them. This situation resembles the degraded broadcast channel, except that we do not assume the channel to be degraded from the outset. If the channel is of degraded type, the previous communication situation reduces to the present one. A random coding theorem is proved which leads to an inner bound on the capacity region.

Our approach in this case is based on the technique introduced by Ahlswede [2], which was further developed by Ulrey [12], for coding for a channel with s senders and r receivers, when all senders send messages simultaneously to all receivers. This technique admits the use of non-stationary sources. Our proof involves also some aspects of the random coding proof given by Bergmans for the degraded broadcast channel.

A comparison is made with the results obtained by Bergmans for the degraded channel. In particular, it is shown that our Theorem 5 incorporates Theorem 1 of [4] as a special result. Also, a limiting expression for the capacity region in this case has been found.

In Section V a random coding theorem is proved for the situation in which two different messages are sent to the two receivers, and, in addition, a third common message is sent to both. Our method of proof is again based on the Ahlswede-Ulrey approach for the situation in which all senders send messages simultaneously to all receivers, but involves aspects of other multi-way channels as well. A comparison is made with the random coding inner bounds found for the previous communication situations. It is shown that these can be obtained from the results of Section V as special cases. Finally, a limiting expression for the capacity region is found.

In this paper we have devoted relatively little attention to converses, except when deriving limiting expressions for the various capacity regions. In this connection we should like to point out that in general there is not for *any* multi-way channel with two independent receivers a simple outer bound on the capacity region known which coincides with the single-digits random coding inner bound. Therefore, it is unlikely to find satisfactory outer bounds on the capacity region of the general broadcast channel, as long as the precise capacity region of the channel with two senders and two receivers, each one located at a different terminal, remains unknown. The study of converses and outer bounds on the capacity region of the general broadcast channel might however be the subject of future investigations.

Finally, in this paper we have stated various conjectures. They should be regarded as open problems, the solution of which is presently unknown to the author.

II. DEFINITIONS AND PRELIMINARIES

A. Broadcast Channels

A general broadcast channel with two receivers is depicted in Fig. 1. It consists of three terminals, labeled 1, 2, and 3, which are connected to a noisy channel K. At terminal 1 there is a sender (also called encoder) S_y, and at terminal 2 and 3 there are receivers (also called decoders or users) U_{z1} and U_{z2} respectively. It is the task of S_y to communicate information over the channel to U_{z1} and U_{z2} as effectively as possible. The information to be transmitted consists of different messages which are presented to S_y by different sources. Various communication situations are possible in this context, since one may wish to send separate information and also common information to both receivers. The specific communication problems which we consider in this paper shall be made precise in Section IIB.

The operation of the broadcast channel may be described as follows. Once each second an input letter y is transmitted to the channel at terminal 1, after which output letters z_1 and z_2 are received at terminal 2 and 3, respectively, according to a transition probability $p(z_1,z_2|y)$. We restrict ourselves throughout this paper to discrete memoryless (d.m.) broadcast channels with two receivers.

Formally, a discrete broadcast channel with two receivers, denoted by $(A,p(z_1,z_2|y),B_1\times B_2)$, or by $p(z_1,z_2|y)$, consists of three finite sets A,B_1, and B_2, having $a \geqq 2$, $b_1 \geqq 2$, and $b_2 \geqq 2$ elements, respectively, and a collection of probability distributions $p(z_1,z_2|y)$ on $B_1\times B_2$, one for each $y \in A$. The set A is called the input alphabet for the sender S_y

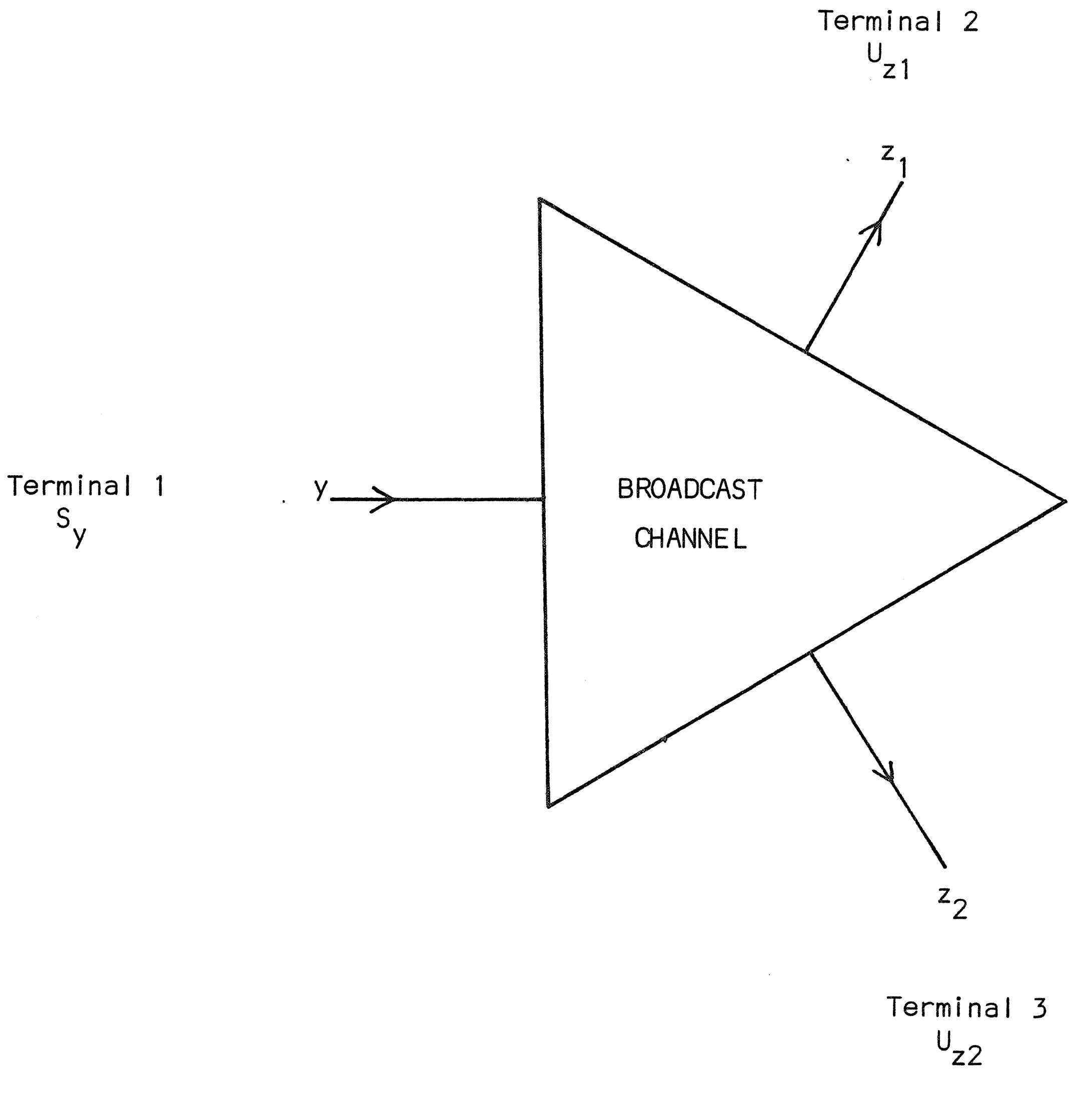
Terminal 2
U_{z1}
z_1
Terminal 1
S_y
y
BROADCAST
CHANNEL
z_2
Terminal 3
U_{z2}

Fig. 1

at terminal 1, whereas B_1 and B_2 are the output alphabets for the receivers U_{z1} and U_{z2} at terminals 2 and 3, respectively; $p(z_1,z_2|y)$ is interpreted as the probability of receiving output letters z_1 and z_2 at terminals 2 and 3, respectively, given that input letter y was transmitted at terminal 1.

For any positive integer n and any set A we denote by A^n the set of all n-tuples $(y_1,\dots,y_n)$ with each $y_i \in A$. A discrete broadcast channel $(A,p(z_1,z_2|y),B_1\times B_2)$ is said to be memoryless if

$$(1) \qquad P(Z_1,Z_2|Y) = \prod_{k=1}^{n} p(z_{1k},z_{2k}|y_k)$$

for all $Y = (y_1,\dots,y_n) \in A^n$, $Z_1 = (z_{11},\dots,z_{1n}) \in B_1^n$, $Z_2 = (z_{21},\dots,z_{2n}) \in B_2^n$, and $n \geqq 1$. $P(Z_1,Z_2|Y)$ is interpreted as the probability of receiving the n-tuples Z_1 and Z_2 at terminals 2 and 3 respectively, given that the n-tuple Y has been transmitted at terminal 1. $P(Z_1,Z_2|Y)$ is called the memoryless n-extension of $p(z_1,z_2|y)$.

Cleary, every d.m. broadcast channel $p(z_1,z_2|y)$ factors out into two marginal d.m. one-way channels defined by

$$(2) \qquad p(z_1|y) = \sum_{z_2\in B_2} p(z_1,z_2|y)$$

and

$$(3) \qquad p(z_2|y) = \sum_{z_1\in B_1} p(z_1,z_2|y).$$

Like Cover [5] and Bergmans [4], we impose a no-collaboration restriction between U_{z1} and U_{z2}. This implies that when one considers the broadcast channel $p(z_1,z_2|y)$, one may restrict attention to the marginal channel transition probabilities defined by (2) and (3).

Ahlswede [1] has developed a convenient notation for multi-way channels. According to his terminology, a d.m. broadcast channel with two receivers can be denoted by a pair (P,T_{12}), where P refers to the transition probabilities as defined in (1), and T_{12} indicates that the channel has one sender and two receivers, each located at a different terminal. We shall either use the notation (P,T_{12}) or simply write $(1s,2r)$ to indicate that we are dealing with a channel with one sender and two receivers.

B. *Communication Situations*

In this paper we consider three communication situations. In order to formulate these properly we first introduce two different communication systems. In the communication system shown in Fig. 2 two sources S_1 and S_2 present two statistically independent messages i and j to the sender S_y for transmission over the channel. Here $1 \leqq i \leqq M_1$, $1 \leqq j \leqq M_2$, and each message pair (i,j) has the same chance $\frac{1}{M_1M_2}$ of being selected. Sender S_y maps the message pair (i,j) into an input sequence $Y \epsilon A^n$ by means of a mapping $f(i,j) = Y$. Subsequently, the sequence Y is transmitted over the channel, and output sequences Z_1 and Z_2 are received by U_{z1} and U_{z2} respectively, with transition probability $P(Z_1,Z_2|Y)$ defined by (1). For the present communication system we consider two communication problems.

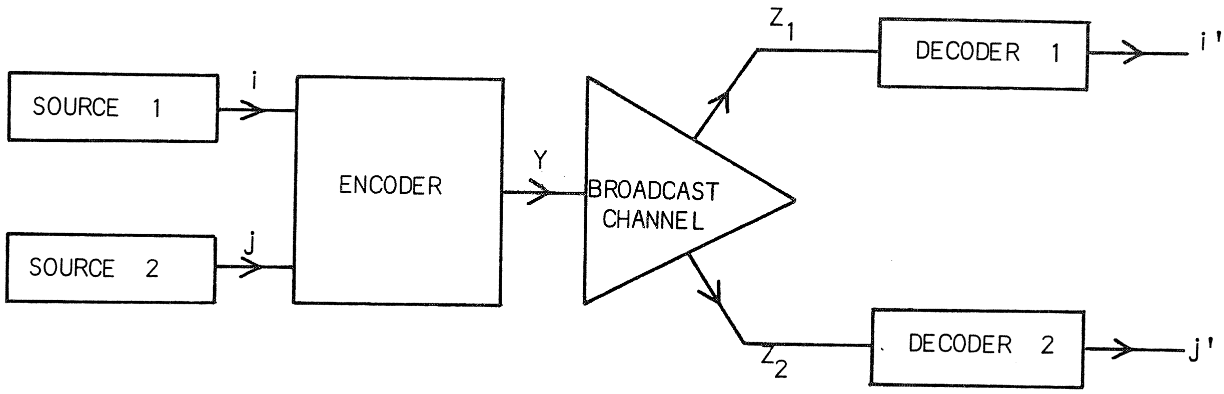
SOURCE 1
SOURCE 2
i
j
ENCODER
Y
BROADCAST CHANNEL
Z1
Z2
DECODER 1
DECODER 2
i'
j'

Fig. 2

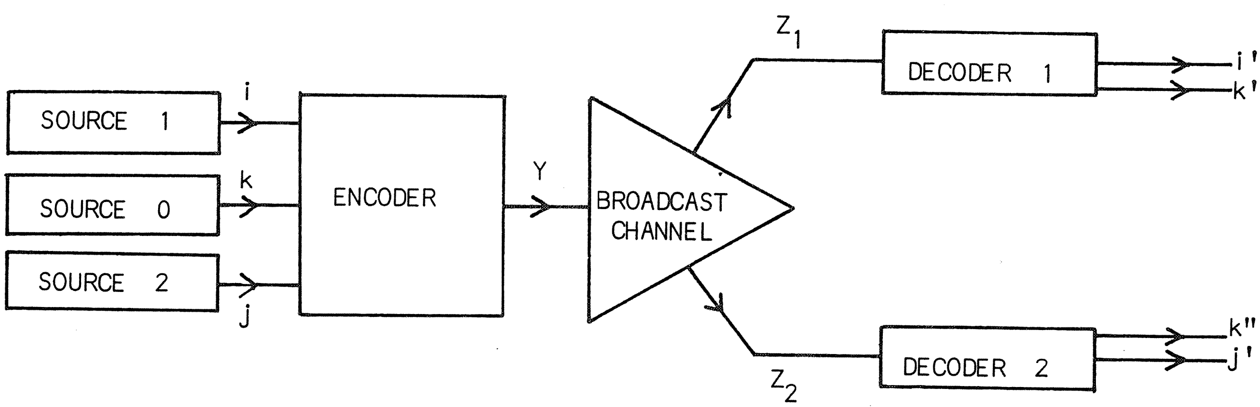
SOURCE 1
SOURCE 0
SOURCE 2
i
k
j
ENCODER
Y
BROADCAST CHANNEL
Z1
Z2
DECODER 1
DECODER 2
i'
k'
k"
j'

Fig. 3

I. S_y sends two different messages to U_{z1} and U_{z2}. Thus, given that the message pair (i,j) has been presented by the sources to S_y for transmission, U_{z1} must distinguish i, and U_{z2} must distinguish j.

II. One source output (message j say) is meant for both U_{z1} and U_{z2}, whereas the other source output (message i) is meant for U_{z1} only. In other words, given that the message pair (i,j) has been presented for transmission, U_{z1} must estimate both i and j, whereas U_{z2} needs to decode only message j correctly.

Next we consider the communication system shown in Fig. 3. Now three sources present three statistically independent messages i, j, and k to the encoder S_y for transmission over the channel. In this case $1 \leqq i \leqq M_1$, $1 \leqq j \leqq M_2$, $1 \leqq k \leqq M_0$, and each message triple (i,j,k) has a chance of $\frac{1}{M_1 M_2 M_0}$ of being selected. Encoder S_y maps the message triple (i,j,k) into an input sequence $Y \in A^n$ by means of a mapping $g(i,j,k) = Y$. This input is transmitted over the channel and received as the random sequence Z_1 by U_{z1} and as the random sequence Z_2 by U_{z2}. For this communication system we consider the following communication problem.

III. The output from source 0 is meant for U_{z1} and U_{z2}, whereas the output from source 1 is meant for U_{z1} only, and the output from source 2 is meant for U_{z2} only. Thus, given that the message triple (i,j,k) has been presented for transmission, U_{z1} must estimate the pair (i,k), and U_{z2} must estimate the pair (j,k).

We denote the above three communication situations by (P,T_{12},I), (P,T_{12},II), and (P,T_{12},III). The problem then is to establish for each case the regions of attainable rate pairs (R_1,R_2) or attainable rate triples (R_1,R_2,R_0). In the present paper, though, we will determine mostly inner bounds on these regions.

The three situations just defined are clearly interrelated. Situation (P,T_{12},I) can be regarded as a special case of (P,T_{12},III) by taking in the latter one $M_0 = 1$ $(R_0 = 0)$. Similarly, situation (P,T_{12},II) can be looked upon as a special case of (P,T_{12},III) by taking now $M_2 = 1$ $(R_2 = 0)$. It therefore would suffice to derive a coding theorem only for (P,T_{12},III) and then obtain the corresponding results for (P,T_{12},I) and (P,T_{12},II) by setting the rates R_0 or R_2 equal to zero. However for reasons of clarity we have judged it more instructive to derive the results of each situation separately, and then comment on their interrelationship in the end. Moreover this approach will enable us to bring out more clearly the relationship with other existing results in the literature.

We remark that communication situation (P,T_{12},I) is not always feasible as a separate case, since the structure of the channel may not allow us to distinguish between (P,T_{12},I) and (P,T_{12},II). For example, the installation of a noiseless feedback link from terminal 3 to terminal 2 makes (P,T_{12},I) coincide with (P,T_{12},II). Actually, all that is needed to change (P,T_{12},I) into (P,T_{12},II) is that the channel input-output statistics for receiver U_{z2} are available to U_{z1}. This is the case if the marginal channel $p(z_2|y)$ is a degraded version of the marginal channel $p(z_1|y)$. The degraded broadcast channel was the main channel under consideration by Cover [5] and Bergmans [4]. Thus, although one generally needs to distinguish between attainable rate pairs for (P,T_{12},I) and (P,T_{12},II) these two concepts coincide by a degraded broadcast channel.

We will derive separate random coding theorems for (P,T_{12},I), (P,T_{12},II), and (P,T_{12},III). Since for a degraded broadcast channel the first two situations coincide, the coding theorems derived for these situations can both be applied, but the one for (P,T_{12},II) will usually yield better results.

We notice that in situation (P,T_{12},II) the channel is in principle not assumed to be degraded. On the contrary, the results obtained for (P,T_{12},II) do not only apply to degraded broadcast channels, but also to non-degraded broadcast channels. Our results for the general d.m. broadcast channel for communication situation (P,T_{12},II) incorporate as a special case the random coding theorem obtained by Bergmans [4] for the degraded broadcast channel with two components.

We conclude by remarking that even for the two communication systems considered, one can conceive of various other communication problems. However, we have chosen to concentrate on the three problems selected above, since we believe that these are conceptually the most important and interesting ones. We also remark that communication situation (P,T_{12},III) resembles the problem considered by Slepian and Wolf [11], except that these authors assume two encoders and one decoder, whereas we consider the reverse situation of one encoder and two decoders.

C. Cascades Of Multi-Way Channels

For the development of this paper we shall need the use of other multi-way channels. Following Ahlswede [1], we denote by (P,T_{sr}) a d.m. channel with s senders and r receivers, each one located at a different terminal, and with transition probability matrix P. The broadcast channel with two components is denoted by (P,T_{12}). In addition we shall need the channels (P,T_{21}), (P,T_{22}), (P,T_{31}), and (P,T_{32}), which will be discussed now.

(P,T_{21}) stands for a d.m. channel with two senders and one receiver. Alternatively we may write $(2s,1r)$. This channel was investigated by Ahlswede in [1] and [2], and by the author in [15]. The main communication situation under consideration for (P,T_{21}) is the one in which both senders send information simultaneously to the single receiver. Ahlswede ([1] and [2]) has found two simple characterizations of the capacity region of this channel.

A d.m. channel with two senders and two receivers each located at a different terminal is denoted by (P,T_{22}) or by $(2s,2r)$. Various communication situations can be considered for this channel. Communication situation (P,T_{22},I) denotes the case in which each sender sends to a different receiver. (P,T_{22},II) denotes the case where each sender sends information simultaneously to both receivers. (P,T_{22},III) stands for the situation where one sender sends to both receivers, and the other sender sends to only one receiver. Channel (P,T_{22}) was introduced by Shannon [10], whose work on the two-way channel suggested inner and outer bounds on the capacity region of (P,T_{22},I). These bounds were later made precise independently by Ahlswede [1] and the author [15]. A complete and simple characterization of the capacity region of (P,T_{22},II) was given by Ahlswede [2]. The case (P,T_{22},III) has bot been studied yet, but will play a role in our investigations of situation (P,T_{12},II).

Consider now the cascade of a channel of type (P,T_{21}) followed by a channel of type (P,T_{12}), as shown in Fig. 4. Here the output of the $(2s,1r)$-channel is the input to the broadcast channel. The resulting channel is of type (P,T_{22}).

Mathematically this can be written as follows. Let $(A,p(z_1,z_2|y),B_1\times B_2)$ be a d.m. channel of type (P,T_{12}), denoted by K. Let $(A_1\times A_2,q(y|x_1,x_2),A)$ be a d.m. channel of type (P,T_{21}), denoted by E_q. Thus the output alphabet of E_q equals the input alphabet of K. The cascade of E_q followed by K is defined to be the $(2s,2r)$-channel $(A_1\times A_2,p(z_1,z_2|x_1,x_2),B_1\times B_2)$ whose transition probabilities are given by

$$(4)\qquad p(z_1,z_2|x_1,x_2) = \sum_{y\in A} p(z_1,z_2|y)q(y|x_1,x_2).$$

We denote this cascaded channel by E_qK.

Since E_q and K are memoryless, so is E_qK. More precisely, since the n-extension of K is given by (1), and the n-extension of E_q is defined by

$$(5)\qquad Q(Y|X_1,X_2) = \prod_{k=1}^{n} q(y_k|x_{1k},x_{2k}),$$

the n-extension of E_qK satisfies

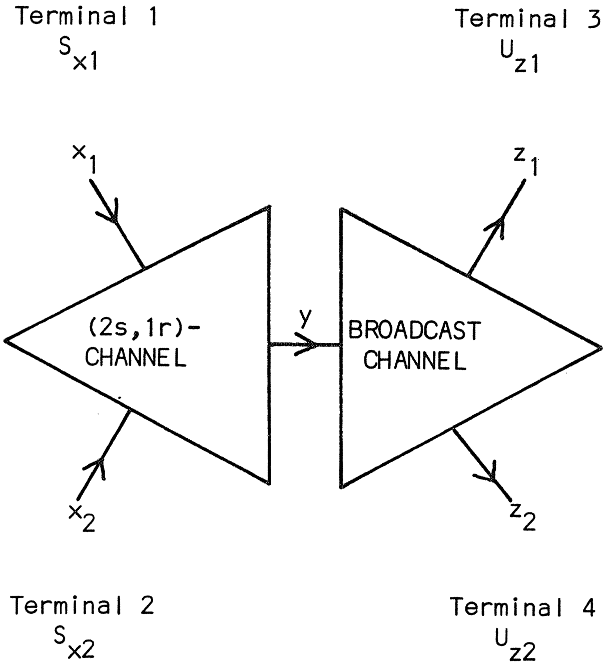
Terminal 1
S_{x1}
Terminal 3
U_{z1}
x_1
z_1
(2s,1r)-
CHANNEL
y
BROADCAST
CHANNEL
x_2
z_2
Terminal 2
S_{x2}
Terminal 4
U_{z2}

Fig. 4

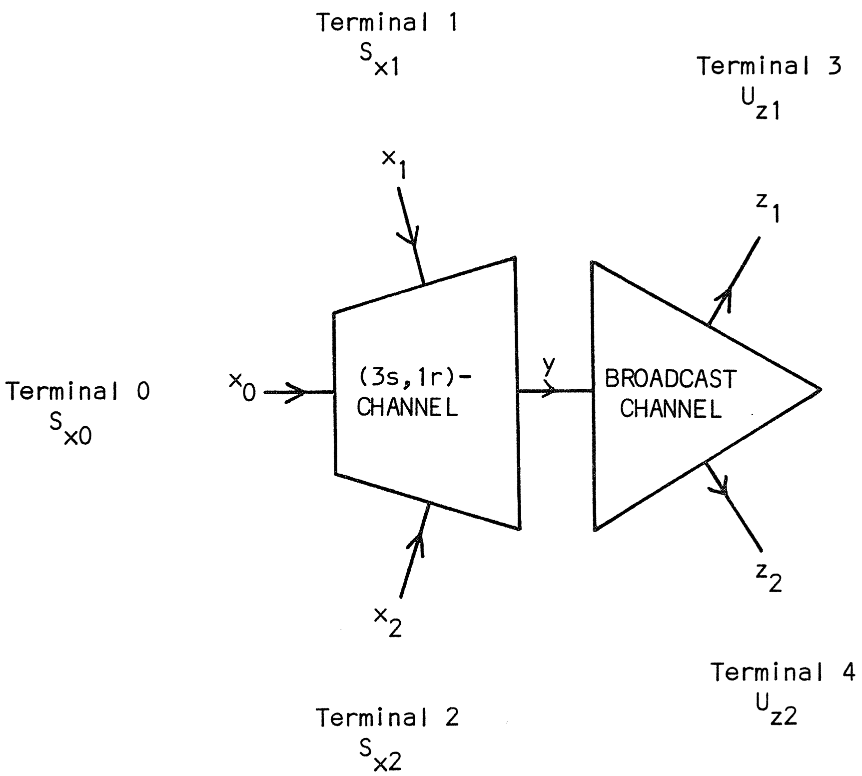
Terminal 1
S_{x1}
Terminal 3
U_{z1}
x_1
z_1
Terminal 0
S_{x0}
x_0
(3s,1r)-
CHANNEL
y
BROADCAST
CHANNEL
x_2
z_2
Terminal 4
U_{z2}
Terminal 2
S_{x2}

Fig. 5

$$(6) \qquad P(Z_1,Z_2|X_1,X_2) = \sum_{y \in A^n} P(Z_1,Z_2|Y)Q(Y|X_1,X_2) = \prod_{k=1}^{n} p(z_{1k},z_{2k}|x_{1k},x_{2k}).$$

Clearly, $E_q K$ factors out into two marginal d.m. channels of type (P,T_{21}) with probability functions $p(z_1|x_1,x_2)$ and $p(z_2|x_1,x_2)$ respectively.

(P,T_{31}) denotes a d.m. channel with three senders and one receiver, each located at a different terminal. Alternatively we write $(3s,1r)$ to denote this channel. Channel (P,T_{31}) was first investigated by Ahlswede [1], who found a simple characterization of its capacity region for the communication situation denoted by (P,T_{31},I) in which all senders send messages simultaneously to all receivers. Recently, Ulrey [12] characterized the capacity region of the general channel (P,T_{sr}) for the situation in which all senders send messages simultaneously to all receivers. His results apply in particular to channel (P,T_{31}) and yield an alternative characterization of the capacity region of (P,T_{31},I). In [17], the author has given a canonical approach to finding weak converses for (P,T_{sr}).

(P,T_{32}) stands for a d.m. channel with three senders and two receivers, and is also denoted by $(3s,2r)$. We distinguish two communication situations for this channel. Situation (P,T_{32},I) denotes the case in which each sender sends information simultaneously to each receiver. Ulrey's results on the general channel (P,T_{sr}) yield as a special case a simple characterization of the capacity region of (P,T_{32},I). Communication situation (P,T_{32},II) stands for the case in which two of the three senders send separate information to the two receivers, whereas the third sender sends common information to both receivers. To our knowledge, situation (P,T_{32},II) has not been studied yet, but it resembles situation (P,T_{12},III), and as such it plays a role in our investigations.

Consider now the cascade of a channel of type (P,T_{31}) followed by a channel of type (P,T_{12}), as shown in Fig. 5. The output of the $(3s,1r)$-channel is the input to the broadcast channel. The resulting channel is of type (P,T_{32}).

More precisely we can describe this cascading process as follows. As before, let $(A,p(z_1,z_2|y),B_1\times B_2)$ be a d.m. channel of type (P,T_{12}), denoted by K. Let $(A_1\times A_0\times A_2,q(y|x_1,x_0,x_2),A)$ be a d.m. channel of type (P,T_{31}), denoted by F_q. Thus the output alphabet of F_q equals the given input alphabet of K. The cascade of F_q followed by K is denoted by F_qK, and is defined as the $(3s,2r)$-channel $(A_1\times A_0\times A_2,p(z_1,z_2|x_1,x_0,x_2),B_1\times B_2)$ whose transition probabilities are given by

$$(7) \qquad p(z_1,z_2|x_1,x_0,x_2) = \sum_{y\in A} p(z_1,z_2|y)q(y|x_1,x_0,x_2).$$

As before, F_qK is memoryless, because F_q and K are assumed memoryless. More precisely, the n-extension of K is given by (1), and the n-extension of F_q is defined by

$$(8) \qquad Q(Y|X_1,X_0,X_2) = \prod_{k=1}^{n} q(y_{1k}|x_{1k},x_{0k},x_{2k}).$$

Therefore the n-extension of F_qK satisfies

$$(9)\qquad P(Z_1,Z_2|X_1,X_0,X_2) = \sum_{y\in A^n} P(Z_1,Z_2|Y)Q(Y|X_1,X_0,X_2)$$

$$= \prod_{k=1}^{n} p(z_{1k},z_{2k}|x_{1k},x_{0k},x_{2k}).$$

The cascaded $(3s,2r)$-channel F_qK factors out into two marginal d.m. channels of type (P,T_{31}) whose probability functions $p(z_1|x_1,x_0,x_2)$ and $p(z_2|x_1,x_0,x_2)$ are obtained by summing in (7) over z_2 or z_1, respectively.

A new notion used here is that of a cascade of two multi-way channels, whereas ordinarily one considers only cascades of one-way channels. These cascades turn out to be very useful in proving random coding theorems for situations (P,T_{12},I), (P,T_{12},II), and (P,T_{12},III). The main idea used in proving a random coding theorem for situation (P,T_{12},I) is that we have placed a "merging" $(2s,1r)$-channel in front of the broadcast channel, rather than a satellizing one-way channel, as was done by Cover [5] and Bergmans [4]. For communication situation (P,T_{12},II) the use of a merging channel leads to the same results as the use of a satellizing channel. In proving a random coding theorem for situation (P,T_{12},III) we consider cascades of the type F_qK.

D. Codes And Rates

We now give the definitions of a code and a capacity region for each of the three communication situations considered.

(i) Communication situation (P,T_{12},I). A code (n,M_1,M_2) for a channel (P,T_{12},I) whose transmission probabilities are defined by (1) is a system

(10) $$\{(w_{ij},B_i,D_j) \mid i=1,\ldots,M_1;\ j=1,\ldots,M_2\},$$

where $w_{ij} \in A^n$, $B_i \subset \mathcal{B}_1^n$, $D_j \subset \mathcal{B}_2^n$ for all $i=1,\ldots,M_1$; $j=1,\ldots,M_2$, and $B_i \cap B_{i'} = \phi$ for $i \neq i'$, $D_j \cap D_{j'} = \phi$ for $j \neq j'$. A code (n,M_1,M_2) is an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for (P,T_{12},I) if

(11) $$\frac{1}{M_1M_2} \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} P(B_i|w_{ij}) \geqq 1-\lambda_1$$

and

(12) $$\frac{1}{M_1M_2} \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} P(D_j|w_{ij}) \geqq 1-\lambda_2.$$

A pair of non-negative real numbers (R_1,R_2) is called a pair of achievable rates for (P,T_{12},I) if for any (λ_1,λ_2), $0 < \lambda_1,\lambda_2 < 1$, and any $\varepsilon > 0$ there exists an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code such that $\frac{1}{n}\log_2 M_1 \geqq R_1-\varepsilon$ and $\frac{1}{n}\log_2 M_2 \geqq R_2-\varepsilon$ for all sufficiently large n. The capacity region of channel (P,T_{12},I) is the set of all pairs of achievable rates for this channel, and is denoted by $G(P,T_{12},\mathrm{I})$. In Section III we derive a random coding inner bound on $G(P,T_{12},\mathrm{I})$.

(ii) Communication situation (P,T_{12},II). A code (n,M_1,M_2) for channel (P,T_{12},II) with transmission probabilities defined by (1) is a system

(13) $$\{(w_{ij},B_{ij},D_j) \mid i=1,\ldots,M_1;\ j=1,\ldots,M_2\},$$

where $w_{ij} \in A^n$, $B_{ij} \subset \mathcal{B}_1^n$, $D_j \subset \mathcal{B}_2^n$ for $i=1,\ldots,M_1$; $j=1,\ldots,M_2$, and $B_{ij} \cap B_{i'j'} = \phi$ for $(i,j) \neq (i',j')$, $D_j \cap D_{j'} = \phi$ for $j \neq j'$. A code

(n,M_1,M_2) is an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for (P,T_{12},II) if

(14) $$\frac{1}{M_1M_2}\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} P(B_{ij}|w_{ij}) \geqq 1-\lambda_1$$

and

(15) $$\frac{1}{M_1M_2}\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} P(D_j|w_{ij}) \geqq 1-\lambda_2.$$

A pair of non-negative real numbers (R_1,R_2) is called a pair of achievable rates for (P,T_{12},II) if for any (λ_1,λ_2), $0 < \lambda_1,\lambda_2 < 1$, and any $\varepsilon > 0$ there exists an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code such that $\frac{1}{n}\log_2 M_1 \geqq R_1-\varepsilon$ and $\frac{1}{n}\log_2 M_2 \geqq R_2-\varepsilon$ for all sufficiently large n. The capacity region of channel (P,T_{12},II) is denoted by $G(P,T_{12},\mathrm{II})$ and is defined as the set of all pairs of achievable rates for this channel. In Section IV we derive a random coding inner bound on $G(P,T_{12},\mathrm{II})$.

(iii) Communication situation (P,T_{12},III). A code (n,M_1,M_2,M_0) for channel (P,T_{12},III) with transmission probabilities defined by (1) is a system

(16) $$\{(w_{ijk},B_{ik},D_{jk}) \mid i=1,\dots,M_1;\ j=1,\dots,M_2;\ k=1,\dots,M_0\}$$

where $w_{ijk} \in A^n$, $B_{ik} \subset \mathcal{B}_1^n$, $D_{jk} \subset \mathcal{B}_2^n$ for $i=1,\dots,M_1$; $j=1,\dots,M_2$; $k=1,\dots,M_0$, and $B_{ik} \cap B_{i'k'} = \phi$ for $(i,k) \neq (i',k')$, and $D_{jk} \cap D_{j'k'} = \phi$ for $(j,k) \neq (j',k')$. A code (n,M_1,M_2,M_0) is an $(n,M_1,M_2,M_0,\lambda_1,\lambda_2)$-code for (P,T_{12},III) if

$$(17) \qquad \frac{1}{M_1 M_2 M_0} \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} \sum_{k=1}^{M_0} P(B_{jk} | w_{ijk}) \geq 1-\lambda_1$$

and

$$(18) \qquad \frac{1}{M_1 M_2 M_0} \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} \sum_{k=1}^{M_0} P(D_{jk} | w_{ijk}) \geq 1-\lambda_2.$$

A triple of non-negative real numbers (R_1, R_2, R_0) is called a triple of achievable rates for (P, T_{12}, III) if for any (λ_1, λ_2), $0 < \lambda_1, \lambda_2 < 1$, and any $\varepsilon > 0$ there exists an $(n, M_1, M_2, M_0, \lambda_1, \lambda_2)$-code such that $\frac{1}{n}\log_2 M_1 \geq R_1 - \varepsilon$, $\frac{1}{n}\log_2 M_2 \geq R_2 - \varepsilon$, and $\frac{1}{n}\log_2 M_0 \geq R_0 - \varepsilon$ for all sufficiently large n. The capacity region of channel (P, T_{12}, III) is defined as the set of all triples (R_1, R_2, R_0) of achievable rates for this channel, and is denoted by $G(P, T_{12}, \text{III})$. In Section V we shall derive a random coding inner bound on the region $G(P, T_{12}, \text{III})$.

III. RANDOM CODING THEOREM FOR (P, T_{12}, I)

A. *Mutual Information Functions*

Let $(A, p(z_1, z_2|y), B_1 \times B_2)$ be a d.m. channel of type (P, T_{12}), whose transition probabilities for operating with blocks of length n are defined by (1), and which is denoted by K. Let $(A_1 \times A_2, q(y|x_1, x_2), A)$ be a d.m. channel of type (P, T_{21}) denoted by E_q. Let $(A_1 \times A_2, p(z_1, z_2|x_1, x_2), B_1 \times B_2)$ be the cascade of E_q followed by K as defined by (4), and denoted by E_qK. E_q can be regarded as a parameter-channel for the given broadcast channel K, since by varying E_q we can generate a whole class of channels E_qK, each one being of type (P, T_{22}).

Let again $E_q = (A_1 \times A_2, q(y|x_1, x_2), A)$ be fixed, and let $p_1(x_1)$ be a probability distribution on A_1, and $p_2(x_2)$ be a probability distribution on A_2. We define mutual information functions $J_{13}(p_1, p_2, q)$ and $J_{24}(p_1, p_2, q)$ as follows. On $A_1 \times B_1$ we define the probability distribution

$$(19) \qquad p(x_1, z_1) = \sum_{x_2 \in A_2} p(z_1|x_1, x_2) p_1(x_1) p_2(x_2),$$

and on $A_2 \times B_2$ we define the probability distribution

$$(20) \qquad p(x_2, z_2) = \sum_{x_1 \in A_1} p(z_2|x_1, x_2) p_1(x_1) p_2(x_2).$$

We define the conditional probabilities $p(z_1|x_1)$ and $p(z_2|x_2)$, and the marginal probabilities $p(z_1)$ and $p(z_2)$ in the usual way in accordance with (19) and (20).

Now let

$$(21)\qquad J_{13}(p_1,p_2,q) = E\left[\log_2 \frac{p(z_1|x_1)}{p(z_1)}\right]$$

where the expectation E is taken with respect to (19). Similarly, let

$$(22)\qquad J_{24}(p_1,p_2,q) = E\left[\log_2 \frac{p(z_2|x_2)}{p(z_2)}\right]$$

where the expectation is taken with respect to (20). Thus J_{13} and J_{24} are functions of the parameters p_1,p_2, and q.

Now, by letting p_1 and p_2 vary, we define for each fixed parameter-channel $E_q = (A_1 \times A_2, q(y|x_1,x_2), A)$ the collection

$$(23)\qquad C_I(q) = \{(J_{13}(p_1,p_2,q), J_{24}(p_1,p_2,q)): \; p_1 \text{ a p.d. on } A_1,\ p_2 \text{ a p.d. on } A_2\}.$$

Next letting q vary we define the set

$$(24)\qquad C_I = \bigcup_q C_I(q)$$

where the union is taken over the collection of all d.m. $(2s,1r)$-channels E_q with given output alphabet A.

Finally let

$$(25)\qquad G_I = \mathrm{co}(C_I)$$

where $\mathrm{co}(A)$ means the convex hull of the set A.

B. Pure Parameter-Channels

A discrete channel is said to be deterministic or *pure* if only zeros and ones occur as its transition probabilities (cf [3], p. 51). Thus the parameter-channel $E_q = (A_1 \times A_2, q(y|x_1,x_2), A)$ is pure if and only if $q(y|x_1,x_2) = 0$ or 1 for all x_1, x_2, and y. We now show that the set G_{I} defined in (25) remains unchanged if in (24) we take the union only over the collection of pure parameter-channels E_q which have A as given output alphabet.

Let us define

$$\mathcal{D}_{\mathrm{I}} = \bigcup_q{}' \, C_{\mathrm{I}}(q) \tag{26}$$

where $\bigcup_q{}'$ denotes the union over all *pure* $(2s, 1r)$-channels E_q with given output alphabet A.

Then we have

Theorem 1: It suffices in (24) to take the union over all *pure* channels E_q without altering G_{I}. Thus

$$G_{\mathrm{I}} = \mathrm{co}(\mathcal{D}_{\mathrm{I}}). \tag{27}$$

Proof: Let $E_q = (A_1 \times A_2, q(y|x_1,x_2), A)$ be any parameter-channel for K. Let a_1 and a_2 denote the size of A_1 and A_2 respectively, and let $t = a^{a_1 a_2}$. According to the discussion at the bottom of p. 392 of [9], the transition probability matrix $\|q(y|x_1,x_2)\|$ can be written as a finite weighted sum of the transition probability matrices of t pure channels. More precisely, there exist t pure channels $E_\alpha = (A_1 \times A_2, q_\alpha(y|x_1,x_2), A)$; $\alpha=1,\ldots,t$; and a probability distribution $\{g_\alpha : \alpha=1,\ldots,t\}$ such that

(28) $$q(y|x_1,x_2) = \sum_{\alpha=1}^{t} g_\alpha q_\alpha(y|x_1,x_2).$$

For each such E_α let

(29) $$p_\alpha(z_1,z_2|x_1,x_2) = \sum_{y\in A} p(z_1,z_2|y)q_\alpha(y|x_1,x_2).$$

Clearly, from (4) and (28) we have

(30) $$p(z_1,z_2|x_1,x_2) = \sum_{\alpha=1}^{t} g_\alpha p_\alpha(z_1,z_2|x_1,x_2).$$

Now, let $p_1(x_1)$ be a probability distribution on A_1 and $p_2(x_2)$ a probability distribution on A_2. Define, for $\alpha=1,\ldots,t$,

(31) $$p_\alpha(z_1,z_2,x_1,x_2) = p_\alpha(z_1,z_2|x_1,x_2)p_1(x_1)p_2(x_2),$$

and derive from it $p_\alpha(z_1|x_1)$, $p_\alpha(z_2|x_2)$, $p_\alpha(z_1)$, and $p_\alpha(z_2)$ in the usual way. Then we have

$$p(z_i|x_i) = \sum_{\alpha=1}^{t} g_\alpha p_\alpha(z_i|x_i) \qquad i=1,2$$

and

$$p(z_i) = \sum_{\alpha=1}^{t} g_\alpha p_\alpha(z_i) \qquad i=1,2.$$

Moreover, it follows from (21) and (22) that

$$(32) \qquad J_{13}(p_1,p_2,q_\alpha) = E\left[\log_2 \frac{p_\alpha(z_1|x_1)}{p_\alpha(z_1)}\right]$$

and

$$(33) \qquad J_{24}(p_1,p_2,q_\alpha) = E\left[\log_2 \frac{p_\alpha(z_2|x_2)}{p_\alpha(z_2)}\right].$$

It is well-known that $J_{13}(p_1,p_2,q_\alpha)$ and $J_{24}(p_1,p_2,q_\alpha)$ are convex functions of the transition probabilities $p_\alpha(z_1|x_1)$ and $p_\alpha(z_2|x_2)$ respectively (see [6], p. 90). Therefore we have

$$(34) \qquad J_{13}(p_1,p_2,q) \leqq \sum_{\alpha=1}^{t} g_\alpha J_{13}(p_1,p_2,q_\alpha)$$

and

$$(35) \qquad J_{24}(p_1,p_2,q) \leqq \sum_{\alpha=1}^{t} g_\alpha J_{24}(p_1,p_2,q_\alpha).$$

Consequently,

$$(36) \qquad (J_{13}(p_1,p_2,q),J_{24}(p_1,p_2,q)) \in \mathrm{co}(\mathcal{D}_I)$$

and hence

$$(37) \qquad C_I \subset \mathrm{co}(\mathcal{D}_I)$$

which completes the proof of the theorem.

Conjecture 1: We conjecture that (27) remains valid if in (26) the union is taken over only those pure channels E_q for which $a_1 = \min(a,b_1)$, and $a_2 = \min(a,b_2)$. A proof of this conjecture might possibly be given along the lines of Gallager ([6], p. 96) or Gallager ([7], Lemma 1).

C. The Coding Theorem

Theorem 2: (i) The region G_{I} is a closed convex region in the Euclidean plane.

(ii) The region G_{I} is contained in the capacity region $G(P,T_{12},\mathrm{I})$. Thus

$$G_{\mathrm{I}} \subset G(P,T_{12},\mathrm{I}). \tag{38}$$

Proof: (i) This part follows from Theorem 3.10 of Valentine [13, p. 40]. (ii) By Theorem 1 it suffices to show that for each pure channel $E_q = (A_1 \times A_2, q(y|x_1,x_2), A)$ every point in $C_{\mathrm{I}}(q)$ is a pair of attainable rates for (P,T_{12},I). By concatenation it will then follow that each point in G_{I} is an attainable pair.

For each q the channel E_qK, whose probability function is given by (4), is of type (P,T_{22}), and therefore one may apply to it known results for the d.m. $(2s,2r)$-channel. It follows from results of Ahlswede ([1], section 2), or alternatively from results of the author ([15], section 6), that for each q every point in $C_{\mathrm{I}}(q)$ is an attainable pair of rates for E_qK in communication situation (P,T_{22},I).

An $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for E_qK in situation (P,T_{22},I) is a system

(39) $$\{u_i,v_j,B_i,D_j|i=1,\ldots,M_1;\ j=1,\ldots,M_2\},$$

where $u_i \in A_1^n$, $v_j \in A_2^n$, $B_i \subset B_1^n$, $D_j \subset B_2^n$ for $i=1,\ldots,M_1$; $j=1,\ldots,M_2$, and $B_i \cap B_{i'} = \phi$ for $i \neq i'$, $D_j \cap D_{j'} = \phi$ for $j \neq j'$, such that

(40) $$\frac{1}{M_1M_2}\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} P(B_i|u_i,v_j) \geqq 1-\lambda_1$$

and

(41) $$\frac{1}{M_1M_2}\sum_{i=1}^{M_1}\sum_{j=1}^{M_2} P(D_j|u_i,v_j) \geqq 1-\lambda_2.$$

Here the error probabilities are based on $P(Z_1,Z_2|X_1,X_2)$ defined in (6).

If E_q is pure, then any $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for E_qK in communication situation (P,T_{22},I) can be translated into an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for K in communication situation (P,T_{12},I) as follows. Let E_q be pure. Then Q, the memoryless n-extension of q, contains only zeros and ones. Let be given the code (39). For all pairs (u_i,v_j) belonging to (39) define $w_{ij}=Y$ if $Q(Y|u_i,v_j)=1$. Then

$$P(B_i|u_i,v_j) = \sum_{Y\in A^n} P(B_i|Y)Q(Y|u_i,v_j) = P(B_i|w_{ij}).$$

Similarly

$$P(D_j|u_i,v_j) = P(D_j|w_{ij}).$$

Therefore the system

$$(42) \qquad \{w_{ij}, B_i, D_j \mid i=1,\ldots,M_1;\ j=1,\ldots,M_2\}$$

forms an $(n, M_1, M_2, \lambda_1, \lambda_2)$-code for K for communication situation (P, T_{12}, I). It follows that every point in $C_{\mathrm{I}}(q)$ is a pair of attainable rates for K in situation (P, T_{12}, I) whenever E_q is pure. This completes the proof.

D. Comparison With The Cover-Bergmans Scheme

Cover [5] and Bergmans [4] exhibited a random coding scheme for respectively binary symmetric broadcast channels and degraded broadcast channels in communication situation (P, T_{12}, II). For the Cover-Bergmans scheme the expected value of the average probability of error goes to zero in both directions simultaneously as the block length n tends to infinity. Motivated by the proof of Theorem 2(ii) we now exhibit a random coding scheme for the general d.m. broadcast channel in communication situation (P, T_{12}, I), which has the same property. Our random coding scheme can be regarded as the analogue of the Cover-Bergmans scheme in the case of non-degraded broadcast channels.

Suppose in (26) we take the union over only those pure $(2s, 1r)$-channels E_q for which $a_1 = a_2 = a$, i.e. such that E_q is of the form $(A \times A, q(y|x_1, x_2), A)$. Then, after taking the convex hull as in (25) we obtain a region which is contained in G_{I} and which by Theorem 2 is contained in $G(P, T_{12}, \mathrm{I})$.

More precisely, denote by q^* the transition probability function of any pure parameter-channel $E_{q^*} = (A\times A, q^*(y|x_1,x_2), A)$ for which the input alphabets are equal to the given output alphabet A. It follows from the proof of Theorem 2(ii) that for every such q^* every point in $C_I(q^*)$ is a pair of attainable rates for $E_{q^*}K$ in communication situation (P,T_{22},I), and a fortiori is a pair of attainable rates for K in communication situation (P,T_{12},I).

Let $E_{q^*} = (A\times A, q^*(y|x_1,x_2), A)$ be given, and let $p_1(x_1)$ and $p_2(x_2)$ be two probability distributions on A. Let $\varepsilon > 0$, and $0 < \lambda_1,\lambda_2 < 1$. Denote the point $(J_{13}(p_1,p_2,q^*), J_{24}(p_1,p_2,q^*))$ in $C_I(q^*)$ by (J_1,J_2). According to the proof of Theorem 2(ii) there exists for all sufficiently large n an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for $E_{q^*}K$ in situation (P,T_{12},I) such that

$$(43)\qquad M_1 \geqq 2^{n(J_1-\varepsilon)} \quad \text{and} \quad M_2 \geqq 2^{n(J_2-\varepsilon)}.$$

This code is a system which is described by (39) except that $u_i \in A^n$ and $v_j \in A^n$. It can be translated into an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for K in situation (P,T_{12},I) by the rule which designates $Y \in A^n$ to be the codeword w_{ij} if $Q^*(Y|u_i,v_j)=1$, where Q^* is the memoryless n-extension of q^*. The resulting code for K is described by the system (42). We are now ready to formulate our random coding scheme.

Choose at random $M_1 \geqq 2^{n(J_1-\varepsilon)}$ horizontal cloud centers $u_1,\ldots,u_{M_1}$ in A^n with letters independently drawn according to $p_1(x_1)$. At the same time, choose $M_2 \geqq 2^{n(J_2-\varepsilon)}$ vertical cloud centers $v_1,\ldots,v_{M_2}$ in A^n with letters independently drawn according to $p_2(x_2)$. The horizontal and vertical cloud centers are depicted in Fig. 6. Their meaning will become apparent shortly.

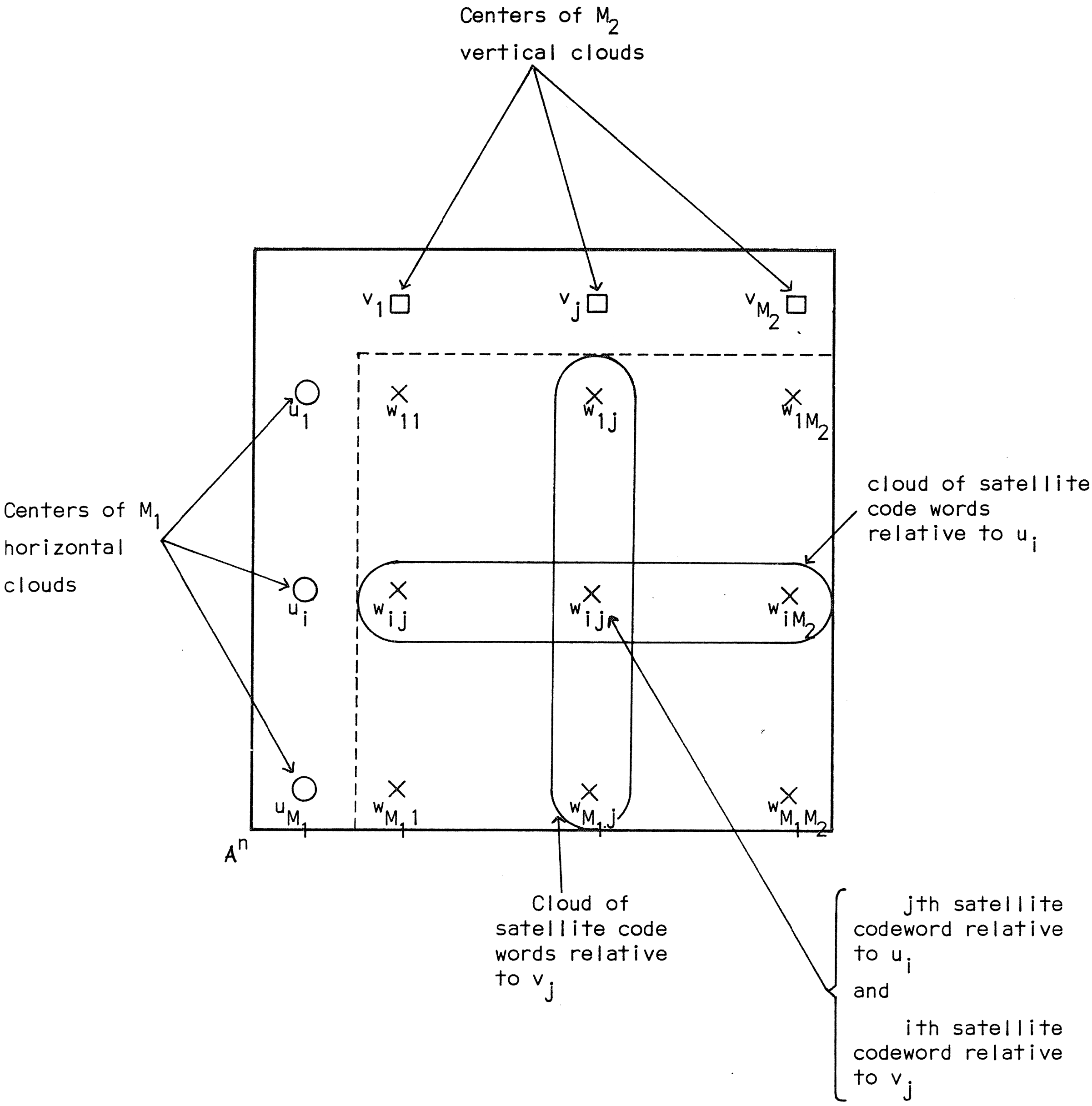

Fig. 6. Clouds, cloud centers, and satellite code words of a code for a non-degraded broadcast channel.

Suppose the sources S_1 and S_2 present the message pair (i,j) to S_y for transmission over K according to (P,T_{12},I). Then the pair (u_i,v_j), consisting of one horizontal and one vertical cloud center, is mapped into the common satellite codeword w_{ij} determined by $Q^*(w_{ij}|u_i,v_j)=1$. Subsequently, the codeword w_{ij} is transmitted over the broadcast channel K. User U_{z1} must decode index i correctly, while user U_{z2} should decode index j correctly.

The set of codewords w_{ij} with same index i constitutes a horizontal cloud of points in A^n, which is represented by the cloud center u_i. It is sufficient for U_{z1} to determine to which horizontal cloud the transmitted codeword w_{ij} belongs, or, in other words, its representative u_i. The different w_{ij} in a given horizontal cloud can be regarded as satellite codewords relative to u_i. Namely, the codewords $w_{i1},\ldots,w_{iM_2}$ can be thought of as obtained by running u_i M_2 times through an artificial channel with transition probability function

$$(44)\qquad \eta(y|x_1) = \sum_{x_2\in A} q^*(y|x_1,x_2)p_2(x_2).$$

Similarly, the set of codewords w_{ij} with same index j forms a vertical cloud of points in A^n, represented by the center v_j. It is sufficient for U_{z2} to determine the vertical cloud to which the transmitted codeword w_{ij} belongs, i.e. its representative v_j. The different w_{ij} in a given vertical cloud can be regarded as satellite codewords relative to v_j. Namely the codewords $w_{1j},\ldots,w_{M_1j}$ can be thought of as obtained by running v_j M_1 times through an artificial channel with transition probability function

(45) $$\rho(y|x_2) = \sum_{x_1 \in A} q^*(y|x_1,x_2)p_1(x_1).$$

Thus our random coding scheme has the simultaneous effect of running each horizontal cloud center u_i M_2 times through an artificial channel $\eta(y|x_1)$, and running each vertical cloud center v_j M_1 times through an artificial channel $\rho(y|x_2)$. The codeword w_{ij} is a common satellite belonging to both the horizontal cloud with center u_i, and the vertical cloud with center v_j.

The results of Ahlswede [1] and the author [15] which are referred to in the proof of Theorem 2(ii) are based on a random coding argument for (P,T_{22},I). From these proofs it follows that for the random coding scheme just described (when supplied with the appropriate maximum likelihood decoding sets) the expected value of the average probability of error goes to zero for both directions simultaneously as n tends to infinity. From this result follows the more precise statement about the existence of an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for K in situation (P,T_{12},I) for sufficiently large n.

An alternative proof of Theorem 2(ii) can be given along the lines of the proof of Theorem 1 of [4], if one replaces expression (26) of [4] by an unconditional decoding set, and modifies the corresponding random coding proof accordingly.

For a comparison of our random coding scheme with the Cover-Bergmans scheme, suppose that the codeword w_{ij} has been transmitted. In the Cover-Bergmans scheme U_{z2} should decode the cloud center v_j to which w_{ij} belongs, whereas U_{z1} should decode w_{ij} *conditional* on the cloud center v_j. In our random coding scheme U_{z2} should also decode the cloud center v_j to which w_{ij} belongs, but U_{z1} should decode w_{ij} *averaged* over all vertical cloud centers v_j, or, in other words, U_{z1} should decode the horizontal cloud center u_i. Our procedure is of course symmetric in i and j. This discussion shows that there are striking similarities but also major differences between our random coding scheme and the Cover-Bergmans random coding scheme.

E. An Example By Blackwell

Blackwell, in 1963, in a course on information theory at Berkeley, introduced the broadcast channel through the following example.

Consider the $(1s,2r)$-channel $K = (A,p(z_1,z_2|y),B_1{\times}B_2)$ with $A = \{0,1,2\}$, $B_1 = B_2 = \{0,1\}$, and the transition probabilities defined by

(46) $$P(z_1=0,z_2=1|y=0) = P(z_1=1,z_2=0|y=1) = P(z_1=1,z_2=1|y=2) = 1.$$

The marginal one-way channels of K, denoted by K_1 and K_2, have transition probabilities $p(z_1|y)$ and $p(z_2|y)$ given by Table Ia and Ib respectively.

TABLE Ia

y \ z_1	0	1
0	1	0
1	0	1
2	0	1

TABLE Ib

y \ z_2	0	1
0	0	1
1	1	0
2	0	1

It is easily verified that there does not exist a post-multiplying channel K_3 such that K_2 can be represented as the cascade of K_1 followed by K_3. In other words, K_2 is not a degraded version of K_1, and neither is K_1 a degraded version of K_2. Therefore, the present example of a broadcast channel does not fall in the class of degraded broadcast channels considered in [4] and [5]. We observe though, that K_1 and K_2 can be obtained from each other by pre-multiplication by a third channel. Thus, K_1 and K_2 are equivalent one-way channels in the sense defined by Shannon [9], but they are not degraded versions of each other.

Clearly, the capacities of the d.m. one-way channels K_1 and K_2 are both equal to one. Therefore, by time-sharing, all pairs (R_1,R_2) such that $R_1 \geqq 0$, $R_2 \geqq 0$, and $R_1+R_2 \leqq 1$ are pairs of attainable rates for K in situation (P,T_{12},I).

Blackwell proposed as a problem to find the capacity region $G(P,T_{12},I)$ of the present channel. He noted that always $R_1+R_2 \leqq \log_2 3$, so that the point (.793,.793) is outside the capacity region.

Using Theorem 2 we found that all points within and on the boundary of the shaded region shown in Fig. 7 are pairs of attainable rates for K in situation (P,T_{12},I). This shaded region is the convex hull of the points $(R_1,R_2) = (H(p),C(p))$ and of the points $(R_1,R_2) = (C(p),H(p))$ as p ranges between zero and one, where

(47) $$H(p) = -p\log_2 p-(1-p)\log_2(1-p)$$

and

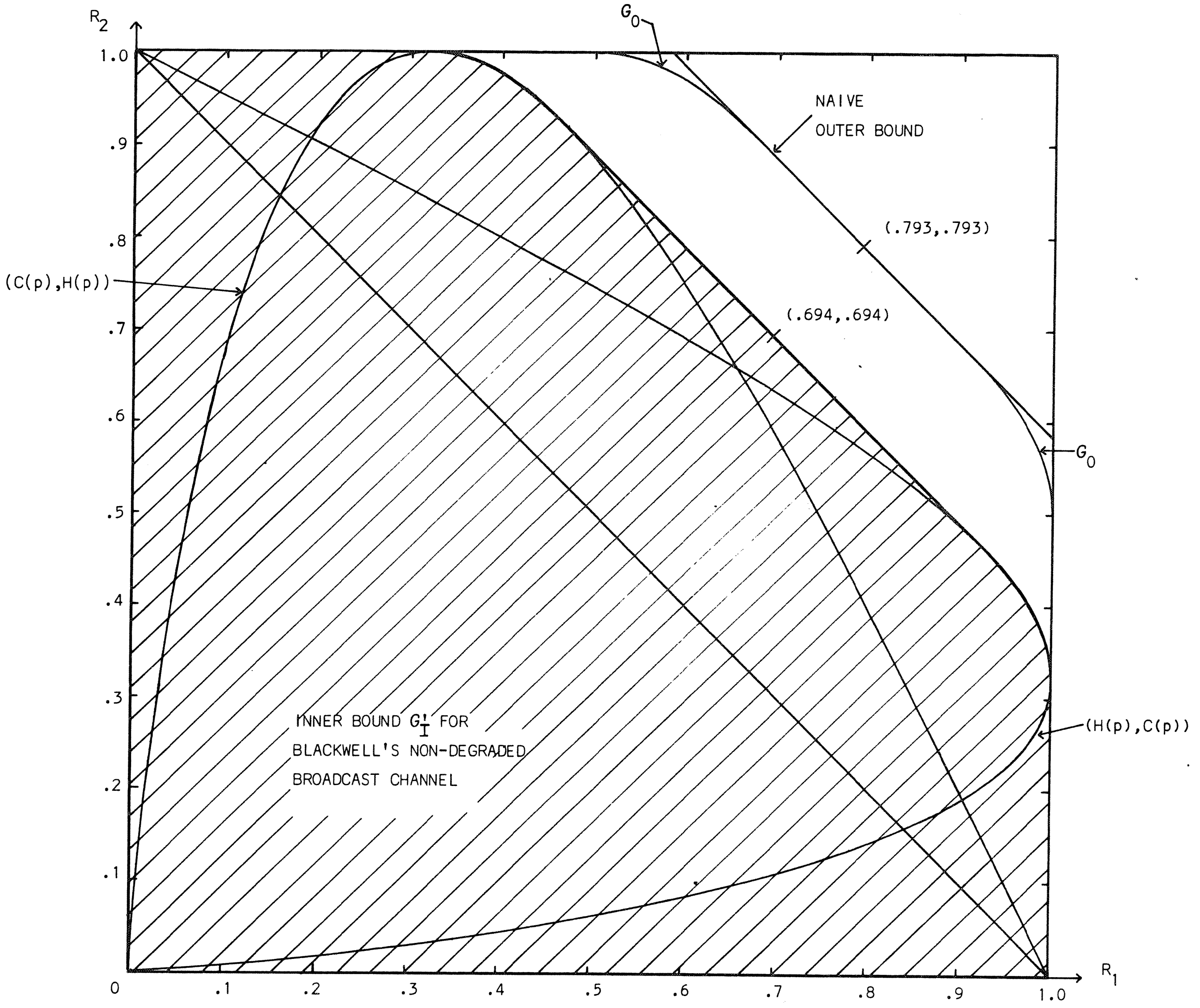
R2
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
R1
G0
NAIVE
OUTER BOUND
(.793,.793)
(.694,.694)
(C(p),H(p))
G0
(H(p),C(p))
INNER BOUND G'I FOR
BLACKWELL'S NON-DEGRADED
BROADCAST CHANNEL

Fig. 7

$$(48)\qquad C(p) = \begin{cases} \log_2 \left[1+\exp_2(-\frac{H(p)}{1-p}) \right] & \text{if } 0 \leq p < 1 \\ 0 & \text{if } p = 1. \end{cases}$$

In applying Theorem 2 we have three parameters q, p_1, and p_2 at our disposal. First we make a particular choice, $E_{q'}$ say, for the pure parameter-channel $E_q = (A_1 \times A_2, q(y|x_1,x_2), A)$. Let $A_1 = A_2 = \{0,1\}$ and let the transition probabilities $q'(y|x_1,x_2)$ be given by Table II.

TABLE II

x_1, x_2 \ y	0	1	2
0 0	1	0	0
0 1	1	0	0
1 0	0	1	0
1 1	0	0	1

Next we form the cascade of $E_{q'}$ followed by K with channel probability function defined by (4). The transition probabilities of the corresponding marginal $(2s,1r)$-channels $p(z_1|x_1,x_2)$ and $p(z_2|x_1,x_2)$ are given by Table IIIa and IIIb respectively.

TABLE IIIa

$x_1 x_2$ \ z_1		0	1
0	0	1	0
0	1	1	0
1	0	0	1
1	1	0	1

TABLE IIIb

$x_1 x_2$ \ z_2		0	1
0	0	0	1
0	1	0	1
1	0	1	0
1	1	0	1

With this choice of q expressions (21) and (22) reduce to

(49) $$J_{13}(p_1,p_2,q') = H(p_{10})$$

and

(50) $$J_{24}(p_1,p_2,q') = H(p_{20}-p_{10}p_{20})-p_{20}H(p_{10})$$

where $p_{10}=p_1(0)$ and $p_{20}=p_2(0)$. We observe that $C(p)$ is the capacity of the binary channel whose transition matrix is

$$\begin{pmatrix} p & 1-p \\ 1 & 0 \end{pmatrix}.$$

(See Ash [3], p. 85.) It follows that, for each fixed choice of p_{10},

$$\max_{p_{20}} \left(H(p_{20}-p_{10}p_{20})-p_{20}H(p_{10})\right) = C(p_{10}). \qquad (51)$$

It therefore suffices to plot the points $(R_1,R_2) = (H(p),C(p))$ for $0 \leqq p \leqq 1$, together with the point $(1,0)$, and to take their envelope, in order to obtain the region $C_{\mathrm{I}}(q')$.

Next we choose a different parameter-channel, denoted by $E_{q''}$, whose transition probability matrix is obtained by changing the first row of Table II into the assignment $(0,1,0)$ and leaving the other rows the same. By symmetry we find that $C_{\mathrm{I}}(q'')$ is the envelope of the points $(R_1,R_2) = (C(p),H(p))$ as p ranges from zero to one, together with the point $(0,1)$.

By taking the convex hull of all pairs so obtained we get the region

$$G_{\mathrm{I}}' = \mathrm{co}(C_{\mathrm{I}}(q') \cup C_{\mathrm{I}}(q'')) \qquad (52)$$

which is depicted as the shaded region of Fig. 7. Clearly $G_{\mathrm{I}}' \subset G_{\mathrm{I}} \subset G(P,T_{12},\mathrm{I})$. In view of conjecture 1 we believe that $G_{\mathrm{I}}=G_{\mathrm{I}}'$. We do not know whether G_{I}' is also the capacity region $G(P,T_{12},\mathrm{I})$ of the present channel.

A brief inspection of the weak converse leads us to believe that the following conjecture might be true.

Conjecture 2: An outer bound on the capacity region $G(P,T_{12},\mathrm{I})$ of the present example is provided by the region

$$G_O = \mathrm{co}(\underline{G}_O \cup \overline{G}_O) \qquad (53)$$

where $\underline{G}_0 = \{(p,H(p)): 0 \leqq p \leqq 1\}$

and

$$\overline{G}_0 = \{(H(p),p): 0 \leqq p \leqq 1\}.$$

The contours of the conjectured outer bound G_0 and of the naive outer bound $R_1+R_2 \leqq \log_2 3$ are sketched in Fig. 7. They are seen to have a line segment in common.

F. A Limiting Expression For $G(P,T_{12},I)$

We proceed as Shannon did in section 15 of [10]. Let be given the d.m. broadcast channel $K = (A,p(z_1,z_2|y),B_1\times B_2)$, and consider its memoryless n-extension $K^n = (A^n,P(Z_1,Z_2|Y),B_1^n\times B_2^n)$ with transmission probabilities defined by (1). K^n is also a (1s,2r)-channel. For each $n \geqq 1$ let

(54) $$E_Q^n = (A^n\times A^n,Q_n(Y|X_1,X_2),A^n)$$

be a pure parameter-channel of type (P,T_{21}) with input and output alphabets all equal to the given set A^n. Thus, the matrix $\|Q_n(Y|X_1,X_2)\|$ contains only zeros and ones, but is not necessarily a product-channel. Consider the cascade

(55) $$E_Q^nK^n = (A^n\times A^n,P^n(Z_1,Z_2|X_1,X_2),B_1^n\times B_2^n)$$

where

(56) $$P^n(Z_1,Z_2|X_1,X_2) = \sum_{Y\in A^n} P(Z_1,Z_2|Y)Q_n(Y|X_1,X_2).$$

Let $P_{1n}(X_1)$ and $P_{2n}(X_2)$ be two probability distributions on A^n. Define

(57) $$P^n(X_1,X_2,Z_1,Z_2) = P^n(Z_1,Z_2|X_1,X_2)P_{1n}(X_1)P_{2n}(X_2)$$

and derive from it $P^n(Z_1|X_1)$, $P^n(Z_2|X_2)$, $P^n(Z_1)$, and $P^n(Z_2)$ in the usual way. Let

(58) $$J^n_{13}(P_{1n},P_{2n},Q_n) = E\left[\log_2 \frac{P^n(Z_1|X_1)}{P^n(Z_1)}\right]$$

and

(59) $$J^n_{24}(P_{1n},P_{2n},Q_n) = E\left[\log_2 \frac{P^n(Z_2|X_2)}{P^n(Z_2)}\right] ,$$

where the expectations are taken with respect to (57). Define

(60) $$C^n_I(Q_n) = \{(J^n_{13}(P_{1n},P_{2n},Q_n),J^n_{24}(P_{1n},P_{2n},Q_n)): P_{1n} \text{ and } P_{2n} \text{ are p.d.'s on } A^n\}.$$

Next define

(61) $$C^n_I = \bigcup_{Q_n} C^n_I(Q_n)$$

where the union is taken with respect to all pure $(2s,1r)$-channels E_Q^n of the form (54). Let

$$G_I^n = \mathrm{co}(C_I^n). \tag{62}$$

Thus, G_I^n is essentially the inner bound of K^n, as given by Theorem 2, except that in (61) we have restricted the union over those Q_n for which the input alphabets are equal to the given set A^n. In view of conjecture 1 this may make no difference.

Next let

$$K_I^n = \frac{G_I^n}{n} = \{(R_1,R_2):(nR_1,nR_2) \in G_I^n\} \tag{63}$$

and finally define

$$G_I^\infty = \bigcup_{n=1}^{\infty} K_I^n. \tag{64}$$

Then we have

Theorem 3: **(i)** The region G_I^∞ is a closed convex region in the Euclidean plane.

(ii) The region G_I^∞ is the capacity region $G(P,T_{12},I)$.

Proof: **(i)** This part is fairly standard. The convexity is immediate, and a precise proof of the fact that G_I^∞ is closed can be given along the lines of the proof of Lemma 2 of [12].

(iia) Let $(R_1,R_2) \in K_I^m$ for some $m \geqq 1$. Then $(mR_1,mR_2) \in G_I^m$. Let $\varepsilon > 0$, $0 < \lambda_1,\lambda_2 < 1$. By Theorem 2 there exists for k sufficiently large a $(k,M_1,M_2,\lambda_1,\lambda_2)$-code for K^n in situation (P,T_{12},I) such that

(65) $$M_i \geqq 2^{k(mR_i-\varepsilon)} \qquad i=1,2.$$

This code is directly translated into a $(km,M_1,M_2,\lambda_1,\lambda_2)$-code for K. It follows that, for any $\varepsilon > 0$, there exists for $n=km$ sufficiently large an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for K in situation (P,T_{12},I) such that

(66) $$M_i \geqq 2^{n(R_i-\varepsilon)} \qquad i=1,2.$$

The statement for general n is proven along the lines of Theorem 5.5.1 of [18] or Theorem 8.1 of [15]. Hence $G_I^\infty \subset G(P,T_{12},I)$.

(iib) Let $(R_1,R_2) \in G(P,T_{12},I)$. Let $\varepsilon > 0$, $0 < \lambda_1,\lambda_2 < 1$. Then there exists for n sufficiently large an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for K in situation (P,T_{12},I) such that

(67) $$\frac{1}{n}\log_2 M_i \geqq R_i-\varepsilon \qquad i=1,2.$$

We denote a code like this by the system (10). We can find letter-sequences $u_1,\ldots,u_{M_1}$; $v_1,\ldots,v_{M_2}$; all in A^n, and a pure parameter-channel

(68) $$E_{Q^*}^n = (A^n \times A^n, Q_n^*(Y|X_1,X_2), A^n)$$

such that $Q_n^*(Y|u_i,v_j) = 1$ whenever $Y = w_{ij}$. Consider the cascade $E_Q^n * K^n$. The system

(69) $$\{u_i,v_j,B_i,D_j | i=1,\ldots,M_1;\ j=1,\ldots,M_2\}$$

is a $(1,M_1,M_2,\lambda_1,\lambda_2)$-code for $E_Q^n * K^n$ in situation (P,T_{22},I). Let $P_{1n}^*(X_1) = \frac{1}{M_1}$ if $X_1=u_i; i=1,\ldots,M_1$ and let $P_{2n}^*(X_2) = \frac{1}{M_2}$ if $X_2=v_j; j=1,\ldots,M_2$. It follows from Fano's Lemma applied to (P,T_{22},I), as is shown in Theorem 7.1 of [15], that

(70) $$\log_2 M_1 \leqq \frac{J_{13}^n(P_{1n}^*,P_{2n}^*,Q_n^*) + 1}{1-\lambda}$$

and

(71) $$\log_2 M_2 \leqq \frac{J_{24}^n(P_{1n}^*,P_{2n}^*,Q_n^*) + 1}{1-\lambda} .$$

Hence for all $\delta > 0$, $(R_1-\delta,R_2-\delta) \in G_{\mathrm{I}}^{\infty}$. Since G_{I}^{∞} is closed, $(R_1,R_2) \in G_{\mathrm{I}}^{\infty}$. Therefore, $G_{\mathrm{I}}(P,T_{12},\mathrm{I}) \subset G_{\mathrm{I}}^{\infty}$ which completes the proof.

Conjecture 3: We conjecture that in (68) the letter-sequences $u_1,\ldots,u_{M_1}$; $v_1,\ldots,v_{M_2}$ can be chosen in such a way that Q_n^* is a product-channel, or at least that it suffices to restrict attention to parameter-channels of this kind. If this is so, one may proceed to derive from inequalities (70) and (71) an outer bound on $G_{\mathrm{I}}(P,T_{12},\mathrm{I})$ in terms of single inputs and outputs to the channel only, in the same way as it was done in [15] for the case (P,T_{22},I). (See also [1], [2], and [17] in this regard.)

Such an outer bound will generally differ from the inner bound G_I derived in section IIIC, because there is not a simple expression known for the capacity region of (P,T_{22},I).

IV. RANDOM CODING THEOREM FOR (P,T_{12},II)

A. *Mutual Information Functions*

Let again be given the d.m. broadcast channel $K = (A,p(z_1,z_2|y),B_1\times B_2)$, a parameter-channel $E_q = (A\times A,q(y|x_1,x_2),A)$, and the cascade $E_qK = (A\times A,p(z_1,z_2|x_1,x_2),B_1\times B_2)$ whose transmission probabilities are defined as in (4). Let $p_1(x_1)$ and $p_2(x_2)$ be two probability distributions on A. Define

$$p(x_1,x_2,z_1,z_2) = p(z_1,z_2|x_1,x_2)p_1(x_1)p_2(x_2) \tag{72}$$

and derive from it $p(z_1|x_1,x_2)$, $p(z_1|x_1)$, $p(z_1|x_2)$, $p(z_2|x_2)$, $p(z_1)$, and $p(z_2)$ in the usual way. Define the following mutual information functions.

$$R_1(p_1,p_2,q;B_1) = E\left[\log_2 \frac{p(z_1|x_1,x_2)}{p(z_1|x_2)}\right] \tag{73}$$

$$R_2(p_1,p_2,q;B_1) = E\left[\log_2 \frac{p(z_1|x_1,x_2)}{p(z_1|x_1)}\right] \tag{74}$$

$$R_{12}(p_1,p_2,q;B_1) = E\left[\log_2 \frac{p(z_1|x_1,x_2)}{p(z_1)}\right] \tag{75}$$

and

$$R^1_{12}(p_1,p_2,q;B_2) = J_{24}(p_1,p_2,q) \tag{76}$$

as defined in (22). Here, all expectations are taken with respect to (72).

Let $\sigma = (P,Q)$ be a finite collection of triples

(77) $\{(p_1^\alpha, p_2^\alpha, q_\alpha) : \alpha = 1,\dots,d\}$

where $(A\times A, q_\alpha(y|x_1,x_2), A)$ is a parameter-channel, and p_1^α and p_2^α are probability distributions on A. Also, let $\nu = \{\nu(\alpha) : \alpha = 1,\dots,d\}$ be a probability distribution on σ. We associate with every pair (σ,ν) a triple

(78) $$\vec{R}(\sigma,\nu) = (\tilde{R}_1(\sigma,\nu), \tilde{R}_2(\sigma,\nu), \tilde{R}(\sigma,\nu))$$

where

(79) $$\tilde{R}_1(\sigma,\nu) = \sum_{\alpha=1}^{d} \nu(\alpha) R_1(p_1^\alpha, p_2^\alpha, q_\alpha; \mathcal{B}_1)$$

(80) $$\tilde{R}_2(\sigma,\nu) = \min\left[\sum_{\alpha=1}^{d} \nu(\alpha) R_2(p_1^\alpha, p_2^\alpha, q_\alpha; \mathcal{B}_1), \sum_{\alpha=1}^{d} \nu(\alpha) R_{12}^1(p_1^\alpha, p_2^\alpha, q_\alpha; \mathcal{B}_2)\right]$$

and

(81) $$\tilde{R}(\sigma,\nu) = \sum_{\alpha=1}^{d} \nu(\alpha) R_{12}(p_1^\alpha, p_2^\alpha, q_\alpha; \mathcal{B}_1).$$

Set

(82) $$F_{II}(\mathcal{B}_1,\mathcal{B}_2) = \{\vec{R} \,|\, \vec{R} = \vec{R}(\sigma,\nu) \text{ for some } (\sigma,\nu)\}.$$

For every $\vec{R} = (\tilde{R}_1, \tilde{R}_2, \tilde{R}) \in F_{II}(\mathcal{B}_1,\mathcal{B}_2)$ define

(83) $$G_{II}(\vec{R}) = \{(R_1,R_2) \mid \sum_{s=1}^{2} R_s \le \tilde{R},\ R_s \le \tilde{R}_s \text{ for } s=1,2\}.$$

Finally define

$$(84) \qquad G_{II} = \bigcup_{\vec{R} \in F_{II}(B_1,B_2)} G_{II}(\vec{R}).$$

We remark that G_{II}, like G_I, is a closed convex region in the Euclidean plane.

B. Pure Parameter-Channels

We may specialize the collection σ to contain only pure parameter-channels. More precisely, let $\sigma^* = (P,Q^*)$ be a finite collection of triples $\{(p_1^\alpha,p_2^\alpha,q_\alpha^*):\alpha=1,\ldots,d\}$ as defined in (77), but now such that each q_α^* is a pure parameter-channel. Define

$$(85) \qquad F_{II}^*(B_1,B_2) = \{\vec{R}|\vec{R} = \vec{R}(\sigma^*,\nu) \text{ for some } (\sigma^*,\nu)\}.$$

Then we have

Theorem 4: It suffices to take in (84) the union over all $\vec{R}$ belonging to $F_{II}^*(B_1,B_2)$. Thus

$$(86) \qquad G_{II} = \bigcup_{\vec{R} \in F_{II}^*(B_1,B_2)} G_{II}(\vec{R}).$$

Proof: We need to show that every $G_{II}(\vec{R}(\sigma,\nu))$ is contained in some $G_{II}(\vec{R}(\sigma^*,\nu'))$. This is indeed so, because there corresponds to every (σ,ν) some pair (σ^*,ν') such that $\vec{R}(\sigma,\nu) \le \vec{R}(\sigma^*,\nu')$. This follows from the convexity of $R_1(p_1,p_2,q;B_1)$, $R_2(p_1,p_2,q;B_1)$, $R_{12}(p_1,p_2,q;B_1)$, and

$R^1_{12}(p_1,p_2,q;\mathcal{B}_2)$ as functions of their transmission probabilities in the same way as in section IIIB.

C. The Main Theorem

Theorem 5: The region G_{II} is contained in the capacity region $G(P,T_{12},\text{II})$. Thus

$$(87)\qquad G_{II} \subset G(P,T_{12},\text{II}).$$

Proof: Our proof combines aspects of the random coding proof given by Ahlswede [2] for (P,T_{22},II), and the one given by Bergmans [4] for the degraded broadcast channel. Ahlswede's approach is based on the use of non-stationary sources (see also Ulrey [12] for generalizations). We shall carry this approach over to the non-degraded broadcast channel. Our random coding proof is really one for (P,T_{22},III).

Let $(R_1,R_2) \in G_{II}$. Then there exists a triple $\vec{R}(\sigma^*,\nu) \in F^*_{II}(\mathcal{B}_1,\mathcal{B}_2)$ such that

$$(88)\qquad R_s \le \tilde{R}_s(\sigma^*,\nu) \qquad \text{for } s=1,2;$$

and

$$(89)\qquad R_1+R_2 \le \tilde{R}(\sigma^*,\nu).$$

Let $\varepsilon > 0$, $\delta = \varepsilon/4$, $0 < \lambda_1,\lambda_2 < 1$. We can find a positive integer

$n=n_\delta$, and a collection of triples

$$\sigma_\delta^* = \{(p_1^t,p_2^t,q_t^*):t=1,\dots,n\}$$

such that

$$\Delta(\vec{R}(\sigma^*,\nu),\vec{R}(\sigma_\delta^*,\nu')) < \delta$$

where $\nu'(t)=1/n;t=1,\dots,n$; and $\Delta(.,.)$ denotes the Euclidean distance in three-space. This implies that

(90) $$R_1 < \frac{1}{n}\sum_{t=1}^{n} R_1(p_1^t,p_2^t,q_t^*;B_1) + \delta$$

(91) $$R_2 < \min\left[\frac{1}{n}\sum_{t=1}^{n} R_2(p_1^t,p_2^t,q_t^*;B_1),\frac{1}{n}\sum_{t=1}^{n} R_{12}^1(p_1^t,p_2^t,q_t^*;B_2)\right] + \delta$$

and

(92) $$R_1+R_2 < \frac{1}{n}\sum_{t=1}^{n} R_{12}(p_1^t,p_2^t,q_t^*;B_1) + \delta.$$

Choose the integers M_1 and M_2 such that

(93) $$2^{n(R_s-\varepsilon)} \le M_s \le 2^{n(R_s-\varepsilon)} + 1 \qquad s=1,2.$$

Define

(94) $$P_{1n}(X_1) = \prod_{t=1}^{n} p_1^t(x_1^t)$$

(95) $$P_{2n}(X_2) = \prod_{t=1}^{n} p_2^t(x_2^t),$$

and

(96) $$Q_n^*(Y|X_1,X_2) = \prod_{t=1}^{n} q_t^*(y^t|x_1^t,x_2^t)$$

for $X_1=(x_1^1,\ldots,x_1^n) \in A^n$, $X_2=(x_2^1,\ldots,x_2^n) \in A^n$, $Y=(y^1,\ldots,y^n) \in A^n$.

Define

(97) $$P_n(Z_1,Z_2|X_1,X_2) = \sum_{Y\in A^n} P(Z_1,Z_2|Y)Q_n^*(Y|X_1,X_2)$$

where $P(Z_1,Z_2|Y)$ is defined by (1). From it derive

(98) $$P_n(Z_2|X_2) = \sum_{X_1\in A^n} P_n(Z_1|X_1,X_2)P_{1n}(X_1)$$

and

(99) $$P_n(Z_2) = \sum_{X_2\in A^n} P_n(Z_2|X_2)P_{2n}(X_2).$$

Next define

(100) $$I_n(X_2;Z_2) = \frac{1}{n}\log_2 \frac{P_n(Z_2|X_2)}{P_n(Z_2)} .$$

Also define

$$(101) \qquad I^t(x_2;z_2) = \log_2 \frac{p^t(z_2|x_2)}{p^t(z_2)} \qquad t=1,\ldots,n;$$

where

$$(102) \qquad p^t(z_2,x_1,x_2) = \sum_{y\in A} p(z_2|y)q_t^*(y|x_1,x_2)p_1^t(x_1)p_2^t(x_2).$$

Clearly

$$(103) \qquad I_n(X_2;Z_2) = \frac{1}{n}\sum_{t=1}^{n} I^t(x_2^t;z_2^t)$$

for $X_2=(x_2^1,\ldots,x_2^n) \in A^n$, and $Z_2=(z_2^1,\ldots,z_2^n) \in B_2^n$.

Moreover

$$(104) \qquad E[I_n(X_2;Z_2)] = \frac{1}{n}\sum_{t=1}^{n} R_{12}^1(p_1^t,p_2^t,q_t^*;B_2).$$

Consider now the following random coding scheme. Select M_1 cloud centers $u_1,\ldots,u_{M_1}$ (all in A^n) independently drawn according to $P_{1n}(X_1)$. Also select M_2 cloud centers $v_1,\ldots,v_{M_2}$ (all in A^n) independently drawn (from each other and from the u_i's) according to $P_{2n}(X_2)$. If the message pair (i,j) is presented for transmission, and the set of cloud centers $(u_1,\ldots,u_{M_1};\ v_1,\ldots,v_{M_2})$ is randomly generated, and $Q_n^*(w_{ij}|u_i,v_j)=1$, then the codeword w_{ij} is transmitted over the channel.

The decoding set for user U_{z1} if message pair (i,j) is sent, is denoted by B_{ij}, and defined by

(105) $$B_{ij} = \{Z_1 \epsilon B_1^n \mid P(Z_1|w_{ij}) > P(Z_1|w_{kl}) \text{ for all } (k,l) \neq (i,j)\}.$$

The probability of error committed by U_{z1} in decoding message pair (i,j) is, for the given set of cloud centers, denoted and defined by

$$\mu_1(i,j) = 1-P(B_{ij}|w_{ij}) = 1-P_n(B_{ij}|u_i,v_j).$$

The expected value, over all sets of cloud centers, of the arithmetic average probability of decoding error made by U_{z1} equals the expected value of $\mu_1(1,1)$, which is denoted by $\overline{\mu_1(1,1)}$. It is an immediate consequence of the results of Ahlswede [2], and those of Ulrey [12], that $\overline{\mu_1(1,1)}$ tends to zero as n tends to infinity, whereby n is an integer-multiple of n_δ. The result for general n follows easily.

Next we investigate the probability of error for sending to U_{z2}. Let

(106) $$S(Z_2) = \left\{X_2 \epsilon A^n \mid I_n(X_2;Z_2) > \frac{R_2+E[I_n(X_2;Z_2)]-3\delta}{2}\right\}$$

and define

(107) $$d(X_2,Z_2) = \begin{cases} 1 & \text{if } X_2 \notin S(Z_2) \\ 0 & \text{otherwise.} \end{cases}$$

Define the decoding set

(108) $$D_j = \{Z_2 \epsilon B_2^n \mid v_j \epsilon S(Z_2);\ v_l \notin S(Z_2) \text{ for all } l \neq j\}.$$

The probability of error in decoding message j by U_{z2} is for the given set of cloud centers denoted by

$$(109) \qquad \mu_2(i,j) = 1-P(D_j|w_{ij}) = 1-P_n(D_j|u_i,v_j).$$

Clearly, $\mu_2(i,j)$ is bounded above by $P_e^{(1)}(i,j)+P_e^{(2)}(i,j)$ where

$$(110) \qquad P_e^{(1)}(i,j) = \sum_{Z_2\in B_2^n} P(Z_2|w_{ij})d(v_j,Z_2)$$

and

$$(111) \qquad P_e^{(2)}(i,j) = \sum_{Z_2\in B_2^n} P(Z_2|w_{ij}) \sum_{\substack{l\neq j\\ l=1}}^{M_2} (1-d(v_l,Z_2)).$$

The expected value over all sets of cloud centers of the arithmetic average probability of decoding error made by U_{z2} is equal to the expected value of $\mu_2(i,j)$, which is denoted by $\overline{\mu_2(i,j)}$. Clearly,

$$(112) \qquad \overline{\mu_2(i,j)} \leqq \overline{P_e^{(1)}(i,j)} + \overline{P_e^{(2)}(i,j)}.$$

Now, proceeding as Bergmans in [4], we obtain that

$$(113) \qquad \overline{P_e^{(1)}(i,j)} \leqq P\left[I_n(X_2,Z_2) \leqq \frac{1}{n} \sum_{t=1}^{n} R_{12}^1(p_1^t,p_2^t,q_t^*;B_2)-\delta\right].$$

The righthand side of (113) goes to zero as n tends to infinity, by the weak law of large numbers for independent, not necessarily identically distributed, random variables. The variance of $I^t(x_2;z_2)$ is bounded uniformly in t. This follows from the remark on p. 123 of Wolfowitz [18]. Similarly we obtain

$$(114) \qquad \overline{P_e^{(2)}(i,j)} \leqq 2^{n(2\delta-\varepsilon)} = 2^{-n\varepsilon/2}.$$

Hence $\mu_2(i,j)$ tends to zero as n tends to infinity, when n is an integer-multiple of n_δ. The result for general n follows readily. This shows the existence of an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for (P,T_{12},II) for n sufficiently large with M_1 and M_2 satisfying (93). Hence $(R_1,R_2) \in G(P,T_{12},\text{II})$ and $G_{\text{II}} \subset G(P,T_{12},\text{II})$. This completes the proof.

Conjecture 4: In defining σ (see (77)) we considered triples (p_1,p_2,q) such that $E_q = (A\times A, q(y|x_1,x_2), A)$ is a parameter-channel with input alphabets equal to the given output alphabet A. We conjecture that the region G_{II} does not change if in (84) we take the union over those $\vec{R}$ which arise from pairs (σ,ν) such that some $E_q = (A_1\times A_2, q(y|x_1,x_2), A)$ have general input alphabets A_1 and A_2. The fact that it suffices to restrict to $a_2=\min(a,b_1,b_2)$ seems to follow in the same way as in Gallager [7]. A possible proof of the conjecture that it also suffices to restrict to $a_1=\min(a,b_1)$ may be given along the same lines.

D. Comparison With The Degraded Broadcast Channel

It is interesting to note what happens to expression (86) when the broadcast channel K is assumed to be degraded. According to Bergmans [4] hereby the following is meant. Let K_1 and K_2 denote the marginal channels whose transition probabilities are defined by (2) and (3) respectively. We say that K_2 is a degraded version of K_1, if there exists a third channel K_3 such that K_2 can be represented as the cascade of K_1 followed by K_3, in which case K is called a degraded broadcast channel.

If K_2 is a degraded version of K_1, expression (80) reduces to

$$\tilde{R}_2^{\mathcal{D}}(\sigma,\nu) = \sum_{\alpha=1}^{d} \nu(\alpha) R_{12}^1(p_1^\alpha, p_2^\alpha, q_\alpha; \mathcal{B}_2). \tag{115}$$

Namely, in this case we have

$$R_{12}^1(p_1,p_2,q;\mathcal{B}_2) \le R_{12}^1(p_1,p_2,q;\mathcal{B}_1) = E\left[\log_2 \frac{p(z_1|x_2)}{p(z_1)}\right] \tag{116}$$

by the cascading process, and

$$R_{12}^1(p_1,p_2,q;\mathcal{B}_1) \le R_2(p_1,p_2,q;\mathcal{B}_1) \tag{117}$$

by convexity. Moreover, one always has

$$R_1(p_1,p_2,q;\mathcal{B}_1) + R_{12}^1(p_1,p_2,q;\mathcal{B}_1) = R_{12}(p_1,p_2,q;\mathcal{B}_1) \tag{118}$$

so that in the case of degradation

$$R_1(p_1,p_2,q;\mathcal{B}_1) + R_{12}^1(p_1,p_2,q;\mathcal{B}_2) \le R_{12}(p_1,p_2,q;\mathcal{B}_1). \tag{119}$$

With every pair (σ^*,ν) we now associate the pair

(120) $\vec{R}^{\mathcal{D}}(\sigma^*,\nu) = (\tilde{R}_1(\sigma^*,\nu),\tilde{R}_2^{\mathcal{D}}(\sigma^*,\nu))$.

Next we set

(121) $F_{II}^{\mathcal{D}}(\mathcal{B}_1,\mathcal{B}_2) = \{\vec{R}^{\mathcal{D}} \mid \vec{R}^{\mathcal{D}}=\vec{R}^{\mathcal{D}}(\sigma^*,\nu) \text{ for some } (\sigma^*,\nu)\}$.

For every $\vec{R}^{\mathcal{D}}=(\tilde{R}_1,\tilde{R}_2^{\mathcal{D}}) \in F_{II}^{\mathcal{D}}(\mathcal{B}_1,\mathcal{B}_2)$ we define

(122) $G_{II}^{\mathcal{D}}(\vec{R}^{\mathcal{D}}) = \{(R_1,R_2) \mid R_1 \leq \tilde{R}_1, R_2 \leq \tilde{R}_2^{\mathcal{D}}\}$.

Finally we define the region

(123) $$G_{II}^{\mathcal{D}} = \bigcup_{\vec{R}^{\mathcal{D}} \in F_{II}^{\mathcal{D}}(\mathcal{B}_1,\mathcal{B}_2)} G_{II}^{\mathcal{D}}(\vec{R}^{\mathcal{D}}).$$

It follows from (115) and (119) that in the case of a degraded broadcast channel G_{II} reduces to $G_{II}^{\mathcal{D}}$. Another way of characterizing $G_{II}^{\mathcal{D}}$ is as follows. For every pure parameter-channel $E_{q^*} = (A\times A, q^*(y|x_1,x_2), A)$, define

(124) $$C_{II}(q^*) = \{(R_1(p_1,p_2,q^*;\mathcal{B}_1), R_{12}^1(p_1,p_2,q^*;\mathcal{B}_2)): p_1 \text{ and } p_2 \text{ are p.d.'s on } A\}.$$

Next let

(125) $$C_{II} = \cup_{q^*} C_{II}(q^*)$$

where the union is taken over all pure parameter-channels E_{q^*}. Clearly

(126) $$G^{\mathcal{D}}_{II} = co(C_{II}).$$

We claim that the region $co(C_{II})$ is precisely the region of attainable rates obtained by Bergmans [4] for the degraded broadcast channel with two components. This can be seen as follows. Every triple (p_1,p_2,q^*) with $E_{q^*} = (A\times A, q^*(y|x_1,x_2), A)$ determines a pair (p_2,ρ) whereby ρ is the transition probability of an artificial channel $A_\rho = (A,\rho(y|x_2),A)$ with $\rho(y|x_2)$ given by (45). Conversely, every pair (p_2,ρ) of this type can be written as a triple (p_1,p_2,q^*) according to the procedure described at the bottom of p. 392 of [9], whereby $E_{q^*} = (A_1\times A, q^*(y|x_1,x_2), A)$ and $a_1 = a^a$. The pair (p_2,ρ) can be considered as representing the two parameters of the Cover-Bergmans random coding scheme, with ρ being the transition probability of an artificial satellizing channel A_ρ, and p_2 being a probability distribution on the inputs of A_ρ. It follows that every pair of rates which is attainable according to the Cover-Bergmans procedure is attainable with our procedure, and conversely, with the proviso that in the definition of $G_{II}, G^{\mathcal{D}}_{II}$, and C_{II} we allow parameter-channels $E_{q^*} = (A_1\times A, q^*(y|x_1,x_2), A)$ such that $a_1 = a^a$. However, in view of Conjecture 4, we believe this makes no difference.

E. A Limiting Expression For $G(P,T_{12},II)$

In analogy with the case (P,T_{12},I) we now derive a limiting expression for $G(P,T_{12},II)$. As in section IIIF let K^n denote the memoryless

n-extension of the broadcast channel $K = (A, p(z_1, z_2|y), B_1 \times B_2)$, and recall expressions (54), (55), (56), and (57). We define mutual information functions as follows. Let

$$(127) \qquad R_1(P_{1n}, P_{2n}, Q_n; B_1^n) = E\left[\log_2 \frac{P^n(Z_1|X_1, X_2)}{P^n(Z_1|X_2)}\right]$$

$$(128) \qquad R_2(P_{1n}, P_{2n}, Q_n; B_1^n) = E\left[\log_2 \frac{P^n(Z_1|X_1, X_2)}{P^n(Z_1|X_1)}\right]$$

$$(129) \qquad R_{12}^1(P_{1n}, P_{2n}, Q_n; B_2^n) = E\left[\log_2 \frac{P^n(Z_2|X_2)}{P^n(Z_2)}\right]$$

and

$$(130) \qquad R_{12}(P_{1n}, P_{2n}, Q_n; B_1^n) = E\left[\log_2 \frac{P^n(Z_1|X_1, X_2)}{P^n(Z_1)}\right]$$

where the expectations are taken with respect to (57).

When the underlying broadcast channel is K^n instead of K, the analogues of expressions (77) to (83) are easily derived, and therefore will be omitted here. We define G_{II}^n as being the inner bound of the capacity region of the d.m. broadcast channel K^n in situation (P, T_{12}, II) obtained according to Theorem 5.

Let

$$(131) \qquad K_{II}^n = \frac{G_{II}^n}{n} = \{(R_1, R_2) \mid (nR_1, nR_2) \in G_{II}^n\},$$

and

(132) $$G_{II}^{\infty} = \bigcup_{n=1}^{\infty} K_{II}^{n}.$$

Then we have

Theorem 6: The capacity region $G(P,T_{12},II)$ equals the region G_{II}^{∞}.

Proof: (a) The proof of the fact that $G_{II}^{\infty} \subset G(P,T_{12},II)$ is completely similar to part (iia) of the proof of Theorem 3.

(b) The fact that $G(P,T_{12},II) \subset G_{II}^{\infty}$ is proven as follows. Let $(R_1,R_2) \in G(P,T_{12},II)$. Let $\varepsilon > 0$, $0 < \lambda_1,\lambda_2 < 1$. Then there exists for n sufficiently large an $(n,M_1,M_2,\lambda_1,\lambda_2)$-code for K in situation (P,T_{12},II) such that

(133) $$\frac{1}{n}\log_2 M_i \geqq R_i - \varepsilon \qquad i=1,2.$$

We denote a code like this by the system (13). In the same way as in part (iib) of the proof of Theorem 3 we can translate (13) into a system

(134) $$\{u_i,v_j,B_{ij},D_j \mid i=1,\ldots,M_1;\ j=1,\ldots,M_2\}$$

by means of a pure parameter-channel $E_Q^n{*}$ for K^n such that $Q_n^*(w_{ij}|u_i,v_j) = 1$. The system (134) is a $(1,M_1,M_2,\lambda_1,\lambda_2)$-code for the cascade $E_Q^n{*}K^n$ in situation (P,T_{22},III). As before, let P_{1n}^* and P_{2n}^* denote the uniform distributions on the sets $\{u_1,\ldots,u_{M_1}\}$ and $\{v_1,\ldots,v_{M_2}\}$ respectively. It follows from standard results on weak converses for multi-way channels (see [1], [2], [12], [15], and [17]) that

(135) $(1-\lambda)\log_2 M_1 \leqq R_1(P^*_{1n},P^*_{2n},Q^*_n;B^n_1) + 1$

(136) $(1-\lambda)\log_2 M_2 \leqq \min\left[R_2(P^*_{1n},P^*_{2n},Q^*_n;B^n_1)+R^1_{12}(P^*_{1n},P^*_{2n},Q^*_n;B^n_2)\right] + 1$

and

(137) $(1-\lambda)\log_2 M_1M_2 \leqq R_{12}(P^*_{1n},P^*_{2n},Q^*_n;B^n_1) + 1.$

It is easily concluded that, for any $\delta > 0$, and for all n sufficiently large, and all ε and λ sufficiently small,

(138) $(nR_1-n\delta,nR_2-n\delta) \in G^n_{\text{II}}.$

This implies that $(R_1-\delta,R_2-\delta) \in G^\infty_{\text{II}}$ for all $\delta > 0$. Since G^∞_{II} is closed, it follows that $(R_1,R_2) \in G^\infty_{\text{II}}$. Therefore $G(P,T_{12},\text{II}) \subset G^\infty_{\text{II}}$, which completes the proof.

V. RANDOM CODING THEOREM FOR (P,T_{12},III)

A. Mutual Information Functions

Let be given the d.m. broadcast channel $K = (A,p(z_1,z_2|y),B_1\times B_2)$, a parameter-channel $F_q = (A_1\times A_0\times A_2,q(y|x_1,x_0,x_2),A)$ of type (P,T_{31}), and the cascade $F_qK = (A_1\times A_0\times A_2,p(z_1,z_2|x_1,x_0,x_2),B_1\times B_2)$ whose transmission probabilities are as defined in (7). Assume $A_1 = A_0 = A_2 = A$, and let $p_1(x_1)$, $p_0(x_0)$, and $p_2(x_2)$ be three probability distributions on A. Define

$$p(x_1,x_0,x_2,z_1,z_2) = p(z_1,z_2|x_1,x_0,x_2)p_1(x_1)p_0(x_0)p_2(x_2) \tag{139}$$

and derive from it $p(z_1|x_1,x_0)$, $p(z_2|x_0,x_2)$, $p(z_1|x_1)$, $p(z_1|x_0)$, $p(z_2|x_0)$, $p(z_2|x_2)$, $p(z_1)$, and $p(z_2)$ in the usual way. Define the following mutual information functions.

$$R^2_{21}(p_1,p_0,p_2,q;B_1) = E\left[\log_2 \frac{p(z_1|x_1,x_0)}{p(z_1|x_0)}\right] \tag{140}$$

$$R^2_{20}(p_1,p_0,p_2,q;B_1) = E\left[\log_2 \frac{p(z_1|x_1,x_0)}{p(z_1|x_1)}\right] \tag{141}$$

$$R^1_{10}(p_1,p_0,p_2,q;B_2) = E\left[\log_2 \frac{p(z_2|x_0,x_2)}{p(z_2|x_2)}\right] \tag{142}$$

$$R^1_{12}(p_1,p_0,p_2,q;B_2) = E\left[\log_2 \frac{p(z_2|x_0,x_2)}{p(z_2|x_0)}\right] \tag{143}$$

$$R^2_{210}(p_1,p_0,p_2,q;B_1) = E\left[\log_2 \frac{p(z_1|x_1,x_0)}{p(z_1)}\right] \tag{144}$$

and

$$(145) \qquad R^1_{102}(p_1,p_0,p_2,q;\mathcal{B}_2) = E\left[\log_2 \frac{p(z_2|x_0,x_2)}{p(z_2)}\right].$$

Here, all expectations are taken with respect to (139).

Let $\sigma = (\mathcal{P},\mathcal{Q})$ be a finite collection of quadruples

$$(146) \qquad \{(p_1^\alpha,p_0^\alpha,p_2^\alpha,q_\alpha):\alpha=1,\dots,d\}$$

where $(A\times A\times A,q_\alpha(y|x_1,x_0,x_2),A)$ is a parameter-channel of type (P,T_{31}), and p_1^α, p_0^α, and p_2^α are probability distributions on A. Also, let $\nu =$ $= \{\nu(\alpha):\alpha=1,\dots,d\}$ be a probability distribution on σ. With every pair (σ,ν) we associate a quintuple

$$(147) \qquad \overleftarrow{R}(\sigma,\nu) = (\bar{R}_1(\sigma,\nu),\bar{R}_0(\sigma,\nu),\bar{R}_2(\sigma,\nu),\bar{R}_{10}(\sigma,\nu),\bar{R}_{02}(\sigma,\nu))$$

where

$$(148) \qquad \bar{R}_1(\sigma,\nu) = \sum_{\alpha=1}^{d} \nu(\alpha)R^2_{21}(p_1^\alpha,p_0^\alpha,p_2^\alpha,q_\alpha;\mathcal{B}_1)$$

$$(149) \qquad \bar{R}_0(\sigma,\nu) = \min\left[\sum_{\alpha=1}^{d} \nu(\alpha)R^2_{20}(p_1^\alpha,p_0^\alpha,p_2^\alpha,q_\alpha;\mathcal{B}_1),\ \sum_{\alpha=1}^{d} \nu(\alpha)R^1_{10}(p_1^\alpha,p_0^\alpha,p_2^\alpha,q_\alpha;\mathcal{B}_2)\right]$$

$$(150) \qquad \bar{R}_2(\sigma,\nu) = \sum_{\alpha=1}^{d} \nu(\alpha)R^1_{12}(p_1^\alpha,p_0^\alpha,p_2^\alpha,q_\alpha;\mathcal{B}_2)$$

(151) $$\bar{R}_{10}(\sigma,\nu) = \sum_{\alpha=1}^{d} \nu(\alpha) R^2_{210}(p_1^\alpha, p_0^\alpha, p_2^\alpha, q_\alpha; \mathcal{B}_1)$$

and

(152) $$\bar{R}_{02}(\sigma,\nu) = \sum_{\alpha=1}^{d} \nu(\alpha) R^1_{102}(p_1^\alpha, p_0^\alpha, p_2^\alpha, q_\alpha; \mathcal{B}_2).$$

Set

(153) $$\mathcal{F}_{III}(\mathcal{B}_1,\mathcal{B}_2) = \{\overleftarrow{R} \mid \overleftarrow{R} = \overleftarrow{R}(\sigma,\nu) \text{ for some } (\sigma,\nu)\}.$$

For every $\overleftarrow{R} = (\bar{R}_1, \bar{R}_0, \bar{R}_2, \bar{R}_{10}, \bar{R}_{02}) \in \mathcal{F}_{III}(\mathcal{B}_1,\mathcal{B}_2)$ define

(154) $$G_{III}(\overleftarrow{R}) = \{(R_1,R_2,R_0) \mid \sum_{s=0}^{1} R_s \le \bar{R}_{10}, \sum_{s=0}^{1} R_{2s} \le \bar{R}_{02}, R_s \le \bar{R}_s \text{ for } s=0,1,2\}.$$

Finally define

(155) $$G_{III} = \bigcup_{\overleftarrow{R} \in \mathcal{F}_{III}(\mathcal{B}_1,\mathcal{B}_2)} G_{III}(\overleftarrow{R}).$$

Clearly, G_{III} is a closed convex region in Euclidean three-space.

B. Pure Parameter-Channels

Let $\sigma^* = (\mathcal{P},\mathcal{Q}^*)$ be a finite collection of quadruples $\{(p_1^\alpha, p_0^\alpha, p_2^\alpha, q_\alpha^*) : \alpha = 1,\ldots,d\}$ as defined in (146), but now such that each q_α^* is a pure parameter-channel of type (P,T_{31}). Define

(156) $$F^*_{III}(\mathcal{B}_1,\mathcal{B}_2) = \{\overleftarrow{R} \mid \overleftarrow{R} = \overleftarrow{R}(\sigma^*,\nu) \text{ for some } (\sigma^*,\nu)\}.$$

Then we have

Theorem 7:

(157) $$G_{III} = \bigcup_{\overleftarrow{R}\in F^*_{III}(\mathcal{B}_1,\mathcal{B}_2)} G_{III}(\overleftarrow{R}).$$

Proof: Follows from convexity as in the proof of Theorem 4.

C. The Main Theorem

Theorem 8:

(158) $$G_{III} \subset G(P,T_{12},\text{III}).$$

Proof: Our proof is a direct application of the random coding proofs given by Ahlswede [2] and Ulrey [12] for (P,T_{22},II) and (P,T_{32},I), respectively. In addition, the proof of this theorem involves some aspects of the random coding proofs for (P,T_{21}) and (P,T_{22},I) given by Ahlswede [1] and [2]. (See also [15] in this regard.) We shall only sketch the proof and omit any details, because of its length and the complexity of the notation involved. We remark here that our random coding proof can also be viewed as one for (P,T_{32},II).

Let $(R_1,R_2,R_0) \in G_{III}$. Let $\varepsilon > 0$, and $0 < \delta < \varepsilon$. There exists a positive integer $n=n_\delta$, and a collection of quadruples

(159) $$\sigma_\delta^* = \{(p_1^t, p_0^t, p_2^t, q_t^*) : t=1,\ldots,n\}$$

such that

(160) $$R_1 < \frac{1}{n} \sum_{t=1}^{n} R_{21}^2(p_1^t, p_0^t, p_2^t, q_t^*; B_1) + \delta$$

(161) $$R_0 < \min\left[\frac{1}{n} \sum_{t=1}^{n} R_{20}^2(p_1^t, p_0^t, p_2^t, q_t^*; B_1), \frac{1}{n} \sum_{t=1}^{n} R_{10}^1(p_1^t, p_0^t, p_2^t, q_t^*; B_2)\right] + \delta$$

(162) $$R_2 < \frac{1}{n} \sum_{t=1}^{n} R_{12}^1(p_1^t, p_0^t, p_2^t, q_t^*; B_2) + \delta$$

(163) $$R_1+R_0 < \frac{1}{n} \sum_{t=1}^{n} R_{210}^2(p_1^t, p_0^t, p_2^t, q_t^*; B_1) + \delta$$

and

(164) $$R_0+R_2 < \frac{1}{n} \sum_{t=1}^{n} R_{102}^1(p_1^t, p_0^t, p_2^t, q_t^*; B_2) + \delta.$$

Choose integers M_1, M_2, and M_0 such that

(165) $$2^{n(R_s-\varepsilon)} \le M_s \le 2^{n(R_s-\varepsilon)} + 1 \qquad s=0,1,2.$$

Define

(166) $$P_{sn}(X_s) = \prod_{t=1}^{n} p_s^t(x_s^t) \qquad s=0,1,2;$$

and

$$(167) \qquad Q_n^*(Y|X_1,X_0,X_2) = \prod_{t=1}^{n} q_t^*(y^t|x_1^t,x_0^t,x_2^t)$$

for $X_s = (x_s^1,\ldots,x_s^n) \in A^n$; $s=0,1,2$; and $Y = (y^1,\ldots,y^n) \in A^n$.

Consider now the following random coding scheme. Select M_1 cloud centers $u_{11},\ldots,u_{1M_1}$ in A^n, independently drawn according to $P_{1n}(X_1)$. At the same time, select M_0 cloud centers $u_{01},\ldots,u_{0M_0}$ in A^n, independently drawn according to $P_{0n}(X_0)$. Also, select M_2 cloud centers $u_{21},\ldots,u_{2M_2}$ in A^n independently drawn according to $P_{2n}(X_2)$. Moreover, choose the three sets of cloud centers independently from each other. If the message triple (i,j,k) is presented for transmission; $1 \le i \le M_1$, $1 \le j \le M_2$, $1 \le k \le M_0$; and $Q_n^*(Y|u_{1i},u_{0k},u_{2j}) = 1$ for some $Y \in A^n$, then the codeword $w_{ijk} = Y$ is transmitted over the channel. (There are various interpretations of this random coding scheme possible in terms of a satellization process, depending on different choices of the satellizing channel.) The decoding sets corresponding to this random coding scheme are defined as follows.

The decoding set for U_{z1} if message triple (i,j,k) is sent is denoted by B_{ik} and defined by

$$(168) \qquad B_{ik} = \{Z_1 \in \mathcal{B}_1^n | \sum_{j=1}^{M_2} P(Z_1|w_{ijk})P_{2n}(u_{2j}) > \sum_{j=1}^{M_2} P(Z_1|w_{i'jk'})P_{2n}(u_{2j})$$

$$\text{for all } (i',k') \neq (i,k)\}.$$

Similarly, the decoding set for U_{z2} is denoted and defined by

$$(169) \qquad D_{jk} = \{Z_2 \in \mathcal{B}_2^n | \sum_{i=1}^{M_1} P(Z_2|w_{ijk})P_{1n}(u_{1i}) > \sum_{i=1}^{M_1} P(Z_2|w_{ij'k'})P_{1n}(u_{1i})$$

$$\text{for all } (j',k') \neq (j,k)\}.$$

Let

$$\mu_1(i,j,k) = 1-P(B_{ik}|w_{ijk}) \tag{170}$$

and

$$\mu_2(i,j,k) = 1-P(D_{jk}|w_{ijk}). \tag{171}$$

Denote by $\overline{\mu_s(i,j,k)}$ the expected value over all sets of cloud centers of $\mu_s(i,j,k)$; $s=1,2$. It is an immediate consequence of the results of Ahlswede [2] and Ulrey [12] that $\overline{\mu_s(i,j,k)}$ tends to zero as n tends to infinity. This completes the proof.

D. Comparison With G_I And G_{II}.

We now comment on how Theorem 2 (expression (38)) and Theorem 5 (expression (87)) can be obtained directly from Theorem 8.

Suppose in communication situation (P,T_{12},III) we set $R_0 = 0$. Then it is not possible for U_{z1} and U_{z2} to decode conditionally on the message k received. This corresponds to using in the mutual information expressions (147) through (152) only those pairs (σ,ν) such that p_0 assigns probability zero or one to every single input letter x_0. Let $\sigma_0 = \{P_0,Q\}$ be a finite collection of quadruples as defined in (146), but now such that $p_0^{\alpha} = 0$ or 1 for $\alpha=1,\ldots,d$. Let

$$F_{III}^{(0)}(B_1,B_2) = \{\overleftarrow{R}_0 | \overleftarrow{R}_0 = \overleftarrow{R}(\sigma_0,\nu) \text{ for some } (\sigma_0,\nu)\} \tag{172}$$

and

$$(173) \qquad G^{(0)}_{\mathrm{III}} = \bigcup_{\overleftarrow{R}_0 \in F^{(0)}_{\mathrm{III}}(\mathcal{B}_1,\mathcal{B}_2)} G_{\mathrm{III}}(\overleftarrow{R}_0).$$

Then it is easily verified that

$$(174) \qquad G_{\mathrm{I}} = \{(R_1,R_2) \mid (R_1,R_2,0) \in G^{(0)}_{\mathrm{III}}\}.$$

Similarly, we can derive the expression for G_{II} from the results of the present section. Suppose, in situation (P,T_{12},III) we set $R_2 = 0$, and put $R_1' = R_1$ and $R_2' = R_0$. Let $\sigma_2 = \{P_2,Q\}$ be a finite collection of quadruples as defined in (146), but now such that $p_2^{\alpha} = 0$ or 1 for $\alpha=1,\ldots,d$. Let

$$(175) \qquad F^{(2)}_{\mathrm{III}}(\mathcal{B}_1,\mathcal{B}_2) = \{\overleftarrow{R}_2 | \overleftarrow{R}_2 = \overleftarrow{R}(\sigma_2,\nu) \text{ for some } (\sigma_2,\nu)\}$$

and

$$(176) \qquad G^{(2)}_{\mathrm{III}} = \bigcup_{\overleftarrow{R}_2 \in F^{(2)}_{\mathrm{III}}(\mathcal{B}_1,\mathcal{B}_2)} G_{\mathrm{III}}(\overleftarrow{R}_2).$$

It is easily verified that

$$(177) \qquad G_{\mathrm{II}} = \{(R_1',R_2') \mid (R_1',0,R_2') \in G^{(2)}_{\mathrm{III}}\}.$$

E. A Limiting Expression For $G(P,T_{12},\text{III})$.

We can derive a limiting expression for $G(P,T_{12},\text{III})$ similar to those obtained for $G(P,T_{12},\text{I})$ and $G(P,T_{12},\text{II})$. Omitting the definitions of the mutual information functions involved, we define directly G^n_{III} to be the inner bound of the capacity region of the d.m. broadcast channel K^n in situation (P,T_{12},III) obtained according to Theorem 8. Let

(178) $$K^n_{\text{III}} = \frac{G^n_{\text{III}}}{n} = \{(R_1,R_2,R_0) \mid (nR_1,nR_2,nR_0) \in G^n_{\text{III}}\}$$

and

(179) $$G^\infty_{\text{III}} = \bigcup_{n=1}^{\infty} K^n_{\text{III}}.$$

Then we have

Theorem 9:

$$G(P,T_{12},\text{III}) = G^\infty_{\text{III}}.$$

Proof: The proof of this theorem is omitted, as it is completely analogous to the proof of Theorem 6.

ACKNOWLEDGEMENT

I wish to thank Mrs. Tona Bays for her careful typing of the manuscript. Thanks are also due to Mr. J. Hillebrand and Mr. T. Baanders for their drawing of the figures. Finally, my thanks are extended to Mr. F. van Dijk and Mr. A. Veldkamp of the Numerical Mathematics Department of the Mathematical Centre for their computer programming.

REFERENCES

[1] R. Ahlswede, "Multi-way communication channels," presented at the *Second International Symposium on Information Theory* at Tsahkador, Armenian S.S.R., 1971. To appear in *Problems of Control and Information Theory*.

[2] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," manuscript, 1972. Submitted for publication.

[3] R.B. Ash, *"Information Theory."* New York: Interscience, 1965.

[4] P.P. Bergmans, "Random coding theorems for broadcast channels with degraded components," *IEEE Trans. Inform. Theory,* vol. IT-19, pp. 197-207, March 1973.

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

[6] R.G. Gallager, *"Information Theory and Reliable Communication."* New York: Wiley, 1968.

[7] R.G. Gallager, "Coding and capacity for degraded broadcast channels," presented at the *Third International Symposium on Information Theory* at Tallin, Estonian S.S.R., 1973.

[8] C.E. Shannon, "Certain results in coding theory for noisy channels," *Inform. Contr.,* vol. 1, pp. 6-25, 1957.

[9] C.E. Shannon, "A note on a partial ordering for communication channels," *Inform. Contr.*, vol. 1, pp. 390-397, 1958.

[10] C.E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, vol. 1, pp. 611-644, 1961.

[11] D. Slepian and J.K. Wolf, "A coding theorem for multiple access channels with correlated sources," manuscript, 1973. To appear in *Bell Syst. Tech. J.*.

[12] M.L. Ulrey, "A coding theorem for a channel with s senders and r receivers," manuscript, 1973. To appear in *Inform. Contr.*.

[13] F.A. Valentine, *"Convex Sets."* New York: McGraw-Hill, 1964.

[14] E.C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol 3, pp. 120-154, Spring 1971.

[15] E.C. van der Meulen, "The discrete memoryless channel with two senders and one receiver," presented at the *Second International Symposium on Information Theory* at Tsahkadsor, Armenian S.S.R., 1971. To appear in *Problems of Control and Information Theory*.

[16] E.C. van der Meulen, "Multi-terminal communication channels," *Adv. Appl. Prob.*, vol. 5, pp. 32-33, April 1973.

[17] E.C. van der Meulen, "On a problem by Ahlswede regarding the capacity region of certain multi-way channels," Mathematical Centre Report SW 19/73, Mathematisch Centrum, Amsterdam, 1973. Submitted for publication.

[18] J. Wolfowitz, *"Coding Theorems of Information Theory."* Second Edition. New York: Springer-Verlag, 1964.