# PACITY OF THE PRODUCT AND SUM OF TWO UNMATCHED BROADCAST CHANNELS

A. A. El Gamal

UDC 621.391.1

Assume that $p(y_1, z_1|x_1) = p(y_1|x_1)p(z_1|y_1)$ and $p(y_2, z_2|x_2) = p(z_2|x_2)p(y_2|z_2)$ are two degraded broadcast channels. A broadcast channel with transmitter $(X_1, X_2)$ and receivers $(Y_1, Y_2)$ and $(Z_1, Z_2)$ is called a product of two unmatched degraded broadcast channels. Similarly, the sum of two unmatched degraded broadcast channels is defined as a broadcast channel with transmitter $Y_1 \cup Y_2$ and receivers $(X_1 \cup X_2)$ and $Z_1 \cup Z_2$. Unlike the original broadcast channels, neither the sum nor the product is a degraded broadcast channel. The region of the capacity is established for the following: i) the product of two unmatched degraded discrete memoryless broadcast channels; ii) a spectral Gaussian broadcast channel; iii) the sum of two unmatched degraded discrete memoryless broadcast channels. These theorems concerning the capacity contain particular results regarding the product of discrete memoryless channels that were obtained by Poltyrev, as well as results concerning spectral Gaussian broadcast channels obtained by Hughes-Hartogs. They also demonstrate that the region of rates obtained by Hughes-Hartogs is optimal for zero overall rate.

## 1. Introduction

A discrete degraded memoryless broadcast channel with two outputs $(\mathscr{X}, p(y|x)p(z|y), \mathscr{Y}\times\mathscr{Z})$ consists of three finite alphabets $\mathscr{X}$, $\mathscr{Y}$, $\mathscr{Z}$ and two transition-probability matrices $p(y|x)$ and $p(z|y)$. As was established in [1, 2], the capacity region of this channel consists of all triples of rates $(R_0, R_1, R_2)$ such that

$$R_0+R_2<I(U; Z), \quad R_1<I(X; Y|Z)$$

some joint probability distribution of the form $p(u, x, y, z) = p(u)p(x|u)p(y|x)p(z|y)$.

Assume that two degraded broadcast channels $(\mathscr{X}_1, p(y_1|x_1)p(z_1|y_1), \mathscr{Y}_1\times\mathscr{Z}_1)$ and $(\mathscr{X}_2, p(z_2|x_2)p(y_2|z_2), \mathscr{Y}_2\times\mathscr{Z}_2)$ are specified, as shown in Fig. 1. We will define the product of unmatched degraded broadcast channels as a broadcast channel $(\mathscr{X}, p(y, z|x), \mathscr{Y}\times\mathscr{Z})$ with two outputs, where

$$\begin{gathered}\mathscr{X}=\mathscr{X}_1\times\mathscr{X}_2,\ \mathscr{Y}=\mathscr{Y}_1\times\mathscr{Y}_2,\ \mathscr{Z}=\mathscr{Z}_1\times\mathscr{Z}_2,\\ p(y, z|x)=p(y_1|x_1)p(z_1|y_1)p(z_2|x_2)p(y_2|z_2).\end{gathered} \tag{1}$$

When $\mathscr{X}_1\cap\mathscr{X}_2=\mathscr{Y}_1\cap\mathscr{Y}_2=\mathscr{Z}_1\cap\mathscr{Z}_2=\varnothing$, we also define the sum of unmatched broadcast channels as a broadcast channel with two outputs $(\mathscr{X}, p(y, z|x), \mathscr{Y}\times\mathscr{Z})$, where $\mathscr{X}=\mathscr{X}_1\cup\mathscr{X}_2$, $\mathscr{Y}=\mathscr{Y}_1\cup\mathscr{Y}_2$, $\mathscr{Z}=\mathscr{Z}_1\cup\mathscr{Z}_2$,

$$p(y, z|x)=\begin{cases}p_1(y_1|x_1)p_1(z_1|y_1), & \text{if} \quad (x, y, z)\in\mathscr{X}_1\times\mathscr{Y}_1\times\mathscr{Z}_1,\\ p_2(z_2|x_2)p(y_2|z_2), & \text{if} \quad (x, y, z)\in\mathscr{X}_2\times\mathscr{Y}_2\times\mathscr{Z}_2,\\ 0 \quad \text{otherwise.}\end{cases} \tag{2}$$

The notion of sum and product of channels was first introduced by Shannon in [3]. It is well known that the capacity of the product of two discrete memoryless channels with capacities $C_1$ and $C_2$ is $C = C_1 + C_2$, whereas the capacity of the sum of these channels is $C = \log_2 (2^{C_1} + 2^{C_2})$.

An example that gave an impetus to the study of the product of degraded broadcast channels is the spectral Gaussian broadcast channel shown in Fig. 2a. It is known that every broadcast channel with additive Gaussian noise is degraded. It is easy to show that if the noise spectrum $Z_1$ lies everywhere below the noise spectrum $Z_2$ (Fig. 2b), then the broadcast channel is degraded, and the capacity region for a two-component spectrum (Fig. 2c), as established by Hughes-Hartogs in [4], consists of all triples of rates $(R_0, R_1, R_2)$ such that

Translated from Problemy Peredachi Informatsii, Vol. 16, No. 1, pp. 3-23, January-March, 1980. Original article submitted May 3, 1978; revision submitted April 6, 1979.

 

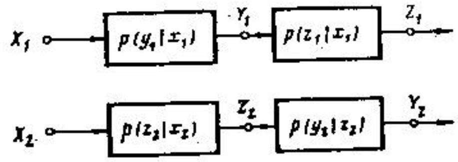


Fig. 1. Two degraded broadcast channels.

$$R_0+R_2<\frac{1}{2}\ln\left(1+\frac{\alpha_1\beta P}{\bar\alpha_1\beta P+N_1}\right)+\frac{1}{2}\ln\left(1+\frac{\alpha_2\beta P}{\bar\alpha_2\beta P+N_2}\right),$$
$$R_1<\frac{1}{2}\ln\left(1+\frac{\bar\alpha_1\beta P}{N_{11}}\right)+\frac{1}{2}\ln\left(1+\frac{\bar\alpha_2\beta P}{N_{21}}\right)$$

for some $\alpha_1$, $\alpha_2$, $\beta$ from the segment [0, 1].

If spectra $Z_1$ and $Z_2$ overlap (Fig. 2d and e), then the channel is not degraded and only the region of permissible rates is known (see [5, Problem 26]).

The product of discrete memoryless broadcasting channels as specified by formula (1) was investigated by Poltyrev [6], who determined the capacity region for independent rates ($R_0 \equiv 0$) and provided internal and external boundaries for the overall capacity region. Sum (2) of broadcast channels was introduced by Cover in [7] as a problem whose solution was unknown. Further studies of these issues may be found in [8,9].

In this paper we will offer theorems regarding the capacity for the following cases:

i) product of channels defined by formula (1);

ii) spectral Gaussian broadcast channel;

iii) sum of channels defined by formula (2).

We will also show that the region of permissible transmission rates given in [4] is optimal in the case of independent rates ($R_0 \equiv 0$).

In this paper we will employ the following standard definitions. An n-fold expansion of a broadcast channel will be understood to mean a broadcast channel ($\mathscr{X}^n$, $p(\mathbf{y}, \mathbf{z}|\mathbf{x})$, $\mathscr{Y}^n\times\mathscr{Z}^n$), where

$$p(\mathbf{y},\mathbf{z}|\mathbf{x})=\prod_{i=1}^{n}p(y_i,z_i|x_i). \tag{3}$$

An $[(M_0, M_1, M_2), n]$-code for a broadcast channel consists of three sets of integers:

$$\mathscr{M}_0=\{1,\ldots,M_0\},\ \mathscr{M}_1=\{1,\ldots,M_1\},\ \mathscr{M}_2=\{1,\ldots,M_2\}, \tag{4}$$

a coding mapping

$$\mathbf{x}:\mathscr{M}_0\times\mathscr{M}_1\times\mathscr{M}_2\to\mathscr{X}^n, \tag{5}$$

and two decoding mappings

$$\begin{aligned}&\mathbf{g}_1:\mathscr{Y}^n\to\mathscr{M}_0\times\mathscr{M}_1,\ \mathbf{g}_1(\mathbf{Y})=(\hat W_0,\ \hat W_1),\\&\mathbf{g}_2:\mathscr{Z}^n\to\mathscr{M}_0\times\mathscr{M}_2,\ \mathbf{g}_2(\mathbf{Z})=(\hat{\hat W}_0,\ \hat W_2).\end{aligned} \tag{6}$$

Set $\{\mathbf{x}(w_0, w_1, w_2):(w_0, w_1, w_2)\in\mathscr{M}_0\times\mathscr{M}_1\times\mathscr{M}_2\}$ is called the set of code words. Integer $w_0$ is interpreted as the common part of the message, while integers $w_1$ and $w_2$ are called the independent parts of the message. Assuming that a uniform distribution is specified on the set of messages $\mathscr{M}_0\times\mathscr{M}_1\times\mathscr{M}_2$, we define the mean error probabilities for decoders $\mathbf{g}_1$ and $\mathbf{g}_2$ by the following formulas, respectively:

$$\begin{aligned}P_{e_1}{}^n&=\frac{1}{M_0M_1M_2}\sum_{w_0,w_1,w_2\in\mathscr{M}_0\times\mathscr{M}_1\times\mathscr{M}_2}P\{\mathbf{g}_1(\mathbf{Y})\neq(w_0,w_1)\,|\,\text{for transmission of }(w_0,w_1,w_2)\},\\P_{e_2}{}^n&=\frac{1}{M_0M_1M_2}\sum_{w_0,w_1,w_2\in\mathscr{M}_0\times\mathscr{M}_1\times\mathscr{M}_2}P\{\mathbf{g}_2(\mathbf{Z})\neq(w_0,w_2)\,|\,\text{for transmission of }(w_0,w_1,w_2)\}.\end{aligned} \tag{7}$$

We also define the triple of rates $(R_0, R_1, R_2)$ for an $[(M_0, M_1, M_2), n]$-code by the formulas

$$R_0=n^{-1}\log M_0,\ R_1=n^{-1}\log M_1,\ R_2=n^{-1}\log M_2. \tag{8}$$

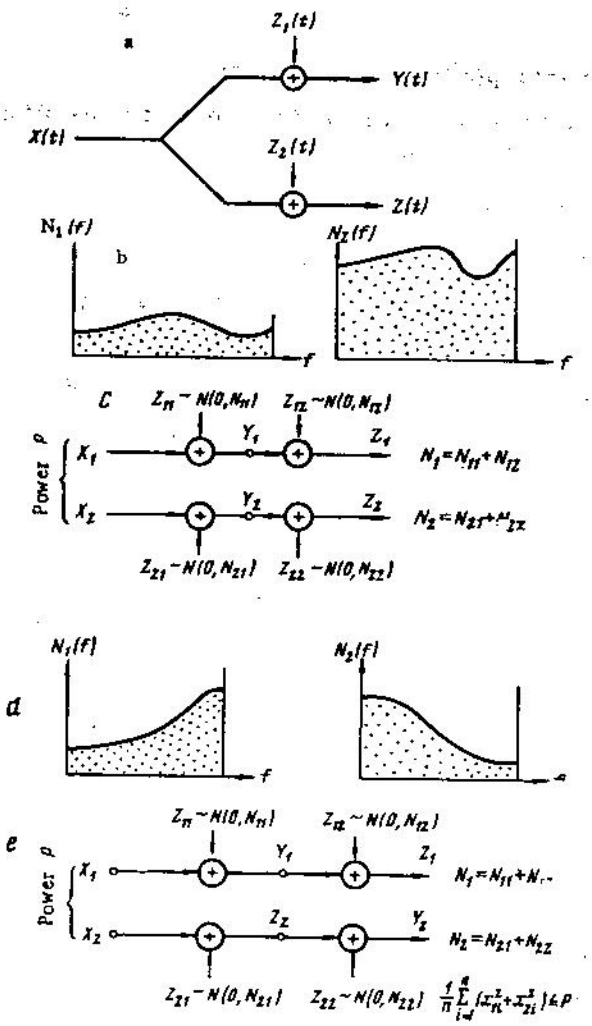


Fig. 2. a) Spectral Gaussian broadcast channel; b) noise spectrum for $N_1(f) < N_2(f)$; c) two-component model with spectral noise for degraded channel [$N_1(f) < N_2(f)$ for all f]; d) undegraded noise spectrum; e) two-component model with spectral noise with unmatched degraded component.

The triple of rates $(R_0, R_1, R_2)$ is called permissible for a broadcast channel if for every $\varepsilon > 0$ and for all sufficiently large n there exists an $[(M_0, M_1, M_2), n]$-code with

$$M_0 \geqslant 2^{nR_0},\ M_1 \geqslant 2^{nR_1},\ M_2 \geqslant 2^{nR_2}, \tag{9}$$

for which $\max\{P^n_{e_1}, P^n_{e_2}\} < \varepsilon$.

The region of the capacity C of a broadcast channel is defined as the set of all permissible triples of rates $(R_0, R_1, R_2)$.

## 2. Capacity of Product of Channels

The capacity $C_\pi$ of the product of two unmatched degraded broadcast channels, defined by formula (1), is given by the following theorem.

THEOREM 1. Assume that $(\mathscr{X}, p(y, z|x), \mathscr{Y}\times\mathscr{Z})$ is the broadcast channel defined in (1), and assume that $U_1$ and $U_2$ are two arbitrary random variables whose number of values $\|\mathscr{U}_i\|$ is bounded by the inequalities

$$\|\mathscr{U}_i\| \leqslant \min(\|\mathscr{X}_i\|, \|\mathscr{Y}_i\|, \|\mathscr{Z}_i\|),\ i=1,\ 2. \tag{10}$$

Then the capacity region is defined as follows:

$$\begin{aligned} C_\pi \triangleq \{(R_0, R_1, R_2) : R_0 &< I(U_1; Y_1) + I(U_2; Y_2), \\ R_0 &< I(U_1; Z_1) + I(U_2; Z_2), \\ R_0 + R_1 &< I(U_2; Y_2) + I(X_1; Y_1), \\ R_0 + R_2 &< I(U_1; Z_1) + I(X_2; Z_2), \\ R_0 + R_1 + R_2 &< I(X_1; Y_1) + I(U_2; Y_2) + I(X_2; Z_2|U_2), \\ R_0 + R_1 + R_2 &< I(X_2, Z_2) + I(U_1; Z_1) + I(X_1; Y_1|U_1),\ \text{for some}\ p\in\mathscr{P}\}, \end{aligned} \tag{11}$$

where $\mathcal{P}$ is the set of probability distributions of the form

$$p(u_1, u_2, x, y, z) = p(u_1)p(u_2)p(x_1|u_1)p(x_2|u_2)p(y, z|x). \tag{12}$$

It is easy to show that for $R_0 \equiv 0$ this region reduces to the region of independent rates obtained by Poltyrev. This is formally contained in the corollary that follows.

COROLLARY (Poltyrev [6]). Let $R_0 \equiv 0$; we set

$$c_1 = \max_{p(x_1)} I(X_1; Y_1), \quad c_2 = \max_{p(x_2)} I(X_2; Y_2), \tag{13}$$

the maximum values being attained on distributions $p^*(x_1)$ and $p^*(x_2)$, respectively. Then, for independent rates, the capacity is given by the following formula:

$$\begin{gathered} C_1 = \{(R_1, R_2) : R_1 < I(U; Y_1) + c_1,\ R_1 + R_2 < I(U; Y_1) + c_1 + I(X_2; Z_2|U) \quad \text{for some} \quad p(u, x, y, z) \\ = p(u)p(x_2|u)p^*(x_1)p(y, z|x)\} \cap \{(R_1, R_2) : R_2 < I(U; Z_1) + c_2,\ R_1 + R_2 < I(U; Z_1) + c_2 \\ + I(X_1; Y_1|U) \quad \text{for some} \quad p(u, x, y, z) = p(u)p(x_1|u)p^*(x_2)p(y, z|x)\}. \end{gathered} \tag{14}$$

It is easy to see that region (14) can be obtained by summing all possible pairs of points from the regions of channel capacities $p(y_1|x_1)p(z_1|y_1)$ and $p(z_2|x_2)p(y_2|z_2)$, respectively.

If $R_1 \equiv R_2 \equiv 0$, we obtain

$$R_0 < \max_{p(x_1)p(x_2)} \min\{I(X_1; Y_1) + I(X_2; Y_2),\ I(X_1; Z_1) + I(X_2; Z_2)\} \triangleq C_0, \tag{15}$$

this being the maximum capacity of the channel.

This capacity region for degraded sets of messages ($R_2 \equiv 0$) is specified by the following formula:

$$\begin{aligned} C_2 = \{(R_0, R_1) : {} & R_0 < I(U; Y_1) + I(X_2; Y_2), \\ & R_0 < I(U; Z_1) + I(X_2; Z_2), \\ & R_0 + R_1 < I(X_1; Y_1) + I(X_2; Y_2), \\ & R_0 + R_1 < I(X_1; Y_1|U) + I(U; Z_1) + I(X_2; Z_2) \\ & \text{for some} \quad p(u, x, y, z) = p(u)p(x_1|u)p(x_2)p(y, z|x)\}. \end{aligned} \tag{16}$$

Remarks Concerning $C_\pi$. 1. It is important to note that, if we admit on $\mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{X}_1 \times \mathcal{X}_2$ more general probability distributions than those described in (12), we do not increase the region as compared to $C_\pi$. The reason for this is that every information quantity in (11) depends either on $p(u_1, x_1)$ or on $p(u_2, x_2)$, but not on the entire distribution $p(u_1, u_2, x_1, x_2)$. 2. $C_\pi$ is a convex region (see Appendix 1). 3. An estimate for the number of values of auxiliary random variables $U_1$ and $U_2$ can be obtained by using the customary procedures (see, e.g., [10]). 4. Assume that $p_{X_1}p_{X_2}$ is a probability distribution on $\mathcal{X}_1 \times \mathcal{X}_2$ that is the product of distributions on $\mathcal{X}_1$ and $\mathcal{X}_2$; we set

$$C_\pi(p_{x_1}p_{x_2}) = \{(R_0, R_1, R_2) : R_0, R_1, R_2 \text{ satisfy the six inequalities in (11) for } p(x_1)p(x_2) = p_{x_1}p_{x_2}\}.$$

Then

$$C_\pi = \bigcup_{p_{x_1} p_{x_2}} C_\pi(p_{x_1}p_{x_2}). \tag{17}$$

Figure 3 shows the region $C_\pi(p_{X_1}p_{X_2})$. The external boundary for $C_\pi(p_{X_1}p_{X_2})$ is a combination of the following three surfaces:

1. Cylindrical surface specified by the equation

$$\begin{aligned} & R_0 + R_1 = I(X_1; Y_1) + I(U_2; Y_2), \\ & R_2 = I(X_2; Z_2|U_2), \\ & R_0 \leqslant I(U_2; Y_2),\ U_1 = \phi. \end{aligned} \tag{18}$$

2. Cylindrical surface specified by the equations

$$\begin{aligned} & R_0 + R_2 = I(X_2; Z_2) + I(U_1; Z_1), \\ & R_1 = I(X_1; Y_1|U_1), \\ & R_0 \leqslant I(U_1; Z_1),\ U_2 = \phi. \end{aligned} \tag{19}$$

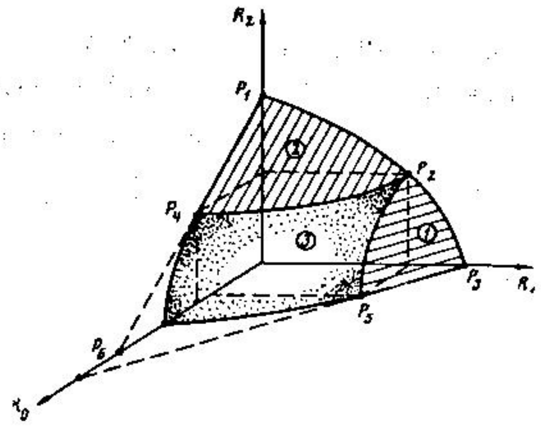


Fig. 3. Region $C_\pi(p_{X_1}p_{X_2})$.

3. Convex surface specified by the equations

$$
\begin{aligned}
&R_1 = I(X_1;\, Y_1 | U_1),\\
&R_2 = I(X_2;\, Z_2 | U_2),\\
&R_0 = \min\,\{I(U_1;\, Y_1) + I(U_2;\, Y_2),\, I(U_1;\, Z_1) + I(U_2;\, Z_2)\}.
\end{aligned} \tag{20}
$$

This reformulation of the assertion of Theorem 1 makes the proof of the direct part of Theorem 1 trivial, since we need only prove that all triples $(R_0, R_1, R_2)$ lying below the boundary of some region $C_\pi(p_{X_1}, p_{X_2})$ are permissible.

To relate the proof of permissibility to Bergmans' coding theorem for a discrete degraded memoryless broadcast channel [1], we note that

1. $U_1$ and $U_2$ are restored by both receivers, and hence can be interpreted as the common information; however, $U_1$ may incorporate particular information for receiver Z, while $U_2$ may incorporate such information for receiver Y.
2. $X_1$ is restored by receiver Y, so that for specified $U_1$ we can interpret $X_1$ as the particular information for Y.
3. $X_2$ is restored by receiver Z, so that for specified $U_2$ we can interpret $X_2$ as the particular information for Z.

Triples of rates on surfaces 1 and 2 can be obtained using one of two degraded channels in a superposition mode, while simultaneously transmitting information over the other channel at the Shannon rate. Triples of rates on surface 3 are obtained using both channels in a superposition mode.

In the next section we will show that this scheme is in fact optimal.

## 3. Converse Part of Theorem 1

The weak converse of Theorem 1 says that if $(R_0, R_1, R_2) \notin C_\pi$, then there exists a $\lambda > 0$ such that for all n we have

$$\max\,\{P^n_{e,1}, P^n_{e,2}\} > \lambda.$$

Proof. Fano's inequality states that

$$P^n_{e,1} < \varepsilon \Rightarrow H(W_1, W_0 | Y_1, Y_2) \leqslant n(R_1 + R_0)\varepsilon + 1 \triangleq n\varepsilon_{1n},$$

$$P^n_{e,2} < \varepsilon \Rightarrow H(W_2, W_0 | Z_1, Z_2) \leqslant n(R_2 + R_0)\varepsilon + 1 \triangleq n\varepsilon_{2n},$$

Now

$$nR_0 = H(W_0) \leqslant I(W_0;\, Y_1, Y_2) + n\varepsilon_{1n} = I(W_0;\, Y_2) + I(W_0;\, Y_1 | Y_2) + n\varepsilon_{1n} \leqslant I(W_0, W_1;\, Y_2) + I(W_0, W_2, Y_2;\, Y_1) + n\varepsilon_{1n}. \tag{21}$$

Similarly,

$$nR_0 \leqslant I(W_0, W_2;\, Z_1) + I(W_0, W_1, Z_1;\, Z_2) + n\varepsilon_{2n}. \tag{22}$$

Furthermore,

$$n(R_0+R_1)=H(W_0)+H(W_1)\leqslant I(W_0, W_1; \mathbf{Y}_1, \mathbf{Y}_2)+n\varepsilon_{1n}$$

$$=I(W_0, W_1; \mathbf{Y}_2)+I(W_0, W_1; \mathbf{Y}_1|\mathbf{Y}_2)+n\varepsilon_{1n}\leqslant I(W_0, W_1; \mathbf{Y}_2)+I(\mathbf{X}_1; \mathbf{Y}_1)+n\varepsilon_{1n}, \tag{23}$$

where the last inequality follows from the fact that we are considering our discrete memoryless channel. Similarly we can show that

$$n(R_0+R_2)\leqslant I(W_0, W_2; \mathbf{Z}_1)+I(\mathbf{X}_2; \mathbf{Z}_2)+n\varepsilon_{2n}. \tag{24}$$

Furthermore, let us consider the triple

$$\begin{aligned}
n(R_0+R_1+R_2)&=H(W_0)+H(W_1)+H(W_2)\leqslant I(W_0, W_1; \mathbf{Y}_1, \mathbf{Y}_2)+I(W_2; \mathbf{Z}_1, \mathbf{Z}_2)+n(\varepsilon_{1n}+\varepsilon_{2n})\\
&\text{1)}\ \leqslant I(W_0, W_1; \mathbf{Y}_1, \mathbf{Y}_2)+I(W_2; \mathbf{Z}_1, \mathbf{Z}_2|W_0, W_1)+n(\varepsilon_{1n}+\varepsilon_{2n})\\
&=I(W_0, W_1; \mathbf{Y}_2)+I(W_0, W_1; \mathbf{Y}_1|\mathbf{Y}_2)+I(W_2; Z_1|W_0, W_1)\\
&\text{2)}\ =[I(W_0, W_1, \mathbf{Z}_1; \mathbf{Y}_2)-I(\mathbf{Z}_1; \mathbf{Y}_2|W_0, W_1)]+I(W_2; \mathbf{Z}_2|W_1, W_0, \mathbf{Z}_1)+n(\varepsilon_{1n}+\varepsilon_{2n})=\\
&+I(W_0, W_1; \mathbf{Y}_1|\mathbf{Y}_2)+I(W_2; \mathbf{Z}_1|W_1, W_0)+I(W_2; \mathbf{Z}_2|W_0, W_1, \mathbf{Z}_1)+n(\varepsilon_{1n}+\varepsilon_{2n})
\end{aligned} \tag{25}$$

$$\begin{aligned}
&\text{3)}\ \leqslant[I(W_0, W_1, \mathbf{Z}_1; \mathbf{Y}_2)+I((W_2; \mathbf{Z}_2|W_1, W_0, \mathbf{Z}_1)]+I(W_0, W_1;\\
&\mathbf{Y}_1|\mathbf{Y}_2)+I(W_2; \mathbf{Z}_1|W_1, W_0, \mathbf{Y}_2)+n(\varepsilon_{1n}+\varepsilon_{2n}),
\end{aligned} \tag{26}$$

where inequalities 1, 2, and 3 are obtained as follows:

1: $I(W_2; \mathbf{Z}_1, \mathbf{Z}_2|W_0, W_1)\geqslant I(W_2, Z_1, Z_2)$,

since $W_0$, $W_1$, and $W_2$ are independent;

2: we add and subtract $I(\mathbf{Z}_1; \mathbf{Y}_2|W_0, W_1)$ to the first term in (25);

3: the inequality

$$I(W_2; \mathbf{Z}_1|W_1, W_0)\leqslant I(W_2; \mathbf{Z}_1, \mathbf{Y}_2|W_1, W_0)$$

and regrouping of terms are employed.

Let us now consider the last two information terms. As a result of the degraded nature of the channel we have

$$\begin{aligned}
&I(W_0, W_1; \mathbf{Y}_1|\mathbf{Y}_2)+I(W_2; \mathbf{Z}_1|W_1, W_0, \mathbf{Y}_2)\leqslant I(W_0, W_1; \mathbf{Y}_1|\mathbf{Y}_2)\\
&+I(W_2; \mathbf{Y}_1|W_1, W_0, \mathbf{Y}_2)=H(\mathbf{Y}_1|\mathbf{Y}_2)-H(\mathbf{Y}_1|W_0, W_1, \mathbf{Y}_2)\\
&+H(\mathbf{Y}_1|W_1, W_0, \mathbf{Y}_2)-H(\mathbf{Y}_1|W_1, W_0, W_2, \mathbf{Y}_2).
\end{aligned} \tag{27}$$

The second and third terms cancel, and we obtain that expression (27) does not exceed

$$H(\mathbf{Y}_1)-H(\mathbf{Y}_1|W_1, W_0, W_2, \mathbf{Y}_2). \tag{28}$$

Since the channel is discrete and memoryless, expression (28) does not exceed

$$H(\mathbf{Y}_1)-H(\mathbf{Y}_1|\mathbf{X}_1)=I(\mathbf{X}_1; \mathbf{Y}_1). \tag{29}$$

Now we obtain from (26) that

$$n(R_0+R_1+R_2)\leqslant[I(W_0, W_1, \mathbf{Z}_1; \mathbf{Y}_2)+I(W_2; \mathbf{Z}_2|W_1, W_0, \mathbf{Z}_1)]+I(\mathbf{X}_1; \mathbf{Y}_1)+n(\varepsilon_{1n}+\varepsilon_{2n}). \tag{30}$$

Similarly, we can show that

$$n(R_0+R_1+R_2)\leqslant[I(W_0, W_2, \mathbf{Y}_2; \mathbf{Z}_1)+I(W_1; \mathbf{Y}_1|W_2, W_0, \mathbf{Y}_2)]+I(\mathbf{X}_2; \mathbf{Z}_2)+n(\varepsilon_{1n}+\varepsilon_{2n}). \tag{31}$$

To bound the right sides in (21)-(24), (30), and (31), we require the following lemma.

<u>LEMMA 1.</u> For every i, $1 \le i \le n$, we set

$$U_{1i}\triangleq(\mathbf{Y}_2, W_2, W_0, Y_1^{i-1})\text{ and }U_{2i}\triangleq(\mathbf{Z}_1, W_1, W_0, Z_2^{i-1}),\text{ where }Y^{i-1}\triangleq(Y_1,\ldots,Y_{i-1}).$$

Then

$$\begin{aligned}
&p(u_{1i}, x_{1i}, y_{1i}, z_{1i})=p(u_{1i})p(x_{1i}|u_{1i})p(y_{1i}, z_{1i}|x_{1i}),\\
&p(u_{2i}, x_{2i}, z_{2i}, y_{2i})=p(u_{2i})p(x_{2i}|u_{2i})p(z_{2i}, y_{2i}|x_{2i})
\end{aligned} \tag{32}$$

and

$$I(W_0, W_2; Z_1) \leqslant \sum_{i=1}^{n} I(U_{1i}; Z_{1i}), \tag{33a}$$

$$(i)\ I(W_0, W_1; Y_2) \leqslant \sum_{i=1}^{n} I(U_{2i}; Y_{2i}), \tag{33b}$$

$$(ii)\ I(W_0, W_1, Y_2; Y_1) \leqslant \sum_{i=1}^{n} I(U_{1i}; Y_{1i}), \tag{34a}$$

$$I(W_0, W_2, Z_1; Z_2) \leqslant \sum_{i=1}^{n} I(U_{2i}; Z_{2i}), \tag{34b}$$

$$(iii)\ I(W_1; Y_1 | W_2, W_0, Y_2) \leqslant \sum_{i=1}^{n} I(X_{1i}; Y_{1i} | U_{1i}), \tag{35a}$$

$$I(W_2; Z_2 | W_1, W_0, Z_1) \leqslant \sum_{i=1}^{n} I(X_{2i}; Z_{2i} | U_{2i}). \tag{35b}$$

Proof. The proof is analogous to that of Lemma 2 in [2] (see Appendix 2 for details). If we apply Lemma 1 directly to (21)-(24), (30), and (31), we obtain that

$$
\begin{gathered}
nR_0 \leqslant \sum_{i=1}^{n} \left(I(U_{1i}; Y_{1i}) + I(U_{2i}; Y_{2i})\right) + n\varepsilon_{1n}, \\
nR_0 \leqslant \sum_{i=1}^{n} \left(I(U_{1i}; Z_{1i}) + I(U_{2i}; Z_{2i})\right) + n\varepsilon_{2n}, \\
n(R_0+R_1) \leqslant \sum_{i=1}^{n} \left(I(U_{2i}; Y_{2i})\right) + I(X_1, Y_1) + n\varepsilon_{1n} \leqslant \sum_{i=1}^{n} \left(I(U_{2i}; Y_{2i}) + I(X_{1i}; Y_{1i})\right) + n\varepsilon_{1n}, \\
n(R_0+R_2) \leqslant \sum_{i=1}^{n} \left(I(U_{1i}; Z_{1i}) + I(X_{2i}; Z_{2i})\right) + n\varepsilon_{2n}, \\
n(R_0+R_1+R_2) \leqslant \sum_{i=1}^{n} \left(I(X_{1i}; Y_{1i}) + I(U_{2i}; Y_{2i}) + I(X_{2i}; Z_{2i} | U_{2i})\right) + n(\varepsilon_{1n}+\varepsilon_{2n}), \\
n(R_0+R_1+R_2) \leqslant \sum_{i=1} \left(I(X_{2i}; Z_{2i}) + I(U_{1i}; Z_{1i}) + I(X_{1i}; Y_{1i} | U_{1i})\right) + n(\varepsilon_{1n}+\varepsilon_{2n}).
\end{gathered} \tag{36}
$$

Now the converse part of Theorem 1 follows directly from the fact that $C_\pi$ is convex.

## 4. Spectral Gaussian Broadcast Channel

We define a spectral Gaussian broadcast channel with two unmatched degraded components in Fig. 2. The following theorem describes the capacity region of this channel.

Assume that $C(P/N) = (1/2)\ln(1 + p/N)$ is the capacity of a channel with additive white Gaussian noise (AWGN), for which the signal-to-noise ratio is $P/N$.

THEOREM 2. The capacity region of the channel in Fig. 2e is specified as follows:

$$
\begin{aligned}
C_G = \{(R_0, R_1, R_2):\ & R_0 \leqslant C(\alpha_1\beta P/(N_{11}+\bar\alpha_1\beta P)) + C(\alpha_2\beta P/(N_2+\bar\alpha_2\beta P)), \\
& R_0 \leqslant C(\alpha_1\beta P/(N_1+\bar\alpha_1\beta P)) + C(\alpha_2\beta P/(N_{21}+\bar\alpha_2\beta P)), \\
& R_0+R_1 \leqslant C(\alpha_2\beta P/(N_2+\bar\alpha_2\beta P)) + C(\beta P/N_{11}), \\
& R_0+R_2 \leqslant C(\alpha_1\beta P/(N_1+\bar\alpha_1\beta P)) + C(\beta P/N_{21}), \\
& R_0+R_1+R_2 \leqslant C(\beta P/N_{11}) + C(\alpha_2\beta P/(N_2+\bar\alpha_2\beta P)) + C(\bar\alpha_2\beta P/N_{21}), \\
& R_0+R_1+R_2 \leqslant C(\beta P/N_{21}) + C(\alpha_1\beta P/(N_1+\bar\alpha_1\beta P)) + C(\bar\alpha_1\beta P/N_{11}) \\
& \text{for some } \alpha_1, \alpha_2, \beta \text{ from the segment } [0, 1]\}.
\end{aligned} \tag{37}
$$

Remarks. 1. Region $C_G$ is convex.

2. For fixed $\alpha_1$, $\alpha_2$, $\beta$ we can achieve any triple $(R_0, R_1, R_2) \in C_G(\alpha_1, \alpha_2, \beta)$ if we first divide the entire power P into the part $\beta P$ used in the first channel and $\bar{\beta} P$ used in the second. Then $\beta P$ is divided into the power $\alpha_1 \beta P$ used to transmit the common information to Y and $\bar{\alpha}_1 \beta P$ used to transmit the particular information to Y; $\bar{\beta} P$ is divided into the power $\alpha_2 \bar{\beta} P$ used to transmit the common information to Z and $\bar{\alpha}_2 \bar{\beta} P$ used to transmit the particular information to Z. The details of the proof of permissibility are standard and will be omitted.

3. In the case of independent rates (i.e., for $R_0 \equiv 0$), region $C_G$ leads to a plane region $C_H$, where

$$\begin{aligned} C_H = \{ & (R_1, R_2): R_1 \leqslant C(\alpha_2 \beta P/(N_2 + \bar{\alpha}_2 \beta P)) + C(\beta P/N_{11}), \\ & R_2 \leqslant C(\alpha_1 \beta P/(N_1 + \bar{\alpha}_1 \beta P)) + C(\beta P/N_{21}), \\ & R_1 + R_2 \leqslant C(\beta P/N_{11}) + C(\alpha_2 \beta P/(N_2 + \bar{\alpha}_2 \beta P)) + C(\bar{\alpha}_2 \beta P/N_{21}), \\ & R_1 + R_2 \leqslant C(\beta P/N_{21}) + C(\alpha_1 \beta P/(N_1 + \bar{\alpha}_1 \beta P)) + C(\bar{\alpha}_1 \beta P/N_{11}), \\ & 0 \leqslant \alpha_1, \alpha_2, \beta \leqslant 1 \}. \end{aligned} \tag{38}$$

It is easy to show that $C_H$ coincides with the Hughes-Hartogs permissible region [4].

Now if we prove the converse, we obtain that Theorem 2 does indeed specify the capacity region.

Converse of Theorem 2. Assume that we are given an arbitrary sequence of codes that leads to the triple of rates $(R_0, R_1, R_2)$. We will show that there exist $\alpha_1$, $\alpha_2$, and $\beta$ such that $(R_0, R_1, R_2)$ satisfy the conditions of Theorem 2. First let us recall Fano's inequality:

$$H(W_0, W_1 | \mathrm{Y}) \leqslant P_{e,1}^n n(R_0 + R_1) + 1 \triangleq n\varepsilon_{1n},$$

$$H(W_0, W_2 | \mathrm{Z}) \leqslant P_{e,2}^n n(R_0 + R_2) + 1 \triangleq n\varepsilon_{2n},$$

Furthermore,

$$\begin{aligned} nR_0 = H(W_0) & \leqslant I(W_0; \mathrm{Y}_1, \mathrm{Y}_2) + n\varepsilon_{1n} \leqslant I(W_0, W_1, \mathrm{Z}_1; \mathrm{Y}_2) + I(W_0, W_2, \mathrm{Y}_2; \mathrm{Y}_1) + n\varepsilon_{1n} \\ & = H(\mathrm{Y}_2) - H(\mathrm{Y}_2 | W_0, W_1, \mathrm{Z}_1) + H(\mathrm{Y}_1) - H(\mathrm{Y}_1 | W_0, W_2, \mathrm{Y}_2) + n\varepsilon_{1n}. \end{aligned} \tag{39}$$

But

$$H(\mathrm{Y}_2) \leqslant (n/2) \ln(\beta P + N_2), \tag{40}$$

where

$$\beta P \triangleq (n/2) \sum_{i=1}^{n} \mathrm{M} X_{2i}^2 \tag{41}$$

and

$$H(\mathrm{Y}_1) \leqslant (n/2) \ln(\bar{\beta} P + N_{11}), \tag{42}$$

where, in view of the constraint on the power,

$$P \geqslant \frac{1}{n} \sum_{i=1}^{n} \mathrm{M}(X_{1i}^2 + X_{2i}^2) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{M} X_{1i}^2 + \beta P. \tag{43}$$

It is easy to see that there exists $\alpha_1$, $\alpha_2 \in [0, 1]$ for which

$$\exp\{(2/n) H(\mathrm{Z}_2 | W_0, W_1, \mathrm{Z}_1)\} = 2\pi e(\alpha_2 \beta P + N_{21}), \tag{44}$$

$$\exp\{(2/n) H(\mathrm{Y}_1 | W_0, W_2, \mathrm{Y}_2)\} = 2\pi e(\alpha_1 \bar{\beta} P + N_{11}). \tag{45}$$

Let us consider the following provisional form of the entropy inequality with the power [11]:

$$\exp\{(2/n) H(\mathrm{Y}_2 | W_0, W_1, \mathrm{Z}_1)\} \geqslant \exp\{(2/n) H(\mathrm{Z}_2 | W_0, W_1, \mathrm{Z}_1)\} + 2\pi e N_{22},$$
$$\exp\{(2/n) H(\mathrm{Z}_1 | W_0, W_2, \mathrm{Y}_2)\} \geqslant \exp\{(2/n) H(\mathrm{Y}_1 | W_0, W_2, \mathrm{Y}_2)\} + 2\pi e N_{11}.$$

From this we have

$$H(\mathrm{Y}_2 | W_0, W_1, \mathrm{Z}_1) \geqslant (n/2) \ln 2\pi e(\alpha_2 \beta P + N_2), \tag{46}$$

$$H(\mathrm{Z}_1 | W_0, W_2, \mathrm{Y}_2) \geqslant (n/2) \ln 2\pi e(\alpha_1 \bar{\beta} P + N_1). \tag{47}$$

Substituting (40), (42), (45), and (46) into (39), we obtain

$$R_0 \leq C\left(\frac{\alpha_1\beta P}{N_{11}+\bar{\alpha}_1\beta P}\right)+C\left(\frac{\alpha_2\beta P}{N_2+\bar{\alpha}_2\beta P}\right)+\varepsilon_{1n}. \quad (48)$$

Similarly, we can show that

$$R_0 \leq C\left(\frac{\alpha_1\beta P}{N_1+\bar{\alpha}_1\beta P}\right)+C\left(\frac{\alpha_2\beta P}{N_{21}+\bar{\alpha}_2\beta P}\right)+\varepsilon_{2n}. \quad (49)$$

Furthermore,

$$\begin{aligned} n(R_0+R_1) &\leq I(W_0, W_1; \mathbf{Y}_1, \mathbf{Y}_2)+n\varepsilon_{1n} \\ 1)\ &\leq I(W_0, W_1, \mathbf{Z}_1; \mathbf{Y}_2)+I(\mathbf{X}_1; \mathbf{Y}_1)+n\varepsilon_{1n} \\ 2)\ &\leq nC\left(\frac{\alpha_2\beta P}{N_2+\bar{\alpha}_2\beta P}\right)+nC\left(\frac{\beta P}{N_{11}}\right)+n\varepsilon_{1n}, \end{aligned} \quad (50)$$

where step 1 was proved in (23), while step 2 follows from the upper bound for $I(W_0, W_1, \mathbf{Z}_1; \mathbf{Y}_2)$ from (40) and (46). Moreover,

$$I(\mathbf{X}_1; \mathbf{Y}_1)=H(\mathbf{Y}_1)-H(\mathbf{Y}_1|\mathbf{X}_1)\leq(n/2)\ \ln 2\pi e(\beta P+N_{11})-(n/2)\ \ln 2\pi eN_{11}=nC(\beta P/N_{11}).$$

Similarly, we can show that

$$R_0+R_2 \leq C\left(\frac{\alpha_1\beta P}{N_1+\bar{\alpha}_1\beta P}\right)+C\left(\frac{\beta P}{N_{21}}\right)+\varepsilon_{2n}. \quad (51)$$

Let us now bound the sum of rates from above. Let us assume that for some $\delta > 0$ we have either

$$R_0+R_1+R_2 > C\left(\frac{\alpha_2\beta P}{N_2+\bar{\alpha}_2\beta P}\right)+C\left(\frac{\beta P}{N_{11}}\right)+C\left(\frac{\bar{\alpha}_2\beta P}{N_{21}}\right)+\delta, \quad (52)$$

or

$$R_0+R_1+R_2 > C\left(\frac{\alpha_1\beta P}{N_1+\bar{\alpha}_1\beta P}\right)+C\left(\frac{\beta P}{N_{21}}\right)+C\left(\frac{\bar{\alpha}_1\beta P}{N_{11}}\right)+\delta. \quad (53)$$

Combining (30) and (50), we obtain that

$$\begin{aligned} n(R_0+R_1+R_2) &\leq nC\left(\frac{\alpha_2\beta P}{N_2+\bar{\alpha}_2\beta P}\right)+nC\left(\frac{\beta P}{N_{11}}\right)+I(W_2; \mathbf{Z}_2|W_0, W_1, \mathbf{Z}_1)+n(\varepsilon_{1n}+\varepsilon_{2n}) \\ &\leq \left(\frac{\alpha_2\beta P}{N_2+\bar{\alpha}_2\beta P}\right)+nC\left(\frac{\beta P}{N_{11}}\right)+H(\mathbf{Z}_2|W_0, W_1, \mathbf{Z}_1)-(n/2)\ln(2\pi eN_{21})+n(\varepsilon_{1n}+\varepsilon_{2n}). \end{aligned} \quad (54)$$

Now if we assume that inequality (52) is satisfied, we obtain that

$$H(\mathbf{Z}_2|W_0, W_1, \mathbf{Z}_1) > (n/2)\ \ln 2\pi e(N_{21}+\bar{\alpha}_2\beta P)-n(\varepsilon_{1n}+\varepsilon_{2n})+n\delta. \quad (55)$$

However, this contradicts (44). Similarly, assuming that (53) is satisfied, we obtain from (31) and (51),

$$H(\mathbf{Y}_1|W_0, W_2, \mathbf{Y}_2) > (n/2)\ \ln 2\pi e(N_{11}+\bar{\alpha}_1\beta P)-n(\varepsilon_{1n}+\varepsilon_{2n})+n\delta, \quad (56)$$

and this contradicts (45). Thus the triple $(R_0, R_1, R_2)$ could satisfy the six inequalities from Theorem 2 for some $\alpha_1$, $\alpha_2$, and $\beta$. The proof of Theorem 2 is now complete.

## 5. Capacity of a Sum of Channels

In the next theorem we will define the capacity region of the sum of two unmatched degraded broadcast channels, which was defined in (2).

THEOREM 3. Assume that $(\mathscr{X}, p(y, z|x), \mathscr{Y}\times\mathscr{Z})$ is the broadcast channel defined by (2), and assume that $U_1$ and $U_2$ are two arbitrary random variables for which the number of values is bounded by inequalities (3). Then the capacity region is defined as follows:

$$\begin{aligned} C_s=\{(R_0, R_1, R_2):\ & R_0<\alpha I(U_1; Y_1)+\bar{\alpha}I(U_2; Y_2)+h(\alpha), \\ & R_0<\alpha I(U_1; Z_1)+\bar{\alpha}I(U_2; Z_2)+h(\alpha), \\ & R_0+R_1<\alpha I(X_1; Y_1)+\bar{\alpha}I(U_2; Y_2)+h(\alpha), \end{aligned} \quad (57)$$

$$R_0+R_2<\alpha I(U_1;Z_1)+\bar\alpha I(X_2;Z_2)+h(\alpha),$$
$$R_0+R_1+R_2<\alpha I(X_1;Y_1)+\bar\alpha[I(U_2;Y_2)+I(X_2;Z_2|U_2)]+h(\alpha),$$
$$R_0+R_1+R_2<\alpha[I(U_1;Z_1)+I(X_1;Y_1|U_1)]+\bar\alpha I(X_2;Z_2)+h(\alpha)$$

for some $p\in\mathcal{P}$ and some $\alpha\in[0,1]$.

Remarks. 1. Region $C_S$ is convex (Appendix 1).

2. If we consider the case $R_1\equiv R_2\equiv 0$, then expression (57) becomes

$$R_0<\max_{\substack{p(x_1)p(x_2),\\ \alpha\in[0,1]}}\min\{\alpha I(X_1;Y_1)+\bar\alpha I(X_2;Y_2)+h(\alpha),\ \alpha I(X_1;Z_1)+\bar\alpha I(X_2;Z_2)+h(\alpha)\}. \tag{58}$$

We should note that there is an analogy between (58) and Shannon's sum of channels [3].

3. By analogy with $C_\pi(p_{X_1}p_{X_2})$, we define $C_S(\alpha,\ p_{X_1}p_{X_2})$ as the set of all triples $(R_0, R_1, R_2)$ that satisfy the six inequalities of Theorem 3 with $p(x_1)=p_{X_1}$, $p(x_2)=p_{X_2}$ for fixed $\alpha$. The outer boundary of $C_S(\alpha,\ p_{X_1}p_{X_2})$ can be represented as the union of the following three surfaces:

i) the cylindrical surface

$$\begin{aligned}&R_0+R_1=\alpha I(X_1;Y_1)+\bar\alpha I(U_2;Y_2)+h(\alpha),\\&R_2=\bar\alpha I(X_2;Z_2|U_2),\\&R_0\leqslant\bar\alpha I(U_2;Y_2)+h(\alpha),\quad U_1=\phi;\end{aligned} \tag{59}$$

ii) the cylindrical surface

$$\begin{aligned}&R_0+R_2=\bar\alpha I(X_2;Z_2)+I(U_1;Z_1)+h(\alpha),\\&R_1=\alpha I(X_1;Y_1|U_1),\\&R_0\leqslant\alpha I(U_1;Z_1)+h(\alpha),\quad U_2=\phi;\end{aligned} \tag{60}$$

iii) the convex surface

$$\begin{aligned}&R_1\leqslant\alpha I(X_1;Y_1|U_1)+h(\alpha),\\&R_2\leqslant\bar\alpha I(X_2;Z_2|U_2)+h(\alpha),\\&R_0+R_1+R_2=\min\{\alpha I(X_1;Y_1)+\bar\alpha[I(U_2;Y_2)+\\&+I(X_2;Z_2|U_2)],\ \bar\alpha I(X_2;Z_2)+\alpha[I(U_1;Z_1)+I(X_1;Y_1|U_1)]\}+h(\alpha),\\&R_0\leqslant\min\{\alpha I(U_1;Y_1)+\bar\alpha I(U_2;Y_2),\ \alpha I(U_1;Z_1)+\bar\alpha I(U_2;Z_2)\}+h(\alpha).\end{aligned} \tag{61}$$

Let us now sketch the proof of attainability of $C_S$. First we fix $\alpha\in[0,1]$ and $p\in\mathcal{P}$.

Random Code. 1. We choose $2^{nh(\alpha)}$ independent identically distributed (IID) sequences $\mathbf{v}\in\{0,1\}^n$ with probability

$$p(\mathbf{v})=\prod_{k=1}^{n}p(v_k), \tag{62}$$

where

$$p(1)=\alpha,\quad p(0)=1-\alpha. \tag{63}$$

Numbering the chosen $\mathbf{v}$, we obtain $\mathbf{v}(i)$, $i\in[1, 2^{nh(\alpha)}]$.

2. For every $\mathbf{v}(i)$ we choose $2^{n(R_0'+R_1'-\beta h(\alpha))}$, $\beta\in[0,1]$, IID sequences $\mathbf{u}_1$ with probability*

$$p(\mathbf{u}_1|\mathbf{v}(i))=\prod_{k=1}^{w(\mathbf{v}(i))}p(u_{1i}) \tag{64}$$

each. Numbering them, we obtain $\mathbf{u}_1(j', i)$, $j'\in[1, 2^{n(R_0'+R_1'-\beta h(\alpha))}]$.

---

*We denote by $w(\mathbf{v})$ the Hamming weight of sequence $\mathbf{v}$, $w(\mathbf{v})=\sum_{l=1}^{n}v_l$.

3. For every $\mathbf{v}(i)$ we choose $2^{n(R_0''+R_2'-\bar\beta h(\alpha))}$ IID sequences $\mathbf{u}_2$ with probability

$$p(\mathbf{u}_2|\mathbf{v}(i)) = \prod_{k=1}^{n-w(\mathbf{v}(i))} p(u_{2k})$$

each. Numbering them, we obtain $\mathbf{u}_2(j'', i)$, $j'' \in [1, 2^{n(R_0''+R_2'-\bar\beta h(\alpha))}]$.

4. For every $\mathbf{u}_1(j', i)$ we choose $2^{nR_1''}$ IID sequences $\mathbf{x}_1$ with probability

$$p(\mathbf{x}_1|\mathbf{u}_1(j', i)) = \prod_{k=1}^{w(\mathbf{v}(i))} p(x_{1k}|u_{1k}(j', i)) \tag{65}$$

each. Numbering them, we obtain $\mathbf{x}_1(m_1, j', i)$, $m_1 \in [1, 2^{nR_1''}]$.

5. For every $\mathbf{u}_2(j'', i)$ we choose $2^{nR_2''}$ IID sequences $\mathbf{x}_2$ with probability

$$p(\mathbf{x}_2|\mathbf{u}_2(j'', i)) = \prod_{k=1}^{n-w(\mathbf{v}(i))} p(x_{2k}|u_{2k}(j'', i)) \tag{66}$$

each. Numbering them, we obtain $\mathbf{x}_2(m_2, j'', i)$, $m_2 \in [1, 2^{nR_2''}]$.

Now we obtain

$$R_0 = R_0' + R_0'', \; R_1 = R_1' + R_1'', \; R_2 = R_2' + R_2''. \tag{67}$$

The code book

$$\{\mathbf{x}(m_1, m_2, j', j'', i) : 1 \leqslant i \leqslant 2^{nh(\alpha)},\; j \leqslant j' \leqslant 2^{n(R_0'+R_1'-\beta h(\alpha))},$$
$$1 \leqslant j'' \leqslant 2^{n(R_0''+R_2'-\bar\beta h(\alpha))},\; 1 \leqslant m_1 \leqslant 2^{nR_1''},\; 1 \leqslant m_2 \leqslant 2^{nR_2''}\}$$

can be obtained from the sequences $\mathbf{x}_1(\cdot)$, $\mathbf{x}_2(\cdot)$ as follows. For every combination $(m_1, m_2, j', j'', i)$ and every $1 \le k \le n$, we set

$$x_k(m_1, m_2, j', j'', i) = \begin{cases} x_{1k_1}, & \text{if} \quad v_k(i) = 1, \\ x_{2k_2}, & \text{if} \quad v_k(i) = 0, \end{cases}$$

where $k_1 = \sum_{l=1}^{k} v_l(i)$, $\quad k_2 = k - k_1$.

Now a direct check establishes that this random code can be used to attain any triple of rates lying on the outer boundary of the region $C_S(\alpha, p_{X_1}p_{X_2})$.

The converse part of Theorem 3 is less obvious than the direct part; it will be taken up in the next section.

## 6. Converse Part of Theorem 3

We will show again that if $(R_0, R_1, R_2) \notin C_S$, then there exists an $\varepsilon > 0$ such that $\max(P_{e,1}^n, P_{e,2}^n) > \varepsilon$ for all n. If code book $\mathscr{B} = \{\mathbf{x}(w_0, w_1, w_2),\ 1 \leqslant w_0 \leqslant 2^{nR_0},\ 1 \leqslant w_1 \leqslant 2^{nR_1},\ 1 \leqslant w_2 \leqslant 2^{nR_2}\}$ with mean error probabilities $(P_{e,1}^n, P_{e,2}^n)$ is specified, then for at least half the code words $\mathbf{x}(w_0, w_1, w_2) \in \mathscr{B}$ we have

$$P\{g_1(\mathbf{y}) \neq (w_0, w_1) \mid \text{was transmitted } (w_0, w_1, w_2)\} \leqslant 2P_{e,1}^n. \tag{68}$$

Thus if

$$\mathscr{C}_1 = \{\mathbf{x}(w_0, w_1, w_2) : \mathbf{x}(w_0, w_1, w_2) \in \mathscr{B} \text{ and condition (68) is satisfied}\}, \tag{69}$$

then $\|\mathscr{C}_1\| \geqslant (1/2)\, 2^{n(R_0+R_1+R_2)}$. Moreover,

$$\frac{1}{\|\mathscr{C}_1\|} \sum_{\substack{(w_0, w_1, w_2): \\ \mathbf{x}(w_0, w_1, w_2) \in \mathscr{C}_1}} p\{g_2(\mathbf{z}) \neq (w_0, w_2) \mid \text{was transmitted}$$
$$(w_0, w_1, w_2)\} \leqslant 2P_{e,2}^n$$

and for half the code words $\mathbf{x}(w_0, w_1, w_2) \in \mathscr{C}_1$ the following condition is satisfied:

$$P\{g_2(\mathbf{z}) \neq (w_0, w_1) \mid \text{was transmitted } (w_0, w_1, w_2)\} \leqslant 4P^*_{e,2}. \tag{70}$$

Therefore, setting

$$\mathscr{C}\{\mathbf{x}(w_0, w_1, w_2) : \mathbf{x}(w_0, w_1, w_2) \in \mathscr{C}_1 \text{ and condition (70), was satisfied}\}, \tag{71}$$

we have

$$\|\mathscr{C}\| \geqslant (1/4)\, 2^{n(R_0+R_1+R_2)}. \tag{72}$$

Now let us set

$$N_1(\mathbf{x}(w_0, w_1, w_2)) = \|\{i: x_i(w_0, w_1, w_2) \in \mathscr{X}_1, 1 \leqslant i \leqslant n\}\|,$$

where $\|A\|$ denotes the cardinality of the set A. The quantity $N_1(\mathbf{x}(w_0, w_1, w_2))$ indicates how many times channel 1 is employed in transmitting code word $\mathbf{x}(w_0, w_1, w_2)$. For $0 \leq n_1 \leq n$ we set

$$\mathscr{C}(n_1) = \{\mathbf{x}(w_0, w_1, w_2) : \mathbf{x}(w_0, w_1, w_2) \in \mathscr{C}, \; N_1(\mathbf{x}(w_0, w_1, w_2)) = n_1\}. \tag{73}$$

Then

$$\mathscr{C} = \bigcup_{n_1=0}^{n} \mathscr{C}(n_1) \tag{74}$$

and

$$\frac{1}{n}\log\|\mathscr{C}\| = \frac{1}{n}\log\sum_{n_1=0} \|\mathscr{C}(n_1)\| \leqslant (1/n)\log(n+1) \max_{0\leqslant n_1 \leqslant n} \|\mathscr{C}(n_1)\|, \tag{75}$$

Let us assume that the maximum in (75) is attained for some $n_1^*$. Then

$$(1/n)\log\|\mathscr{C}\| \leqslant (1/n)\log\|\mathscr{C}(n_1^*)\| + (1/n)\log(n+1) < (1/n)\log\|\mathscr{C}(n_1^*)\| + \delta_n', \tag{76}$$

where

$$\delta_n' = (1/n)\, O(\log n). \tag{77}$$

Now let us consider a new code book $\mathscr{C}(n_1^*)$ of cardinality

$$\|\mathscr{C}(n_1^*)\| > 2^{n(R_0+R_1+R_2-\delta_n)}, \tag{78}$$

where $\delta_n = \delta_n' + 2/n$. This code book $\mathscr{C}$ specifies an empirical distribution on $\mathscr{X}^n$. Assume that $\mathbf{V} = (V_1, \ldots, V_n)$ is a random vector such that for $1 \leq i \leq n$ we have

$$V_i \in \{0, 1\}, \quad V_i = \begin{cases} 1, & \text{if } X_i \in \mathscr{X}_1, \\ 0, & \text{if } X_i \in \mathscr{X}_2, \end{cases} \tag{79}$$

so that $\sum_{i=1}^{n} V_i = n_1^*$. We also define the random vectors $\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Z}_1, \mathbf{Z}_2$ as follows: $\mathbf{X}_1 = (X_{11}, X_{12}, \ldots, X_{1n_1^*})$ is a random vector consisting of components X that belong to $\mathscr{X}_1$, ordered by increasing number. The random vectors $\mathbf{X}_2, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Z}_1, \mathbf{Z}_2$ are similarly defined.

Now let us define the following joint probability distribution on $\mathscr{W}_0 \times \mathscr{W}_1 \times \mathscr{W}_2 \times \mathscr{V}^n \times \mathscr{X}^n \times \mathscr{Y}^n \times \mathscr{Z}^n$:

$$p(w_0, w_1, w_2, \mathbf{v}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1, \mathbf{z}_2)$$
$$= p(w_0, w_1, w_2)\, p(\mathbf{v}, \mathbf{x}_1, \mathbf{x}_2 | w_0, w_1, w_2) \prod_{i=1}^{n_1^*} p(y_{1i}|x_{1i})\, p(z_{1i}|y_{1i}) \prod_{i=1}^{(n-n_1^*)} p(z_{2i}|x_{2i})\, p(y_{2i}|z_{2i}), \tag{80}$$

where

$$p(w_0, w_1, w_2) = \begin{cases} 1/\|\mathscr{C}(n_1^*)\|, & \text{if } \mathbf{x}(w_0, w_1, w_2) \in \mathscr{C}(n_1^*), \\ 0 \text{ otherwise.} \end{cases} \tag{81}$$

Note that all the $(w_0, w_1, w_2)$ for which $\mathbf{x}(w_0, w_1, w_2) \in \mathscr{C}(n_1^*)$ satisfy conditions (68) and (70). Thus, employing Fano's inequality, we have

$$H(W_0, W_1 | \mathbf{Y}) \leqslant 4P^*_{e,1} n(R_0+R_1) + 1 \triangleq n\varepsilon_{1n}, \tag{82a}$$

$$H(W_0, W_2|Z) \leqslant 4P^*_{e,2} n(R_0+R_2)+1 \triangleq n\varepsilon_{2n}. \tag{82b}$$

From (78) and (81) we obtain

$$H(W_0, W_1, W_2) = \|\mathscr{C}(n_1^*)\| > n(R_0+R_1+R_2-\delta_n). \tag{83}$$

But

$$H(W_0) \leqslant nR_0, \quad H(W_1) \leqslant nR_1, \quad H(W_2) \leqslant nR_2, \tag{84}$$

so that

$$\begin{aligned} &n(R_0+R_1+R_2-\delta_n) < H(W_0, W_1, W_2) \leqslant n(R_0+R_1+R_2), \\ &H(W_0) \geqslant n(R_0-\delta_n), \quad H(W_1|W_0, W_2) \geqslant n(R_1-\delta_n), \\ &H(W_2|W_0, W_1) \geqslant n(R_2-\delta_n), \quad H(W_0, W_1) \geqslant n(R_0+R_1-\delta_n), \\ &H(W_0, W_2) \geqslant n(R_0+R_2-\delta_n). \end{aligned} \tag{85}$$

Now we will proceed as in the proof of the converse in Theorem 1. We have

$$\begin{aligned} &i)\ nR_0 \leqslant H(W_0)+n\delta_n = I(W_0; \mathbf{V}, \mathbf{Y})+n(\varepsilon_{1n}+\delta_n) = I(W_0; \mathbf{V}) \\ &+I(W_0; \mathbf{Y}_1, \mathbf{Y}_2|\mathbf{V})+n(\varepsilon_{1n}+\delta_n) \leqslant H(\mathbf{V})+I(W_0; \mathbf{Y}_2|\mathbf{V}) \\ &+I(W_0; \mathbf{Y}_1|\mathbf{V}, \mathbf{Y}_2)+n(\varepsilon_{1n}+\delta_n) \leqslant H(\mathbf{V})+I(W_0, W_1, \mathbf{V}; \mathbf{Y}_2) \\ &+I(W_0, W_2, \mathbf{Y}_2, \mathbf{V}; \mathbf{Y}_1)+n(\varepsilon_{1n}+\delta_n). \end{aligned} \tag{86}$$

Moreover,

$$nR_0 \leqslant H(\mathbf{V})+I(W_0, W_1, \mathbf{V}; \mathbf{Z}_1)+I(W_0, W_2, \mathbf{Z}_1, \mathbf{V}; \mathbf{Z}_2)+n(\varepsilon_{2n}+\delta_n). \tag{87}$$

Furthermore,

$$\begin{aligned} &ii)\ n(R_0+R_1) \leqslant I(W_0, W_1; \mathbf{Y})+n(\varepsilon_{1n}+\delta_n) = I(W_0, W_1; \mathbf{V}) \\ &+I(W_0, W_1; \mathbf{Y}_1, \mathbf{Y}_2|\mathbf{V})+n(\varepsilon_{1n}+\delta_n) \leqslant H(\mathbf{V})+I(W_0, W_1, \mathbf{V}; \mathbf{Y}_2) \\ &+I(\mathbf{X}_1; \mathbf{Y}_1)+n(\varepsilon_{1n}+\delta_n), \end{aligned} \tag{88}$$

and

$$n(R_0+R_2) \leqslant H(\mathbf{V})+I(W_0, W_2, \mathbf{V}; \mathbf{Z}_1)+I(\mathbf{X}_2; \mathbf{Z}_2)+n(\varepsilon_{1n}+\delta_n). \tag{89}$$

Finally,

$$\begin{aligned} &iii)\ n(R_0+R_1+R_2) \leqslant H(W_0, W_1)+H(W_2|W_0, W_1)+n(\varepsilon_{1n}+\varepsilon_{2n} \\ &+2\delta_n) \leqslant I(W_0, W_1; \mathbf{Y})+I(W_2; \mathbf{Z}|W_0, W_1)+n(\varepsilon_{1n}+\varepsilon_{2n}+2\delta_n) \\ &\leqslant I(W_0, W_1; \mathbf{V})+I(W_0, W_1; \mathbf{Y}_1, \mathbf{Y}_2|\mathbf{V})+I(W_2; \mathbf{V}|W_0, W_1) \\ &+I(W_2; \mathbf{Z}_1, \mathbf{Z}_2|W_0, W_1, \mathbf{V})+n(\varepsilon_{1n}+\varepsilon_{2n}+2\delta_n) \leqslant H(\mathbf{V}) \\ &+I(W_0, W_1; \mathbf{Y}_1, \mathbf{Y}_2|\mathbf{V})+I(W_2; \mathbf{Z}_1, \mathbf{Z}_2|W_0, W_1, \mathbf{V})+n(\varepsilon_{1n}+\varepsilon_{2n}+2\delta_n). \end{aligned} \tag{90}$$

Similarly to (30), we can readily show that

$$n(R_0+R_1+R_2) \leqslant H(\mathbf{V})+[I(W_0, W_1, \mathbf{Z}_1, \mathbf{V}; \mathbf{Y}_2)+I(W_2; \mathbf{Z}_2|W_1, W_0, \mathbf{Z}_1, \mathbf{V})]+I(\mathbf{X}_1; \mathbf{Y}_1)+n(\varepsilon_{1n}+\varepsilon_{2n}+2\delta_n). \tag{91}$$

Moreover,

$$n(R_0+R_1+R_2) \leqslant H(\mathbf{V})+[I(W_0, W_2, \mathbf{Y}_2, \mathbf{V}; \mathbf{Z}_1)+I(W_1; \mathbf{Y}_1|W_2, W_0, \mathbf{Y}_2, \mathbf{V})]+I(\mathbf{X}_2; \mathbf{Z}_2)+n(\varepsilon_{1n}+\varepsilon_{2n}+2\delta_n). \tag{92}$$

Now we require an analog of Lemma 1 in order to bound the right sides in (86)-(92).

LEMMA 2. We set $U_{1i} \triangleq (\mathbf{Y}_2, W_2, W_0, Y_1^{i-1}, \mathbf{V})$, $1 \leqslant i \leqslant n_1^*$;

$$U_{2i} \triangleq (\mathbf{Z}_1, W_1, W_0, Z_2^{i-1}, \mathbf{V}),\ 1 \leqslant i \leqslant (n-n_1^*).$$

Then

$$p(u_{1i}, x_{1i}, y_{1i}, z_{1i}) = p(u_{1i})\,p(x_{1i}|u_{1i})\,p(y_{1i}, z_{1i}|x_{1i}),$$

$$p(u_{2i}, x_{2i}, y_{2i}, z_{2i}) = p(u_{2i})\,p(x_{2i}|u_{2i})\,p(y_{2i}, z_{2i}|x_{2i})$$

and

$$\text{i)}\quad I(W_0, W_1, \mathbf{V}; \mathbf{Y}_2) \leqslant \sum_{i=1}^{n-n_1^*} I(U_{2i}; Y_{2i}),$$

$$I(W_0, W_2, \mathbf{V}; \mathbf{Z}_2) \leqslant \sum_{i=1}^{n_1^*} I(U_{1i}; Z_{1i}),$$

ii) $I(W_0, W_1, \mathbf{Y}_2, \mathbf{V}; \mathbf{Y}_1) \leqslant \sum_{i=1}^{n_1^*} I(U_{1i}; Y_{1i})$,

$$I(W_0, W_2, \mathbf{Z}_1, \mathbf{V}; \mathbf{Z}_2) \leqslant \sum_{i=1}^{n-n_1^*} I(U_{2i}; Z_{2i});$$

iii) $I(W_1; \mathbf{Y}_1 | W_2, W_0, \mathbf{Y}_2, \mathbf{V}) \leqslant \sum_{i=1}^{n_1^*} I(X_{1i}; Y_{1i} | U_{1i})$,

$$I(W_2; \mathbf{Z}_2 | W_1, W_0, \mathbf{Z}_1, \mathbf{V}) \leqslant \sum_{i=1}^{n-n_1^*} I(X_{2i}; Z_{2i} | U_{2i}).$$

The proof of this lemma is analogous to that of Lemma 1, and will therefore be omitted.

Now we will bound H(V) from above. Using standard reasoning involving complex functions, we obtain

$$H(\mathbf{V}) \leqslant \sum_{i=1}^{n} H(V_i) = \sum_{i=1}^{n} h(\mathbf{M}V_i) = n\left(\frac{1}{n}\sum_{i=1}^{n} h(\mathbf{M}V_i)\right) \leqslant nh(n_1^*/n), \tag{93}$$

since $n_1^* = \sum_{i=1} V_i = \mathbf{M}\sum_{i=1} V_i$. Combining (86)–(92), Lemma 2, and (93), and using the fact that $C_S$ is convex, we obtain that there exists an $\alpha = n_1^*/n$ and random variables $U_1$ and $U_2$ such that $(R_0, R_1, R_2)$ satisfies the inequalities of Theorem 3.

## 7. Conclusions

In this paper we have obtained the capacity of the sum and product of degraded broadcast channels. The product of channels is a model for a channel with spectral noise. The sum of channels is of interest in terms of its mathematical duality with the product of channels. The capacity regions for both channels have a relatively similar description. The proofs of the converses in Theorems 1-3 are comparatively similar. In the case of a Gaussian spectral channel, the entropy power inequality is used to bound the entropy of a noisy channel. In the case of a sum of channels, we reduce the volume of the specified code book by a factor of n, obtaining a code that is not a product and has an exponential volume as before. Then we apply Fano's inequality and other bounds to the resultant code, these yielding bounds for the rate of the original code.

The author wishes to thank T. Cover and J. van Kampenhout for discussions in the course of writing the paper.

## APPENDIX 1

Convexity of $C_\pi$. Assume that $(U_{1i}, U_{2i}, X_{1i}, X_{2i}, Y_{1i}, Y_{2i}, Z_{1i}, Z_{2i})$, $i = 1, 2$, are two sets of random variables whose distributions lie in $\mathcal{P}$. Assume, moreover, that T is a random variable that assumes the values 1 and 2 with probabilities $\alpha$ and $\bar{\alpha}$, respectively. We set

$$U_{1T} = U_{1i} \quad U_{2T} = U_{2i}, \quad X_1 = X_{1i}, \quad X_2 = X_{2i},$$
$$Y_1 = Y_{1i}, \quad Y_2 = Y_{2i}, \quad Z_1 = Z_{1i}, \quad Z_2 = Z_{2i}$$

for T = i.

Then the sequences

$$(T, U_{1T}) \to X_1 \to Y_1 \to Z_1$$

and

$$(T, U_{2T}) \to X_2 \to Y_2 \to Z_2$$

form two Markov chains in the order indicated. We have

$$\begin{aligned} &\alpha I(U_{11}; Y_{11}) + \bar{\alpha} I(U_{12}; Y_{12}) + \alpha I(U_{21}; Y_{21}) + \bar{\alpha} I(U_{22}; Y_{22}) \\ &= I(U_{1T}; Y_1 | T) + I(U_{2T}; Y_2 | T) \leqslant I(T, U_{1T}; Y_1) + I(T, U_{2T}; Y_2) \\ &= I(U_1; Y_1) + I(U_2; Y_2), \end{aligned}$$

where the distribution $U_i = (T_i, U_{iT_i})$ coincides with $(T_i, U_{iT})$, $i = 1, 2$, while $T_1$ and $T_2$ are independent. We should note that we have employed Remark 1 from Sec. 2 here.

Similarly we have

$$\begin{aligned}&\alpha I(U_{21}; Y_{21})+\bar{\alpha} I(U_{22}; Y_{22})+\alpha I(X_{11}; Y_{11})+\bar{\alpha} I(X_{12}; Y_{12})\\&=I(U_{2T}; Y_2|T)+I(X_1; Y_1|T)\leqslant I(U_{2T}, T; Y_2)+I(X_1; Y_1)\\&=I(U_2; Y_2)+I(X_1; Y_1).\end{aligned}$$

Finally,

$$\begin{aligned}&\alpha I(X_{11}; Y_{11})+\bar{\alpha} I(X_{12}; Y_{12})+\alpha I(U_{21}; Y_{21})+\bar{\alpha} I(U_{22}; Y_{22})\\&+\alpha I(X_{21}; Z_{21}|U_{21})+\bar{\alpha} I(X_{22}; Z_{22}|U)=I(X_1; Y_1|T)+I(U_{2T}; Y_2|T)\\&+I(X_2; Z_2|U_{2T}, T)\leqslant I(X_1; Y_1)+I(U_2; Y_2)+I(X_2; Z_2|U_2).\end{aligned}$$

## APPENDIX 2

Proof of Lemma 1. We will prove only (33a) and (35a). The other two inequalities can be proved similarly. We have

$$I(W_0, W_1; \mathbf{Y}_2)=\sum_{i=1}^{n} I(W_0, W_1; Y_{2i}|Y_2^{i-1})$$

$$\leqslant \sum_{i=1}^{n} I(W_0, W_1, Y_2^{i-1}, \mathrm{Z}_i; Y_{2i}) \leqslant \sum_{i=1}^{n} I(W_0, W_1, Z_2^{i-1}, \mathrm{Z}_i; Y_{2i})=\sum_{i=1}^{n} I(U_{2i}; Y_{2i}).$$

Now let us consider (35a). We have

$$I(W_1; \mathbf{Y}_1|W_0, W_2, \mathbf{Y}_2)=\sum_{i=1}^{n} I(W_1; Y_{1i}|W_0, W_2, \mathbf{Y}_2, Y_1^{i-1}) \leqslant \sum_{i=1}^{n} I(X_{1i}; Y_{1i}|U_{1i}).$$

To prove Markov property (32) we write the joint distribution of the channel:

$$p(w_0, w_1, w_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1, \mathbf{z}_2)=p(w_0, w_1, w_2)\ p(\mathbf{x}_1, \mathbf{x}_2|w_0, w_1, w_2) \sum_{i=1}^{n} p(y_{1i}|x_{1i})p(z_{1i}|y_{1i})p(z_{2i}|x_{2i})p(y_{2i}|z_{2i}).$$

Summing first over $w_1$ and $\mathbf{z}_2$, we obtain

$$p(w_0, w_2, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1)=p(w_0, w_2)p(\mathbf{x}_1, \mathbf{x}_2|w_0, w_2)\prod_{i=1}^{n} p(y_{1i}|x_{1i})p(z_{1i}|y_{1i})p(y_{2i}|x_{2i})$$

$$=p(w_0, w_2)p(\mathbf{x}_1|w_0, w_2)p(\mathbf{x}_2|w_0, w_2, \mathbf{x}_1)p(\mathbf{y}_2|\mathbf{x}_2)\prod_{i=1}^{n} p(y_{1i}|x_{1i})p(z_{1i}|y_{1i}).$$

Now summing over $\mathbf{x}_2$, we obtain

$$p(w_0, w_2, \mathbf{x}_1, \mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1)=p(w_0, w_1)p(\mathbf{y}_2, \mathbf{x}_1|w_0, w_2)\prod_{i=1}^{n} p(y_{1i}|x_{1i})p(z_{1i}|y_{1i}).$$

Now we sum over $(y_{1j+1}, y_{1n})$ and over all $z_{1i}$ except for $z_{1j}$. We obtain

$$p(w_0, w_2, \mathbf{x}_1, y_{11}, \ldots, y_{1j}, \mathbf{y}_2, z_{1j})=p(w_0, w_2)\ p(\mathbf{y}_2, \mathbf{x}_1|w_0, w_2)p(y_{1j}|x_{1j})p(z_{1j}|y_{1j})\prod_{i=1}^{j-1} p(y_{1i}|x_{1i}).$$

Finally, summing over all $x_{1i}$ except for $x_{1j}$, we obtain

$$p(w_0, w_2, \mathbf{y}_2, y_1^{j-1}, x_{1j}, y_{1j}, z_{1j})=p(w_0, w_2)p(\mathbf{y}_2, x_{1j}, y_1^{j-1}|w_0, w_1)p(y_{1j}|x_{1j})p(z_{1j}|y_{1j}).$$

## APPENDIX 3

Convexity of $C_S$. The proof is analogous to what is contained in Appendix 1, so we will point out only the additional necessary steps. Let $\beta \in [0, 1]$. Consider the following expression:

$$\beta\alpha_1 I(U_{11};Y_{11})+\beta\alpha_2 I(U_{12};Y_{12})+\beta h(\alpha_1)+\beta h(\alpha_2)$$

$$=(\beta\alpha_1+\beta\alpha_2)\left[\frac{\beta\alpha_1}{(\beta\alpha_1+\beta\alpha_2)}I(U_{11};Y_{11})+\frac{\beta\alpha_2}{(\beta\alpha_1+\beta\alpha_2)}I(U_{12};Y_{12})\right]+\beta h(\alpha_1)+\beta h(\alpha_2). \qquad (A3.1)$$

We set $\alpha \overset{\triangle}{=} \beta\alpha_1+\beta\alpha_2$ and define T as before, but with $P(T=1)=\beta\alpha_1/(\beta\alpha_1+\bar{\beta}\alpha_2)$. Then the left side in (A3.1) does not exceed

$$\alpha I(U_{1T};Y_1|T)+h(\alpha)\leqslant\alpha I(U_{1T},T;Y_1)+h(\alpha).$$

## LITERATURE CITED


1. P. Bergmans, "Coding theorem for broadcast channels with degraded components," IEEE Trans. Inf. Theory, 19, No. 2, 197-207 (1973).
2. R. G. Gallager, "Capacity and coding for some broadcast channels," Probl. Peredachi Inf., 10, No. 3, 3-14 (1974).
3. C. E. Shannon, "Certain results in coding theory for noisy channels," Inf. Control, 1, No. 1, 6-25 (1957).
4. D. Hughes-Hartogs, "The capacity of a degraded spectral Gaussian broadcast channel," Technical Report No. 7002-2, Information Systems Laboratory, Center for Systems Research, Stanford University, July 1975.
5. E. van der Meulen, "A survey of multiway channels in information theory," IEEE Trans. Inf. Theory, 23, No. 1, 1-37 (1977).
6. G. Sh. Poltyrev, "Capacity for parallel broadcast channels with degraded components," Probl. Peredachi Inf., 13, No. 2, 23-35 (1977).
7. T. M. Cover, "Open problems in information theory," Proc. of the 1975 IEEE-USSR Joint Workshop on Information Theory, Moscow, USSR, December, 1975, IEEE Press, New York (1976), pp. 35-36.
8. M. S. Pinsker and G. Sh. Poltyrev, "Capacity of a Gaussian broadcast channel with continuous time," in: Proc. Seventh All-Union Symposium on Problems of Redundancy in Information Systems [in Russian], Vol. 1, Leningrad (1977), pp. 42-46.
9. G. Sh. Poltyrev, "Capacity for a sum of certain broadcast channels," Probl. Peredachi Inf., 15, No. 2, 40-44 (1979).
10. R. F. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channels," IEEE Trans. Inf. Theory, 21, No. 6, 629-637 (1975).
11. P. Bergmans, "A simple converse for broadcast channels with additive Gaussian noises," IEEE Trans. Inf. Theory, 20, No. 2, 279-280 (1974).