Cooperative Broadcasting

PATRICK P. BERGMANS, MEMBER, IEEE, AND THOMAS M. COVER, FELLOW, IEEE

***Abstract*—This paper shows that several transmitters operating in an additive white Gaussian noise environment can send at rates strictly dominating time-multiplex and frequency-multiplex rates by use of a superposition scheme that pools the time, bandwidth, and power allocations of the transmitters. This pooling can be achieved without cooperative action, except for agreement on the actual rate of transmission each transmitter will allow itself. The superposition scheme involves subtraction from the received signal of the estimated signals sent by the other transmitters, followed by decoding of the intended signal. This scheme has been shown to be optimal. We conclude that present methods of allocating different frequency bands to different transmitters are necessarily suboptimal.**

I. Introduction

CONSIDER two radio transmitters with total allotted power P and total available bandwidth W. Suppose that transmitter i, $i = 1,2$, is transmitting to receiver i in the presence of additive white Gaussian noise of one-sided power spectral density N_i, $N_1 \leq N_2$. The capacity of each channel operating *alone* (utilizing all of the available power P and the total bandwidth W) is then given by (Shannon [1], [2])

$$C_i = W \ln \left(1 + \frac{P}{N_i W}\right), \qquad \text{nats/s, } i = 1,2. \tag{1}$$

However, if one transmitter uses all the power and all the bandwidth, the other channel is being used at zero rate. We ask what set of rates (R_1,R_2) are simultaneously achievable. Clearly, the previous comments imply that $(C_1,0)$ and $(0,C_2)$ are achievable.

The first logical candidate for a scheme of channel sharing is that of time sharing. That is, channel 1 is used at full power P and full bandwidth W, a proportion τ_1 of the time, and channel 2 a proportion $\tau_2 = 1 - \tau_1$ of the time. This allows any rate pair

$$R_1 = \tau_1 C_1$$
$$R_2 = \tau_2 C_2 \tag{2}$$

to be achieved, as shown in Fig. 1. We shall refer to this as the naive time-sharing scheme, as opposed to the variable-power time-sharing scheme we will discuss in Section IV.

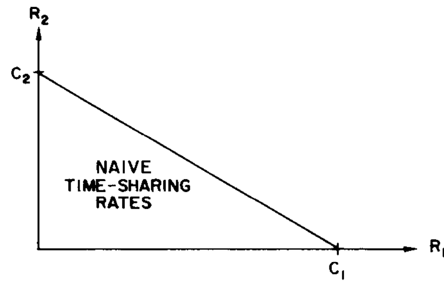

Fig. 1. Rates achievable by naive time sharing.

The second logical approach is the standard frequency division or band allocation approach. Here, let transmitter i operate with power P_i in a band of width W_i, $P_1 + P_2 = P$, $W_1 + W_2 = W$. The set of achievable rates is then given by

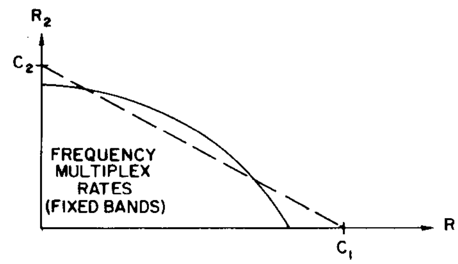

Fig. 2. Rates achievable for fixed frequency division and variable power allocation.

$$R_1 = W_1 \ln \left(1 + \frac{P_1}{N_1 W_1}\right)$$
$$R_2 = W_2 \ln \left(1 + \frac{P_2}{N_2 W_2}\right) \tag{3}$$

where P_i, W_i vary within the constraints just given. Fig. 2 depicts the set of (R_1,R_2) generated for a fixed frequency allocation W_1, W_2 as the power $P_1 = P - P_2$ is allowed to vary in $0 \leq P_1 \leq P$. As will be proved in Section II, allocation of power proportional to bandwidth achieves a point on the naive time-sharing line. Also, a certain range of power allocations dominates naive time sharing. Finally, the envelope of these curves, as $W_1 = W - W_2$ varies over $0 < W_1 < W$, strictly dominates the naive time-sharing bound as shown in Fig. 3. We show this in Section IV.

We now turn our attention to superposition schemes in which the transmitters make independent use of the entire band allocated for both.

The obvious decoding scheme, which we shall call naive superposition, decodes each signal separately as if the other signals were noise with respect to it. This yields rates

$$\tilde{R}_1 = W \ln \left(1 + \frac{P_1}{WN_1 + P_2}\right)$$
$$\tilde{R}_2 = W \ln \left(1 + \frac{P_2}{WN_2 + P_1}\right) \tag{4}$$

Manuscript received February 26, 1973; revised December 5, 1973. This work was supported in part by Stanford University, Stanford, Calif., under Contract AF-49(638)1517 and the National Science Foundation under Grants GK-34363 and GK-32240. Portions of this work were done at the Massachusetts Institute of Technology, Cambridge, Mass., and Harvard University, Cambridge, Mass.
P. P. Bergmans is with Cornell University, Ithaca, N.Y. 14850.
T. M. Cover is with Stanford University, Stanford, Calif. 94305.

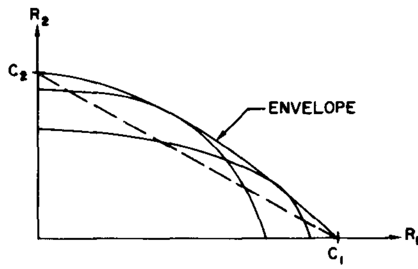

Fig. 3. Rates achievable by frequency division and variable power time sharing.

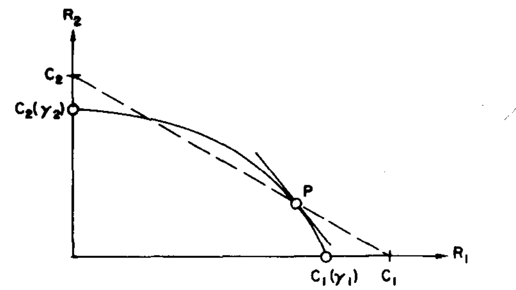

Fig. 4. Dominance of frequency division over naive time division.

as achievable for power allocation (P_1,P_2), $P_1 + P_2 = P$. See, for example, Stefanyuk [3], [4] and Costas [5]. However, this decoding scheme is strictly dominated by a scheme that subtracts out estimates of the other signal. Thus we are led to superposition codes achieving rates

$$R_1^* = W \ln \left(1 + \frac{P_1}{WN_1}\right)$$

$$R_2^* = W \ln \left(1 + \frac{P_2}{WN_2 + P_1}\right) \tag{5}$$

for $N_1 < N_2$. Note that (5) dominates (4) in the sense that $R_2^* = \tilde{R}_2$ and $R_1^* > \tilde{R}_1$. The achievability of this rate pair is shown in Cover [6] and is a special case of the subsequent results on continuous degraded channels in Bergmans [7], [8]. The optimality is proved in Bergmans [10].

II. The Achievability of the Superposition Bound

A heuristic motivation of the achievability of (5) is now given. Since $N_1 < N_2$, user 1 can also receive correctly all the information transmitted to user 2. Consequently, user 1 can subtract out that information from the received message, and decode its own information as if no transmission to user 2 was present. This justifies the expression for R_1^*. User 2, however, cannot receive all the information intended for user 1, since $N_2 > N_1$. The power P_1 of the first transmission is only noise to it, and must be added to the noise power WN_2 in its own channel. Hence the expression for R_2^*. A more thorough discussion can be found in [6]–[8].

III. Applications

We envision the use of this idea in several ways. For example, suppose that a transmitter wishes to send a stereo audio signal if the reception is good (i.e., the noise power is N_1), and wishes to send monaural if the reception is poor (i.e., the noise power is $N_2 > N_1$). Then, if the noise power is less than or equal to N_1, stereo information can be received at the rate $R_1^* + R_2^*$ given in (5). If, however, the noise power is between N_1 and N_2, only monaural information will be received—at rate R_2^*. Finally, if the noise power exceeds N_2, no information is received.

Another example is that of TV information transmission, in which black-and-white information can be sent at rate R_2^*, and color can be sent at the rate $R_1^* + R_2^*$, in the noise regions where $N_1 \le N \le N_2$ and $N \le N_1$, respectively.

As a final example, suppose that a deep-space probe must transmit all of its information before it plunges into some foreign planet. Suppose also that there is some probability π_1 that the additive Gaussian noise will be low ($N = N_1$) and some probability $\pi_2 = 1 - \pi_1$ that the noise will be high ($N = N_2 > N_1$). Then (5) implies that an expected rate $\pi_1 R_1^* + R_2^*$ can be achieved.

In all three examples just cited, these rates are strictly greater than the theoretically achievable rates achievable by frequency band allocation.

IV. Frequency Multiplexing

In this section, we formalize the results on frequency multiplexing presented in the introduction. Let $\gamma_i = W_i/W$, $i = 1,2$, be the fraction of the total bandwidth allocated to each transmitter, and $\alpha_i = P_i/P$, $i = 1,2$, the fraction of the total power. Let $\boldsymbol{\alpha} = (\alpha_1,\alpha_2)$ and $\gamma = (\gamma_1,\gamma_2)$ be two points in $\mathscr{S}_2 = \{(s_1,s_2): s_i \ge 0, s_1 + s_2 = 1\}$. All rate points of the form

$$R_1 = \gamma_1 W \ln \left(1 + \frac{\alpha_1 P}{\gamma_1 W N_1}\right)$$

$$R_2 = \gamma_2 W \ln \left(1 + \frac{\alpha_2 P}{\gamma_2 W N_2}\right) \tag{6}$$

can be achieved by simple frequency multiplexing. Without loss of generality, we can set P and W equal to 1, to simplify the notation. This is really equivalent to substituting R_i for R_i/W and N_i for $N_i W/P$. The capacity C_i of each channel is now given by $C_i = \ln(1 + 1/N_i)$, and (6) becomes

$$R_1 = \gamma_1 \ln \left(1 + \frac{\alpha_1}{\gamma_1 N_1}\right) = R_1(\gamma_1,\alpha_1)$$

$$R_2 = \gamma_2 \ln \left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right) = R_2(\gamma_2,\alpha_2) \tag{7}$$

with

$$\boldsymbol{\alpha},\gamma \in \mathscr{S}_2 = \{(s_1,s_2): s_i \ge 0, s_1 + s_2 = 1\}.$$

By varying the power distribution between transmitters 1 and 2, we generate a curve $C(\gamma)$ as illustrated in Fig. 4. The curve intersects the axes at the points $(C_1(\gamma_1),0)$ and $(0,C_2(\gamma_2))$, corresponding to the cases where, with the given bandwidth partition, the full power is allocated to transmitter 1 or transmitter 2. The "reduced-bandwidth" capacities $C_1(\gamma_1)$ and $C_2(\gamma_2)$ are defined by

$$C_1(\gamma_1) = R_1(\gamma_1,1) = \gamma_1 \ln \left(1 + \frac{1}{\gamma_1 N_1}\right)$$

$$C_2(\gamma_2) = R_2(\gamma_2,1) = \gamma_2 \ln\left(1 + \frac{1}{\gamma_2 N_2}\right). \tag{8}$$

$C_i(\gamma_i)$ is an increasing function of γ_i, and is equal to C_i for $\gamma_i = 1$.

We now prove the following propositions.

Proposition 1: A point on the naive time-sharing line is achieved by letting the power allocation be proportional to the bandwidth allocation. This point is labeled P on Fig. 4.

Proposition 2: Points above the naive time-sharing line are achieved for a certain range of power allocations.

To prove Propositions 1 and 2, we define

$$e = \frac{R_1}{C_1} + \frac{R_2}{C_2} - 1. \tag{9}$$

The quantity e is proportional to the distance from a given rate point (R_1,R_2) to the time-sharing line; e is positive for points above the time-sharing line, and negative for points below it.

Let $\boldsymbol{\alpha} = \boldsymbol{\gamma}$. Thus

$$\begin{aligned} e &= \frac{R_1(\gamma_1,\gamma_1)}{C_1} + \frac{R_2(\gamma_2,\gamma_2)}{C_2} - 1 \\ &= \frac{\gamma_1 \ln\left(1 + \frac{1}{N_1}\right)}{\ln\left(1 + \frac{1}{N_1}\right)} + \frac{\gamma_2 \ln\left(1 + \frac{1}{N_2}\right)}{\ln\left(1 + \frac{1}{N_2}\right)} - 1 \\ &= \gamma_1 + \gamma_2 - 1 = 0. \end{aligned} \tag{10}$$

Hence, the point $(R_1(\gamma_1,\gamma_1),R_2(\gamma_2,\gamma_2))$ is on the naive time-sharing line.

To prove that there are points represented by (7) above the naive time-sharing line, we first prove that $e(\gamma)$ is concave. We then differentiate e with respect to α_1 and α_2, subject to the constraint $\alpha_1 + \alpha_2 = 1$. We shall find that the distance e is not maximized for the rate point achieving time-sharing rates, and that there must be a range of the $\boldsymbol{\alpha}$ giving points above the naive time-sharing line.

We have

$$\begin{aligned} \frac{dR_1}{d\alpha_1} &= \frac{\gamma_1}{1 + \alpha_1/\gamma_1 N_1} \cdot \frac{1}{\gamma_1 N_1} = \frac{\gamma_1}{\gamma_1 N_1 + \alpha_1} \\ \frac{dR_2}{d\alpha_1} &= \frac{-\gamma_2}{\gamma_2 N_2 + (1 - \alpha_1)} \\ \frac{dR_2}{dR_1} &= -\frac{\gamma_1}{\gamma_2} \cdot \frac{\gamma_1 N_1 + \alpha_1}{\gamma_2 N_2 + (1 - \alpha_1)} \\ \frac{d^2 R_2}{dR_1^2} &= \frac{d}{d\alpha_1}\left(\frac{dR_2}{dR_1}\right) \cdot \frac{d\alpha_1}{dR_1} \\ &= -\frac{\gamma_2}{\gamma_1} \cdot \frac{\gamma_2 N_2 + (1 - \alpha_1) + (\gamma_1 N_1 + \alpha_1)}{(\gamma_2 N_2 + (1 - \alpha_1))^2} \\ &\quad \cdot \frac{\gamma_1 N_1 + \alpha_1}{\gamma_1} \\ &= -\frac{\gamma_2}{\gamma_1} \cdot \frac{\gamma_2 N_2 + \gamma_1 N_1 + 1}{(\gamma_2 N_2 + (1 - \alpha_1))^2} \cdot \frac{\gamma_1 N_1 + \alpha_1}{\gamma_1} \\ &< 0. \end{aligned} \tag{11}$$

Therefore, the curve $e(\gamma)$ is concave. Further,

$$\frac{\partial}{\partial \alpha_1}[e + \lambda(\alpha_1 + \alpha_2)] = \left(\frac{1}{C_1}\right)\left(\frac{\gamma_1}{\gamma_1 N_1 + \alpha_1}\right) + \lambda = 0. \tag{12}$$

$$\frac{\partial}{\partial \alpha_2}[e + \lambda(\alpha_1 + \alpha_2)] = \left(\frac{1}{C_2}\right)\left(\frac{\gamma_2}{\gamma_2 N_2 + \alpha_2}\right) + \lambda = 0. \tag{13}$$

Equating (12) and (13), we find

$$\gamma_2(\gamma_1 N_1 + \alpha_1) \ln\left(1 + \frac{1}{N_1}\right) = \gamma_1(\gamma_2 N_2 + \alpha_2) \ln\left(1 + \frac{1}{N_2}\right). \tag{14}$$

If (14) were satisfied for $\alpha_1 = \gamma_1$ and $\alpha_2 = \gamma_2$, we would have

$$(N_1 + 1) \ln\left(1 + \frac{1}{N_1}\right) = (N_2 + 1) \ln\left(1 + \frac{1}{N_2}\right) \tag{15}$$

which is true if and only if $N_1 = N_2$. This is in accordance with general conclusions obtained in Bergmans [7] on the impossibility of dominating time sharing when the channels are identical.

In general, (14) will not be satisfied for $\boldsymbol{\alpha} = \boldsymbol{\gamma}$, and the stationary point of e will be reached when e is positive, since $e = 0$ when $\boldsymbol{\alpha} = \boldsymbol{\gamma}$. This, together with the concavity of $e(\gamma)$, allows us to conclude that there is a nontrivial portion of that curve above the time-sharing line.

It is possible to show that

$$\left.\frac{de}{d\alpha_1}\right|_{\alpha_1 = \gamma_1} < 0 \tag{16}$$

and hence that the point corresponding to $\boldsymbol{\alpha} = \boldsymbol{\gamma}$ is the point P on Fig. 4. Rates dominating time-sharing rates will be achieved for

$$\begin{gathered} \delta_1 < \alpha_1 < \gamma_1 \\ 1 - \delta_1 = \delta_2 > \alpha_2 > \gamma_2 \end{gathered} \tag{17}$$

where $\boldsymbol{\delta}$ is the second solution of the equation $e(\boldsymbol{\alpha}) = 0$ (for a given γ). In conclusion, we should give the noisier channel a fraction of the total power which is slightly greater than the fraction of the total bandwidth it was allocated. Also, as a consequence of the fact that time sharing cannot be dominated for equal channels, it is evident that the closer the S/N ratios of the two channels, the smaller the interval $[\delta_1,\gamma_1]$ will be.

The curve of Fig. 4 was drawn for a given frequency partition γ. By letting γ vary, we shall generate a continuous set of similar curves. The envelope of all these curves strictly dominates the time-sharing line, since all the curves will have some portion of them above the time-sharing line.

Points on the envelope will be achieved for an optimal proportion of power as a function of bandwidth. There will be some tradeoff between power and bandwidth for rate points below the envelope.

V. Duality of Time and Frequency

The dominance of frequency multiplex over naive time multiplex has been established. However, in the naive approach we fixed the power of transmitter i to be a constant P over a certain proportion τ_i of the available time and let the power be zero for the remaining proportion of the time. It is clear that, for transmitter i, the average power over the *entire* time is actually $\tau_i P$ instead of P. An equivalence of the time and frequency rate curves can be established when the power is calculated with respect to the entire time interval. However, this equivalence is illusory from a practical standpoint, because it can only be achieved by lengthy power deviations from the average power constraints in the time-varying case.

To see this consider the frequency multiplex case, where the set of achievable rates is given by

$$R_1 = \gamma_1 W \ln\left(1 + \frac{\alpha_1 P}{\gamma_1 W N_1}\right)$$
$$R_2 = \gamma_2 W \ln\left(1 + \frac{\alpha_2 P}{\gamma_2 W N_2}\right), \qquad \boldsymbol{\alpha}, \boldsymbol{\gamma} \in \mathscr{S}_2 \qquad (18)$$

where $\boldsymbol{\alpha}$ is the power division and $\boldsymbol{\gamma}$ is the frequency division. Similarly the set of time-sharing rates is given by

$$R_1 = \tau_1 W \ln\left(1 + \frac{P_1}{W N_1}\right)$$
$$P_2 = \tau_2 W \ln\left(1 + \frac{P_2}{W N_2}\right), \qquad \boldsymbol{\tau} \in \mathscr{S}_2 \qquad (19)$$

where $\boldsymbol{\tau}$ is the division in time, and P_i is the power of transmission during communication to user i. Clearly $\tau_1 T P_1 + \tau_2 T P_2 = TP$ where T is the total duration of the waveform. For the two transmitters we let the power $P_1(t)$ and $P_2(t)$ be

$$P_1(t) = \begin{cases} P_1, & 0 \le t \le \tau_1 T \\ 0, & \tau_1 T < t \le T \end{cases}$$

and

$$P_2(t) = \begin{cases} 0, & 0 \le t \le \tau_1 T \\ P_2, & \tau_1 T < t \le T. \end{cases} \qquad (20)$$

Letting $\tau_i P_i / P = \lambda_i$, $\boldsymbol{\lambda} \in \mathscr{S}_2$, we obtain

$$R_1 = \tau_1 W \ln\left(1 + \frac{\lambda_1 P}{\tau_1 W N_1}\right)$$
$$R_2 = \tau_2 W \ln\left(1 + \frac{\lambda_2 P}{\tau_2 W N_2}\right), \qquad \boldsymbol{\tau}, \boldsymbol{\lambda} \in \mathscr{S}_2 \qquad (21)$$

which is, indeed, completely equivalent to the set of rates achievable by frequency multiplex. One major objection is the following. To achieve these rates in a coding sense, the time T must tend to infinity. Although it is still true that over the *total* time T the average power of the transmitter is P, over one of the two intervals the power P_i is larger than P. This may be inadmissible because of power limitations in the transmitter. An obvious solution to this problem would be to divide the total length of time T into n subintervals, and to let $n \to \infty$ as $T \to \infty$.

The argument against this is that the signal which results from this scheme does not have a vanishing power content outside $(-0,W)$ when $T \to \infty$, because of the switching between signals. If there are only two intervals ($\tau_1 T$ and $\tau_2 T$), the fraction of the total power of the resulting signal outside $(-0,W)$ will go to zero when $T \to \infty$, provided the original signals in $\tau_1 T$ and $\tau_2 T$ satisfy this condition.

In conclusion, time sharing is equivalent to frequency multiplex only if very long deviations from the average power P are allowed.

The situation is very different in frequency multiplex, because in this case, the spectrum of the transmitter need not be flat; i.e., the (single sided) power spectral density $P_i(f)$ is given by

$$P_1(f) = \begin{cases} \dfrac{\alpha_1 P}{\gamma_1 W}, & 0 \le f \le \gamma_1 W \\ 0, & \gamma_1 W < f \le W \end{cases}$$

and

$$P_2(f) = \begin{cases} 0, & 0 \le f \le \gamma_1 W \\ \dfrac{\alpha_2 P}{\gamma_2 W}, & \gamma_1 W < f \le W. \end{cases} \qquad (22)$$

However, there will not be a *time* discontinuity in power. Since the transmitter is operating in the time domain, this is the important factor. One can consider time-sharing equivalent (dual) to frequency multiplex only if time variations in input power are allowed, just as spectral variations of power are allowed in the frequency multiplex case.

VI. Superposition Coding

With the simplified notation introduced in Section II, the rates achieved by superposition coding are given by

$$R_1^* = \ln\left(1 + \frac{\alpha_1^*}{N_1}\right)$$
$$R_2^* = \ln\left(1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*}\right)$$
$$\boldsymbol{\alpha}^* \in \mathscr{S}_2. \qquad (23)$$

It can be easily shown that the curve described by (23) is concave, and that it dominates time sharing. The question, of course, is whether it dominates the envelope of frequency multiplexing curves described in the Section IV. This section establishes this dominance.

Superposition coding will dominate frequency multiplexing if, for any given $\boldsymbol{\alpha},\boldsymbol{\gamma} \in \mathscr{S}_2$, we can always find an $\boldsymbol{\alpha}^* \in \mathscr{S}_2$ such that

$$\ln\left(1 + \frac{\alpha_1^*}{N_1}\right) \ge \gamma_1 \ln\left(1 + \frac{\alpha_1}{\gamma_1 N_1}\right)$$
$$\ln\left(1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*}\right) \ge \gamma_2 \ln\left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right) \qquad (24)$$

or, equivalently

$$\left(1 + \frac{\alpha_1^*}{N_1}\right) \ge \left(1 + \frac{\alpha_1}{\gamma_1 N_1}\right)^{\gamma_1}$$
$$\left(1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*}\right) \ge \left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right)^{\gamma_2}. \qquad (25)$$

The proof of inequality (25) is deceptively difficult, and we are grateful to D. Hughes-Hartogs for his contribution of the following argument. We shall need the following lemmas, which we give without their straightforward algebraic verification.

Lemma I:

$$\max_{\gamma \in [0,1]} \left(1 + \frac{a}{\gamma}\right)^{\gamma} = 1 + a, \qquad a > 0.$$

The maximum is achieved for $\gamma = 1$.

Lemma II:

$$\max_{\gamma \in [0,1]} \left(1 + \frac{a}{\gamma}\right)^{\gamma} \left(1 + \frac{b}{1-\gamma}\right)^{1-\gamma} = 1 + a + b, \qquad a, b > 0.$$

The maximum is achieved for $\gamma = a/(a + b)$.

Proposition 3—(Hughes-Hartogs): Given $\boldsymbol{\alpha},\boldsymbol{\gamma} \in \mathcal{S}_2$, i.e., $\alpha_1 + \alpha_2 = 1; \gamma_1 + \gamma_2 = 1; \alpha_i, \gamma_i \geq 0;$ and $N_2 > N_1 > 0$; then there exists an $\boldsymbol{\alpha}^* \in \mathcal{S}_2$ such that

$$\left(1 + \frac{\alpha_1^*}{N_1}\right) \geq \left(1 + \frac{\alpha_1}{\gamma_1 N_1}\right)^{\gamma_1} \tag{26}$$

$$\left(1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*}\right) \geq \left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right)^{\gamma_2}. \tag{27}$$

Proof: Using condition (26), which is then automatically satisfied, set

$$\alpha_1^* = N_1 \left[\left(1 + \frac{\alpha_1}{\gamma_1 N_1}\right)^{\gamma_1} - 1\right]. \tag{28}$$

Since $\alpha_1 \geq 0$, $\gamma_1 \geq 0$, and $N_1 \geq 0$

$$\left(1 + \frac{\alpha_1}{\gamma_1 N_1}\right)^{\gamma_1} \geq 1$$

and

$$\alpha_1^* \geq 0. \tag{29}$$

By Lemma I

$$\alpha_1^* \leq N_1 \left[\left(1 + \frac{\alpha_1}{N_1}\right) - 1\right] = \alpha_1 \leq 1. \tag{30}$$

Let $\alpha_2^* = 1 - \alpha_1^*$, and hence $\boldsymbol{\alpha}^* \in \mathcal{S}_2$.

We now verify condition (27)

$$\begin{aligned} 1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*} &= \frac{N_2 + \alpha_1^* + \alpha_2^*}{N_2 + \alpha_1^*} \\ &= \frac{N_2 + 1}{(N_2 - N_1) + N_1 \left(1 + \frac{\alpha_1}{\gamma_1 N_1}\right)^{\gamma_1}} \\ &= \frac{N_2 + 1}{(N_2 - N_1)\left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right)^{\gamma_2} + N_1 \left(1 + \frac{\alpha_1}{\gamma_1 N_1}\right)^{\gamma_1} \left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right)^{\gamma_2}} \left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right)^{\gamma_2} \\ &\geq \frac{N_2 + 1}{(N_2 - N_1)\left(1 + \frac{\alpha_2}{N_2}\right) + N_1 \left(1 + \frac{\alpha_1}{N_1} + \frac{\alpha_2}{N_2}\right)} \left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right)^{\gamma_2}. \end{aligned} \tag{31}$$

In (31), we invoked both Lemma I and Lemma II and the fact that $N_2 > N_1 > 0$. We now have

$$\begin{aligned} 1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*} &\geq \frac{N_2 + 1}{N_2 - N_1 + \alpha_2 - \frac{\alpha_2 N_1}{N_2} + N_1 + \alpha_1 + \frac{\alpha_2 N_1}{N_2}} \cdot \left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right)^{\gamma_2} \\ &= \left(1 + \frac{\alpha_2}{\gamma_2 N_2}\right)^{\gamma_2}. \end{aligned} \tag{32}$$

Equation (32) proves that condition (27) is also satisfied. The inequalities in the proof are strict for nondegenerate cases, i.e., for $\gamma_i \neq 0$.

A set of $e(\gamma)$ curves, together with the superposition coding curve, is given in Fig. 5.

We now investigate the performance of multiband mixed modes of operation, and show that any such mode of operation is dominated by a superposition coding using the full available bandwidth. The following propositions are proved in the Appendix.

Proposition 4: If a continuous AWGN channel is used to transmit information from a single source, the best performance is achieved by using a single code operating in the full available band, rather than by partitioning the total band in smaller bands and using separate codes in the smaller bands. We shall refer to this last mode of operation as the multiband mode for simple channels. Let

$$\mathcal{S}_n = \left\{(s_1, s_2, \cdots, s_n): s_i \geq 0, \sum_{i=1}^{n} s_i = 1\right\} \tag{33}$$

and $\boldsymbol{\alpha},\boldsymbol{\gamma} \in \mathcal{S}_n$ be the power and frequency proportion in subband i. The total rate R_T is given by

$$R_T = \sum_{i=1}^{n} \gamma_i \ln \left(1 + \frac{\alpha_i}{\gamma_i N}\right). \tag{34}$$

The Appendix shows that $R_T \leq C = \ln(1 + 1/N)$, with equality iff $\boldsymbol{\alpha} = \boldsymbol{\gamma}$.

Proposition 5: If a continuous AWGN broadcast channel is used to transmit information to two users using superposition codes exclusively, the best performance is achieved

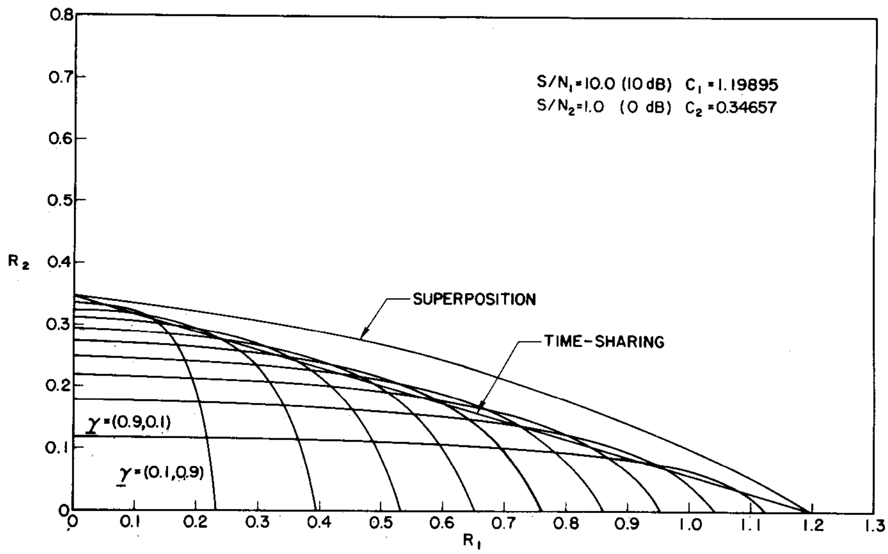

Fig. 5. Cooperative broadcasting.

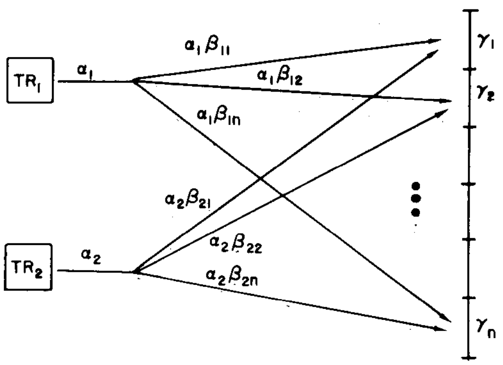

Fig. 6. Arbitrary frequency division.

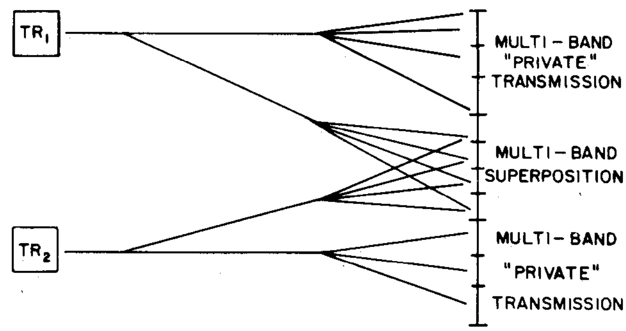

Fig. 7. Mixed multiband channel.

by using a superposition code operating in the full available band, rather than by using separate superposition codes operating in disjoint smaller subbands (multiband superposition operation).

To see this, let $\gamma \in \mathcal{S}_n$ represent the (relative) width of the ith subband, and $\boldsymbol{\alpha} \in \mathcal{S}_2$ be the power repartition between transmitter 1 and transmitter 2. Finally, the repartition of the power for transmitter j ($j = 1,2$) in each subband is done proportionally to $\boldsymbol{\beta}_j \in \mathcal{S}_n$, $j = 1,2$ (Fig. 6). We have

$$R_{1T} = \sum_{i=1}^{n} \gamma_i \ln\left(1 + \frac{\alpha_1\beta_{1i}}{\gamma_i N_1}\right)$$
$$R_{2T} = \sum_{i=1}^{n} \gamma_i \ln\left(1 + \frac{\alpha_2\beta_{2i}}{\gamma_i N_2 + \alpha_1\beta_{1i}}\right). \qquad (35)$$

The Appendix shows that there exists an $\boldsymbol{\alpha}^* \in \mathcal{S}_2$ such that

$$R_{1T} \le \ln\left(1 + \frac{\alpha_1^*}{N_1}\right)$$
$$R_{2T} \le \ln\left(1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*}\right) \qquad (36)$$

with equality iff $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \gamma$.

The following procedure describes a method for finding a superposition transmission mode which uses the full available bandwidth and which dominates any multiband mixed mode of operation. In Fig. 7, we represent a mode of operation involving both frequency multiplex (in so-called "private" bands) and superposition (in the common band). Moreover, each mode is multiband, as defined previously.

Step 1: Using Proposition 4, we first replace the various private bands for each transmitter by a single private band.

Step 2: Using Proposition 3, we pool the powers and bandwidths of the two private bands and replace them by one single band with a superposition code over the width of the two private bands.

Step 3: We are now operating with multiband superposition coding, and using Proposition 5, we replace this by a single superposition code using the full available bandwidth.

At each step, the achievable rates have been increased (except in trivial cases). This finally proves the dominance of superposition coding.

For Gaussian channels, the signals can be "mixed in the air." That is

$$s(t) = s_1(t) + s_2(t)$$
$$y_i(t) = s(t) + n_i(t), \qquad i = 1,2. \qquad (37)$$

User 1 first determines $s_2(t)$ from $y_1(t)$, then finds $s_1(t)$ from $y_1(t) - s_2(t)$. Thus no *active* cooperation is needed—the channels only have to share the band.

The situation is different when the various bands are preassigned, because of receiver limitations or regulations. If the bands are disjoint, we are in the case of frequency multiplexing, and the results of Section IV apply. If there is some overlap between the bands, we shall use superposition coding in the common band, and the resulting maximization problem is awkward.

The arguments of this paper can easily be generalized to a situation with N transmitters and N receivers. As shown in [8], the boundary of the set of rates achievable by superposition coding is given by

$$R_i^* = \ln\left(1 + \frac{\alpha_i^*}{N_i + \sum_{j<i} \alpha_j^*}\right) \tag{38}$$

for $\boldsymbol{\alpha}^* \in \mathscr{S}_n$, as opposed to

$$\tilde{R}_i = \ln\left(1 + \frac{\alpha_i^*}{N_i + \sum_{j\neq i} \alpha_j^*}\right) \tag{39}$$

for naive superposition [3], [4].

VII. Summary

First, we have shown that frequency multiplexing is equivalent to naive time multiplexing when the power allocation is proportional to the bandwidth allocation, and that it is better than time multiplexing when the power allocation is slightly biased over the bandwidth allocation in favor of the noisier channel. Second, we have proved that superposition coding (or cooperative broadcasting) dominates frequency multiplexing, unless very long deviations from average power are allowed. Finally, we have concluded that cooperative broadcasting strictly dominates any type of mixed-mode transmission, involving both frequency multiplexing (possibly multiband) and superposition coding (possibly multiband). The absolute optimality of the cooperative broadcasting scheme achieving rates as given by (38) is established in Bergmans [10].

Moreover, we have observed that the cooperation needed for superposition codes need not be active—the signals can be independently generated by the two transmitters and "mixed in the air." These results hold for Shannon-type encoding. It is to be hoped that the general result that "superposition dominates frequency multiplexing which in turn dominates time multiplexing" also holds for modulation schemes and that improvements on existing transmission methods may be achieved by using clever modulation schemes with pooled powers and bandwidths.

Acknowledgment

We wish again to thank D. Hughes-Hartogs for his contribution of the inequality of (20). We have also benefited from the suggestions of C. Keilers on schemes for sending information over Gaussian channels. We would also like to thank J. Thorp for helpful suggestions concerning some of the inequalities. We were led to discuss the duality and possible lack thereof of the time and frequency multiplexing schemes as a consequence of remarks by T. Gaarder and M. Hellman.

Appendix

Proposition 4: If $\boldsymbol{\alpha}, \boldsymbol{\gamma} \in \mathscr{S}_n$ then

$$R_T = \sum_{i=1}^{n} \gamma_i \ln\left(1 + \frac{\alpha_i}{\gamma_i N}\right) \le C = \ln\left(1 + \frac{1}{N}\right) \tag{40}$$

with equality if and only if $\boldsymbol{\alpha} = \boldsymbol{\gamma}$.

Proof: Consider R_T as a function of $\boldsymbol{\alpha}$ for a fixed $\boldsymbol{\gamma}$. We wish to maximize R_T subject to $\boldsymbol{\alpha} \in \mathscr{S}_n$. Hence, define

$$J = \sum_{i=1}^{n} \gamma_i \ln\left(1 + \frac{\alpha_i}{\gamma_i N}\right) + \lambda \sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \left(\gamma_i \ln\left(1 + \frac{\alpha_i}{\gamma_i N}\right) + \lambda \alpha_i\right)$$

$$\frac{\partial J}{\partial \alpha_i} = \frac{\gamma_i}{1 + \dfrac{\alpha_i}{\gamma_i N}} \cdot \frac{1}{\gamma_i N} + \lambda = \frac{\gamma_i}{\gamma_i N + \alpha_i} + \lambda = 0. \tag{41}$$

This is only possible if $\gamma_i/(\gamma_i N + \alpha_i)$ does not depend on i, which requires $\boldsymbol{\alpha} = \boldsymbol{\gamma}$. The convexity of the logarithm guarantees that the extremum is a maximum. This maximum is given by

$$\sum_{i=1}^{n} \gamma_i \ln\left(1 + \frac{\gamma_i}{\gamma_i N}\right) = C. \tag{42}$$

Proposition 5: If $\boldsymbol{\gamma}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathscr{S}_n$, $\boldsymbol{\alpha} \in \mathscr{S}_2$, then the rate point

$$R_{1T} = \sum_{i=1}^{n} \gamma_i \ln\left(1 + \frac{\alpha_1 \beta_{1i}}{\gamma_i N_1}\right) \tag{43a}$$

$$R_{2T} = \sum_{i=1}^{n} \gamma_i \ln\left(1 + \frac{\alpha_2 \beta_{2i}}{\gamma_i N_2 + \alpha_1 \beta_{1i}}\right) \tag{43b}$$

is strictly dominated by the rate point

$$R_1^* = \ln\left(1 + \frac{\alpha_1^*}{N_1}\right) \tag{44a}$$

$$R_2^* = \ln\left(1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*}\right) \tag{44b}$$

for some $\boldsymbol{\alpha}^* \in \mathscr{S}_2$, unless $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\gamma}$, in which case we have equality.

Proof: We equate (43b) and (44b) to solve for α_1^*

$$\left(1 + \frac{\alpha_2^*}{N_2 + \alpha_1^*}\right) = \prod_{i=1}^{n} \left(1 + \frac{\alpha_2 \beta_{2i}}{\gamma_i N_2 + \alpha_1 \beta_{1i}}\right)^{\gamma_i}$$

$$\alpha_1^* = (N_2 + 1) \prod_{i=1}^{n} \left(1 + \frac{\alpha_2 \beta_{2i}}{\gamma_i N_2 + \alpha_1 \beta_{1i}}\right)^{-\gamma_i} - N_2. \tag{45}$$

From (43b), and by Proposition 4

$$0 \le R_{2T} \le \sum_{i=1}^{n} \gamma_i \ln\left(1 + \frac{\alpha_2 \beta_{2i}}{\gamma_i N_2}\right) \le \ln\left(1 + \frac{1}{N_2}\right) = C_2$$

and, since R_2^* is a continuous function of α_2^*, with $R_2^* = 0$ for $\alpha_2^* = 0$ and $R_2^* = C_2$ for $\alpha_2^* = 1$, equating (43b) and (44b) yields $\boldsymbol{\alpha}^* \in \mathscr{S}_2$. Substituting α_1^* from (45) in exp (R_1^*), yields

$$\exp(R_1^*) = 1 + \frac{\alpha_1^*}{N_1} = 1 - \frac{N_2}{N_1} + \left(\frac{N_2+1}{N_1}\right)\prod_{i=1}^{n}\left(1 + \frac{\alpha_2\beta_{2i}}{\gamma_i N_2 + \alpha_1\beta_{1i}}\right)^{-\gamma_i}. \tag{46}$$

To find a lower bound for $\exp(R_1^*)$, we minimize

$$\prod_{i=1}^{n}\left(1 + \frac{\alpha_2\beta_{2i}}{\gamma_i N_2 + \alpha_1\beta_{1i}}\right)^{-\gamma_i} = \exp\left(\sum_{i=1}^{n}\gamma_i \ln\left(\frac{\gamma_i N_2 + \alpha_1\beta_{1i}}{\gamma_i N_2 + \alpha_1\beta_{1i} + \alpha_2\beta_{2i}}\right)\right)$$

subject to $\sum \beta_{2i} = 1$. Thus, taking the logarithm

$$\frac{\partial}{\partial\beta_{2i}}\sum_{i=1}^{n}\left(\gamma_i \ln\left(\frac{\gamma_i N_2 + \alpha_1\beta_{1i}}{\gamma_i N_2 + \alpha_1\beta_{1i} + \alpha_2\beta_{2i}}\right) + \lambda\beta_{2i}\right) = \gamma_i \frac{\gamma_i N_2 + \alpha_1\beta_{1i} + \alpha_2\beta_{2i}}{\gamma_i N_2 + \alpha_1\beta_{1i}} \frac{-\alpha_2(\gamma_i N_2 + \alpha_1\beta_{1i})}{(\gamma_i N_2 + \alpha_1\beta_{1i} + \alpha_2\beta_{2i})^2} + \lambda = 0$$

or

$$\frac{\gamma_i\alpha_2}{\gamma_i N_2 + \alpha_1\beta_{1i} + \alpha_2\beta_{2i}} = \lambda \tag{47}$$

which requires

$$\beta_{2i} = \frac{\gamma_i - \alpha_1\beta_{1i}}{\alpha_2}. \tag{48}$$

Since $\ln(k_1/(k_2 + x))$ is convex for $x \geq 0$, $\boldsymbol{\beta}_2$ of (48) achieves a minimum, given by

$$\prod_{i=1}^{n}\left(1 + \frac{\alpha_2\beta_{2i}}{\gamma_i N_2 + \alpha_1\beta_{1i}}\right)^{-\gamma_i} \geq \prod_{i=1}^{n}\left(\frac{\gamma_i N_2 + \alpha_1\beta_{1i}}{\gamma_i(N_2+1)}\right)^{\gamma_i} \tag{49}$$

and

$$\exp(R_1^*) \geq 1 - \frac{N_2}{N_1} + \left(\frac{N_2+1}{N_1}\right)\prod_{i=1}^{n}\left(\frac{\gamma_i N_2 + \alpha_1\beta_{1i}}{\gamma_i(N_2+1)}\right)^{\gamma_i} = 1 - \frac{N_2}{N_1} + \frac{N_2}{N_1}\prod_{i=1}^{n}\left(1 + \frac{\alpha_1\beta_{1i}}{\gamma_i N_2}\right)^{\gamma_i}. \tag{50}$$

Finally,

$$\begin{aligned}\exp(R_1^*) - \exp(R_{IT}) &\geq 1 - \frac{N_2}{N_1} + \frac{N_2}{N_1}\prod_{i=1}^{n}\left(1 + \frac{\alpha_1\beta_{1i}}{\gamma_i N_2}\right)^{\gamma_i} - \prod_{i=1}^{n}\left(1 + \frac{\alpha_1\beta_{1i}}{\gamma_i N_1}\right)^{\gamma_i} \\ &\geq \frac{1}{N_1}\left[N_2\left(\prod_{i=1}^{n}\left(1 + \frac{\alpha_1\beta_{1i}}{\gamma_i N_2}\right)^{\gamma_i} - 1\right) - N_1\left(\prod_{i=1}^{n}\left(1 + \frac{\alpha_1\beta_{1i}}{\gamma_i N_1}\right)^{\gamma_i} - 1\right)\right] \\ &= \frac{1}{N_1}[f(N_2) - f(N_1)]\end{aligned} \tag{51}$$

where

$$f(x) \triangleq x\left[\prod_{i=1}^{n}\left(1 + \frac{\alpha_1\beta_{1i}}{\gamma_i x}\right)^{\gamma_i} - 1\right] = \prod_{i=1}^{n}\left(x + \frac{\alpha_1\beta_{1i}}{\gamma_i}\right)^{\gamma_i} - x. \tag{52}$$

We now show that $f(x)$ is strictly increasing in x, unless $\boldsymbol{\beta}_1 = \boldsymbol{\gamma}$, in which case $f(x)$ is constant. We have

$$\begin{aligned}f'(x) &= \sum_{i=1}^{n}\gamma_i\left(x + \frac{\alpha_1\beta_{1i}}{\gamma_i}\right)^{\gamma_i - 1}\prod_{j\neq i}\left(x + \frac{\alpha_1\beta_{1j}}{\gamma_j}\right)^{\gamma_j} - 1 \\ &= \prod_{j=1}^{n}\left(x + \frac{\alpha_1\beta_{1j}}{\gamma_j}\right)^{\gamma_j}\sum_{i=1}^{n}\gamma_i\left(x + \frac{\alpha_1\beta_{1i}}{\gamma_i}\right)^{-1} - 1 \\ &\geq \prod_{j=1}^{n}\left(x + \frac{\alpha_1\beta_{1j}}{\gamma_j}\right)^{\gamma_j}\prod_{i=1}^{n}\left(\left(x + \frac{\alpha_1\beta_{1i}}{\gamma_i}\right)^{-1}\right)^{\gamma_i} - 1 \\ &= \left(\prod_{j=1}^{n}\left(x + \frac{\alpha_1\beta_{1j}}{\gamma_j}\right)^{\gamma_j}\right)\left(\prod_{i=1}^{n}\left(x + \frac{\alpha_1\beta_{1i}}{\gamma_i}\right)^{\gamma_i}\right)^{-1} - 1 \\ &= 0.\end{aligned} \tag{53}$$

The inequality in (53) is a consequence of the theorem of the arithmetic and geometric means ([9], pp. 16 et seq.). We shall have equality in (53) iff $(x + \alpha_1\beta_{1i}/\gamma_i)^{-1}$ is not a function of i, i.e., iff $\alpha_1\beta_{1i}$ is proportional to γ_i, which implies $\boldsymbol{\beta}_1 = \boldsymbol{\gamma}$, since $\boldsymbol{\beta}_1, \boldsymbol{\gamma} \in \mathscr{S}_n$. Hence $f(x)$ is increasing in x (unless $\boldsymbol{\beta}_1 = \boldsymbol{\gamma}$), and $f(N_2) \geq f(N_1)$. In conclusion

$$\exp(R_1^*) - \exp(R_{IT}) \geq \frac{1}{N_1}[f(N_2) - f(N_1)] \geq 0 \tag{54}$$

and the proof of the existence of a dominating point is completed. Finally, the inequalities in this proof are strict, unless $\boldsymbol{\gamma} = \boldsymbol{\beta}_1$ and (48) are satisfied, in which case $\boldsymbol{\gamma} = \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948; also, reprinted in C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication.* Urbana, Ill.: Univ. of Illinois Press, 1949.

[2] ——, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.

[3] V. L. Stefanyuk, "Kratkie Soobshcheniia o 'Vzaimopomoshchi' v Kollektivie Radiostantsii," *Probl. Peredach. Inform.*, vol. VII, pp. 103–107, 1971.

[4] V. L. Stefanyuk and M. L. Tsetlin, "Power regulation in a group of radio stations," *Probl. Peredach. Inform.*, vol. 3, pp. 49–57, 1967.

[5] J. P. Costas, "Poisson, Shannon, and the radio amateur," *Proc. IRE*, vol. 47, pp. 2058–2068, Dec. 1959.

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

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

[8] ——, "Degraded broadcast channels," Ph.D. dissertation, Dep. Elec. Eng., Stanford University, Stanford, Calif., 1972.

[9] G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities.* New York: Cambridge, 1934.

[10] P. P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory* (Corresp.), vol. IT-20, pp. 279–280, Mar. 1974.