Comments on Broadcast Channels

Thomas M. Cover, *Fellow, IEEE*

(Invited Paper)

***Abstract*— The key ideas in the theory of broadcast channels are illustrated by discussing some of the progress toward finding the capacity region. The capacity region is still unknown.**

***Index Terms*— Binning, broadcast channel, capacity, degraded broadcast channel, feedback capacity, Slepian–Wolf, superposition.**

I. Introduction

A broadcast channel has one sender and many receivers. The object is to broadcast information to the receivers. The information may be independent or nested. We shall treat broadcast channels with two receivers as shown in Fig. 1. Multiple receiver broadcast channels are defined similarly.

Definition: A *broadcast channel* consists of an input alphabet $\mathcal{X}$ and two output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ and a probability transition function $p(y_1, y_2|x)$. The broadcast channel is said to be *memoryless* if

$$p(y_1^n, y_2^n|x^n) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}|x_i).$$

A $((2^{nR_1}, 2^{nR_2}), n)$ code for a broadcast channel with independent information consists of an encoder

$$x^n: 2^{nR_1} \times 2^{nR_2} \to \mathcal{X}^n,$$

and two decoders

$$\hat{W}_1: \mathcal{Y}_1^n \to 2^{nR_1}$$
$$\hat{W}_2: \mathcal{Y}_2^n \to 2^{nR_2}.$$

The probability of error $P_e^{(n)}$ is defined to be the probability the decoded message is not equal to the transmitted message, i.e.,

$$P_e^{(n)} = P(\hat{W}_1(Y_1^n) \neq W_1 \text{ or } \hat{W}_2(Y_2^n) \neq W_2)$$

where the message (W_1, W_2) is assumed to be uniformly distributed over $2^{nR_1} \times 2^{nR_2}$.

Definition: A rate pair (R_1, R_2) is said to be *achievable* for the broadcast channel if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \to 0$.

Manuscript received December 15, 1997; revised May 1, 1998. This work was supported in part by NSF under Grant #NCR-9628193, by ARPA under Contract #J-FBI-94-218-2, and by JSEP under Grant #ARMY DAAG55-97-1-0115.

The author is with the Departments of Electrical Engineering and Statistics, Stanford University, Stanford, CA 94305 USA (e-mail: cover@isl.stanford.edu).

Publisher Item Identifier S 0018-9448(98)06081-7.

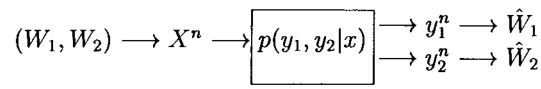

Fig. 1. Broadcast channel.

Definition: The *capacity region* of the broadcast channel is the closure of the set of achievable rates.

It is often the case in practice that one received signal is a degraded, or corrupted, version of the other. One receiver may be farther away or "downstream." When X, Y_1, Y_2 forms a Markov chain, i.e., when $p(y_1, y_2|x) = p(y_1|x)p(y_2|y_1)$ we say that Y_2 is a *physically degraded* version of Y_1 and that $p(y_1, y_2|x)$ is a *physically degraded* broadcast channel. We note that the probabilities of error $P(\hat{W}_1 \neq W_1)$ and $P(\hat{W}_2 \neq W_2)$ depend only on the marginals $p(y_1|x)$ and $p(y_2|x)$ and not on the joint. Thus we define a weaker notion of degraded.

Definition: A broadcast channel $p(y_1, y_2|x)$ is said to be *degraded* if there exists a distribution $\tilde{p}(y_2|y_1)$ such that

$$p(y_2|x) = \sum_{y_1} p(y_1|x)\tilde{p}(y_2|y_1).$$

II. Capacity Region for the Degraded Broadcast Channel

Achievable rate regions for Gaussian broadcast channels, cascades of binary-symmetric channels (a special case of degraded broadcast channels), the push-to-talk channel, orthogonal broadcast channels, and product broadcast channels were found in Cover [16]. Surveys of multiuser theory, including broadcast channels, can be found in [19], [22], [23], [26], [35], [62], [69], [98], [99], [100], [107], and [108].

We first consider sending independent information over a degraded broadcast channel (Fig. 2) at rates R_1 to Y_1 and R_2 to Y_2. The capacity region, conjectured in [16], was proved to be achievable by Bergmans [9], and the converse was established by Bergmans [10] and Gallager [41].

Theorem 1: The capacity region for the degraded broadcast channel $X \to Y_1 \to Y_2$ is the convex hull of the closure of all (R_1, R_2) satisfying

$$R_2 \leq I(U; Y_2)$$
$$R_1 \leq I(X; Y_1|U)$$

for some joint distribution $p(u)p(x|u)p(y, z|x)$, where the auxiliary random variable U has cardinality bounded by $|\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|\}$.

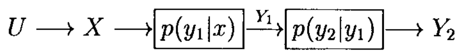

Fig. 2. Degraded broadcast channel with auxiliary input U.

Proof (Outline of Achievability): We first give an outline of the basic idea of superposition coding for the broadcast channel. The auxiliary random variable U will serve as a cloud center distinguishable by both receivers Y_1 and Y_2. Each cloud consists of 2^{nR_1} codewords X^n distinguishable by receiver Y_1. The worst receiver Y_2 can only see the clouds, while the better receiver can see the individual codewords within the clouds.

Fix $p(u)$ and $p(x|u)$.

Random Codebook Generation. Generate 2^{nR_2} independent codewords of length n, $u^n(w_2)$, $w_2 \in \{1, 2, \cdots, 2^{nR_2}\}$, according to $\prod_{i=1}^n p(u_i)$.

For each codeword $u^n(w_2)$, generate 2^{nR_1} independent codewords $x^n(w_1, w_2)$ according to the conditional probability mass function $\prod_{i=1}^n p(x_i|u_i(w_2))$.

Here $u^n(w_2)$ plays the role of the cloud center understandable to both Y_1 and Y_2, while $x^n(w_1, w_2)$ is the w_1th satellite codeword in the w_2th cloud. The cloud center $u^n(w_2)$ is never actually sent.

Encoding: To transmit the pair (W_1, W_2), send the corresponding codeword $x^n(W_1, W_2)$.

Decoding: Receiver Y_2 determines the unique $\hat{\hat{W}}_2$ such that $(u^n(\hat{\hat{W}}_2), y_2^n)$ is jointly typical. If there are none such or more than one such, an error is declared.

Receiver Y_1 looks for the unique $(\hat{W}_1, \hat{W}_2)$ such that $(u^n(\hat{W}_2), x^n(\hat{W}_1, \hat{W}_2), y_1^n)$ is jointly typical. If there are none such or more than one such, an error is declared.

Error Analysis (Outline): The condition

$$R_2 < I(U; Y_2)$$

guarantees that $\hat{\hat{W}}_2 = W_2$ with high probability because there are $2^{nI(U;Y_2)}$ distinguishable u^n's as observed by Y_2. The extra information in $x^n \sim p(x^n|u^n)$ is viewed as noise by Y_2. The condition $R_1 < I(X; Y_1|U)$ guarantees that receiver Y_1 can decode $\hat{W}_1 = W_1$ with high probability, given that the receiver has already decoded W_2. □

Note that the proof uses a "subtract-off" or conditioning idea for receiver Y_1. Let Y_1 first determine $u^n(W_2)$. This can be done, because the inferior receiver Y_2 can also determine W_2. Then condition on u^n (or subtract it from the received signal for the Gaussian channel) and decode the refined message W_1 given u^n and Y_1^n.

This subtract-off method can also be used for the multiple-access channel, and its implementation is one of the challenges of code-division multiple access (CDMA). A treatment of code-division broadcasting (one sender and m receivers) and code division multiple access (m senders and one receiver) for the bandlimited additive white Gaussian noise channel is given in Bergmans and Cover [11], where it is proved that the CDMA rate region is strictly larger than the rate regions achievable by frequency-division multiple access (FDMA) and time-division multiple access (TDMA).

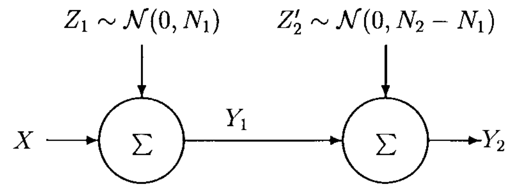

Fig. 3. Gaussian broadcast channel.

We now consider the Gaussian broadcast channel

$$Y_1 = X + Z_1$$
$$Y_2 = X + Z_2$$

where $Z_1 \sim \mathcal{N}(0, N_1)$ and $Z_2 \sim \mathcal{N}(0, N_2)$. This is a particular example of a degraded broadcast channel because the channel can be recharacterized as shown in Fig. 3, where $Z_2' \sim \mathcal{N}(0, N_2 - N_1)$.

Let

$$C(x) = \frac{1}{2} \log(1 + x)$$

denote the capacity in bits per transmission of a memoryless Gaussian channel with signal-to-noise ratio x.

Theorem 2: The capacity region for the Gaussian broadcast channel, with signal power constraint P, is given by

$$R_1 \le C\left(\frac{\alpha P}{N_1}\right)$$
$$R_2 \le C\left(\frac{(1-\alpha)P}{\alpha P + N_2}\right), \qquad \text{for } 0 \le \alpha \le 1.$$

This region is achieved by the coding scheme described in [16]. Choose 2^{nR_2} Gaussian codewords $u^n(i)$ independent and identically distributed (i.i.d.) $\sim \mathcal{N}(0, (1-\alpha)P)$. For each of these codewords $u^n(i)$, generate 2^{nR_1} satellite Gaussian codewords $v^n(j)$ of power αP and add them to form codewords $x^n(i, j) = u^n(i) + v^n(j)$. Thus, the fine information $v^n(j)$ is "superimposed" on the coarse information $u^n(i)$. Bermans [9], [10] proved the converse.

The achievability of the region in Theorem 1 for general degraded broadcast channels was established by Bergmans [9]. There followed a year of intense activity trying to prove the converse, i.e., to prove that the natural achievable rate region, was indeed the capacity region. Correspondences were exchanged between Aaron Wyner (Bell Labs), Patrick Bergmans (then at Cornell), and Robert Gallager (MIT). Finally, one day at the end of the year, Wyner received proofs of the converse by Bergmans [10] and by Gallager [41]. Gallager's proof successfully defined the role of the auxiliary random variable U in terms of the collection of all the outputs up to the current time. Bergmans' proof, on the other hand, held for the Gaussian channel. Gallager's proof did not apply to the Gaussian channel with a power constraint, nor did Bergmans' proof apply to the general unrestricted broadcast channel. Bergmans' proof, instead, used a conditional entropy power inequality, the first use of this inequality since Shannon (1948).

So the key ideas in the early papers were superposition coding, subtracting off (or conditioning on) message information layer by layer, identification of the superposition variable in the converse, and the use of the entropy power inequality.

III. The Deterministic Broadcast Channel

Van der Meulen [97] and Cover [18] established an achievable rate region for sending common information at rate R_0 to both receivers and conditionally independent information at rates R_1 and R_2 to the two receivers. (Jahn [58] considered the arbitrarily varying broadcast channel counterpart.) The region was soon enlarged by ingenious work by Gelfand [42], Pinsker [43], [82], and Marton [74], [75]. Gelfand looked at a particular deterministic broadcast channel, known as the Blackwell channel, given by

$$Y_1 = \begin{cases} 1, & x = 1 \\ 0, & x = 2 \text{ or } 3 \end{cases}$$

$$Y_2 = \begin{cases} 1, & x = 1 \text{ or } 2 \\ 0, & x = 3. \end{cases}$$

Here one sees that one can send at one bit per transmission to receiver Y_1 or to receiver Y_2, but not simultaneously to both. What, then, is the set of achievable (R_1, R_2) pairs?

Gelfand found the capacity region in [42]. Soon thereafter, Marton [74] and Pinsker [82] independently established the capacity region for general deterministic broadcast channels. The extra ingredient in the deterministic broadcast channel investigation is the use of the Slepian–Wolf theorem [94] and a binning argument [17] used in its proof.

In the Slepian–Wolf theorem, one has two correlated random variables U and V, and i.i.d. copies (U_i, V_i) all drawn according to $p(u, v)$. How many bits of information R_1 does one need to say about U and how many bits R_2 does one need to say about V so that the combined description will recover U and V with negligible probability of error?

Theorem 3 (Slepian and Wolf [94]): Let $(U_i, V_i) i = 1, 2, \cdots,$ be i.i.d. discrete random variables. There exist maps $i_n\colon \mathcal{U}^n \longrightarrow 2^{nR_1}$, $j_n\colon \mathcal{V}^n \longrightarrow 2^{nR_2}$, $|i_n(\cdot)| = 2^{nR_1}$, $|j_n(\cdot)| = 2^{nR_2}$, and reconstruction functions $\hat{u}^n(i_n, j_n)$, $\hat{v}^n(i_n, j_n)$, such that

$$\Pr\{(\hat{U}^n, \hat{V}^n) \neq (U^n, V^n)\} \longrightarrow 0$$

if and only if

$$\begin{aligned} R_1 &> H(U|V) \\ R_2 &> H(V|U) \\ R_1 + R_2 &> H(U, V). \end{aligned} \tag{1}$$

One can achieve a rate pair in this region by a random binning argument. Suppose that one randomly throws all u^n sequences into 2^{nR_1} bins. Similarly, one randomly throws the v^n sequences into 2^{nR_2} bins. Describe U^n by its bin number $i(U^n)$ and V^n by its bin number $j(V^n)$, where $|i(\cdot)| = 2^{nR_1}$, and $|j(\cdot)| = 2^{nR_2}$. Then a common receiver will be given the bin numbers of U^n and V^n. If there is only one jointly typical (U^n, V^n) pair in that bin, the receiver will make no mistake in reconstructing U^n and V^n. So the idea is to form a product partition of $2^{nR_1} \times 2^{nR_2}$ bins that is fine enough to isolate the typical (U^n, V^n) pairs. Rates (R_1, R_2) satisfying (1) suffice.

For the proof of the capacity of the deterministic broadcast channel, we use a product partition that is coarse enough so that with high probability any product bin will contain at least one typical (y_2^n, y_2^n) receiver sequence. To see how this is done we consider a channel in which $y_1 = f_1(x)$, $y_2 = f_2(x)$, where f_1 and f_2 are deterministic functions.

Suppose one wishes to send a pair of indices i and j to receivers 1 and 2, respectively. Fix a probability distribution $p(x)$, thus inducing a joint distribution $p(x, y_1, y_2)$. From this we can calculate the marginal distribution $p(y_1, y_2)$. The object here is to control Y_1 and Y_2 simultaneously by use of X. We first do a product binning of y_1^n and y_2^n, 2^{nR_1} bins for the y_1^n sequences and 2^{nR_2} bins for y_2^n. For what set of rates R_1 and R_2 will these bins contain at least one jointly typical (y_1^n, y_2^n)? Once we have answered that question, the problem is solved, because y_1^n and y_2^n are deterministic functions of x^n, so if there exists a jointly typical (y_1^n, y_2^n) in bin (i, j), say, one merely looks up the sequence x^n which results in y_1^n and y_2^n in order to send information i to Y_1 and j to Y_2. Thus rates R_1 and R_2 are achieved.

The partitioning of $\mathcal{Y}_1^n \times \mathcal{Y}_2^n$ is coarse enough so that a given (i, j) bin contains at least one jointly typical pair (X^n, Y^n), with high probability, if $R_1 < H(Y_1)$, $R_2 < H(Y_2)$, and $R_1 + R_2 < H(Y_1, Y_2)$. Thus we have the following theorem:

Theorem 4 [74], [82]: The capacity region of the deterministic memoryless broadcast channel with $y_1 = f_1(x)$, $y_2 = f_2(x)$, is given by the convex closure of the union of the rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &< H(Y_1) \\ R_2 &< H(Y_2) \\ R_1 + R_2 &< H(Y_1, Y_2). \end{aligned}$$

Comment: Here $R_1 < H(Y_1)$ ensures that there is at least one typical y_1^n per bin, and $R_1 + R_2 \leq H(Y_1, Y_2)$ ensures there is at least one jointly typical (y_1^n, y_2^n) per product bin. We note the interesting complementary relationship of this rate region to the Slepian–Wolf region in Fig. 4.

Marton [75] then generalized this result to arbitrary broadcast channels by setting up a kind of determinism by selecting a subset of distinguishable input sequences. Soon thereafter El Gamal and Van der Meulen [36] gave a simpler proof.

In the following theorem, we outline a proof of a special case of Marton's general result, where it is assumed that the information is independent and there is no common message. This special case isolates a new coding idea involving a pair of auxiliary random variables. This, together with superposition, yields Marton's theorem. Papers referring to Marton's region include Gelfand [43], Hajek [51], Han [52], Heegard [56], and Jahn [58], as well as [22], [23], [35], and [36]. The outline of the proof of the following theorem is due to El Gamal and Van der Meulen [36].

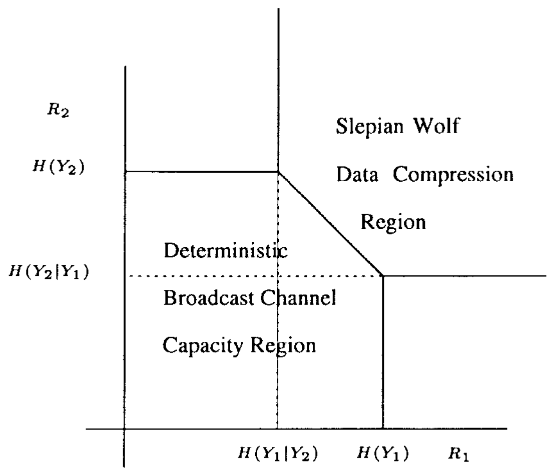

Fig. 4. The capacity regions for Slepian–Wolf data compression and for the deterministic broadcast channel, for the joint probability mass function $p(y_1, y_2)$ induced by $p(x)$.

Theorem 5 (Marton [75]): The rates (R_1, R_2) are achievable for the broadcast channel $\{\mathcal{X}, p(y_1, y_2|x), \mathcal{Y}_1 \times \mathcal{Y}_2\}$ if

$$\begin{aligned} R_1 &\le I(U; Y_1) \\ R_2 &\le I(V; Y_2) \\ R_1 + R_2 &\le I(U; Y_1) + I(V; Y_2) - I(U; V) \end{aligned} \tag{2}$$

for some $p(u, v, x)$ on $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$.

Comment: This achievable region is the capacity region if the broadcast channel has one deterministic component [75].

Proof (Outline): Fix $p(u, v)$, $p(x|u, v)$. The channel $p(y_1, y_2|x)$ is given. The idea is to send u to y_1 and v to y_2.

Random Coding: Generate $2^{nI(U;Y_1)}$ typical u's $\sim p(u)$. Generate $2^{nI(V;Y_2)}$ typical v's $\sim p(v)$. Randomly throw the u's into 2^{nR_1} bins and the v's into the 2^{nR_2} bins. For each product bin, find a *jointly* typical (u, v) pair. This can be done if

$$R_1 + R_2 < I(U; Y_1) + I(V; Y_2) - I(U; V).$$

To see this, recall that independent choices of u^n and v^n result in a jointly typical (u^n, v^n) with probability $2^{-nI(U;V)}$. Now there are $2^{n(I(U;Y_1)-R_1)}$ u^n's in any U bin, and $2^{n(I(V;Y_2)-R_2)}$ v^n's in any V bin. Thus the expected number of jointly typical (u^n, v^n) pairs in a given product $U \times V$ bin is

$$2^{n(I(U;Y_1)-R_1)}2^{n(I(V;Y_2)-R_2)}2^{-nI(U;V)}.$$

The desired jointly typical (u^n, v^n) pair can be found if this expected number is much greater than 1, which follows if (R_1, R_2) satisfies (2).

Continuing with the coding, for each $U \times V$ bin and its designated jointly typical (u^n, v^n) pair, generate $x^n(u^n, v^n)$ according to the conditional distribution $\prod_{k=1}^{n} p(x_k|u_k, v_k)$.

Encoding: To send i to Y_1 and j to Y_2, send $x^n(u^n, v^n)$, where (u^n, v^n) is the designated pair in the product bin (i, j).

Decoding: Receiver Y_1, upon receiving y_1^n, finds the u^n such that (u^n, y_1^n) is jointly typical. Thus it is necessary that $R_1 < I(U; Y_1)$. He then finds the index i of the bin in which u^n lies. Receiver Y_2 finds the v^n such that (v^n, y_2^n) is jointly typical. Thus we need $R_2 < I(V; Y_2)$. He then finds the index j of the bin in which v^n lies. □

IV. Results for Specific Channels

El Gamal [30] showed that feedback cannot increase the capacity of the physically degraded broadcast channel, i.e., broadcast channels for which $p(y_1, y_2|x) = p(y_1|x)p(y_2|y_1)$. It was later shown by Dueck [29] and Ozarow [80], [81] that feedback can in fact increase the capacity of general broadcast channels, in contrast to the single-user channel, where Shannon [91] proved that feedback does not increase capacity.

Ozarow and Leung [81] showed a new way to achieve the capacity region for the Gaussian broadcast channel with feedback using the Kailath–Schalkwijk coding scheme, in which one uses feedback to attempt to correct the misperceptions of (Y_1, Y_2) as seen by the transmitter. Their method, however, does not generalize to more than two receivers. Work on feedback capacity for broadcast channels appears in [29]–[31], [34], [71], [72], [80], and [81].

Poltyrev [84]–[87] looked at the reversely degraded broadcast channel (see also Hughes-Hartogs [57] for the Gaussian channel and Ohkubo [75]). Later, El Gamal [33] furnished a proof of the converse, thus establishing the Poltyrev region for the reversely degraded broadcast channel as the capacity region.

Channels in which one receiver is superior to another and channels with nested information were studied by Marton, Körner, Csiszár, El Gamal, and others [24], [32], [39], [62]–[64].

V. An Alternative View of Capacity

In this section, we illustrate the delicacy of the definition of the capacity region for broadcast channels and multiuser channels in general.

We consider a memoryless broadcast channel with m receivers $Y_1, Y_2, \cdots, Y_m$. If we were to ignore the needs of all receivers but the kth, the sender could communicate to receiver k at capacity

$$C_k = \max_{p(x)} I(X; Y_k).$$

But an optimal code for receiver Y_k generally precludes transmission at capacity to the other receivers. We now argue that a single communication strategy can achieve communication at capacity C_k bits per transmission for all the receivers, $k = 1, 2, \cdots, m$. This seems to violate the known results bounding the capacity region. Nonetheless there is some truth to this assertion. What is going on?

Suppose, for example, that an advanced civilization wishes to transmit its knowledge to other stars. Having little idea of which stars are listening, when they started to listen, or the noise characteristics of the receivers, it is not clear at first what communication strategy to employ.

But the following process seems reasonable. From time to time send a brief beacon signal to get any newcomer's attention. For somewhat longer periods, send a simple description of the language. Then send several years of information. Follow it up with thousands of years of information, including previous information. Then repeat the cycle with longer periods and more information. If the time durations are appropriately chosen, each star can receive all the information at its own capacity from the time it comes on line.

More precisely, use a $(2^{n_{ik}C_k}, n_{ik})$ code, for receiver k, for the kth segment of the ith cycle. Thus $n_{ik}C_k$ bits would be received by Y_k during its segment of n_{ik} transmissions. Let the blocklengths n_{ik} increase rapidly enough so that

$$n_{ik}/N_{ik} \longrightarrow 1, \qquad \text{as } i \longrightarrow \infty$$

where

$$N_{ik} = \sum_{r \le i} \sum_{s \le k} n_{rs}$$

is the total communication time up through segment ik. Thus even if earlier information is discarded, the information rate for receiver k at time N_{ik} is

$$n_{ik}C_k/N_{ik} \longrightarrow C_k.$$

So capacity is achieved.

In fact, these remarks are applicable to time-invariant memoryless communication networks with, say, m senders and n receivers with arbitrary noise and feedback. Let C_{jk} be the capacity from transmitter j to receiver k when all the rest of the resources of the network are devoted to aiding the communication from j to k. The other senders will presumably act as facilitators, relays, or simply get out of the way. Then, by letting the blocklengths grow as before, the capacities C_{jk} are achieved.

By now it should be clear that the resolution of the apparent discrepancy in capacity regions is that the time at which the information becomes available is different for each transmitter–receiver pair. Capacity is ϵ-achieved at a different subset of times for each receiver.

If, however, we had asked for the set of achievable rates $\{R_{jk}\}$ for block n-codes with probability of error $P_e^{(n)}(j, k) \longrightarrow 0$, we would be confined to the classical capacity region. The resolution, then, is that the capacity region is the set of rates that can be achieved simultaneously.

VI. Concluding Remarks

One of the coding ideas used in achieving good rate regions is superposition, in which one layers, or superimposes, the information intended for each of the receivers. The receiver can then peel off the information in layers. To achieve superposition, one introduces auxiliary random variables that act as virtual signals. These virtual signals participate in the construction of the code, but are not actually sent. One useful idea used in the proof of capacity for the deterministic broadcast channel is random binning of the outputs Y_1 and Y_2. Another technique is Marton's introduction of correlated auxiliary random variables. Marton's region is the largest known achievable rate region for the general broadcast channel, but the capacity region remains unknown.

Acknowledgment

The comments of A. El Gamal and J. Thomas have been very helpful.

References

[1] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Inform Theory* (Tsahkadsor, Armenian S.S.R., 1971). Budapest, Hungary: Hungarian Acad. Sciences, 1973, pp. 23–52.
[2] ——, "The capacity region of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, pp. 805–814, Oct. 1974.
[3] R. F. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 629–637, Nov. 1975.
[4] M. Aref, "Information flow in relay networks," Ph.D. dissertation, Stanford Univ., Stanford, CA, Oct. 1980.
[5] L. A. Bassalygo, "Bound for permissible transmission rates in a broadcast channel with errors in both components," *Probl. Pered. Inform.*, vol. 17, no. 4, pp. 19–28, Oct.–Dec. 1981; translated in *Probl. Inform. Transm.*, vol. 17, no. 4, pp. 228–236, Oct.–Dec. 1981.
[6] L. A. Bassalygo and M. S. Pinsker, "Bounds on volume of error-correcting codes in a broadcast channel (generalized Blackwell channel)," *Probl. Pered. Inform.*, vol. 16, no. 3, pp. 3–16, July–Sept. 1980; translated in *Probl. Inform. Transm.*, vol. 16, no. 3, pp. 163–174, July–Sept. 1980.
[7] ——, "A comment on the paper of Kasami, Lin, Wei, and Yamamura, 'Coding for the binary symmetric broadcast channel with two receivers,'" *Probl. Pered. Inform.*, vol. 24, no. 3, pp. 102–106, July–Sept. 1988; translated in *Probl. Inform. Transm.*, vol. 24, no. 3, pp. 253–257, July–Sept. 1988.
[8] R. Benzel, "The capacity region of a class of discrete additive degraded interference channels," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 228–231, Mar. 1979.
[9] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 197–207, Mar. 1973.
[10] ——, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 279–280, Mar. 1974.
[11] P. P. Bergmans and T. M. Cover, "Cooperative broadcasting," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 317–324, May 1974.
[12] R. E. Blahut, *Principles and Practice of Information Theory.* Reading, MA: Addison-Wesley, 1987.
[13] L. Boutsikaris and R. Narayan, "An achievable rate region for a broadcast channel and its implications on signal design" in *IEEE Int. Symp. Information Theory* (St. Jovite, Que., Canada, Sept. 1983), p. 62.
[14] A. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 387–390, May 1977.
[15] S. C. Chang, "Coding for T-user binary symmetric broadcast channels," in *IEEE Int. Symp. Information Theory* (Santa Monica, CA, Feb. 1981), p. 34.
[16] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 2–14, Jan. 1972.
[17] ——, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 226–278, Mar. 1975.
[18] ——, "An achievable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 399–404, July 1975.
[19] ——, "Some advances in broadcast channels," in *Advances in Communication Systems, Vol. 4, Theory and Applications*, A. Viterbi, Ed. New York: Academic, 1975, ch. 4.
[20] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. IT-25, Sept. 1979.
[21] T. M. Cover and S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 292–298, Mar. 1981.
[22] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley, 1991.

[23] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic, 1981.

[24] ——, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 339–348, May 1978.

[25] H. De Pedro, C. Heegard, and J. K. Wolf, "Permutation codes for the Gaussian broadcast channel," in *IEEE Int. Symp. Information Theory* (Ronneby, Sweden, June 1976), pp. 64–65.

[26] R. L. Dobrushin, "Survey of Soviet research in information theory," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 703–724, Nov. 1972.

[27] C. P. Downey and J. K. Karlof, "Group codes for the Gaussian broadcast channel with two receivers," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 406–411, July 1980.

[28] ——, "Group codes for the M-receiver Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 595–597, July 1983; also, in *IEEE Int. Symp. Information Theory* (Santa Monica, CA, Feb. 1981), pp. 34–35.

[29] G. Dueck, "Partial feedback for two-way and broadcast channels," *Inform. Contr.*, vol. 46, no. 1, pp. 1–15, July 1980.

[30] A. El Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 379–381, May 1978.

[31] ——, "Broadcast channels with and without feedback," in *11th Annu. Asilomar Conf. Circuits, Systems and Computers* (Pacific Grove, CA, Nov. 1977), pp. 180–183.

[32] ——, "The capacity of a class of broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 166–169, Mar. 1979.

[33] ——, "Capacity of the product and sum of two unmatched broadcast channels," *Probl. Pered. Inform.*, vol. 16, no. 1, pp. 3–23, Jan.–Mar. 1980; translated in *Probl. Inform. Transm.*, vol. 16, no. 1, pp. 1–16, Jan.–Mar. 1980.

[34] ——, "The capacity of the physically degraded Gaussian broadcast channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-27, p. 508, July 1981.

[35] A. El Gamal and T. M. Cover, "Multiple user information theory," *Proc. IEEE*, vol. 68, pp. 1466–1483, Dec. 1980.

[36] A. El Gamal and E. Van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 120–122, Jan. 1981.

[37] P. Estabrook and B. Jacobs, "Optimization of direct broadcast satellite system channel capacity," in *Int. Conf. Communications: Integrating Communication for World Progress (ICC'83)* (Boston, MA, June 1983), vol. 3, pp. 1416–1420.

[38] G. D. Forney, "Information theory," unpublished course notes, 1972.

[39] P. Gacs and J. Korner, "Common information is much less than mutual information," *Probl. Contr. Inform. Theory*, vol. 2, no. 2, pp. 149–162, 1973.

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

[41] ——, "Capacity and coding for degraded broadcast channels," *Probl. Pered. Inform.*, vol. 10, no. 3, pp. 3–14, July–Sept. 1974; translated in *Probl. Inform. Transm.*, pp. 185–193, July–Sept. 1974.

[42] S. I. Gelfand, "Capacity of one broadcast channel," *Probl. Pered. Inform.*, vol. 13, no. 3, pp. 106–108, July–Sept. 1977; translated in *Probl. Inform. Transm.*, pp. 240–242, July–Sept. 1977.

[43] S. I. Gelfand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Pered. Inform.*, vol. 16, no. 1, pp. 24–34, Jan.–Mar. 1980; translated in *Probl. Inform. Transm.*, vol. 16, no. 1, pp. 17–25, Jan.–Mar. 1980.

[44] ——, "Coding for channel with random parameters," *Probl. Contr. Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[45] L. Georgiadis and P. Papantoni-Kazakos, "Limited feedback sensing algorithms for the broadcast channel," in *IEEE Int. Conf. Communications* (Chicago, IL, June 1985), vol. 1, pp. 52–55.

[46] ——, "Limited feedback sensing algorithms for the packet broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 280–294, Mar. 1985.

[47] A. Goldsmith, "Capacity and dynamic resource allocation in broadcast fading channels," in *Proc. 33rd Annu. Allerton Conf. Communication, Control, and Computing*, (Monticello, IL, Oct. 1995), pp. 915–924.

[48] ——, "Capacity of broadcast fading channels with variable rate and power," in *IEEE GLOBECOM 1996, Communications: The Key to Global Prosperity* (London, UK, Nov. 1996), pp. 92–96.

[49] A. Goldsmith and M. Effros, "The capacity region of broadcast channels with memory," in *Proc. IEEE Int. Symp. Inform. Theory* (Ulm, Germany, June–July 1997), p. 28.

[50] I. Gopal and R. Rom, "Multicasting to multiple groups over broadcast channels," *IEEE Trans. Commun.*, vol. 42, pp. 2423–2431, July 1994.

[51] B. E. Hajek and M. B. Pursley, "Evaluation of an achievable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 36–46, Jan. 1979; also in *Proc. Int. Conf. Communications* (Chicago, IL, June 1977), pp. 249–253.

[52] T. S. Han, "The capacity region for the deterministic broadcast channel with a common message," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 122–125, Jan. 1981.

[53] ——, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 49–60, Jan. 1981.

[54] T. S. Han and M. H. M. Costa, "Broadcast channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 641–650, Sept. 1987.

[55] C. Heegard, H. E. DePedro, and J. K. Wolf, "Permutation codes for the Gaussian broadcast channel with two receivers," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 569–578, Sept. 1978.

[56] C. Heegard and A. El Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 731–739, Sept. 1983.

[57] D. Hughes-Hartogs, "The capacity of the degraded spectral Gaussian broadcast channel," Ph.D. dissertation, Inform. Syst. Lab., Stanford Univ. Stanford, CA, Tech. Rep. 7002-2, July 1975; also in *Proc. Nat. Telecommun. Conf.* (Dallas, TX, Dec. 1976), p. 52.1/1–5.

[58] J. H. Jahn, "Coding of arbitrarily varying multiuser channels," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 212–226, Mar. 1981.

[59] T. Kasami, S. Lin, V. K. Wei, and S. Yamamura, "Coding for the binary symmetric broadcast channel with two receivers," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 616–625, Sept. 1985; also in *Proc. IEEE Int. Symp. Information Theory* (St. Jovite, Que., Canada, Sept. 1983), p. 61.

[60] C. Keilers, "The capacity of the spectral Gaussian multiple-access channel," Ph.D. dissertation, Dept. Elec. Eng., Stanford Univ., Stanford, CA, Apr. 1976.

[61] R. King, "Multiple access channels with generalized feedback," Ph.D. dissertation, Stanford Univ., Stanford, CA, Mar. 1978.

[62] J. Körner, "Some methods in multi-user communication," in *Information Theory, New Trends and Open Problems*, G. Longo, Ed. Vienna, Austria: Springer-Verlag, 1976, pp. 172–176.

[63] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 60–64, Jan. 1977.

[64] ——, "Comparison of two noisy channels," *Coll. Math. Sco. F. Boliya*, vol. 16, in *Topics of Information Theory.* Amsterdam, The Netherlands: North Holland, 1977, pp. 411–423.

[65] J. Körner and A. Sgarro, "Universally attainable error exponents for broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 670–679, Nov. 1980.

[66] B. D. Kudryashov, "Upper bounds for decoding error probability in some broadcast channels," *Probl. Pered. Inform.*, vol. 15, no. 3, pp. 3–17, July–Sept. 1979; translated in *Probl. Inform. Transm.*, pp. 163–174, July–Sept. 1979.

[67] B. D. Kudryashov and G. S. Poltyrev, "Upper bounds for decoding error probability in some broadcast channels," *Probl. Pered. Inform.*, vol. 15, no. 3, pp. 3–17, July–Sept. 1979; translated in *Probl. Inform. Transm.*, vol. 15, no. 3, pp. 163–174, July–Sept. 1979.

[68] A. V. Kuznetsov, "Use of defect-correcting codes to restrict access to a broadcast channel," *Probl. Pered. Inform.*, vol. 17, no. 4, pp. 3–10, Oct.–Dec. 1981; translated in *Probl. Inform. Transm.*, pp. 217–222, Oct.–Dec. 1981.

[69] W. J. Leighton and H. H. Tan, "Broadcast channels: Some recent developments and open problems," in *Proc. Int. Telemetering Conf.* (Los Angeles, CA, Sept. 1976), pp. 284–293.

[70] ——, "Trellis coding for degraded broadcast channels," *Inform. Contr.*, vol. 39, no. 2, pp. 119–134, Nov. 1978; and in *Proc. IEEE Int. Symp. Inform. Theory* (Ithaca, NY, Oct. 1977), p. 54.

[71] ——, "Capacity region of degraded broadcast channels with feedback," *Inform.ation Sci.*, vol. 13, no. 2, pp. 167–177, 1977.

[72] C. Leung, "On the capacity of an n-receiver broadcast channel with partial feedback," in *IEEE Int. Symp. Inform. Theory* (Santa Monica, CA, Feb. 1981), p. 27.

[73] H. Liao, "A coding theorem for multiple access communication," in *Int. Symp. Information Theory* (Asilomar, 1972); also "Multiple access channels," Ph.D. dissertation, Dept. Elec. Eng., Univ. Hawaii, Honolulu, HI, 1972.

[74] K. Marton, "The capacity region of deterministic broadcast channels," in *Trans. Int. Symp. Inform. Theory* (Paris-Cachan, France, 1977).

[75] ——, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 306–311, May 1979.

[76] R. J. McEliece and L. Swanson, "A note on the wide-band Gaussian broadcast channel," *IEEE Trans. Commun.*, vol. COM-35, pp. 452–453, Apr. 1987.

[77] F. Morii, H. Nakanishi, and Y. Tezuka, "Error probabilities for Gaussian broadcast channel," *Electron. and Commun. Japan*, vol. 59, no. 6, pp. 18–26, June 1976.

[78] M. Ohkubo, "A coding theorem for broadcast channels with degraded components," *Trans. Inst. Electron. Commun. Eng. Japan*, Section E (English), vol. E61, no. 12, p. 979, Dec. 1978.

[79] J. M. Ooi and G. W. Wornell, "Decentralized control of a multiple access broadcast channel: Performance bounds," in *Proc. 35th IEEE Conf. Decision and Control* (Kobe, Japan, Dec. 1996), vol. 1, pp. 293–298.

[80] L. Ozarow, "Coding and capacity for additive white Gaussian noise multi-user channels with feedback," Ph.D. dissertation, Dept. Elec. Eng. and Comp. Sci., MIT, Cambridge, MA, 1983.

[81] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 667–671, July 1984.

[82] M. S. Pinsker, "Capacity of noiseless broadcast channels," *Probl. Pered. Inform.*, vol. 14, no. 2, pp. 28–34, Apr.–June 1978; translated in *Probl. Inform. Transm.*, pp. 97–102, Apr.–June 1978.

[83] N. Pippenger, "Bounds on the performance of protocols for a multiple-access broadcast channel," *IEEE Trans. Inform.Theory*, vol. IT-27, pp. 145–151, Mar. 1981.

[84] G. S. Poltyrev, "Carrying capacity for parallel broadcast channels with degraded components," *Probl. Pered. Inform.*, vol. 13, no. 2, pp. 23–35, Apr.–June 1977; translated in *Probl. Inform. Transm.*, pp. 97–107, Apr.–June 1977.

[85] ——, "Capacity for a sum of broadcast channels," *Probl. Inform Transm.*, vol. 15, no. 2, pp. 111–114, Apr.–June 1979 (English translation).

[86] ——, "On accuracy of random coding bounds for broadcast channels," *Probl. Contr. Inform. Theory*, vol. 11, no. 5, pp. 1–11, 1982 (English translation of a Russian paper).

[87] ——, "Random-coding bounds for some broadcast channels," *Probl. Pered. Inform.*, vol. 19, no. 1, pp. 9–20, Jan.–Mar. 1983; translated in *Probl. Inform. Transm.*, vol. 19, no. 1, pp. 6–16, Jan.–Mar. 1983.

[88] H. Sari, G. Karam, and I. Jeanclaud, "Frequency-domain equalization of mobile radio and terrestrial broadcast channels," in *Proc. IEEE GLOBECOM* (San Francisco, CA, Nov.–Dec. 1994), vol. 1, pp. 1–5.

[89] H. Sato, "An outer bound to the capacity region of broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-24, p. 374, May 1978.

[90] S. Shamai, "A broadcast strategy for the Gaussian slowly fading channel," in *Proc. IEEE Int. Symp. Information Theory* (Ulm, Germany, June–July 1997), p. 150.

[91] C. E. Shannon, "The zero-error capacity of a noisy channel," *IRE Trans. Inform. Theory*, vol. IT-2, pp. 8–19, 1956.

[92] ——, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Prob. Stat.* Berkeley, CA: Univ. California Press, 1961, vol. 1, pp. 611–644.

[93] B. D. Sharma and V. Priya, "On broadcast channels with side information under fidelity criteria," *Kybernetika*, vol. 19, no. 1, pp. 27–41, 1983.

[94] D. Slepian and J. K. Wolf, "Noiseless coding of corelated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471–480, July 1973.

[95] B. S. Tsybakov and V. A. Mikhailov, "Free synchronous packet access in a broadcast channel with feedback," *Probl. Pered. Inform.*, vol. 14, no. 4, pp. 32–59, Oct.–Dec. 1978; translated in *Probl. Inform. Transm.*, pp. 259–280, Oct.–Dec. 1978.

[96] D. N. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *Proc. IEEE Int. Symp. Information Theory* (Ulm, Germany, June–July 1997), p. 27.

[97] E. C. Van der Meulen, "Random coding theorems for the general discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 180–190, Mar. 1975.

[98] ——, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 1–37, Jan. 1977.

[99] ——, "Recent coding theorems for multi-way channels. Part I: The broadcast channel (1976–1980)," in *New Concepts in Multi-User Communication (NATO Advanced Study Institute Series)*, J. K. Skwyrzinsky, Ed. Groningen, The Netherlands: Sijthoff & Noordhof, 1981.

[100] ——, "Recent coding theorems and converses for multi-way channels. Part II: The multiple access channel (1976–1985)," Department Wiskunde, Katholieke Universiteit Leuven, Leuven, Belgium, 1985.

[101] M. Van Dijk, "The binary symmetric broadcast channel with confidential messages, with tampering," in *Proc. 1995 IEEE Int. Symp. Information Theory* (Whistler, BC, Canada, Sept. 1995), p. 487.

[102] ——, "On a special class of broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 43, pp. 712–714, Mar. 1997.

[103] P. Vanroose and E. C. Van der Meulen, "A new proof of the zero-error capacity region of the Blackwell broadcast channel," in *Proc. 10th Symp. Information Theory in the Benelux* (Houthalen, Belgium, May 1989), pp. 37–43.

[104] B. Verboven and E. Van der Meulen, "Capacity bounds for identification via broadcast channels that are optimal for the deterministic broadcast channel," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1197–1205, Nov. 1990.

[105] N. D. Vvedenskaya and B. S. Tsybakov, "Stack algorithm in a broadcast channel with capture," *Probl. Pered. Inform.*, vol. 27, no. 2, pp. 25–34, Apr.–June 1991; translated in *Probl. Inform. Transm.*, pp. 107–116, Apr.–June 1991.

[106] F. M. J. Willems, "The maximal-error and average-error capacity region of the broadcast channel are identical," *Probl. Contr. Inform. Theory—Probl. Upravl. I Teorii Inform.*, vol. 19, no. 4, pp. 339–347, 1990.

[107] J. Wolf, "Multi-user communication networks," in *Commun. Syst. and Random Process Theory*, J. K. Skwirzynski, Ed. Alphen aan den Rijn, The Netherlands: Sijthoff & Norrdhoff, 1978, pp. 37–53.

[108] A. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 2–10, Jan. 1974.