Broadcast Channels

Arvind Sridharan

1 Introduction

The broadcast situation as the name suggests involves the simultaneous communication of information from one sender to multiple receivers as shown[1] in Figure 1. The goal is to find the capacity region, i.e., the set of simultaneously achievable rates (R_1, R_2). TV and radio transmissions, lecturing in a classroom are some examples of broadcasting.

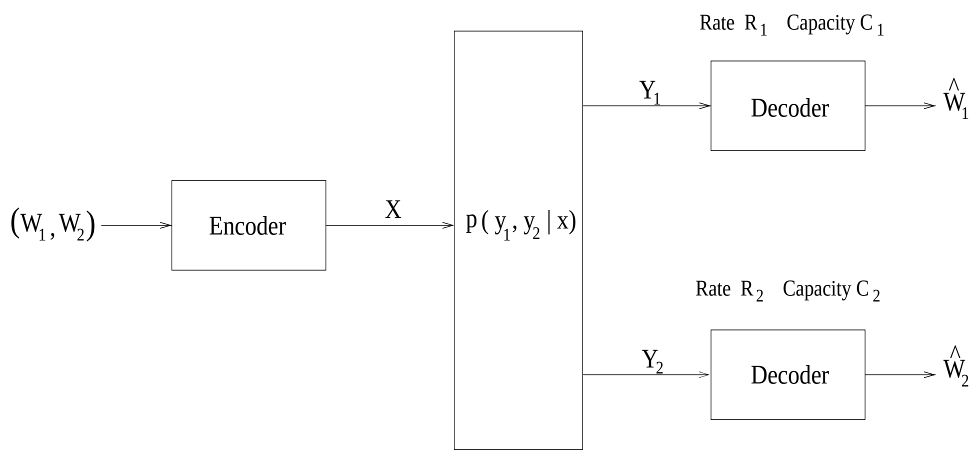

Figure 1: The Broadcast channel

Intuitively, it is clear that it is possible to transmit to both receivers at a rate equal to the minimum of the two capacities, C_1 and C_2, i.e., the transmission rate is limited by the worst channel. At the other extreme we could transmit on the best channel at a rate equal to its capacity and transmit no information on the other channel. With time-sharing, assuming $C_1 \leq C_2$, the rates $R_1 = \alpha C_1$ and $\alpha C_1 + (1-\alpha)C_2$ can be achieved. In [1] Cover showed that it is possible to do better than time-sharing. This idea is discussed in Section 3. The capacity region for the broadcast channel is as yet unknown. The best known achievable rate region was obtained in [2] and is described in Section 4. The formal definitions for the broadcast channel are discussed in the next section. The summary is based on [3] and the review paper [4].

[1] In general, we can have $k \geq 2$ receivers.

2 Broadcast Channel

Definition 1: A *broadcast channel* (BC) consists of an input alphabet $\mathscr{X}$ and two output alphabets $\mathscr{Y}_1$ and $\mathscr{Y}_2$ and a probability transition function $p(y_1, y_2|x)$. For a memoryless BC, $p(y_1^n, y_2^n|x^n) = \prod_{i=1}^n p(y_{1i}, y_{2i}|x_i)$.

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

$$X : (1, 2, \cdots, 2^{nR_1} \times 1, 2, \cdots, 2^{nR_2}) \longrightarrow \mathscr{X}^n \tag{1}$$

and two decoders,

$$g_i : \mathscr{Y}_i^n \longrightarrow 1, 2, \cdots, 2^{nR_i}, \quad i = 1, 2 \tag{2}$$

The average probability of error is defined as

$$P_e^{(n)} = P(g_1(Y_1^n) \neq W_1 \ or \ g_2(Y_2^n) \neq W_2), \tag{3}$$

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

A rate pair (R_1, R_2) is said to be achievable for the BC channel if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \to 0$. The capacity region is the closure of the set of achievable rates.

Often we are also interested in transmitting information common to both receivers in addition to the private information of each receiver. The definition of codes, average probability of error and the set of achievable rate triples (the extra rate component corresponds to the rate at which common information is transmitted) are analogous.

Finally, note that the event $E_1 = \{g_1(Y_1^n) \neq W_1\}$ and the event $E_2 = \{g_2(Y_2^n) \neq W_2\}$ imply the event $E = \{g_1(Y_1^n) \neq W_1 \ or \ g_2(Y_2^n) \neq W_2)\}$, hence $P(E_1) \leq P(E)$ and $P(E_2) \leq P(E)$. Also $P(E) \leq P(E_1) + P(E_2)$ by the union bound. This implies $P(E) \to 0 \Leftrightarrow P(E_1) \to 0$ and $P(E_2) \to 0$. Therefore the capacity region depends only on the conditional distributions $p(y_1|x)$ and $p(y_2|x)$.

3 Physically Degraded BC

Definition 3: A BC channel is said to be *physically degraded* if $p(y_1, y_2|x) = p(y_1|x)p(y_2|x)$.

For the physically degraded BC channel, as shown in Figure 2 the channel from the transmitter to the second receiver is the cascade of the channel from the transmitter to the first receiver and the channel from the first receiver to the second receiver. This provides the motivation for *superposition coding*

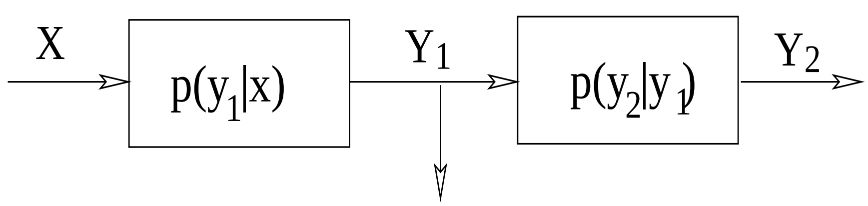

Figure 2: The Broadcast channel

proposed in [1].The main idea is to superpose the message intended for the better receiver on the poorer

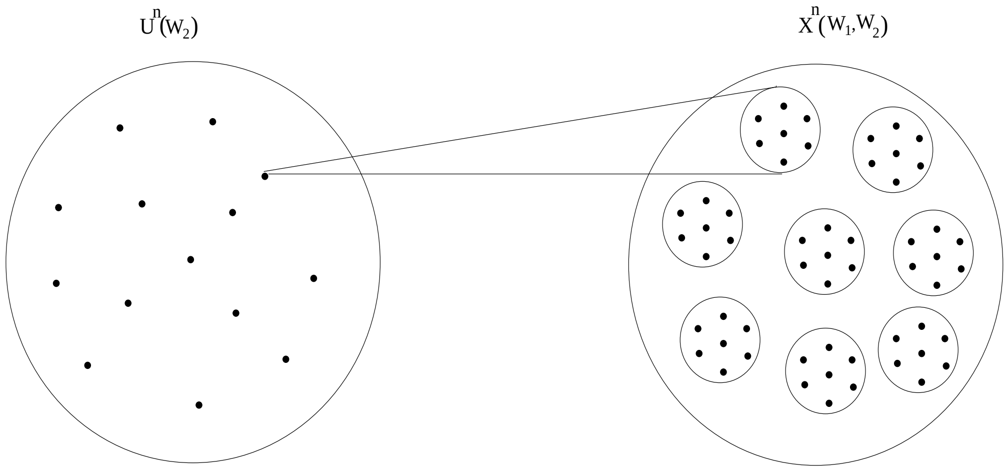

Figure 3: Superposition coding

receiver's message. The better receiver first determines the poorer receiver's message and then its own message. Cover [1] showed that in the case of binary symmetric BC and AWGN BC superposition coding expands the rate region beyond that achievable with time-sharing. We note that a binary symmetric BC and an AWGN BC can always be converted into their physically degraded counterparts. Later, in [5] the same technique was used to improve the achievable rate region for any physically degraded BC. Roughly, a year later the converse was proved [6][7] thus establishing the capacity region for the physically degraded BC channel.

Theorem 1: The capacity region for sending independent information over the degraded BC $X \rightarrow Y_1 \rightarrow Y_2$ is the convex hull of the closure of all (R_1, R_2) satisfying

$$\begin{aligned} R_2 &\leq I(U; Y_2) \\ R_1 &\leq I(X; Y_1|U) \end{aligned} \tag{4}$$

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

Sketch of Proof: The complete details maybe found in [3].

Achievability

The idea of superposition coding is captured in Figure 3. The auxiliary random variable U serves as a cloud center distinguishable by both receivers . Each cloud consists of 2^{nR_1} codewords X^n distinguishable by the receiver Y_1. The worst receiver only sees the clouds while the better receiver can see the individual codewords within the clouds.

Codebook Generation: We first generate 2^{nR_2} cloud centers, $u^n(w_2)$ according to $\prod_{i=1}^{n} p(u_i)$. For each cloud center, we then generate 2^{nR_1} codewords $x^n(w_1, w_2)$ according to $\prod_{i=1}^{n} p(x_i|u_i(w_2))$.

Encoding: To transmit the message (w_1, w_2) the corresponding codeword $x^n(w_1, w_2)$ is sent.

Decoding: The decoders perform the usual typical set decoding based on their received sequences y_1^n and y_2^n respectively. The second decoder tries to determine the codeword $u^n(w_2)$ and the first receiver determines the codeword $x^n(w_1, w_2)$.

The best possible rate for the poorer user, i.e., the second user, is the capacity of the single user channel from U to Y_2. The physically degraded nature of the channel implies that the better i.e., first receiver can determine the cloud center. Hence, the rate achievable is the capacity of the single user channel from X to Y_1, with perfect information of the auxiliary variable U.

Converse

The key element here is to define the auxiliary random variable U in terms of outputs upto the present time. The other steps essentially follow the single user converse. □

Remarks

1. In [5] it was shown that if a set of rates were achievable under the average probability of error criterion then could also be achieved under the maximum probability of error criterion.
2. The bounds on the cardinality of the auxiliary random variable U were established in [7].
3. The idea of superposition appears in several other contexts in information theory for example capacity calculations of MAC channels and in multilevel coding.
4. In [8] it was shown that in the case of multiple transmitter/receivers operating on a bandlimited AWGN channel, superposition coding can do better than both FDMA and TDMA.
5. The idea of superposition is also useful in the case of compound channels, i.e., channels with a randomly varying transition matrices.

Example: The Gaussian Channel

As an application of Theorem 1 consider the physically degraded Gaussian BC channel

$$\begin{aligned} Y_1 &= X + Z_1 \\ Y_2 &= X + Z_2 = Y_1 + Z_2' \end{aligned}$$

where $Z_1 \sim \mathscr{N}(0, N_1)$ and $Z_2' \sim \mathscr{N}(0, N_2 - N_1)$ with a power constraint P. The capacity region for the channel is given by

$$\begin{aligned} R_1 &\leq C(\frac{\alpha P}{N_1}) \\ R_2 &\leq C(\frac{(1-\alpha)P}{\alpha P + N_2}) \end{aligned} \tag{5}$$

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

The poorer receiver with power $(1-\alpha)P$ decodes in the presence of his ambient noise N_2 and also the "corruption" in X due to part of the power being used to communicate to receiver 1. However, the better receiver can decode the message intended for the poorer receiver and hence only has to combat a noise power of N_1.

4 Marton's Lower Bound on the Achievable Region

4.1 The Deterministic Broadcast Channel

Theorem 2: The capacity region of the deterministic memoryless BC with $y_1 = f_1(x),\ y_2 = f_2(x)$, is given by the convex closure of the union of the rate pairs satisfying

$$\begin{aligned} R_1 &\leq H(Y_1) \\ R_2 &\leq H(Y_2) \\ R_1 + R_2 &\leq H(Y_1, Y_2) \end{aligned} \tag{6}$$

Sketch of Proof: The idea is to do a product binning of the y_1^n and y_2^n sequences into 2^{nR_1} and 2^{nR_2} bins respectively. Suppose if we were to find a jointly typical (y_1^n, y_2^n) in bin (i, j) say. Then to send the message (i, j), we find the sequence x^n which results in y_1^n and y_2^n respectively. This is possible because y_1^n and y_2^n are deterministic functions of x^n. Note that the number of typical y_1^n sequences, typical y_2^n sequences and jointly typical sequences (y_1^n, y_2^n) are $nH(Y_1)$, $nH(Y_2)$ and $nH(Y_1, Y_2)$ respectively. The achievability of the given rates now follows. □

Remarks This result was obtained independently by Marton and Pinsker[9]. The capacity region has a complementary relationship to the Slepian-Wolf region. The sketch of the proof above is based on [4].

4.2 Marton's Theorem

Theorem 5: The rates (R_1, R_2) are achievable for the BC channel $\{\mathscr{X},\ p(y_1, y_2|x), \mathscr{Y}_1 \times \mathscr{Y}_2\}$ if

$$\begin{aligned} R_1 &\leq I(U; Y_1) \\ R_2 &\leq I(V; Y_2) \\ R_1 + R_2 &\leq I(U; Y_1) + I(V; Y_2) - I(U, V) \end{aligned} \tag{7}$$

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

Sketch of Proof: The idea is to send the auxiliary variables u to y_1 and v to y_2. As with the deterministic BC we randomly throw the u's into 2^{nR_1} bins and the $v's$ into 2^{nR_2} bins. The final condition of the theorem ensures the existence of atleast one jointly typical (u, v) in each bin. For each bin and the corresponding jointly typical (u^n, v^n) pair the codeword $x^n(u^n, v^n)$ is generated according to $\prod_{k=1}^n p(x_k|u_k, v_k)$. The first two conditions then ensure that we can find jointly typical pairs (u^n, y_1^n) and (v^n, y_2^n) at the two decoders. □

Remarks The above theorem was proved by Marton [2]. Later, a simplified proof was given in [10]. The sketch of the proof above is based on [4].

Further Reading

A comprehensive survey of the broadcast channel can be found in [4]. This review paper also has an exhaustive list of references.

References

[1] T. M. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. IT-18, pp. 2–14, January 1972.

[2] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Transactions on Information Theory*, vol. IT-25, pp. 306–311, May 1979.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley Interscience, 1991.

[4] T. M. Cover, "Comments on broadcast channels," *IEEE Transactions on Information Theory*, vol. IT-44, pp. 2524–2530, October 1998.

[5] P. P. Bergmans, "Random coding theorems for broadcast channels with degraded components," *IEEE Transactions on Information Theory*, vol. IT-19, pp. 197–207, March 1973.

[6] P. P. Bergmans, "A simple converse for broadcast channels with additive white gaussian noise," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 279–280, March 1974.

[7] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Infor. Transm.*, pp. 185–193, July-Sept 1974.

[8] P. P. Bergmans and T. M. Cover, "Cooperative broadcasting," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 279–280, May 1974.

[9] M. S. Pinsker, "Capacity of noiseless broadcast channels," *Probl. Infor. Transm.*, pp. 97–102, Apr-Jun 1978.

[10] A. E. Gamal and E. Van der Meulen, "A proof of marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Transactions on Information Theory*, vol. IT-27, pp. 120–122, January 1981.