The Capacity Region of Multiple Input Erasure Broadcast Channels

Amir F. Dana[1]
Electrical Eng. Department
Caltech, MC 136-93
Pasadena, CA 91125
Email: dana@systems.caltech.edu

Babak Hassibi[1]
Electrical Eng. Department
Caltech, MC 136-93
Pasadena, CA 91125
Email: hassibi@systems.caltech.edu

***Abstract*— In this paper, we look at the capacity region of a special class of broadcast channels with multiple inputs at the transmitter and a number of receivers. The channel between an input of the transmitter and a receiver is modelled as an independent memoryless erasure channel. We assume that the signals coming from different inputs to the receiver do not interfere with each other. Also for each input, the transmitter sends the same signal through the channels outgoing from that input. This class of broadcast channels does not necessarily belong to the class of "more capable". We will show that the capacity region of these broadcast channels is achieved by time-sharing between the receivers at each input. Finally, the implications of these results to the more general network setup are discussed.**

I. Introduction

Determining the capacity region of general multi-terminal networks is still an open problem. Even for the simplest networks such as the single relay channel and the broadcast channel, the capacity (region) is not known in general. Broadcast channels are used for modelling the communication between one sender and a number of receivers. For some special classes of broadcast channels, e.g. "degraded", "more capable", or "less noisy", the capacity region has been determined (see [5],[6] and references therein). Essentially, in these cases the receivers can be sorted according to their "quality of reception". The capacity region then has a single letter characterization in terms of the input, the output and a number of auxiliary random variables. Recently the capacity region of the Gaussian MIMO broadcast channels has been found in [1]. The Gaussian MIMO broadcast channels are not degraded in general.

In this paper, we will look at a class of broadcast channels, called erasure broadcast channels, with multiple inputs at the transmitter and a number of receivers. The channel between each input and each of the receivers is modelled by independent memoryless erasure channels. This broadcast system is not always "degraded" or "more capable". We will find necessary and sufficient conditions so that the channel belongs to one of the known classes of "more capable", "less noisy" or "degraded". In fact we will see that for erasure broadcast channels these definitions coincide. For the case of a single input transmitter, it is shown in [3],[7] that the channel is degraded and the capacity region is given by time-sharing between the receivers. We will show that the capacity region of the general erasure broadcast channel is achieved by time-sharing between the receivers at each input.

[1]This work was supported in part by the National Science Foundation under grant no. CCR-0133818, by David and Lucille Packard Foundation, and by Caltech's Lee Center for Advanced Networking.

II. Problem Formulation

In this paper, we consider erasure broadcast channels with multiple inputs at transmitter. The results of this paper hold for arbitrary alphabet size. However, in this presentation we work with binary alphabets.

Definition 1. *An (m,n)-erasure broadcast channel with erasure matrix $\underline{\epsilon}$ (see Fig. 1) has m inputs at the transmitter and n receivers. Moreover,*

- *The channel between i-th input of the transmitter and receiver j is modelled as a memoryless erasure channel with erasure probability given by the i,j coordinate of the erasure matrix $\underline{\epsilon}$. The output of this channel is denoted by Y_{ij}. Furthermore, different channels are independent from each other.*
- *The transmitter sends out the same signal X_i (chosen from alphabet $\mathcal{X} = \{0,1\}$) through the channels going out from each input i.*
- *There is no interference among the signals coming through different channels to the receivers. $\underline{Y}_j$ denotes the collection of the signals received at receiver i from all its incoming channels, i.e., $\underline{Y}_j = (Y_{1j}, \dots, Y_{mj})$.*

The transition probability of the channel can be written as

$$\begin{aligned} &\Pr(\underline{Y}_1 = \underline{y}_1, \dots, \underline{Y}_n = \underline{y}_n | X_1 = x_1, \dots, X_m = x_m) \\ = &\prod_{i=1}^{m}\prod_{j=1}^{n} \Pr{}_{ij}(Y_{ij} = y_{ij} | X_i = x_i), \end{aligned}$$

where $\Pr_{ij}(\cdot|\cdot)$ is the transition probability of a memoryless erasure channel with probability of erasure ϵ_{ij}.

We are interested in the capacity region of a general (m,n)-erasure broadcast channel with erasure matrix $\underline{\epsilon}$. Each of the receivers request an independent information. A $(\lceil 2^{TR_1} \rceil, \dots, \lceil 2^{TR_n} \rceil, T)$ code for an (m,n)-erasure broadcast channel consists of the following components:

- A set of integers $\mathcal{W}^j = \{1, 2, \ldots, \lceil 2^{TR_j} \rceil\}$ that represent the message indices corresponding to the information that is intended for receiver j. We assume that all the messages are equally likely and independent from each other.
- An encoding function for the transmitter: $f : \prod_{j=1}^{n} \mathcal{W}^{(j)} \to \prod_{i=1}^{m} \mathcal{X}^T$, that gives the signals transmitted from the m inputs for any given set of messages.
- A decoding function g_j at receiver j that maps the received signals to $\mathcal{W}^j$. $g_j(\underline{y}_j^T)$ is the estimate of the message sent from the transmitter based on the received signal $\underline{Y}_j^T$.

We define the probability of error as the probability that the decoded message at one of the receivers is not equal to the transmitted message, i.e.,

$$P_{err} = \Pr\left(\exists 1 \le j \le n : g_j(\underline{y}_j^T) \neq W^{(j)}\right) \quad (1)$$

The set of rates $(R_1, R_2, \ldots, R_n)$ is said to be achievable if there exist a sequence of $(\lceil 2^{TR_1} \rceil, \ldots, \lceil 2^{TR_n} \rceil, T)$ codes such that $P_{err} \to 0$ as $T \to \infty$. The capacity region, $\mathcal{C}_g$, is the set closure of the set of achievable rates.

III. Time-division Achievable Region

In this section, we look at the achievable region of the time-division scheme. In this scheme, the i-th input of the transmitter, allocates α_{ij} of the time to transmit to receiver j. The total amount of information transmitted to the j-th receiver, R_j, is

$$R_j = \sum_{i=1}^{m} \alpha_{ij}(1 - \epsilon_{ij}).$$

Therefore, the achievable rate region of the time-division scheme is given by

$$\mathcal{R}_T = \bigcup_{\underline{\alpha}} \{(R_1, \ldots, R_n) | 0 \le R_j < \sum_{i=1}^{m} \alpha_{ij}(1 - \epsilon_{ij})\}, \quad (2)$$

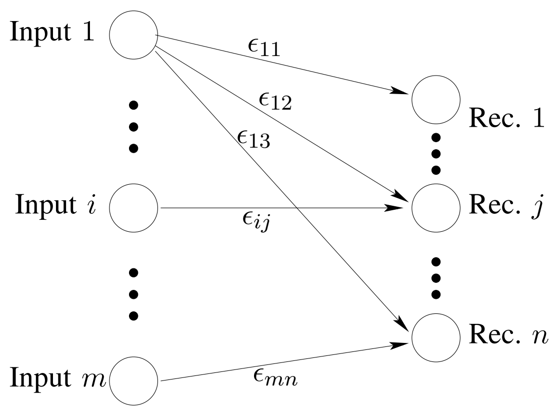

Fig. 1. An (m, n)-erasure broadcast channel with erasure matrix $\underline{\epsilon}$. The channel between input i and receiver j is modelled by a memoryless erasure channel with probability of erasure ϵ_{ij}. The transmitter transmits the same signal to all the receivers. Reception at receivers are assumed to be interference-free.

where the union is over all admissible time-sharing matrices $\underline{\alpha}$ that satisfy:

1) $\alpha_{ij} \ge 0$
2) $\sum_{j=1}^{n} \alpha_{ij} = 1$ for all inputs $1 \le i \le m$.

As mentioned earlier, we will show that the capacity region of the erasure broadcast channel is given by time-sharing schemes, i.e., $\mathcal{C}_g = \mathcal{R}_T$.

IV. Degraded Case

In this section we look at necessary and sufficient conditions so that an (m, n)-erasure broadcast channel with erasure matrix $\underline{\epsilon}$ belongs to the "degraded" class. We have stated these conditions in the following lemma.

Lemma 1. *An (m, n)-erasure channel with erasure matrix $\underline{\epsilon}$ belongs to "degraded" class iff there exists some permutation $\pi(\cdot)$ on $\{1, 2, \ldots, n\}$ such that*

$$\epsilon_{i\pi(j)} \le \epsilon_{i\pi(j+1)}, \quad \forall 1 \le i \le m,\ 1 \le j \le n.$$

Moreover the classes of "more capable", "less noisy" and "degraded" coincide for the erasure channel .

Proof: First note that "more capable" is the weakest class among the above mentioned classes. Therefore we prove that the above condition is necessary for that the channel belongs to "more capable" class.

Recalling from [6] the definition of "more capable" class,

$$I(\underline{X}; \underline{Y}_i) - I(\underline{X}; \underline{Y}_j)$$

should not change sign over all input probability distributions. Now if we consider all the inputs except the k-th to be fixed and X_k to be i.i.d Bernoulli(1/2), then we get that $\epsilon_{kj} - \epsilon_{ki}$ should not change sign for all $k \in \{1, \ldots, m\}$. Hence the condition mentioned in the theorem is necessary.

We can easily see that the condition mentioned in the theorem is sufficient for that the channel belongs to the "degraded" class. However, degradedness implies the two other classes, i.e., "more capable" and "less noisy" and we are done. $\square$

Now we show that the capacity of degraded erasure channels is given by the time-sharing scheme described in Section III. In doing so, we use of the following lemma, the proof of which we omit.

Lemma 2. *Suppose U, V, X, Y are random variables with probability distribution of the form*

$$Pr(U, V, X, Y) = Pr(U, V, X) Pr_{\epsilon}(Y|X)$$

where $Pr_{\epsilon}(.|.)$ is the transition probability of an erasure channel with probability of erasure ϵ. Then

$$I(U; Y|V) = (1 - \epsilon) I(U; X|V)$$

Theorem 1. *The capacity of a degraded (m, n)-erasure channel with erasure matrix $\underline{\epsilon}$ is given by time-sharing between the receivers at each input, i.e.,*

$$\mathcal{C}_g = \mathcal{R}_T,$$

where $\mathcal{R}_T$ is given in (2).

Proof: Without loss of generality let us assume that for $i < j$, receiver j is a degraded version of receiver i. According to [4], the capacity of the degraded broadcast channel is given by the convex hull of the closure of the $(R_1, \ldots, R_n)$ satisfying

$$0 \leq R_j \leq I(U_j; \underline{Y}_j | U_{j+1}, \ldots, U_n)$$

for $j = 1, 2, \ldots, n$, where $\underline{U} = (U_1, \ldots, U_n)$ and $\underline{U} - \underline{X} - \underline{Y}_1 - \cdots - \underline{Y}_n$ forms a Markov chain.

Now consider the following maximization problem

$$S^\star(\underline{\mu}) = \max_{(R_1,\ldots,R_n)\in\mathcal{C}_g} \sum_{j=1}^{n} \mu_j R_j$$

where $\underline{\mu} = (\mu_1, \ldots, \mu_n) \geq 0$. Note that every point on the boundary of the capacity region is the maximizing solution for some $\underline{\mu} \geq 0$. Also the maximizing solution of the above optimization problem corresponds to a boundary point of the capacity region.

Using the chain rule for mutual information we can write

$$\begin{aligned}
S^\star(\underline{\mu}) &\leq \max_{P(\underline{U},\underline{X})} \sum_{j=1}^{n} \mu_j I(U_j; \underline{Y}_j | U_{j+1}, \ldots, U_n) \qquad (3)\\
&= \max_{P(\underline{U},\underline{X})} \sum_{j=1}^{n} \mu_j I(U_j; Y_{1j}, \underline{Y}'_j | U_{j+1}, \ldots, U_n)\\
&= \max_{P(\underline{U},\underline{X})} \Bigg(\sum_{j=1}^{n} \mu_j I(U_j; \underline{Y}'_j | U_{j+1}, \ldots, U_n)\\
&+ \sum_{j=1}^{n} \mu_j I(U_j; Y_{1j} | U_{j+1}, \ldots, U_n, \underline{Y}'_j) \Bigg)\\
&= \max_{P(\underline{U},\underline{X})} \Bigg(\sum_{j=1}^{n} \mu_j I(U_j; \underline{Y}'_j | U_{j+1}, \ldots, U_n) \qquad (4)\\
&+ \sum_{j=1}^{n} \mu_j (1-\epsilon_{1j}) I(U_j; X_1 | U_{j+1}, \ldots, U_n, \underline{Y}'_j) \Bigg)
\end{aligned}$$

where $\underline{Y}'_j = (Y_{2j}, \ldots, Y_{nj})$ for all j and (4) follows from Lemma 2. Defining $\underline{X}' = (X_2, \ldots, X_m)$, it can be easily verified that $(U_1, \ldots, U_n, X_1) - \underline{X} - \underline{Y}'_1 - \cdots - \underline{Y}'_n$ forms a Markov chain.

The above Markov property implies that

$$\begin{aligned}
H(X_1 | U_j, \ldots, U_n, \underline{Y}'_{j-1}) &= H(X_1 | U_j, \ldots, U_n, \underline{Y}'_j) \qquad (5)\\
&- I(\underline{Y}'_{j-1}; X_1 | U_j, \ldots, U_n, \underline{Y}'_j)
\end{aligned}$$

And therefore

$$H(X_1 | U_j, \ldots, U_n, \underline{Y}'_{j-1}) \leq H(X_1 | U_j, \ldots, U_n, \underline{Y}'_j).$$

Using the above inequality one can show that

$$\sum_{j=1}^{n} I(U_j; X_1 | U_{j+1}, \ldots, U_n, \underline{Y}'_j) \leq H(X_1).$$

Therefore, the second weighted sum of (4) is at most $\max_j \mu_j (1-\epsilon_{1j}) H(X_1)$ and replacing this in (4) we have

$$\begin{aligned}
S^\star(\underline{\mu}) &\leq \max_{P(\underline{U},\underline{X}')} \sum_{j=1}^{n} \mu_j I(U_j; \underline{Y}'_j | U_{j+1}, \ldots, U_n) \qquad (6)\\
&+ \max_j \mu_j (1-\epsilon_{1j})
\end{aligned}$$

The first summation on the right hand side of the inequality corresponds to the maximum weighted sum rate of the $(m-1, n)$ degraded broadcast channel obtained by excluding the connections from the first input of the transmitter to all the receivers. Using similar arguments for the new $(m-1, n)$ it can be verified that

$$\max_{P(\underline{U},\underline{X}')} \sum_{j=1}^{n} \mu_j I(U_j; \underline{Y}'_j | U_{j+1}, \ldots, U_n) \leq \sum_{i=2}^{m} \max_j \mu_j (1-\epsilon_{ij}) \qquad (7)$$

Using this in (6), we have

$$S^\star(\underline{\mu}) \leq \sum_{i=1}^{m} \max_j \mu_j (1-\epsilon_{ij}) \qquad (8)$$

Now the right hand side value can be achieved by time-sharing. For that, input i sends to the the receiver with maximum $\mu_j(1-\epsilon_{ij})$. Therefore each boundary point of the capacity region can be achieved by time-sharing and $\mathcal{C}_g = \mathcal{R}_T$. ☐

V. NON-DEGRADED CASE

In this section, we show that the capacity region of the general erasure broadcast channel is given by time-sharing. In [2] the capacity of the product of two individually reversely degraded broadcast channels is characterized. Using this result it can be easily verified that the capacity region of the $(2, 2)$ erasure broadcast channel is given by time-sharing. However, generalizing and then specializing the technique of [2] for the general (m, n) erasure broadcast channel does not seem straightforward (or maybe even possible). Instead in this paper we use another argument to show that every boundary point of the capacity region for the general erasure broadcast channel is achieved by time-sharing.

Consider an (m, n)-erasure broadcast channel with erasure matrix $\underline{\epsilon} = [\epsilon_{ij}]$. let $\mathcal{C}_g$ denote the capacity region of this general erasure broadcast channel. For every boundary point $\underline{R}'$, there exist positive $\mu_1, \ldots, \mu_n$, such that $\underline{R}'$ is the optimal solution of

$$\max_{(R_1,\ldots,R_n)\in\mathcal{C}_g} \sum_{j=1}^{n} \mu_j R_j$$

The idea is to construct for each value of $\mu_1, \ldots, \mu_n$ a degraded (m, n)-erasure broadcast channel whose capacity region, $\mathcal{C}_d$, contains the capacity region of the original channel, $\mathcal{C}_g$. Moreover, we require that the time-sharing region of both channels meet at some specific point(s) on the boundary.

Let us look at the $(2,2)$ case first. Suppose that the channel is not degraded. Without loss of generality (w.l.o.g) assume that

$$\epsilon_{11} \leq \epsilon_{12}, \quad \epsilon_{21} \geq \epsilon_{22}.$$

Consider the following maximization problem,

$$\max_{(R_1,R_2)\in\mathcal{C}_g} \mu_1 R_1 + \mu_2 R_2. \tag{9}$$

We construct a $(2,2)$-degraded erasure broadcast channel that contains the capacity region $\mathcal{C}_g$. For this, we consider the following two cases separately.

- $\frac{\mu_1}{\mu_2} \geq 1$: In this case consider the erasure broadcast channel with the erasure matrix constructed from ϵ as follows:
$$\epsilon^\star = \begin{bmatrix} \epsilon_{11} & \epsilon_{11} \\ \epsilon_{21} & \epsilon_{22} \end{bmatrix}.$$
First note that the above channel is degraded. Moreover, the erasure probabilities on every edge of this new channel is less than or equal to that of the corresponding edge in the original channel. Therefore the capacity region of the degraded channel contains $\mathcal{C}_g$. Now let's look back at the maximization problem in (9). Based on the above discussions we know that the optimal solution is less than
$$\max_{(R_1,R_2)\in\mathcal{C}_d} \mu_1 R_1 + \mu_2 R_2,$$
where $\mathcal{C}_d$ is the capacity region of the degraded channel. Based on Theorem 1, the maximum of the above problem is achieved by time sharing and it equals
$$\max\{\mu_1(1-\epsilon_{21}), \mu_2(1-\epsilon_{22})\} + \max\{\mu_1(1-\epsilon_{11}), \mu_2(1-\epsilon_{11})\}$$
Since $\mu_1 \geq \mu_2$, we can write the above rate as
$$\max\{\mu_1(1-\epsilon_{21}), \mu_2(1-\epsilon_{22})\} + \mu_1(1-\epsilon_{11}).$$
Looking closely at the above rate, we observe that we can achieve the above rate in the original erasure channel by time-sharing as well. Input 1 transmits only to the first receiver and the second input sends to the maximum of the two terms appearing in the above formula. Therefore, the boundary point corresponding to μ_1, μ_2 is achieved by time-sharing.
- $\frac{\mu_1}{\mu_2} < 1$: In this case, consider the $(2,2)$-erasure broadcast channel with the following erasure matrix:
$$\epsilon^\star = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{22} & \epsilon_{22} \end{bmatrix}.$$
Similar to the first case, the broadcast channel represented by $\epsilon^\star$ is degraded and its capacity region contains $\mathcal{C}_g$. Also the solution of (9) coincides with the solution of the same cost function over region $\mathcal{C}_d$. Thus, the boundary point corresponding to μ_1, μ_2 is achieved by time-sharing.

Based on the above results and Theorem 1, we have the following result.

Theorem 2. *The capacity region of any $(2,2)$-erasure broadcast channel is given by time-sharing.*

We can generalize the preceding arguments to general (m,n)-erasure broadcast channels by finding a degraded erasure broadcast channel that contains the capacity region of the original channel and its time-sharing region coincides with the original one at a specific point. We have stated the above result as a theorem.

Theorem 3. *Consider an (m,n)-erasure broadcast channel with channel matrix $\underline{\epsilon}$. The capacity region of this broadcast channel is given by time-sharing between the receivers at each input.*

Proof: We outline the proof here. Similar to previous discussions, we need to show that every boundary point of the capacity region is achieved by time-sharing. For any $\underline{\mu} \geq 0^2$, consider the following maximization problem

$$f_g(\underline{\mu}, \underline{\epsilon}) = \max_{(R_1,\ldots,R_n)\in\mathcal{C}_g} \sum_{j=1}^{n} \mu_j R_j, \tag{10}$$

where the maximization is over the capacity region of the general channel, i.e, channel g. Now we construct an (m,n)-degraded channel from channel g, such that its capacity region $\mathcal{C}_d$ contains $\mathcal{C}_g$ and also the following maximization problem

$$f_d(\underline{\mu}, \underline{\epsilon}) = \max_{(R_1,\ldots,R_n)\in\mathcal{C}_d} \sum_{j=1}^{n} \mu_j R_j, \tag{11}$$

has the same maximizing value over the time-sharing of the original channel g. First, it can be shown that the optimal value of the following maximization problem

$$f_T(\underline{\mu}, \underline{\epsilon}) = \max_{(R_1,\ldots,R_n)\in\mathcal{R}_T} \sum_{j=1}^{n} \mu_j R_j, \tag{12}$$

with R_T defined in (2) is

$$\sum_{i=1}^{n} \max_{1\leq j\leq n} \mu_j(1-\epsilon_{ij}). \tag{13}$$

Next, we claim that there exist at least one input i and two receivers j,k such that

$$\mu_j \leq \mu_k, \quad \epsilon_{ik} < \epsilon_{ij}.$$

If not, then it can be easily verified that for all i's, ϵ_{ij} are ordered in the reverse order that μ_j's are ordered and therefore the erasure matrix $\underline{\epsilon}$ satisfies the constraint of Proposition 1 with the permutation that sorts μ_j's in decreasing order; hence our (m,n)-erasure channel is "more capable" or equivalently "degraded". But in that case we already know from Theorem 1 that the capacity region is achieved by time-sharing. Therefore, let i^*, j^*, k^* be such numbers. Consider a new (m,n)-erasure channel with erasure matrix $\underline{\epsilon}^{(1)}$ derived from $\underline{\epsilon}$ by replacing $\epsilon_{i^*j^*}$ with $\epsilon_{i^*k^*}$ in the i^*k^* coordinate of $\underline{\epsilon}$. In other words

$$\underline{\epsilon}^{(1)}_{ij} = \underline{\epsilon}_{ij} + (\epsilon_{i^*k^*} - \epsilon_{i^*j^*})\delta(i-i^*)\delta(j-j^*)\delta(k-k^*).$$

[2]Here by $\underline{a} \geq 0$, we mean every component of $\underline{a}$ should be greater than or equal to zero.

where $\delta(\cdot)$ denotes the Dirac delta function. This new channel has the following properties:

- It's capacity region, $\mathcal{C}_1$, contains $\mathcal{C}_g$, since a link is replaced by a link with lower probability of erasure.
- The solution to the maximization problem of (10) over the capacity region $\mathcal{C}_1$ is greater than that of (10).
- The value of (13) remains unchanged for the new channel. This is because of the particular choice of i^*, j^*, k^*.

Now we repeat the above process, till we cannot find any input i, receivers j, k with $(\mu_j \le \mu_k)$ and $(\epsilon_{ik} < \epsilon_{ij})$. In that case, we know that the channel is degraded. Furthermore, it's capacity region $\mathcal{C}_d$ contains the capacity region of all previously derived channels (in particular, the original channel) and the value of (13) for it remains unchanged, i.e.,

$$f_T(\underline{\mu}, \underline{\epsilon}_d) = f_T(\underline{\mu}, \underline{\epsilon}).$$

Based on Theorem 1, for the derived degraded channel we have

$$f_d(\underline{\mu}, \underline{\epsilon}_d) = f_T(\underline{\mu}, \underline{\epsilon}_d).$$

Putting these together we get,

$$f_T(\underline{\mu}, \underline{\epsilon}) \le f_g(\underline{\mu}, \underline{\epsilon}) \le f_d(\underline{\mu}, \underline{\epsilon}_d) \le f_T(\underline{\mu}, \underline{\epsilon}_d) = f_T(\underline{\mu}, \underline{\epsilon}),$$

and therefore $f_T(\underline{\mu}, \underline{\epsilon}) = f_g(\underline{\mu}, \underline{\epsilon})$. This completes the proof. $\square$

VI. Discussion

In this paper we considered a special class of broadcast channels, i.e., erasure broadcast channels with some number of inputs at the transmitter and more than one receiver. These broadcast channels are not degraded (and they do not belong to the "more capable" class). We proved that the capacity region for these broadcast channels is achieved by time-sharing between the receivers at each input. This result has implications for more general network models (of the types proposed in [8]) where the transmitter and receiver are part of a network with a channel model similar to this paper, i.e., no interference and erasure memoryless channels. There it is also possible to propose a similar time-sharing scheme (see [8]). Using theorem 3 one can find upper bounds on the capacity region for broadcast problems in such networks. In this work the erasure events across different channels were assumed to be independent. Finding the capacity region of erasure broadcast channels with dependent erasure events is an interesting problem.

References

[1] H. Weingarten, Y. Steinberg and S. Shamai,"The Capacity Region of the Gaussian MIMO Broadcast Channel,"*Proc. 38th CISS*, 2004.

[2] A. El Gamal,"Capacity of the Product and Sum of Two Reversely Degraded Broadcast Channels,"*Prob. of Info. Transm.*, 1980.

[3] S. Boucheron and M. R. Salamatian,"About Priority Encoding Transmission,"*IEEE Trans. on Info. Theory*, vol. 46, 2000.

[4] P. P. Bergmans,"Random Coding Theorem for Broadcast Channels with Degraded Components,"*IEEE Trans. on Info. Theory*, vol. 19, 1973.

[5] T. M. Cover,"Comment on Broadcast Channels,"*IEEE Trans. on Info. Theory*, vol.44, 1998.

[6] A. A. Gamal,"The Capacity of a Class of Broadcast Channel," *IEEE Trans. on Info. Theory*, vol. 25, 1979.

[7] M. Effros, M. Médard, T. Ho, S. Ray, D. Karger, R. Koetter, B. Hassibi, "Linear network codes: a unified framework for source channel, and network Coding,", *DIMACS workshop on network info. theory*, 2003.

[8] A. F. Dana, R. Gowaiker and B. Hassibi,"On the capacity of broadcast over wireless erasure networks,"*Proc. 42th Allerton conference*, 2004.