The Capacity Region of a Class of 3-Receiver Broadcast Channels with Degraded Message Sets

Chandra Nair
Dept. Of Information Engineering
Chinese University of Hong Kong
Sha Tin, Hong Kong
Email: chandra@ie.cuhk.edu.hk

Abbas El Gamal
Dept. Of Electrical Engineering
Stanford University
Stanford, CA 94305
Email: abbas@ee.stanford.edu

***Abstract*— Körner and Marton established the capacity region for the 2-receiver broadcast channel with degraded message sets. Recent results and conjectures suggest that a straightforward extension of the Körner-Marton region to more than 2 receivers is optimal. This paper shows that this is not the case. We establish the capacity region for a class of 3-receiver broadcast channels with 2 degraded message sets and show that it can be strictly larger than the straightforward extension of the Körner-Marton region. The key new idea is indirect decoding, whereby a receiver who cannot directly decode a cloud center, finds it indirectly by decoding satellite codewords. This idea is then used to establish new inner bounds on the capacity region of the general 3-receiver broadcast channel with 2 and 3 degraded message sets. These bounds are tight for some nontrivial cases.**

I. Introduction

A broadcast channel with degraded message sets represents a scenario where a sender wishes to communicate a common message to *all* receivers, a first private message to a first subset of the receivers, a second private message to a second subset of the first subset and so on. Such scenarios can arise, for example, in video or music broadcasting over a wireless network to nested subsets of receivers at varying levels of quality. What is the set of simultaneously achievable rates for communicating such degraded message sets over the network?

This question was first studied by Körner and Marton in 1977 [**?**]. They considered a general 2-receiver discrete-memoryless broadcast channel with sender X and receivers Y_1 and Y_2. A common message $M_0 \in [1, 2^{nR_0}]$ is to be sent to both receivers and a private message $M_1 \in [1, 2^{nR_1}]$ is to be sent only to receiver Y_1. They showed that the capacity region is given by the set of all rate pairs (R_0, R_1) such that [1]

$$R_0 \leq \min\{I(U;Y_1), I(U;Y_2)\}, \qquad (1)$$
$$R_1 \leq I(X;Y_1|U),$$

for some $p(u,x)$. These rates are achieved using superposition coding [**?**]. The common message is represented by the auxiliary random variable U and the private message is superimposed to generate X. The main contribution of [**?**] is proving a strong converse using the technique of images-of-a-set [**?**].

Extending the Körner-Marton result to more than 2 receivers has remained open even for the simple case of 3 receivers Y_1, Y_2, Y_3 with 2 degraded message sets, where a common message M_0 is to be sent to all receivers and a private message M_1 is to be sent only to receiver Y_1. The straightforward extension of the Körner-Marton region to this case yields the achievable rate region consisting of the set of all rate pairs (R_0, R_1) such that

$$R_0 \leq \min\{I(U;Y_1), I(U;Y_2), I(U;Y_3)\}, \qquad (2)$$
$$R_1 \leq I(X;Y_1|U),$$

for some $p(u,x)$. Is this region optimal?

In [**?**], it was shown that the above region (and its natural extension to $k > 3$ receivers) is optimal for a class of product discrete-memoryless and Gaussian broadcast channels, where each of the receivers who decode only the common message is a degraded version of the unique receiver that also decodes the private message. In [**?**], it was shown that a straightforward extension of Körner-Marton region is also optimal for the class of linear deterministic broadcast channels, where the operations are performed in a finite field. In addition to establishing the degraded message set capacity for this class the authors gave an explicit characterization of the optimal auxiliary random variables. In a recent paper Borade et al. [**?**] introduced *multilevel* broadcast channels, which combine aspects of degraded broadcast channels and broadcast channels with degraded message sets. They established an achievable rate region as well as a "mirror-image" outer bound for these channels. Their achievable rate region is again a straightforward extension of the Körner-Marton region to k-receiver multilevel broadcast channels. In particular, Conjecture 5 of [**?**] states that the capacity region of the 3-receiver multilevel broadcast channels depicted in Figure **??** is the set of all rate pairs (R_0, R_1) such that

$$R_0 \leq \min\{I(U;Y_2), I(U;Y_3)\}, \qquad (3)$$
$$R_1 \leq I(X;Y_1|U),$$

for some $p(u,x)$. Note that this region, henceforth referred to as *the BZT region*, is the same as (**??**) because in the multilevel

[1] The Körner-Marton characterization does not include the second term inside the min in the first inequality, $I(U;Y_1)$. Instead it includes the bound $R_0 + R_1 \leq I(X;Y_1)$. It can be easily shown that the two characterizations are equivalent.

broadcast channel Y_3 is a degraded version of Y_1 and therefore $I(U;Y_3) \leq I(U;Y_1)$.

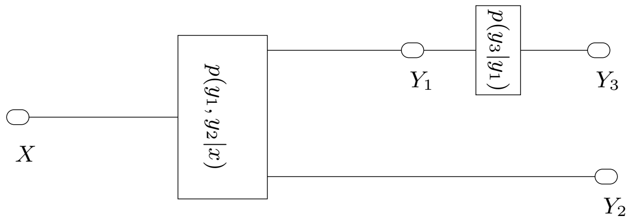

Fig. 1. Multilevel 3-receiver broadcast channels. Message M_0 is to be sent to all receivers and message M_1 is to be sent only to Y_1.

In this paper we show that the straightforward extension of the Körner-Marton region to more than 2 receivers is not in general optimal [2]. In particular, we establish the capacity region of the multilevel broadcast channels depicted in Figure **??** and show that it can be strictly larger than the BZT region (Theorem **??**). We then extend the results on the multilevel broadcast channel to establish inner bounds on the capacity region of the general 3-receiver broadcast channel with 2 and 3 degraded message sets (Proposition **??** and Theorem **??**).

II. Definitions

Consider a discrete-memoryless 3-receiver broadcast channel consisting of an input alphabet $\mathcal{X}$, output alphabets $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{Y}_3$, and a probability transition function $p(y_1, y_2, y_3|x)$.

A $(2^{nR_0}, 2^{nR_1}, n)$ *2-degraded message set* code for a 3-receiver broadcast channel consists of (i) a pair of messages (M_0, M_1) uniformly distributed over $[1, 2^{nR_0}] \times [1, 2^{nR_1}]$, (ii) an encoder that assigns a codeword $x^n(m_0, m_1)$, for each message pair $(m_0, m_1) \in [1, 2^{nR_0}] \times [1, 2^{nR_1}]$, and (iii) three decoders, one that maps each received y_1^n sequence into an estimate $(\hat{m}_{01}, \hat{m}_1) \in [1, 2^{nR_0}] \times [1, 2^{nR_1}]$, a second that maps each received y_2^n sequence into an estimate $\hat{m}_{02} \in [1, 2^{nR_0}]$, and a third that maps each received y_3^n sequence into an estimate $\hat{m}_{03} \in [1, 2^{nR_0}]$.

The probability of error is defined as

$$P_e^{(n)} = \mathrm{P}\{\hat{M}_1 \neq M_1 \text{ or } \hat{M}_{0k} \neq M_0 \text{ for } k = 1, 2, \text{ or } 3\}.$$

A rate tuple (R_0, R_1) is said to be achievable if there exists a sequence of $(2^{nR_0}, 2^{nR_1}, n)$ 2-degraded message set codes with $P_e^{(n)} \to 0$. The capacity region of the broadcast channel is the closure of the set of achievable rates.

A 3-receiver *multilevel* broadcast channel [**?**] is a 3-receiver broadcast channel with 2 degraded message sets where $p(y_1, y_2, y_3|x) = p(y_1, y_2|x)p(y_3|y_1)$ for every $(x, y_1, y_2, y_3) \in \mathcal{X} \times \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3$ (see Figure **??**).

In addition to considering the multilevel 3-receiver broadcast channel and the general 3-receiver broadcast channel with 2 degraded message sets, we shall also consider the 3-receiver broadcast channel with 3 message sets, where M_0 is to be sent to all receivers, M_1 is to be sent to Y_1 and Y_2, and M_2 is to be sent only to Y_1. Definitions of codes, achievability and capacity regions for these cases are straightforward extensions of the above definitions.

[2]A complete version of this paper has been posted on arXiv and submitted to the IEEE Trans. on IT [**?**].

III. Capacity of 3-Receiver Multilevel Broadcast Channel

Theorem 1: The capacity region of the 3-receiver multilevel broadcast channel in Figure **??** is the set of rate pairs (R_0, R_1) such that

$$\begin{aligned} R_0 &\leq \min\{I(U_1;Y_3), I(U_2;Y_2)\}, \qquad (4)\\ R_0 + R_1 &\leq I(U_1;Y_3) + I(X;Y_1|U_1),\\ R_0 + R_1 &\leq I(U_2;Y_2) + I(X;Y_1|U_2), \end{aligned}$$

for some $p(u_1)p(u_2|u_1)p(x|u_2)$, where the cardinalities of the auxiliary random variables satisfy $\|\mathcal{U}_1\| \leq \|\mathcal{X}\| + 4$ and $\|U_2\| \leq \|\mathcal{X}\|^2 + 5\|\mathcal{X}\| + 4$.

Remarks:

1) It is easy to show by setting $U_1 = U_2 = U$ in the above theorem that the BZT region (**??**) is contained in the capacity region (**??**). We show in the next section that the capacity region (**??**) can be strictly larger the BZT region.
2) It is straightforward to show that the above region is convex and therefore there is no need to use a time-sharing auxiliary random variable.

Proof: The interesting part of the proof of Theorem **??** is achievability. Specifically, step 3 of the decoding procedure for Case 2 below describes a key contribution of this paper. We show how Y_2 can find M_0 without directly decoding U_1^n or uniquely decoding U_2^n.

To show achievability of any rate pair (R_0, R_1) in region (**??**), because of its convexity, it suffices to show the achievability of any rate pair (R_0, R_1) such that

$$\begin{aligned} R_0 &= \min\{I(U_1;Y_3), I(U_2;Y_2)\} - \delta\\ R_0 + R_1 &= \min\{I(U_1;Y_3) + I(X;Y_1|U_1),\\ &\qquad I(U_2,Y_2) + I(X;Y_1|U_2)\} - 3\delta, \end{aligned}$$

for some $U_1 \to U_2 \to X$ and any $\delta > 0$.

Rewriting the second equality we obtain

$$\begin{aligned} R_0 + R_1 = I(U_1;Y_3) + \min\{&I(U_2;Y_1|U_1), I(U_2;Y_2)\\ &- I(U_1;Y_3)\} + I(X;Y_1|U_2) - 3\delta. \end{aligned}$$

Now consider the following two cases:

Case 1: $I(U_1;Y_3) > I(U_2;Y_2)$: The rates reduce to

$$\begin{aligned} R_0 &= I(U_2;Y_2) - \delta\\ R_1 &= I(X;Y_1|U_2) - 2\delta. \end{aligned}$$

This pair can be achieved via a simple superposition coding scheme [**?**].

Case 2: $I(U_1;Y_3) \leq I(U_2;Y_2)$: The rates reduce to

$$\begin{aligned} R_0 &= I(U_1;Y_3) - \delta\\ R_1 &= I(X;Y_1|U_2) + \min\{I(U_2;Y_1|U_1),\\ &\qquad I(U_2,Y_2) - I(U_1;Y_3)\} - 2\delta. \end{aligned}$$

Let $S_1 = \min\{I(U_2;Y_1|U_1), I(U_2,Y_2) - I(U_1;Y_3)\} - \delta$ and $S_2 = I(X;Y_1|U_2) - \delta$, then $R_1 = S_1 + S_2$.

Code Generation:

Fix $p(u_1)p(u_2|u_1)p(x|u_2)$ that satisfies the condition of Case 2. Generate $2^{nR_0} = 2^{n(I(U_1;Y_3)-\delta)}$ sequences $U_1^n(1), \ldots, U_1^n(2^{nR_0})$ distributed uniformly at random over the set of ϵ-typical† U_1^n sequences, where $\delta \to 0$ as $\epsilon \to 0$. For each $U_1^n(m_0)$, generate $2^{nS_1} = 2^{n(\min\{I(U_2;Y_1|U_1),I(U_2,Y_2)-I(U_1;Y_3)\}-\delta)}$ sequences $U_2^n(m_0,1)$, $U_2^n(m_0,2)$, $\ldots$, $U_2^n(m_0,2^{nS_1})$ distributed uniformly at random over the set of conditionally ϵ-typical U_2^n sequences. For each $U_2^n(m_0,s_1)$ generate $2^{nS_2} = 2^{n(I(X;Y_1|U_2)-\delta)}$ sequences $X^n(m_0,s_1,1)$, $X^n(m_0,s_1,2), \ldots, X^n(m_0,s_1,2^{nS_2})$ distributed uniformly at random over the set of conditionally ϵ-typical X^n sequences.

Encoding:

To send the message pair $(m_0,m_1) \in [1,2^{nR_0}] \times [1,2^{nR_1}]$, the sender expresses m_1 by the pair $(s_1,s_2) \in [1,2^{nS_1}] \times [1,2^{nS_2}]$ and sends $X^n(m_0,s_1,s_2)$.

Decoding and Analysis of Error Probability:

1) Receiver Y_3 declares that m_0 is sent if it is the unique message such that $U_1^n(m_0)$ and Y_3^n are jointly ϵ-typical. It is easy to see that this can be achieved with arbitrarily small probability of error because $R_0 = I(U_1;Y_3) - \delta$.
2) Receiver Y_1 first declares that m_0 is sent if it is the unique message such that $U_1^n(m_0)$ and Y_1^n are jointly ϵ-typical. This decoding step succeeds with arbitrarily high probability because $R_0 = I(U_1;Y_3) - \delta \le I(U_1;Y_1) - \delta$. It then declares that s_1 is sent if it is the unique index such that $U_2^n(m_0,s_1)$ and Y_1^n are jointly ϵ-typical. This decoding step succeeds with arbitrarily high probability because $S_1 \le I(U_2;Y_1|U_1) - \delta$. Finally it declares that s_2 is sent if it is the unique index such that $X^n(m_0,s_1,s_2)$ and Y_1^n are jointly ϵ-typical. This decoding step succeeds with high probability because $S_2 = I(X;Y_1|U_2) - \delta$.
3) Receiver Y_2 finds m_0 as follows. It declares that $m_0 \in [1,2^{nR_0}]$ is sent if it is the unique index such that $U_2^n(m_0,s_1)$ and Y_2^n are jointly ϵ-typical for some $s_1 \in [1,2^{nS_1}]$. Suppose $(1,1) \in [1,2^{nR_0}] \times [1,2^{nS_1}]$ is the message pair sent, then the probability of error averaged over the choice of codebooks can be upper bounded as follows

$$
\begin{aligned}
P_e^{(n)} &\le \mathrm{P}\{(U_2^n(1,1),Y_2^n) \text{ not jointly } \epsilon\text{-typical}\} \\
&\quad + \mathrm{P}\{(U_2^n(m_0,s_1),Y_2^n) \\
&\qquad \text{jointly } \epsilon\text{-typical for some } m_0 \ne 1\} \\
&\overset{(a)}{<} \delta' + \sum_{m_0 \ne 1}\sum_{s_1} \mathrm{P}\{(U_2^n(m_0,s_1),Y_2^n) \\
&\qquad \text{jointly } \epsilon\text{-typical}\} \\
&\overset{(b)}{\le} \delta' + 2^{n(R_0+S_1)}2^{-n(I(U_2;Y_2)-\delta)} \overset{(c)}{\le} \delta' + 2^{-n\delta},
\end{aligned}
$$

where (a) follows by the union of events bound, (b) follows by the fact that for $m_0 \ne 1$, $U_2^n(m_0,s_1)$ and Y_2^n are generated completely independently and thus each probability term under the sum is upper bounded by $2^{-n(I(U_2;Y_2)-\delta)}$ [?], (c) follows because by construction $R_0 + S_1 \le I(U_2;Y_2) - 2\delta$, $\delta' \to 0$ as $\epsilon \to 0$. Thus with arbitrarily high probability, any jointly ϵ-typical $U_2^n(m_0,s_1)$ with the received Y_2^n sequence must be of the form $U_2^n(1,s_1)$, and receiver Y_2 can correctly decodes M_0 with arbitrarily small probability of error. Note that Y_2 may or may not be able to uniquely decode $U_2^n(1,1)$. However, it finds the correct common message with arbitrarily small probability of error even if its rate $R_0 > I(U_1;Y_2)$!

Thus all receivers can decode their intended messages with arbitrarily small probability of error and hence the rate pair $R_0 = I(U_1;Y_3) - \delta, R_1 = I(X;Y_1|U_2) + \min\{I(U_2;Y_1|U_1), I(U_2,Y_2) - I(U_1;Y_3)\} - 2\delta$ is achievable. This completes the proof of achievability of Theorem **??**.

The converse proof is quite similar to previous weak converse and outer bound proofs for 2-receiver broadcast channels (e.g., see [?], [?], [?]).

The identification of the auxiliary random variables are as follows: Let $U_{1i} = (M_0, Y_1^{i-1})$, $i = 1,\ldots,n$; $U_{2i} = (M_0, Y_1^{i-1}, Y_{2\,i+1}^{\;n})$, $i = 1,\ldots,n$; and let Q be a time-sharing random variable uniformly distributed over the set $\{1,2,...,n\}$ and independent of X^n, Y_1^n, Y_2^n, Y_3^n. We then set $U_1 = (Q, U_{1Q})$ and $U_2 = (Q, U_{2Q})$. Using this identification the converse to Theorem **??** follows with the required Markov structure on the auxiliary random variables. Please refer to [?] for the details.

The bounds on the cardinality of the auxiliary random variables are also established in [?]. ■

Remark: We denote the decoding technique used in step 3 as *indirect decoding* because Y_2 decodes the cloud center U_1 indirectly by decoding possibly a list of satellite codewords.

IV. Multilevel Product Broadcast Channel

In this section we show that the BZT region can be strictly smaller than the capacity region in Theorem **??**.

Consider the product of 3-receiver broadcast channels given by the Markov relationships

$$
\begin{aligned}
&X_1 \to Y_{21} \to Y_{11} \to Y_{31}, \\
&X_2 \to Y_{12} \to Y_{32}.
\end{aligned} \tag{5}
$$

In [?] we derive the following simplified characterizations for the capacity and the BZT regions.

Proposition 1: The BZT region for the above product channel reduces to the set of rate pairs (R_0, R_1) such that

$$
\begin{aligned}
R_0 &\le I(V_1;Y_{31}) + I(V_2;Y_{32}), \\
R_0 &\le I(V_1;Y_{21}), \\
R_1 &\le I(X_1;Y_{11}|V_1) + I(X_2;Y_{12}|V_2),
\end{aligned} \tag{6}
$$

for some $p(v_1)p(v_2)p(x_1|v_1)p(x_2|v_2)$.

†We assume strong typicality throughout this paper [?].

Proposition 2: The capacity region for the product channel reduces to the set of rate pairs (R_0, R_1) such that

$$\begin{aligned} R_0 &\le I(V_{11}; Y_{31}) + I(V_{12}; Y_{32}), \\ R_0 &\le I(V_{21}; Y_{21}), \\ R_0 + R_1 &\le I(V_{11}; Y_{31}) + I(V_{12}; Y_{32}) \\ &\quad + I(X_1; Y_{11}|V_{11}) + I(X_2; Y_{12}|V_{12}), \\ R_0 + R_1 &\le I(V_{21}; Y_{21}) + I(X_1; Y_{11}|V_{21}) + I(X_2; Y_{12}|V_{12}), \end{aligned} \tag{7}$$

for some $p(v_{11})p(v_{21}|v_{11})p(x_1|v_{21})p(v_{12})p(x_2|v_{12})$.

Now we compare these two regions via the following example.

Example:

Consider the multilevel product broadcast channel example in Figure **??**, where: $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y}_{12} = \mathcal{Y}_{21} = \{0, 1\}$, and $\mathcal{Y}_{11} = \mathcal{Y}_{31} = \mathcal{Y}_{32} = \{0, E, 1\}$, $Y_{21} = X_1$, $Y_{12} = X_2$, the channels $Y_{21} \to Y_{11}$ and $Y_{12} \to Y_{32}$ are binary erasure channels (BEC) with erasure probability $\frac{1}{2}$, and the channel $Y_{11} \to Y_{31}$ is given by the transition probabilities: $\mathrm{P}\{Y_{31} = E|Y_{11} = E\} = 1$, $\mathrm{P}\{Y_{31} = E|Y_{11} = 0\} = \mathrm{P}\{Y_{31} = E|Y_{11} = 1\} = 2/3$, $\mathrm{P}(Y_{31} = 0|Y_{11} = 0\} = \mathrm{P}\{Y_{31} = 1|Y_{11} = 1\} = 1/3$. Therefore, the channel $X_1 \to Y_{31}$ is effectively a BEC with erasure probability $5/6$.

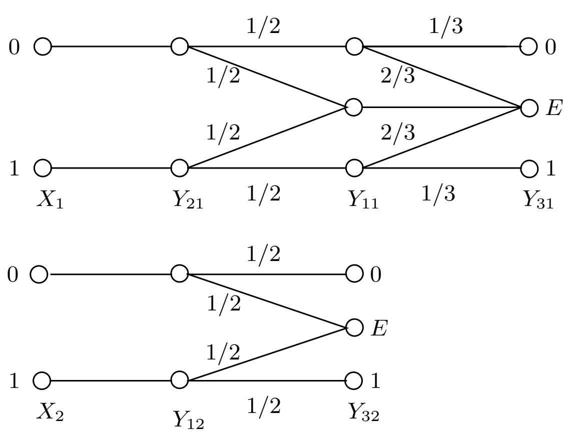

Fig. 2. Product multilevel broadcast channel example.

Proposition 3: The BZT region for the above example reduces to the set of rate pairs (R_0, R_1) satisfying

$$\begin{aligned} R_0 &\le \min\left\{\frac{p}{6} + \frac{q}{2}, p\right\}, \\ R_1 &\le \frac{1-p}{2} + 1 - q. \end{aligned} \tag{8}$$

for some $0 \le p, q \le 1$.

The proof of this proposition is given in [**?**]. It is quite straightforward to see that $(R_0, R_1) = (\frac{1}{2}, \frac{5}{12})$ lies on the boundary of this region.

Proposition 4: The capacity region for the channel in Figure **??** reduces to set of rate pairs (R_0, R_1) satisfying

$$\begin{aligned} R_0 &\le \min\left\{\frac{r}{6} + \frac{s}{2}, t\right\}, \\ R_0 + R_1 &\le \min\left\{\frac{r}{6} + \frac{s}{2} + \frac{1-r}{2} + 1 - s, \frac{1+t}{2} + 1 - s\right\}, \end{aligned} \tag{9}$$

for some $0 \le r \le t \le 1, 0 \le s \le 1$.

The proof of this proposition is also given in [**?**]. Note that substituting $r = t$ yields the BZT region. By setting $r = 0, s = 1, t = 1$ it is easy to see that $(R_0, R_1) = (1/2, 1/2)$ lies on the boundary of the capacity region. Figure **??** plots the BZT region and the capacity region for the example channel showing that the capacity region is strictly larger.

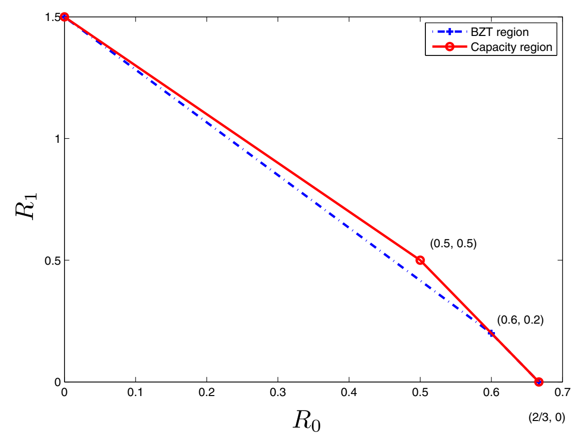

Fig. 3. The BZT and the capacity regions for the channel in Figure **??**.

V. Achievable Rate Regions for 3-Receiver Broadcast Channel with Degraded Message Sets

We use superposition coding, indirect decoding, and the Marton achievability scheme for the general 2-receiver broadcast channels [**?**] to establish the following inner bound for the 3-receiver broadcast channel with 2 degraded message sets.

Proposition 5: A rate pair (R_0, R_1) is achievable in a general 3-receiver broadcast channel with 2 degraded message sets if it satisfies the following inequalities:

$$\begin{aligned} R_0 &\le \min\{I(U_2; Y_2), I(U_3; Y_3)\}, \\ 2R_0 &\le I(U_2; Y_2) + I(U_3; Y_3) - I(U_2; U_3|U_1), \\ R_1 &\le \min\{I(X; Y_1|U_2) + I(X; Y_1|U_3), I(X; Y_1|U_1)\}, \\ R_0 + R_1 &\le \min\{I(X; Y_1), I(U_2; Y_2) + I(X; Y_1|U_2), \\ &\quad I(U_3; Y_3) + I(X; Y_1|U_3)\}, \\ 2R_0 + R_1 &\le I(U_2; Y_2) + I(U_3; Y_3) \\ &\quad + I(X; Y_1|U_2, U_3) - I(U_2; U_3|U_1), \\ 2R_0 + 2R_1 &\le I(U_2; Y_2) + I(X; Y_1|U_2) + I(U_3; Y_3) \\ &\quad + I(X; Y_1|U_3) - I(U_2; U_3|U_1), \end{aligned}$$

for some $p(u_1, u_2, u_3, x) = p(u_1)p(u_2|u_1)p(x, u_3|u_2) = p(u_1)p(u_3|u_1)p(x, u_2|u_3)$ (or in other words, both $U_1 \to U_2 \to (U_3, X)$ and $U_1 \to U_3 \to (U_2, X)$ form Markov chains).

The general idea of the proof is to represent M_0 by U_1, superimpose two independent pieces of information about M_1 to obtain U_2 and U_3, respectively, and then superimpose the remaining information about M_1 to obtain X. Receiver Y_1 decodes U_1, U_2, U_3, X, receivers Y_2 and Y_3 find M_0 via indirect decoding of U_2 and U_3, respectively, as in Theorem **??**. The details of the argument can be found in [**?**].

Remarks: (see [**?**] for proofs) The inner bound is tight when Y_1 is less noisy receiver than Y_3 [**?**]. This is more general class than degradedness and hence the above inner bound is also tight for the multilevel case in Theorem **??**.

The above inner bound can be extended to the case of 3 degraded message sets.

Theorem 2: A rate triple (R_0, R_1, R_2) is achievable in a general 3-receiver broadcast channel with 3 degraded message sets if it satisfies the conditions:

$$
\begin{aligned}
R_0 &\le I(U_3; Y_3) \\
R_1 &\le \min\{I(U_2; Y_2|U_1), I(X; Y_1|U_3)\}, \\
R_2 &\le I(X; Y_1|U_2) \\
R_0 + R_1 &\le \min\big\{I(U_2; Y_2), I(U_2; Y_2|U_1) \\
&\qquad + I(U_3; Y_3) - I(U_2; U_3|U_1)\big\}, \\
2R_0 + R_1 &\le I(U_2; Y_2) + I(U_3; Y_3) - I(U_2; U_3|U_1), \\
R_0 + R_2 &\le I(U_3; Y_3) + I(X; Y_1|U_2, U_3) \\
R_1 + R_2 &\le I(X; Y_1|U_1), \\
R_0 + R_1 + R_2 &\le \min\big\{I(X; Y_1), I(U_3; Y_3) + I(X; Y_1|U_3), \\
&\qquad I(U_2; Y_2|U_1) + I(U_3; Y_3) + \\
&\qquad I(X; Y_1|U_2, U_3) - I(U_2; U_3|U_1)\big\}, \\
2R_0 + R_1 + R_2 &\le I(U_2; Y_2) + I(U_3; Y_3) \\
&\qquad + I(X; Y_1|U_2, U_3) - I(U_2; U_3|U_1), \\
R_0 + 2R_1 + R_2 &\le I(U_2; Y_2|U_1) + I(U_3; Y_3) \\
&\qquad + I(X; Y_1|U_3) - I(U_2; U_3|U_1), \\
2R_0 + 2R_1 + R_2 &\le I(U_2; Y_2) + I(U_3; Y_3) \\
&\qquad + I(X; Y_1|U_3) - I(U_2; U_3|U_1).
\end{aligned}
$$

for some $p(u_1, u_2, u_3, x) = p(u_1)p(u_2|u_1)p(x, u_3|u_2) = p(u_1)p(u_3|u_1)p(x, u_2|u_3)$ (i.e., as before both $U_1 \to U_2 \to (U_3, X)$ and $U_1 \to U_3 \to (U_2, X)$ form Markov chains).

Remarks:

1) The region of Theorem **??** reduces to the inner bound of Proposition **??** by setting $R_1 = 0$.
2) When $R_2 = 0$ this region reduces to straightforward extension of the Körner-Marton scheme!
3) When $R_2 = 0$ the above region (equivalently the straightforward extension of the Körner-Marton scheme) is optimal in the following two non-trivial cases:
   - The broadcast channel is deterministic.
   - Y_1 is a less noisy receiver than Y_3 and Y_2 is a less noisy receiver than Y_3.

VI. Conclusion

We show that the capacity region of the 3-receiver broadcast channels with 2 degraded message sets can be strictly larger than the straightforward extension of the Körner-Marton region. The achievability proof uses the new idea of indirect decoding whereby a receiver decodes a cloud center indirectly through joint typicality with a satellite codeword. Using this idea, we devise new inner bounds to the capacity of the general 3-receiver broadcast channel with 2 and 3 degraded message sets, which are tight in some cases. It would be interesting to investigate applications of indirect decoding to other problems, for example, 3-receiver broadcast channels with confidential message sets [**?**].

References

[1] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Info. Theory*, vol. IT-23, pp. 60–64, Jan, 1977.
[2] T. Cover, "Broadcast channels," *IEEE Trans. Info. Theory*, vol. IT-18, pp. 2–14, January, 1972.
[3] J. Körner and K. Marton, "Images of a set via two channels and their role in mult-user communication," *IEEE Trans. Info. Theory*, vol. IT-23, pp. 751–761, Nov, 1977.
[4] S. Diggavi and D. Tse, "On opportunistic codes and broadcast codes with degraded message sets," *Information theory workshop (ITW)*, 2006.
[5] V. Prabhakaran, S. Diggavi, and D. Tse, "Broadcasting with degraded message sets: A deterministic approach," *Proceedings of the 45th Annual Allerton Conference on Communication, Control and Computing*, 2007.
[6] S. Borade, L. Zheng, and M. Trott, "Multilevel broadcast networks," *International Symposium on Information Theory*, 2007.
[7] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Probl. Peredac. Inform.*, vol. 10(3), pp. 3–14, 1974.
[8] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Info. Theory*, vol. IT-25, pp. 166–169, March, 1979.
[9] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Info. Theory*, vol. IT-53, pp. 350–355, January, 2007.
[10] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley Interscience, 1991.
[11] I. Csizár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Info. Theory*, vol. IT-24, pp. 339–348, May, 1978.
[12] R. F. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Info. Theory*, vol. IT-21(6), pp. 629–637, November, 1975.
[13] J. Körner and K. Marton, "A source network problem involving the comparison of two channels ii," *Trans. Colloquim Inform. Theory, Keszthely, Hungary*, August, 1975.
[14] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Info. Theory*, vol. IT-25, pp. 306–311, May, 1979.
[15] A. El Gamal and E. C. van der Meulen, "A proof of marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 120–121, 1981.
[16] C. Nair and A. El Gamal, "The capacity of a class of 3-receiver discrete memoryless channels with degraded message sets," *http://arxiv.org/abs/0712.3327*.
[17] A. Schrijver, *Theory of Integer and Linear Programming*. John Wiley & sons, 1986.