# Capacity Bounds for Broadcast Channels with Bidirectional Conferencing Decoders

Reza K. Farsani and Wei Yu
Department of Electrical and Computer Engineering, University of Toronto, Canada
E-mail: {rkfarsani, weiyu}@comm.utoronto.ca

***Abstract*—The two-user broadcast channel (BC) with decoders connected by cooperative links of given capacities (known as conferencing decoders) is considered. A novel outer bound on the capacity region is established. This outer bound is derived using multiple applications of the Csiszár-Körner identity. A new achievable rate region for the channel is also presented which is derived by applying Marton's coding as the transmission scheme, and quantize-bin-and-forward at one receiver and decode-and-forward at the other receiver as cooperative strategy. It is proved that the outer bound coincides with the achievable region for a class of semi-deterministic BCs with degraded message sets. This is the first capacity result for the two-user BC with bidirectional conferencing decoders. This result demonstrates that a one-round cooperation scheme is sufficient to achieve capacity for this class of semi-deterministic BCs with degraded message set. A capacity result is also derived for a new class of more capable and semi-deterministic BCs with both common and private messages and one-sided conferencing.**

***Index Terms*—Broadcast channel, conferencing decoders, semi-deterministic channel, Csiszár-Körner identity**

## I. Introduction

This paper investigates the impact of user cooperation via digital links of given capacities, known as conferencing links, on the capacity region of a two-user broadcast channel (BC) with both common and private messages. The two-user BC with conferencing decoders is previously studied in [1]- [7]. In [1], the authors develop communication strategies for the interactive decoding of a common message broadcast to cooperative users. In [2], the capacity region of physically degraded channel is derived and also an achievable rate region is given for the general case. In [3], the problems of communication over physically degraded, state-dependent BCs with one-sided conferencing decoders are investigated. In [4], the capacity region of the semi-deterministic BC with one-sided decoder cooperation is derived and its duality with a source coding problem is addressed. The authors in [5] consider the BC with one-sided cooperating users under the strong secrecy constraints and present capacity results for semi-deterministic and physically degraded cases. In [6], the BC with (one-sided) unreliable cooperating decoders is studied. Lastly, in a recent work [7], the BC with degraded message sets and one-sided cooperation link that may be absent is considered and its capacity region is given.

As reviewed, the existing capacity results for BC with conferencing decoders are all for the case of one-sided cooperation, i.e., only one of the users is connected to the other by a conferencing link. This is due to the lack of outer bounds for the two-sided cooperation case. In this paper, we first establish a novel outer bound on the capacity region of the two-user BC with bidirectional conferencing decoders. The new outer bound is derived using multiple applications of the Csiszár-Körner identity [8, Lemma 7]. It is strictly tighter than the previous outer bounds including that of [2, Prop. 1], which is essentially the cut-set bound.

Further, we propose an achievability strategy for the channel. In [2, Th. 2] an achievable region is derived for the BC with bidirectional cooperation by applying Marton's coding at the transmitter and compress-and-forward cooperative scheme at both receivers. A second achievable region is given in [4, App. B] for the BC with one-sided cooperation between receivers that is derived by applying Marton's coding at the transmitter and decode-and-forward as cooperative protocol. Both of these achievable schemes (which are given for the channel with private messages only) are in general insufficient to either derive new capacity results or approximate capacity results for the Gaussian channel. Instead, this paper presents an achievability scheme for the two-user BC with both common and private messages and bidirectional conferencing receivers, in which we apply Marton's coding as the transmission scheme and quantize-bin-and-forward at one receiver and decode-and-forward at the other receiver as cooperative strategy.

We prove that the novel outer bound coincides with the proposed achievable region for a class of semi-deterministic BCs with degraded message sets. This capacity result is important from two viewpoints. First, it is the first capacity result for the two-user BC with two-sided conferencing receivers (all previously known capacity results are regarding the channel with one-sided cooperation). Second, it is among rare cases in network information theory for which quantize-bin-and-forward is optimal. Our result also demonstrates that a one-round bidirectional cooperation protocol is sufficient to achieve capacity and multi-round strategies similar to those devised in [1] are not needed. Finally, we derive the capacity region of a class of more capable semi-deterministic BCs with both common and private messages and one-sided cooperation.

## II. Preliminaries

We use the following notations. Random variables are denoted by upper case letters (e.g. $X$) and lower case letter are used to show their realization (e.g. $x$). The probability distribution function (PDF) of $X$ is denoted by $P_X(x)$ and the conditional PDF of $X$ given $Y$ is denoted by $P_{X|Y}(x|y)$,

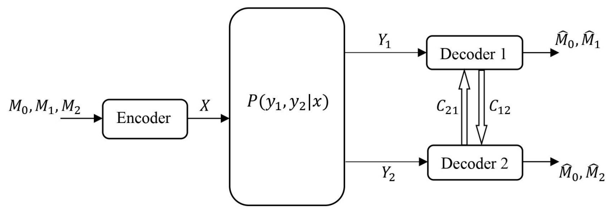


Figure 1. The Two-User BC with Conferencing Decoders.

where the subscripts are omitted occasionally for brevity. The operator $\{a\}^+$ is defined as: $\{a\}^+ = \max\{0, a\}$.

*Definition:* The two-user BC with conferencing decoders is a communication scenario in which a transmitter sends a common message and two private messages to two users. The users are able to exchange information via communication links of finite capacities called conferencing. links. Fig. 1 illustrates the channel model. The channel is given by $(\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, P(y_1, y_2|x), C_{12}, C_{21})$ where $\mathcal{X}$ denote input alphabet, $\mathcal{Y}_1$ and $\mathcal{Y}_2$ denote output alphabets, $P(y_1, y_2|x)$ is the channel transition probability function, and $C_{12}$, and $C_{21}$ are capacities of the conferencing links.

*Encoding:* For the BC with conferencing decoders, a length-$n$ code with $\Gamma$ conferencing rounds is described as follows. The transmitter encodes independent messages $M_0$, $M_1$, and $M_2$, which are uniformly distributed over the sets $[1 : 2^{nR_0}]$, $[1 : 2^{nR_1}]$, and $[1 : 2^{nR_2}]$, respectively, into a codeword and sends over the channel according to the following:

$$\Delta : [1 : 2^{nR_0}] \times [1 : 2^{nR_1}] \times [1 : 2^{nR_2}] \longmapsto \mathcal{X}^n$$
$$X^n = \Delta(M_0, M_1, M_2)$$

*Decoding:* The receiver $Y_i,\ i = 1, 2$, receives a sequence $Y_i^n \in \mathcal{Y}_i^n$. The code consists of two sets of conferencing functions $\{\Xi_{12,\gamma}\}_{\gamma=1}^{\Gamma}$ and $\{\Xi_{21,\gamma}\}_{\gamma=1}^{\Gamma}$ with the corresponding output alphabets $\{\mathcal{J}_{12,\gamma}\}_{\gamma=1}^{\Gamma}$ and $\{\mathcal{J}_{21,\gamma}\}_{\gamma=1}^{\Gamma}$, respectively, which are described as follows.

$$\Xi_{12,\gamma} : \mathcal{Y}_1^n \times \mathcal{J}_{21,1} \times ... \times \mathcal{J}_{21,\gamma-1} \longmapsto \mathcal{J}_{12,\gamma}$$
$$J_{12,\gamma} = \Xi_{12,\gamma}(Y_1^n, J_{21}^{\gamma-1})$$
$$\Xi_{21,\gamma} : \mathcal{Y}_2^n \times \mathcal{J}_{12,1} \times ... \times \mathcal{J}_{12,\gamma-1} \longmapsto \mathcal{J}_{21,\gamma}$$
$$J_{21,\gamma} = \Xi_{21,\gamma}(Y_2^n, J_{12}^{\gamma-1})$$

A conference is said to be $(C_{12}, C_{21})$-permissible if

$$\sum_{\gamma=1}^{\Gamma} \log \|\mathcal{J}_{12,\gamma}\| \leq nC_{12}, \qquad \sum_{\gamma=1}^{\Gamma} \log \|\mathcal{J}_{21,\gamma}\| \leq nC_{21}$$

Before decoding, the receivers exchange information by holding a $(C_{12}, C_{21})$-permissible conference. Thus, the first receiver obtains the sequence $J_{21}^{\Gamma} = (J_{21,1}, J_{21,2}, ..., J_{21,\Gamma})$ and the second one obtains the sequence $J_{12}^{\Gamma} = (J_{12,1}, J_{12,2}, ..., J_{12,\Gamma})$. The receivers then detect their respective messages based on the following decoding functions:

$$\nabla_1 : \mathcal{Y}_1^n \times \mathcal{J}_{21}^{\Gamma} \longmapsto [1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$$
$$(\hat{M}_0, \hat{M}_1) = \nabla_1(Y_1^n \times J_{21}^{\Gamma})$$
$$\nabla_2 : \mathcal{Y}_2^n \times \mathcal{J}_{12}^{\Gamma} \longmapsto [1 : 2^{nR_0}] \times [1 : 2^{nR_2}]$$
$$(\hat{M}_0, \hat{M}_2) = \nabla_2(Y_2^n \times J_{12}^{\Gamma})$$

The capacity region for the channel is defined as usual [9]. Here, we omit the details for brevity.

## III. Main Results

**Theorem 1.** *Consider the two-user BC with conferencing decoders shown in Fig. 1. Let $\mathcal{R}_o$ denote the set of all rate triples $(R_0, R_1, R_2)$ such that:*

$$R_0 + R_1 \leq I(U; Y_1) + C_{21} \quad (1)$$
$$R_1 \leq I(X; Y_1|Y_2, V) + I(X; Y_2) \quad (2)$$
$$R_1 \leq I(X; Y_2|Y_1, V) + I(X; Y_1) \quad (3)$$
$$R_0 + R_2 \leq I(V; Y_2) + C_{12} \quad (4)$$
$$R_2 \leq I(X; Y_2|Y_1, U) + I(X; Y_1) \quad (5)$$
$$R_2 \leq I(X; Y_1|Y_2, U) + I(X; Y_2) \quad (6)$$
$$R_0 + R_1 + R_2 \leq I(X; Y_1|V) + I(V; Y_2) + C_{12} + C_{21} \quad (7)$$
$$R_0 + R_1 + R_2 \leq I(X; Y_2|U) + I(U; Y_1) + C_{12} + C_{21} \quad (8)$$
$$R_0 + R_1 + R_2 \leq I(X; Y_1|Y_2, V) + I(X; Y_2) + C_{12} \quad (9)$$
$$R_0 + R_1 + R_2 \leq I(X; Y_2|Y_1, U) + I(X; Y_1) + C_{21} \quad (10)$$
$$R_0 + R_1 + R_2 \leq I(X; Y_1, Y_2) \quad (11)$$

*for some joint PDFs $P_{U,V,X}$ where $U, V \to X \to Y_1, Y_2$ forms a Markov chain. The set $\mathcal{R}_o$ constitutes an outer bound on the capacity region.*

*Proof.* The proof is given in Appendix A. □

The above outer bound $\mathcal{R}_o$ is derived using the Csiszár-Körner identity and is clearly tighter than that of [2, Prop. 1] which is essentially the cut-set bound. As we demonstrate later, the novel structure of the bound, specially the constraints (9) and (10), are crucial to derive new capacity results.

As mentioned in the introduction, in [2, Th. 2] and [4, App. B], some achievable rate regions are given for the channel. However, those regions (which are given for the channel with private messages only) are in general insufficient to derive new capacity results. We next present a more effective achievable strategy for the two-user BC with both common and private messages and bidirectional conferencing receivers.

**Theorem 2.** *Consider the two-user BC with conferencing decoders shown in Fig. 1. Let $\mathcal{R}_i^{(1)}$ denote the set of all rate triples $(R_0, R_1, R_2)$ such that*

$$R_0 + R_1 \leq \min\{I(U, W; Y_1) + \zeta, I(U, W; Y_1, \hat{Y}_2)\} \quad (12)$$
$$R_0 + R_2 \leq I(V, W; Y_2) + C_{12} \quad (13)$$
$$R_0 + R_1 + R_2 \leq \min\{I(U; Y_1|W) + \zeta, I(U; Y_1, \hat{Y}_2|W)\} + I(V, W; Y_2) + C_{12} - I(U; V|W) \quad (14)$$
$$R_0 + R_1 + R_2 \leq \min\{I(U, W; Y_1) + \zeta, I(U, W; Y_1, \hat{Y}_2)\} + I(V; Y_2|W) - I(U; V|W) \quad (15)$$
$$2R_0 + R_1 + R_2 \leq \min\{I(U, W; Y_1) + \zeta, I(U, W; Y_1, \hat{Y}_2)\} + I(V, W; Y_2) + C_{12} - I(U; V|W) \quad (16)$$

*where*

$$\zeta = \{C_{21} - I(\hat{Y}_2; Y_2|W, U, Y_1)\}^+, \quad (17)$$

*for some joint PDFs $P(u,v,w,x)P(y_1,y_2|x)P(\hat{y}_2|y_2)$. The convex closure of the set $\mathcal{R}_i^{(1)}$ is achievable.*

The cooperation scheme to derive the rate region $\mathcal{R}_i^{(1)}$ is to apply Marton's coding scheme with quantize-bin-and-forward at the second receiver first and then decode-and-forward at the first receiver. We omit the details of the proof of Theorem 2 due to lack of space. It is clear that a second achievable rate region can also be derived by exchanging the cooperative protocols at the users. Moreover, one may consider achievable rates by applying multiple rounds of cooperation at the users. However, we demonstrate below that the region (12) is already sufficient to prove new capacity results for some two-user BCs with cooperative users. Our first capacity result is regarding a class of semi-deterministic BCs with degraded message set.

**Theorem 3.** *Consider the two-user BC with degraded message sets, with a common message for both users and a private message for the first user, and bidirectional conferencing receivers. For the case of semi-deterministic channel where $Y_2 = f(X,Y_1)$, the capacity region is given by:*

$$R_0 \le I(V;Y_2) + C_{12} \tag{18}$$
$$R_0 + R_1 \le I(X;Y_1) + C_{21} \tag{19}$$
$$R_0 + R_1 \le I(X;Y_1|V) + I(V;Y_2) + C_{12} + C_{21} \tag{20}$$
$$R_0 + R_1 \le I(X;Y_1,Y_2|V) + I(V;Y_2) + C_{12} \tag{21}$$
$$R_0 + R_1 \le I(X;Y_1,Y_2) \tag{22}$$

*for some joint PDFs $P(v,x)$.*

*Proof.* The achievability is derived by setting $U \equiv X$, $W \equiv V$, and $\hat{Y}_2 \equiv Y_2$ in $\mathcal{R}_i^{(1)}$. The converse is due to $\mathcal{R}_o$. □

**Remark:** By setting $R_0 = 0$ and $C_{12} = 0$, the result of Theorem 3 is reduced to the capacity of semi-deterministic primitive relay channel derived in [10]. In this case, the capacity is given by: $\max_{P(x)} I(X;Y_1) + C_{21}$. Note that for the case of $R_0 = 0$ and $C_{12} = 0$, it is optimal to set $V \equiv \varnothing$. In fact, the cut-set bound is exactly optimal. This capacity is achieved by quantize-bin-and-forward as the relay strategy.

We now present two interesting examples of semi-deterministic BC with degraded message set.

**Example 1:** Consider the following binary channel,

$$\begin{cases} Y_1 = X \oplus Z \\ Y_2 = Z \end{cases} \tag{23}$$

where $Z$ is a binary noise and $\oplus$ is the XOR operator. For this channel, we have $Y_2 = X \oplus Y_1$, so the channel is semi-deterministic. The capacity region of this channel with degraded message set $(R_0, R_1)$ is given below:

$$R_0 \le C_{12} \tag{24}$$
$$R_0 + R_1 \le I(X;Y_1) + C_{21} \tag{25}$$
$$R_0 + R_1 \le H(X) \tag{26}$$

for some $P(x)$. In this example, the user $Y_2$ does not receive information from the transmitter directly. Instead, it observes the noise $Z$ and relays the noise to the user $Y_1$ by sending a compressed version of it through the digital link $C_{21}$. The user $Y_1$ then decodes both messages and forwards the common message to the user $Y_2$ through the digital link $C_{12}$. Therefore, the user $Y_2$ can still receive information at a positive rate.

**Example 2:** Consider the following binary channel,

$$\begin{cases} Y_1 = Z \\ Y_2 = X \oplus Z \end{cases} \tag{27}$$

where again $Z$ is a binary noise. For this case, the capacity region with degraded message set $(R_0, R_1)$ is given as follows.

$$R_0 \le I(V;Y_2) + C_{12} \tag{28}$$
$$R_0 + R_1 \le C_{21} \tag{29}$$
$$R_0 + R_1 \le H(X|V) + I(V;Y_2) + C_{12} \tag{30}$$
$$R_0 + R_1 \le H(X) \tag{31}$$

for some $P(v,x)$. As shown, unlike the previous case, the capacity region is characterized based on some auxiliary variable $V$. The optimal choice for this variable depends on the values of $C_{12}$ and $C_{21}$.

For this channel, the optimal coding strategy is as follows. Note that the user $Y_1$ (which is supposed to detect both common and private messages) does not receive any information from the transmitter directly. The optimal cooperation strategy is that the user $Y_2$ first sends a compressed version of its received signal to the user $Y_1$ through the digital link $C_{21}$. Next, the user $Y_1$ decodes both messages using the information received (and its own signal which is in fact channel noise) and then forwards part of the common message (using the variable $V$) to the user $Y_2$ through the link $C_{12}$. Lastly, the user $Y_2$ decodes the unknown part of the common message using its received signal.

It is worthwhile to make the following observation. For Example 2, as the user $Y_1$ observes the channel noise only, one would think that a cooperative scheme in which $Y_1$ applies compress-and-forward and $Y_2$ applies decode-and-forward is the right strategy. However, it turns out that such a scheme is not optimal when the message to $Y_2$ is a degraded version of the message to $Y_1$. This is in contrast to Example 1, in which the message set degrades in the opposite direction and performing compress-and-forward at the receiver that observes the noise only is the capacity-achieving strategy.

As far as we know, the above results are the first cases where a combination of quantize-and-forward and decode-and-forward strategies yields an *optimal bidirectional cooperation protocol*. There is no previously known capacity result in the literature for channels with bidirectional cooperation between users. Moreover, our results demonstrate that a one-round cooperation scheme is sufficient to achieve capacity for this class of channels.

We also prove a capacity result for a new class of semi-deterministic and more capable BCs with both common and private messages and one-sided conferencing.

**Theorem 4.** *Consider the two-user semi-deterministic BC with both common and private messages and $Y_2 = f(X,Y_1)$. Moreover, assume that the channel is more-capable, i.e., $I(X;Y_2) \le I(X;Y_1)$ for every input distribution $P(x)$. For* *the channel with one-sided conferencing where only $Y_2$ is connected to $Y_1$ by a digital link of capacity $C_{21}$, the capacity region is given by:*

$$R_0 + R_2 \le I(V; Y_2) \tag{32}$$
$$R_0 + R_1 + R_2 \le I(X; Y_1) + C_{21} \tag{33}$$
$$R_0 + R_1 + R_2 \le I(X; Y_1|V) + I(V; Y_2) + C_{21} \tag{34}$$
$$R_0 + R_1 + R_2 \le I(X; Y_1, Y_2|V) + I(V; Y_2) \tag{35}$$
$$R_0 + R_1 + R_2 \le I(X; Y_1, Y_2) \tag{36}$$

*for some joint PDFs $P(v, x)$.*

*Proof.* The achievability is derived by setting $U \equiv X$, $W \equiv V$, and $\hat{Y}_2 \equiv Y_2$ in $\mathcal{R}_i^{(1)}$. The converse is given by $\mathcal{R}_o$. □

Note that the structure of the novel outer bound $\mathcal{R}_o$ given in Theorem 1, in particular, the constraints (9) and (10), is crucial for deriving the capacity results in Theorems 3 and 4. In ongoing work [11], we show that the bounds can also be used to derive approximate capacity results for Gaussian channels.

## IV. Concluding Remarks

This paper shows that the Csiszár-Körner identity can be used to derive novel outer bounds for the BC with bidirectional conferencing receiver that are tighter than the cut-set bound. Combining with the achievability results based on quantize-bin-and-forward in one direction and decode-and-forward in the other direction, it allows us to derive capacity results for specific classes of semi-deterministic BCs with degraded message set. Two binary channel examples are used to illustrate the capacity-achieving coding strategy.

## Appendix A
## Proof of Theorem 1

We only derive the constraints that include the auxiliary variable $U$, i.e., (1), (5), (6), (8), and (10). The constraints that include $V$ are derived symmetrically, and the last constraint is due to the cut-set bound. Consider a length-$n$ code with vanishing average error probability. Define new auxiliary variables as follows:

$$U_t = (M_0, M_1, Y_2^{t-1}, Y_{1,t+1}^n) \qquad t = 1, ..., n \tag{37}$$

By Fano's inequality, we have

$$\begin{aligned} &n(R_0 + R_1) \\ &\quad\le I(M_0, M_1; Y_1^n, J_{21}^\Gamma) + n\epsilon_n^{(1)} \\ &\quad= I(M_0, M_1; Y_1^n) + I(M_0, M_1; J_{21}^\Gamma | Y_1^n) + n\epsilon_n^{(1)} \\ &\quad\le \sum_{t=1}^n I(M_0, M_1; Y_{1,t}|Y_{1,t+1}^n) + H(J_{21}^\Gamma) + n\epsilon_n^{(1)} \\ &\quad\overset{(a)}{\le} \sum_{t=1}^n I(M_0, M_1, Y_{1,t+1}^n, Y_2^{t-1}; Y_{1,t}) + H(J_{21}^\Gamma) + n\epsilon_n^{(1)} \\ &\quad\le \sum_{t=1}^n I(U_t; Y_{1,t}) + nC_{21} + n\epsilon_n^{(1)} \end{aligned} \tag{38}$$

where inequality $(a)$ holds because conditioning does not increase the entropy.

Next, we derive the constraints on $R_2$. By Fano's inequality,

$$\begin{aligned} nR_2 &\le I(M_2; Y_2^n, J_{12}^\Gamma) + n\epsilon_n^{(2)} \\ &\le I(M_2; Y_1^n, Y_2^n, J_{12}^\Gamma) + n\epsilon_n^{(2)} \\ &\overset{(a)}{=} I(M_2; Y_1^n, Y_2^n) + n\epsilon_n^{(2)} \\ &\le I(M_2; Y_1^n, Y_2^n | M_0, M_1) + n\epsilon_n^{(2)} \end{aligned} \tag{39}$$

where $(a)$ holds because $J_{12}^\Gamma$ is given by a deterministic function of $(Y_1^n, Y_2^n)$.

Now consider the following:

$$\begin{aligned} &I(M_2; Y_1^n, Y_2^n | M_0, M_1) \\ &\quad= I(M_2; Y_2^n|M_0, M_1) + I(M_2; Y_1^n|Y_2^n, M_0, M_1) \\ &\quad\le I(M_0, M_1, M_2; Y_2^n) + I(M_2; Y_1^n|Y_2^n, M_0, M_1) \\ &\quad\overset{(a)}{\le} \sum_{t=1}^n I(X_t; Y_{2,t}) \\ &\qquad + \sum_{t=1}^n I(X_t; Y_{1,t}|Y_{2,t}, M_0, M_1, Y_2^{t-1}, Y_{2,t+1}^n, Y_{1,t+1}^n) \\ &\quad= \sum_{t=1}^n I(X_t; Y_{2,t}) + \sum_{t=1}^n I(X_t; Y_{1,t}|Y_{2,t}, U_t, Y_{2,t+1}^n) \\ &\quad\overset{(b)}{\le} \sum_{t=1}^n I(X_t; Y_{2,t}) + \sum_{t=1}^n I(X_t; Y_{1,t}|Y_{2,t}, U_t) \end{aligned} \tag{40}$$

where $(a)$ holds because $X_t$ is a deterministic function of $(M_0, M_1, M_2)$ and the channel is memoryless, and $(b)$ holds because conditioning does not increase entropy and also given the input signal $X_t$, the outputs $Y_{1,t}, Y_{2,t}$ are independent of other variables.

In the same way, one can also derive:

$$\begin{aligned} &I(M_2; Y_1^n, Y_2^n | M_0, M_1) \\ &\quad= I(M_2; Y_1^n|M_0, M_1) + I(M_2; Y_2^n|Y_1^n, M_0, M_1) \\ &\quad\le \sum_{t=1}^n I(X_t; Y_{1,t}) + \sum_{t=1}^n I(X_t; Y_{2,t}|Y_{1,t}, U_t) \end{aligned} \tag{41}$$

By substituting (40) and (41) in (39), we obtain the desired bounds on $R_2$.

Next, we derive the constraints on the sum-rate. We have

$$\begin{aligned} &n(R_0 + R_1 + R_2) \\ &\quad\le I(M_0, M_1; Y_1^n, J_{21}^\Gamma) + I(M_2; Y_2^n, J_{12}^\Gamma) + n(\epsilon_n^{(1)} + \epsilon_n^{(2)}) \\ &\quad\le I(M_0, M_1; Y_1^n) + I(M_2; Y_2^n) + I(M_0, M_1; J_{21}^\Gamma|Y_1^n) \\ &\qquad\qquad + I(M_2; J_{12}^\Gamma|Y_2^n) + n(\epsilon_n^{(1)} + \epsilon_n^{(2)}) \\ &\quad\le I(M_0, M_1; Y_1^n) + I(M_2; Y_2^n|M_0, M_1) \\ &\qquad\qquad + H(J_{21}^\Gamma) + H(J_{12}^\Gamma) + n(\epsilon_n^{(1)} + \epsilon_n^{(2)}) \\ &\quad\le I(M_0, M_1; Y_1^n) + I(M_2; Y_2^n|M_0, M_1) \\ &\qquad\qquad + n(C_{12} + C_{21}) + n(\epsilon_n^{(1)} + \epsilon_n^{(2)}) \end{aligned} \tag{42}$$

Now consider the following:

$$
\begin{aligned}
&I(M_0, M_1;Y_1^n) + I(M_2; Y_2^n|M_0, M_1)\\
&\quad= \sum_{t=1}^{n} I(M_0, M_1; Y_{1,t}|Y_{1,t+1}^n)\\
&\qquad+ \sum_{t=1}^{n} I(M_2; Y_{2,t}|M_0, M_1, Y_2^{t-1})\\
&\quad\leq \sum_{t=1}^{n} I(Y_2^{t-1}, M_0, M_1; Y_{1,t}|Y_{1,t+1}^n)\\
&\qquad- \sum_{t=1}^{n} I(Y_2^{t-1}; Y_{1,t}|M_0, M_1, Y_{1,t+1}^n)\\
&\qquad+ \sum_{t=1}^{n} I(Y_{1,t+1}^n, M_2; Y_{2,t}|M_0, M_1, Y_2^{t-1})\\
&\quad\leq \sum_{t=1}^{n} I(Y_2^{t-1}, M_0, M_1; Y_{1,t}|Y_{1,t+1}^n)\\
&\qquad- \sum_{t=1}^{n} I(Y_2^{t-1}; Y_{1,t}|M_0, M_1, Y_{1,t+1}^n)\\
&\qquad+ \sum_{t=1}^{n} I(Y_{1,t+1}^n; Y_{2,t}|M_0, M_1, Y_2^{t-1})\\
&\qquad+ \sum_{t=1}^{n} I(M_2; Y_{2,t}|M_0, M_1, Y_2^{t-1}, Y_{1,t+1}^n)\\
&\quad\overset{(a)}{\leq} \sum_{t=1}^{n} I(Y_2^{t-1}, M_0, M_1; Y_{1,t}|Y_{1,t+1}^n)\\
&\qquad+ \sum_{t=1}^{n} I(M_2; Y_{2,t}|M_0, M_1, Y_2^{t-1}, Y_{1,t+1}^n)\\
&\quad\leq \sum_{t=1}^{n} I(Y_2^{t-1}, Y_{1,t+1}^n, M_0, M_1; Y_{1,t})\\
&\qquad+ \sum_{t=1}^{n} I(M_2; Y_{2,t}|M_0, M_1, Y_2^{t-1}, Y_{1,t+1}^n)\\
&\quad\leq \sum_{t=1}^{n} I(U_t; Y_{1,t}) + \sum_{t=1}^{n} I(X_t; Y_{2,t}|U_t)
\end{aligned}
\tag{43}
$$

where $(a)$ is due to the Csiszár-Körner identity. By substituting (43) in (42), we obtain the bound (8) on the sum-rate.

Finally, we can write:

$$
\begin{aligned}
&n(R_0 + R_1 + R_2)\\
&\quad\leq I(M_0, M_1; Y_1^n, J_{21}^\Gamma) + I(M_2; Y_2^n, J_{12}^\Gamma) + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})\\
&\quad\leq I(M_0, M_1; Y_1^n) + I(M_2; Y_1^n, Y_2^n, J_{12}^\Gamma)\\
&\qquad\qquad+ I(M_0, M_1; J_{21}^\Gamma|Y_1^n) + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})\\
&\quad\leq I(M_0, M_1; Y_1^n) + I(M_2; Y_1^n, Y_2^n)\\
&\qquad\qquad+ H(J_{21}^\Gamma) + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})\\
&\quad\leq I(M_0, M_1; Y_1^n) + I(M_2; Y_1^n, Y_2^n|M_0, M_1)\\
&\qquad\qquad+ nC_{21} + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})
\end{aligned}
\tag{44}
$$

Continuing with this chain of inequalities, we have

$$
\begin{aligned}
&n(R_0 + R_1 + R_2)\\
&\quad\leq I(M_0, M_1; Y_1^n) + I(M_2; Y_1^n, Y_2^n|M_0, M_1)\\
&\qquad+ nC_{21} + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})\\
&\quad\leq I(M_0, M_1; Y_1^n) + I(M_2; Y_2^n|M_0, M_1)\\
&\qquad+ I(M_2; Y_1^n|Y_2^n, M_0, M_1) + nC_{21} + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})\\
&\quad\overset{(a)}{\leq} \sum_{t=1}^{n} I(U_t; Y_{1,t}) + \sum_{t=1}^{n} I(X_t; Y_{2,t}|U_t)\\
&\qquad+ \sum_{t=1}^{n} I(X_t; Y_{1,t}|Y_{2,t}, U_t) + nC_{21} + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})\\
&\quad= \sum_{t=1}^{n} I(U_t; Y_{1,t}) + \sum_{t=1}^{n} I(X_t; Y_{1,t}, Y_{2,t}|U_t)\\
&\qquad+ nC_{21} + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})\\
&\quad= \sum_{t=1}^{n} I(U_t; Y_{1,t}) + \sum_{t=1}^{n} I(X_t; Y_{1,t}|U_t)\\
&\qquad+ \sum_{t=1}^{n} I(X_t; Y_{2,t}|Y_{1,t}, U_t) + nC_{21} + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})\\
&\quad= \sum_{t=1}^{n} I(X_t; Y_{1,t}) + \sum_{t=1}^{n} I(X_t; Y_{2,t}|Y_{1,t}, U_t)\\
&\qquad+ nC_{21} + n(\epsilon_n^{(1)} + \epsilon_n^{(2)})
\end{aligned}
\tag{45}
$$

where inequality $(a)$ is derived by following the same line of argument as in (43). By applying a standard time-sharing argument, we derive the desired constraint (10).

Finally, (11) is just the cut-set bound. The proof is thus complete.

## References


[1] S. C. Draper, B. J. Frey, and F. R. Kschischang, “Interactive Decoding of a Broadcast Message,” in *Proc. Allerton Conf. Commun., Control, Computing*, Monticello, IL, 2003.

[2] R. Dabora, and S. D. Servetto, “Broadcast channels with cooperating decoders,” *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5438–5454,Dec. 2006.

[3] L. Dikstein, H. H. Permuter, and Y. Steinberg, “On state-dependent degraded broadcast channels with cooperation,” *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2308–2323, May 2016.

[4] Z. Goldfeld, H. H. Permuter, and G. Kramer, “Duality of a source coding problem and the semi-deterministic broadcast channel with rate-limited cooperation,” *IEEE Trans. Inf. Theory*, vol. 65, no. 5, pp. 2285–2307, May 2016.

[5] Z. Goldfeld, G. Kramer, H. H. Permuter, and P. Cuff, “Strong secrecy for cooperative broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 63, no. 19, pp. 469–495, Jan. 2017.

[6] W. Huleihel, and Y. Steinberg, “Channels with cooperation links that may be absent,” *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5886–5906, Sep. 2017.

[7] D. Itzhak, and Y. Steinberg, “The broadcast channel with degraded message sets and unreliable conference,” *IEEE Trans. Inf. Theory*, vol. 67, no. 9, pp. 5623–5650, Sep. 2021.

[8] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” *IEEE Trans. Inf. Theory,* vol. IT-24, no. 3, pp. 339–348, May 1978.

[9] A. El Gamal, and Y. H. Kim, Network Information Theory. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[10] Y.-H. Kim, “Capacity of a class of deterministic relay channels,” *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1328–1329, Mar. 2008.

[11] R. K. Farsani and Wei Yu, “Capacity Bounds for Broadcast Channels with Bidirectional Conferencing Decoders,” in preparation.