# Leveraging Receiver Message Side Information in Two-Receiver Broadcast Channels: A General Approach †

**Behzad Asadi \*, Lawrence Ong and Sarah J. Johnson**

School of Electrical Engineering and Computing, University of Newcastle, Callaghan, New South Wales 2308, Australia; lawrence.ong@newcastle.edu.au (L.O.); sarah.johnson@newcastle.edu.au (S.J.J.)

\*   Correspondence: behzad.asadi@uon.edu.au; Tel.: +61-249-215-291

†   This work was presented partially at the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015, and the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016.

**Abstract:** We consider two-receiver broadcast channels where each receiver may know a priori some of the messages requested by the other receiver as receiver message side information (RMSI). We devise a general approach to leverage RMSI in these channels. To this end, we first propose a pre-coding scheme considering the general message setup where each receiver requests both common and private messages and knows a priori part of the private message requested by the other receiver as RMSI. We then construct the transmission scheme of a two-receiver channel with RMSI by applying the proposed pre-coding scheme to the best transmission scheme for the channel without RMSI. To demonstrate the effectiveness of our approach, we apply our pre-coding scheme to three categories of the two-receiver discrete memoryless broadcast channel: (i) channel without state; (ii) channel with states known causally to the transmitter; and (iii) channel with states known non-causally to the transmitter. We then derive a unified inner bound for all three categories. We show that our inner bound is tight for some new cases in each of the three categories, as well as all cases whose capacity region was known previously.

**Keywords:** broadcast channel; capacity region; receiver message side information; channel state information

---

## 1. Introduction

Communication over wireless channels motivates the study of broadcast channels [1]. In these channels, a transmitter wishes to send a number of messages to multiple receivers via a noisy shared medium, which can also be time-varying due to, for example, fading or interference. The time-varying factor is commonly referred to as the channel state.

The messages to be sent by the transmitter may already be present in parts at some of the receivers, referred to as receiver message side information (RMSI). This form of side information appears in multimedia broadcasting with packet loss. Suppose that a multimedia file is requested by multiple receivers, and the transmission is subject to packet loss. After a few rounds of transmissions, each receiver may a priori know some of the packets re-demanded (due to packet loss) by other receivers. This form of side information appears also in the downlink phase of applications modeled by the multi-way relay channel [2]. For example, consider two stations exchanging data through a satellite. Since each station is also the source of the message to be transmitted by the satellite to the other station in the downlink phase, this phase can be modeled by a broadcast channel with RMSI. As proper use of RMSI may increase transmission rates over broadcast channels, we investigate the capacity

region of two-receiver broadcast channels with RMSI. We consider three categories of the two-receiver discrete memoryless broadcast channel: (i) channel without state; (ii) channel with causal channel state information at the transmitter (CSIT) including the cases where the channel state may also be available causally or non-causally at each receiver (It does not make a difference whether the channel state is available causally or non-causally at a receiver. This is due to block decoding, where the receiver decodes its requested message(s) at the end of the channel-output sequence.); and (iii) channel with non-causal CSIT including the cases where the channel state may also be available causally or non-causally at each receiver.

As we will see later, we construct transmission schemes of two-receiver broadcast channels with RMSI using a pre-coding scheme and the best transmission schemes for the channels without RMSI. We here present existing capacity results for the aforementioned three categories of both the two-receiver discrete memoryless broadcast channel without and with RMSI.

### 1.1. Broadcast Channel without RMSI

#### 1.1.1. Without State

The capacity region is known for several special types of the discrete memoryless broadcast channels, such as the degraded [3], less noisy [4] (p. 121), more capable [5], deterministic [6,7] and semideterministic [8] broadcast channel. The capacity region is also known for the discrete memoryless broadcast channel with degraded message sets (a special message setup where one of the receivers requests only a common message) [9]. Marton's inner bound with a common message [4] (p. 212) is tight for all cases in this category whose capacity region is known. This inner bound is achieved using superposition coding and Marton coding [4] (p. 207).

#### 1.1.2. With Causal CSIT

Under this category, the capacity region is known for the following two special cases: the degraded broadcast channel with private messages where the channel state is available causally at: (i) only the transmitter [10]; or (ii) the transmitter and the non-degraded receiver [4] (p. 184). Superposition coding and the Shannon strategy [11] were used to characterize the capacity region of these two cases. Superposition coding achieves the capacity region of the degraded broadcast channel without state, and the Shannon strategy is a capacity-achieving strategy for the discrete memoryless point-to-point channel with causal CSIT.

#### 1.1.3. With Non-Causal CSIT

Under this category, the capacity region is known for a few special cases, such as the degraded broadcast channel with private messages where the channel state is available non-causally at the transmitter and the non-degraded receiver [10], and the semideterministic broadcast channel with private messages where the channel state is available non-causally at only the transmitter [12]. Khosravi-Farsani and Marvasti [13] (Theorem 2) used superposition coding, Marton coding and Gelfand–Pinsker coding [14] to derive the best-known inner bound for the discrete memoryless broadcast channel with non-causal CSIT (i.e., it is tight for all cases with known capacity region). Superposition coding plus Marton coding achieves the best-known inner bound for the discrete memoryless broadcast channel without state, and Gelfand–Pinsker coding achieves the capacity region of the discrete memoryless point-to-point channel with non-causal CSIT.

### 1.2. Broadcast Channel with RMSI

#### 1.2.1. Without State

The capacity region is known for the following three special cases: (i) the discrete memoryless channel with complementary RMSI [15,16] where both receivers need to decode all of the source

messages, i.e., all of the messages not known a priori; (ii) the discrete memoryless channel with degraded message sets [17] where one of the receivers needs to decode all of the source messages and one only part of the source messages; and (iii) the less noisy channel with the general message setup [18] (Theorem 3) where each receiver has both common- and private-message requests and knows part of the private message requested by the other receiver.

### 1.2.2. With Causal CSIT

Under this category, the capacity region is known for the two-receiver discrete memoryless broadcast channel with complementary RMSI where the channel state is available causally at only the transmitter [19].

### 1.2.3. With Non-Causal CSIT

Under this category, the capacity region is known for the two-receiver discrete memoryless broadcast channel with complementary RMSI where the channel state is available non-causally at the transmitter and one of the receivers [20].

## 2. Summary of the Main Results

We propose a pre-coding scheme to construct the transmission scheme of a two-receiver broadcast channel with RMSI by applying the pre-coding scheme to the best transmission scheme for the channel without RMSI. We design the pre-coding scheme considering the general message setup where each receiver requests both common and private messages and knows part of the private message requested by the other receiver as RMSI. This provides a general approach for utilizing RMSI in two-receiver broadcast channels. We use our pre-coding scheme and derive a unified inner bound for three categories of the discrete memoryless broadcast channel: (i) channel without state; (ii) channel with causal CSIT (causal category); and (iii) channel with non-causal CSIT (non-causal category). The steps to derive our unified inner bound are shown in Figure 1 in which rectangles with solid sides represent the bounds established in this work. Here, we briefly explain the steps.



**Figure 1.** The steps to derive our unified inner bound (the rightmost rectangle) for the three categories: (i) channel with receiver message side information (RMSI), without state; (ii) channel with RMSI and causal channel state information at the transmitter (CSIT); and (iii) channel with RMSI and non-causal CSIT. The rectangles with dashed sides represent the bounds derived prior to this work (i.e., Marton's inner bound with a common message [4] and non-causal CSIT inner bound [13]). The rectangles with solid sides represent the bounds derived in this work. The arrows provide the key techniques to derive each bound.

Step 1: We first derive an inner bound for the causal category without RMSI. We use superposition coding, Marton coding and the Shannon strategy to construct the transmission scheme. By choosing the common message to be a constant, our scheme reduces to the one used by Kramer [21] for the discrete memoryless broadcast channel without RMSI, with causal CSIT. Our inner bound is tight for all of the cases without RMSI, with causal CSIT, whose capacity region is known (mentioned in the previous section).

Step 2: We then unify our inner bound for the causal category without RMSI, and the best-known inner bound for the non-causal category without RMSI by Khosravi-Farsani and Marvasti [13] (Theorem 2). This result is analogous to the work of Jafar [22] in which a unified capacity-region expression is provided for the point-to-point channel with causal CSIT and the point-to-point channel with non-causal CSIT. Clearly, the capacity region of a non-causal case is larger than or equal to the capacity region of the corresponding causal case (where only the transmitter knows the channel state causally instead of non-causally, and the knowledge of the receivers about the channel state is the same in both cases). This relationship is not necessarily true for their inner bounds especially when one transmission scheme is not a special case of the other. One of the advantages of having a unified inner bound is that it allows us to show that the best-known inner bound ([13] (Theorem 2)) for a non-causal case is larger than or equal to the best inner bound (our inner bound) for the corresponding causal case.

By considering that the channel state is zero with probability one for the channel without state, our unified inner bound reduces to Marton's inner bound with a common message [4] (p. 212) for the discrete memoryless broadcast channel without state. Consequently, our inner bound covers all three categories without RMSI.

Step 3: Moving on to the channel with RMSI, we propose a pre-coding scheme in order to take RMSI into account. We use our pre-coding scheme in conjunction with the schemes achieving the best inner bounds for the causal and non-causal categories without RMSI, and derive a unified inner bound, which is the best inner bound for all three categories with RMSI. This inner bound reduces to the unified inner bound without RMSI by setting RMSI to zero.

Capacity results: Using our inner bound, we obtain the following new capacity results for the discrete memoryless broadcast channel with RMSI. We also show that our inner bound is tight for all of the cases whose capacity region was known prior to this work. This demonstrates the effectiveness of our proposed approach for utilizing RMSI.

- For the channel with RMSI, without state, we establish the capacity region of two new cases, namely the deterministic channel and the more capable channel. In a concurrent work with the preliminary published version of this work [23], Bracher and Wigger [24] established the capacity region of the semideterministic channel without state for which our inner bound is also tight.
- For the channel with RMSI and causal CSIT, we establish the capacity region of the degraded broadcast channel where the channel state is available causally at: (i) only the transmitter; (ii) the transmitter and the non-degraded receiver; or (iii) the transmitter and both receivers.
- For the channel with RMSI and non-causal CSIT, we establish the capacity region of the degraded broadcast channel where the channel state is available non-causally at: (i) the transmitter and the non-degraded receiver; or (ii) the transmitter and both receivers.

## 3. System Model

We consider the two-receiver discrete memoryless broadcast channel with independent and identically distributed (i.i.d.) states $p(y_1, y_2 \mid x_0, s)p(s)$, depicted in Figure 2, where $X_0 \in \mathcal{X}_0$ is the channel input, $Y_1 \in \mathcal{Y}_1$ and $Y_2 \in \mathcal{Y}_2$ are the channel outputs and $S \in \mathcal{S}$ is the channel state. Considering $n$ uses of the channel, $X_0^n = (X_{0,1}, X_{0,2}, \ldots, X_{0,n})$ is the transmitted codeword, and $Y_i^n = (Y_{i,1}, Y_{i,2}, \ldots, Y_{i,n})$, $i = 1, 2$, is the channel-output sequence at receiver $i$.

The source messages $\{M_i\}_{i=0}^4$ are independent, and $M_i$ is uniformly distributed over the set $\mathcal{M}_i = \{1, 2, \ldots, 2^{nR_i}\}$, i.e., transmitted at rate $R_i$ bits per channel use. Receiver 1 requests $\{M_0, M_1, M_3\}$ and knows $M_4$ as RMSI. Receiver 2 requests $\{M_0, M_2, M_4\}$ and knows $M_3$ as RMSI. For Receiver 1,

$M_1$ is the part of the private-message request that is not known a priori to the other receiver, and $M_3$ is the part that is known. For Receiver 2, these are $M_2$ and $M_4$, respectively.
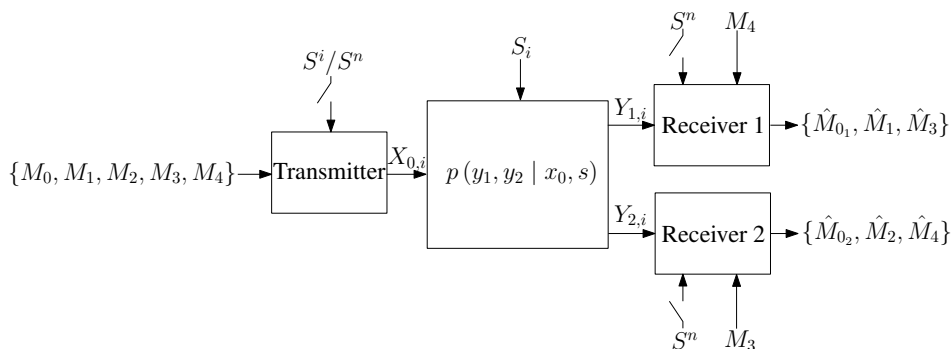


**Figure 2.** The two-receiver discrete memoryless broadcast channel with i.i.d. states $p(y_1, y_2 \mid x_0, s)p(s)$ where the channel state may be available either causally or non-causally at each of the transmitter and receivers. Receiver 1 requests $(M_0, M_1, M_3)$ while it knows $M_4$ as side information. Receiver 2 requests $(M_0, M_2, M_4)$ while it knows $M_3$ as side information. $(\hat{M}_{0_1}, \hat{M}_1, \hat{M}_3)$ is the decoded version of $(M_0, M_1, M_3)$ at Receiver 1, and $(\hat{M}_{0_2}, \hat{M}_2, \hat{M}_4)$ is the decoded version of $(M_0, M_2, M_4)$ at Receiver 2.

The channel without state is a special case of our channel model by considering $\mathcal{S} = \{0\}$, i.e., $p_S(0) = 1$. This implies that the transmitter and receivers know that the channel state is equal to zero at all channel uses. The channel without RMSI is also a special case of our channel model by considering $(M_3, M_4) = (0, 0)$.

A $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR_3}, 2^{nR_4}, n\right)$ causal code for the channel consists of a sequence of maps for the encoding:

$$f_j : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_3 \times \mathcal{M}_4 \times \mathcal{S}^j \to \mathcal{X}_0, \; j = 1, 2, \cdots, n,$$

where $\times$ denotes the Cartesian product and $\mathcal{S}^j$ denotes the $j$-fold Cartesian product of $\mathcal{S}$, i.e., $X_{0,j} = f_j(M_0, M_1, M_2, M_3, M_4, S^j)$ where $S^j = (S_1, S_2, \ldots, S_j)$. This code also consists of two decoding functions:

$$g_1 : \tilde{\mathcal{Y}}_1^n \times \mathcal{M}_4 \to \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_3,$$
$$g_2 : \tilde{\mathcal{Y}}_2^n \times \mathcal{M}_3 \to \mathcal{M}_0 \times \mathcal{M}_2 \times \mathcal{M}_4,$$

where $\tilde{\mathcal{Y}}_i = \mathcal{Y}_i \times \mathcal{S}$, $i = 1, 2$, if the channel state is available at receiver $i$, and $\tilde{\mathcal{Y}}_i = \mathcal{Y}_i$ otherwise. $(\hat{M}_{0_1}, \hat{M}_1, \hat{M}_3) = g_1(\tilde{Y}_1^n, M_4)$ is the decoded $(M_0, M_1, M_3)$ at Receiver 1, and $(\hat{M}_{0_2}, \hat{M}_2, \hat{M}_4) = g_2(\tilde{Y}_2^n, M_3)$ is the decoded $(M_0, M_2, M_4)$ at Receiver 2 where $\tilde{Y}_i^n = (Y_i^n, S^n)$ if the channel state is available at receiver $i$ and $\tilde{Y}_i^n = Y_i^n$ otherwise.

A $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR_3}, 2^{nR_4}, n\right)$ non-causal code for the channel consists of an encoding function:

$$f : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{M}_3 \times \mathcal{M}_4 \times \mathcal{S}^n \to \mathcal{X}_0^n,$$

i.e., $X_0^n = f(M_0, M_1, M_2, M_3, M_4, S^n)$. This code also consists of two decoding functions, which are defined the same as for the causal code.

A $\left(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR_3}, 2^{nR_4}, n\right)$ code for the channel without state is defined by choosing $\mathcal{S} = \{0\}$ in the definition of either the causal code or the non-causal code.

The average probability of error for a code is defined as:

$$P_e^{(n)} = P\left((\hat{M}_{0_1}, \hat{M}_1, \hat{M}_3) \neq (M_0, M_1, M_3) \text{ or } (\hat{M}_{0_2}, \hat{M}_2, \hat{M}_4) \neq (M_0, M_2, M_4)\right).$$

**Definition 1.** *For causal (non-causal) cases, a rate tuple $(R_0, R_1, R_2, R_3, R_4)$ is said to be achievable for the channel if there exists a sequence of $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, 2^{nR_3}, 2^{nR_4}, n)$ causal (non-causal) codes with $P_e^{(n)} \to 0$ as $n \to \infty$.*

**Definition 2.** *The capacity region of the channel is the closure of the set of all achievable rate tuples $(R_0, R_1, R_2, R_3, R_4)$.*

**Definition 3.** *The two-receiver discrete memoryless broadcast channel with states is said to be physically degraded if:*

$$p(y_1, y_2 \mid x_0, s) = p(y_1 \mid x_0, s) \, p(y_2 \mid y_1),$$

*and it is said to be stochastically degraded or degraded if there exists a $Y_2'$ such that $(X_0, S) \to Y_1 \to Y_2'$ form a Markov chain, and we have:*

$$p_{Y_2 \mid X_0, S}(y_2 \mid x_0, s) = \sum_{y_1} p_{Y_1 \mid X_0, S}(y_1 \mid x_0, s) \, p_{Y_2' \mid Y_1}(y_2 \mid y_1).$$

**Definition 4.** *The two-receiver discrete memoryless broadcast channel without state is said to be deterministic if the channel outputs are deterministic functions of the channel input, i.e., $Y_i = \phi_i(X_0)$, $i = 1, 2$.*

**Definition 5.** *The two-receiver discrete memoryless broadcast channel without state is said to be more capable if $I(X_0; Y_1) \geq I(X_0; Y_2)$ for any probability mass function (pmf) $p(x_0)$.*

## 4. Broadcast Channel without RMSI

In this section, we address the two-receiver discrete memoryless broadcast channel without RMSI, i.e., $(M_3, M_4) = (0, 0)$. We first derive an inner bound for the channel with causal CSIT. We then present the best-known inner bound for the channel with non-causal CSIT [13] (Theorem 2). We finally show that we can have a unified inner bound that covers both causal and non-causal cases, although the achievability schemes are different.

### 4.1. With Causal CSIT

We utilize superposition coding, Marton coding and the Shannon strategy to derive an inner bound for the discrete memoryless broadcast channel with causal CSIT, stated as Theorem 1.

**Theorem 1.** *A rate triple $(R_0, R_1, R_2)$ for the two-receiver discrete memoryless broadcast channel with causal CSIT is achievable if it satisfies:*

$$R_0 + R_1 < I(U_0, U_1; \tilde{Y}_1), \tag{1}$$

$$R_0 + R_2 < I(U_0, U_2; \tilde{Y}_2), \tag{2}$$

$$R_0 + R_1 + R_2 < I(U_0, U_1; \tilde{Y}_1) + I(U_2; \tilde{Y}_2 \mid U_0) - I(U_1; U_2 \mid U_0), \tag{3}$$

$$R_0 + R_1 + R_2 < I(U_1; \tilde{Y}_1 \mid U_0) + I(U_0, U_2; \tilde{Y}_2) - I(U_1; U_2 \mid U_0), \tag{4}$$

$$2R_0 + R_1 + R_2 < I(U_0, U_1; \tilde{Y}_1) + I(U_0, U_2; \tilde{Y}_2) - I(U_1; U_2 \mid U_0), \tag{5}$$

*for some pmf $p(u_0, u_1, u_2)$ and some function $x_0 = \gamma(u_0, u_1, u_2, s)$. $\tilde{Y}_i = (Y_i, S)$, $i = 1, 2$, if the channel state is available at receiver i and $\tilde{Y}_i = Y_i$ otherwise.*

The proof of this theorem is similar to the proof of Marton's inner bound with a common message [4] (p. 212) for the discrete memoryless broadcast without state. In the encoding, we just need to add the Shannon strategy, and in the decoding, we need to consider $\tilde{Y}_i^n$ instead of $Y_i^n$ as the channel

output sequence at receiver $i$. We here present the proof as we refer to it in Appendix A and in the next section for the channel with RMSI.

**Proof of Theorem 1.** (codebook construction) The codebook of the transmission scheme is formed from three sub-codebooks constructed using the pmf $p(u_0, u_1, u_2)$. Before constructing the sub-codebooks, using rate splitting, $M_i$, $i = 1, 2$, is divided into the two independent messages $M_{i1}$ of rate $R_{i1}$ and $M_{i2}$ of rate $R_{i2}$, such that $R_i = R_{i1} + R_{i2}$. Sub-codebook 0 consists of i.i.d. codewords:

$$u_0^n(m_0, m_{11}, m_{21}),$$

generated according to $\prod_{j=1}^n p_{U_0}(u_{0,j})$. Sub-codebook $i$, $i = 1, 2$, consists of codewords:

$$u_i^n(m_0, m_{11}, m_{21}, m_{i2}, l_i),$$

generated according to:

$$\prod_{j=1}^n p_{U_i|U_0}(u_{i,j}|u_{0,j}(m_0, m_{11}, m_{21})),$$

where $l_i \in \{1, \ldots, 2^{nR_i'}\}$, i.e., for each $u_0^n(m_0, m_{11}, m_{21})$, $2^{n(R_{i2}+R_i')}$ codewords are generated.

(Encoding) For the encoding, given $\{m_i\}_{i=0}^2$, we first find a pair $(l_1, l_2)$, such that:

$$(U_0^n(\cdot), U_1^n(\cdot, l_1), U_2^n(\cdot, l_2)) \in \mathcal{T}_{\epsilon'}^n,$$

where $\mathcal{T}_{\epsilon'}^n$ is the set of jointly $\epsilon'$-typical $n$-sequences with respect to the considered distribution [4] (p. 29). If there is more than one pair, we arbitrary choose one of them, and if there is no pair, we choose $(l_1, l_2) = (1, 1)$. We then construct the transmitted codeword as $x_{0,j} = \gamma(u_{0,j}(\cdot), u_{1,j}(\cdot), u_{2,j}(\cdot), s_j)$, $j = 1, 2, \ldots, n$.

(Decoding) Receiver 1 decodes $(\hat{m}_{0_1}, \hat{m}_{11}, \hat{m}_{12})$, if it is the unique tuple that satisfies:

$$(U_0^n(\cdot), U_1^n(\cdot, l_1), \tilde{Y}_1^n) \in \mathcal{T}_{\epsilon}^n \quad \text{for some } m_{21} \text{ and } l_1;$$

otherwise, an error is declared. Receiver 2 similarly decodes $(\hat{m}_{0_2}, \hat{m}_{21}, \hat{m}_{22})$, if it is the unique tuple that satisfies:

$$(U_0^n(\cdot), U_2^n(\cdot, l_2), \tilde{Y}_2^n) \in \mathcal{T}_{\epsilon}^n \quad \text{for some } m_{11} \text{ and } l_2;$$

otherwise, an error is declared. $\epsilon$ in the decoding is strictly greater than $\epsilon'$ in the encoding.

To derive sufficient conditions for achievability, we assume without loss of generality that the transmitted messages are each equal to one by the symmetry of code construction, and $(l_1, l_2) = (l_1^*, l_2^*)$ where $1 \leq l_i^* \leq 2^{nR_i'}$. Receiver 1 makes an error only if one or more of the following events happen:

$\mathcal{E}_0 : (U_0^n(1, 1, 1), U_1^n(1, 1, 1, 1, l_1), U_2^n(1, 1, 1, 1, l_2)) \notin \mathcal{T}_{\epsilon'}^n$ for all $l_1$ and $l_2$,

$\mathcal{E}_{11} : (U_0^n(1, 1, 1), U_1^n(1, 1, 1, 1, l_1^*), \tilde{Y}_1^n) \notin \mathcal{T}_{\epsilon}^n$,

$\mathcal{E}_{12} : (U_0^n(1, 1, 1), U_1^n(1, 1, 1, m_{12}, l_1), \tilde{Y}_1^n) \in \mathcal{T}_{\epsilon}^n$ for some $m_{12} \neq 1$ and $l_1$,

$\mathcal{E}_{13} : (U_0^n(1, 1, m_{21}), U_1^n(1, 1, m_{21}, m_{12}, l_1), \tilde{Y}_1^n) \in \mathcal{T}_{\epsilon}^n$ for some $m_{21} \neq 1, m_{12} \neq 1$ and $l_1$,

$\mathcal{E}_{14} : (U_0^n(m_0, m_{11}, m_{21}), U_1^n(m_0, m_{11}, m_{21}, m_{12}, l_1), \tilde{Y}_1^n) \in \mathcal{T}_{\epsilon}^n$

$\qquad\qquad\qquad\qquad\qquad$ for some $(m_0, m_{11}) \neq (1, 1), m_{21}, m_{12}$ and $l_1$.

Events leading to an error at Receiver 2 are written similarly. Based on the error events and using the packing lemma [4] (p. 45) and the mutual covering lemma [4] (p. 208), sufficient conditions for achievability are:

$$R_1' + R_2' > I(U_1; U_2 \mid U_0),$$
$$R_{12} + R_1' < I(U_1; \tilde{Y}_1 \mid U_0),$$
$$R_0 + R_1 + R_{21} + R_1' < I(U_0, U_1; \tilde{Y}_1),$$
$$R_{22} + R_2' < I(U_2; \tilde{Y}_2 \mid U_0),$$
$$R_0 + R_2 + R_{11} + R_2' < I(U_0, U_2; \tilde{Y}_2).$$

We finally perform the Fourier–Motzkin elimination to obtain the conditions in (1)–(5). □

### 4.2. With Non-Causal CSIT

Superposition coding, Marton coding and Gelfand–Pinsker coding were used to derive an inner bound for the discrete memoryless broadcast channel with non-causal CSIT [13] (Theorem 2). This inner bound, stated as Proposition 1, is the best-known inner bound for the non-causal category without RMSI.

**Proposition 1.** *A rate triple* $(R_0, R_1, R_2)$ *for the two-receiver discrete memoryless broadcast channel with non-causal CSIT is achievable if it satisfies:*

$$R_0 + R_1 < I(U_0, U_1; \tilde{Y}_1) - I(U_0, U_1; S), \tag{6}$$

$$R_0 + R_2 < I(U_0, U_2; \tilde{Y}_2) - I(U_0, U_2; S), \tag{7}$$

$$R_0 + R_1 + R_2 < I(U_0, U_1; \tilde{Y}_1) + I(U_2; \tilde{Y}_2 \mid U_0) - I(U_1; U_2 \mid U_0) - I(U_0, U_1, U_2; S), \tag{8}$$

$$R_0 + R_1 + R_2 < I(U_1; \tilde{Y}_1 \mid U_0) + I(U_0, U_2; \tilde{Y}_2) - I(U_1; U_2 \mid U_0) - I(U_0, U_1, U_2; S), \tag{9}$$

$$2R_0 + R_1 + R_2 < I(U_0, U_1; \tilde{Y}_1) + I(U_0, U_2; \tilde{Y}_2) - I(U_1; U_2 \mid U_0) - I(U_0, U_1, U_2; S) - I(U_0; S), \tag{10}$$

*for some pmf* $p(u_0, u_1, u_2 \mid s)$ *and some function* $x_0 = \gamma(u_0, u_1, u_2, s)$. $\tilde{Y}_i = (Y_i, S)$, $i = 1, 2$, *if the channel state is available at receiver* $i$ *and* $\tilde{Y}_i = Y_i$ *otherwise.*

Here, we review the scheme achieving this inner bound in order to highlight its differences with the scheme for the channel with causal CSIT.

(Codebook construction) The transmission scheme is formed from three sub-codebooks. These sub-codebooks are constructed using the pmf $p(u_0, u_1, u_2 \mid s)$ (i.e., $(U_0, U_1, U_2)$ is not independent of $S$). Before constructing the sub-codebook, using rate splitting, $M_i$, $i = 1, 2$, is divided into the two independent messages $M_{i1}$ of rate $R_{i1}$ and $M_{i2}$ of rate $R_{i2}$, such that $R_i = R_{i1} + R_{i2}$. Sub-codebook 0 consists of i.i.d. codewords:

$$u_0^n(m_0, m_{11}, m_{21}, l_0)$$

generated according to $\prod_{j=1}^{n} p_{U_0}(u_{0,j})$ where $l_0 \in \{1, \dots, 2^{nR_0'}\}$, i.e., for each $(m_0, m_{11}, m_{21})$, $2^{nR_0'}$ codewords are generated. Sub-codebook $i$, $i = 1, 2$, consists of codewords:

$$u_i^n(m_0, m_{11}, m_{21}, l_0, m_{i2}, l_i)$$

generated according to:

$$\prod_{j=1}^{n} p_{U_i|U_0}(u_{i,j} \mid u_{0,j}(m_0, m_{11}, m_{21}, l_0)),$$

where $l_i \in \{1, \dots, 2^{nR_i'}\}$.

(Encoding) For the encoding, given $\{m_i\}_{i=0}^{2}$, we first find a triple $(l_0, l_1, l_2)$ such that:

$$(U_0^n(\cdot, l_0), U_1^n(\cdot, l_0, \cdot, l_1), U_2^n(\cdot, l_0, \cdot, l_2), S^n) \in \mathcal{T}_{\epsilon'}^n.$$

If there is more than one triple, we arbitrarily choose one of them, and if there is no triple, we choose $(l_0, l_1, l_2) = (1, 1, 1)$. We then construct the transmitted codeword as $x_{0,j} = \gamma(u_{0,j}(\cdot), u_{1,j}(\cdot), u_{2,j}(\cdot), s_j)$, $j = 1, 2, \ldots, n$.

(Decoding) Receiver 1 decodes $(\hat{m}_{0_1}, \hat{m}_{11}, \hat{m}_{12})$, if it is the unique tuple that satisfies:

$$(U_0^n(\cdot, l_0), U_1^n(\cdot, l_0, \cdot, l_1), \tilde{Y}_1^n) \in \mathcal{T}_{\epsilon}^n \text{ for some } (m_{21}, l_0, l_1);$$

otherwise, an error is declared. Receiver 2 decodes $(\hat{m}_{0_2}, \hat{m}_{21}, \hat{m}_{22})$, if it is the unique tuple that satisfies:

$$(U_0^n(\cdot, l_0), U_2^n(\cdot, l_0, \cdot, l_2), \tilde{Y}_2^n) \in \mathcal{T}_{\epsilon}^n \text{ for some } (m_{11}, l_0, l_2);$$

otherwise, an error is declared. $\epsilon$ in the decoding is strictly greater than $\epsilon'$ in the encoding.

### 4.3. A Unified Inner Bound

Here, we discuss that we can have a unified inner-bound expression for both causal and non-causal cases. The inequalities in (6)–(10) can be used for both the channel with causal CSIT and the channel with non-causal CSIT. This is because, for the channel with causal CSIT, $(U_0, U_1, U_2)$ is independent of $S$, and the terms $I(U_0; S)$, $I(U_0, U_1; S)$, $I(U_0, U_2; S)$ and $I(U_0, U_1, U_2; S)$ are zero. Then, the inequalities in (6)–(10) reduce to the ones in (1)–(5), and we can have a unified inner bound, stated as Corollary 1.

**Corollary 1.** *A rate triple $(R_0, R_1, R_2)$ is achievable for the discrete memoryless broadcast channel with causal CSIT if it satisfies (6)–(10) for some pmf $p(u_0, u_1, u_2)$ and some function $x_0 = \gamma(u_0, u_1, u_2, s)$. It is achievable for the discrete memoryless broadcast channel with non-causal CSIT if it satisfies (6)–(10) for some pmf $p(u_0, u_1, u_2 \mid s)$ and some function $x_0 = \gamma(u_0, u_1, u_2, s)$.*

**Remark 1.** *This unified inner bound is also the best inner bound for the channel without state. This is because it reduces to Marton's inner bound with a common message [4] (p. 212) by considering that the channel state is zero with probability one for the channel without state. Therefore, it is a unified inner bound for all three categories without RMSI.*

As discussed in Section 2, the capacity region of a non-causal case is larger than or equal to the capacity region of the corresponding causal case. However, an inner bound for a non-causal case is not necessarily larger than an inner bound for the corresponding causal case when the scheme for the latter is not a special case of the scheme for the former. By having a unified inner bound, we can show that the inner bound for a non-causal case (Proposition 1) is larger than or equal to the inner bound for the corresponding causal case (Theorem 1). This is because the domain of the unified inner bound for a causal case (including all pmfs $p(u_0, u_1, u_2)$ and all functions $x_0 = \gamma(u_0, u_1, u_2, s)$) is a subset of the domain of the unified inner bound for the corresponding non-causal case (including all pmfs $p(u_0, u_1, u_2 \mid s)$ and all functions $x_0 = \gamma(u_0, u_1, u_2, s)$).

In Appendix A, we show that the scheme for the causal category, described in Section 4.1, is not a special case of the scheme for the non-causal category, described in Section 4.2. However, we show that by considering some special cases of the parameters in the scheme for the non-causal category, it asymptotically almost surely has the same codebook construction, encoding and decoding as the scheme for the causal category.

## 5. Broadcast Channel with RMSI

In this section, we address the two-receiver discrete memoryless broadcast channel with RMSI. We first propose a pre-coding scheme designed to construct the transmission scheme of a channel with RMSI based on the transmission scheme of the channel without RMSI. By applying this pre-coding scheme to the transmission schemes achieving the best inner bounds for the causal and non-causal categories without RMSI, we then derive a unified inner bound that covers all three categories with RMSI. This inner bound includes the unified inner bound for the three categories without RMSI, stated as Corollary 1, as a special case.

### 5.1. Moving from without RMSI to with RMSI

We construct a pre-coding scheme for a channel with RMSI by considering $M_{\mathrm{m}} = (M_0, M_3, M_4)$ as a new common message and treating only $M_1$ and $M_2$ as the private messages. $M_{\mathrm{m}}$, $M_1$ and $M_2$ are then fed to the transmission scheme of the same channel without RMSI. Although Receiver 1 need not decode $M_4$, having $M_4$ as a part of the common message does not impose any extra constraint. This is because Receiver 1 knows $M_4$ a priori. The same argument applies to $M_3$ for Receiver 2. Since Receiver 1 knows $M_4$ a priori and Receiver 2 knows $M_3$ a priori, Receiver 1 decodes $M_{\mathrm{m}}$ over a set $2^{n(R_0 + R_3)}$ candidates, and Receiver 2 decodes it over a set of $2^{n(R_0 + R_4)}$ candidates.

**Conjecture 1.** *We conjecture that our pre-coding scheme is an optimal pre-coding scheme in the sense that if a transmission scheme achieves the capacity region of a channel without RMSI, then the transmission scheme, constructed by applying our pre-coding scheme to that transmission scheme, also achieves the capacity region of the same channel with RMSI.*

### 5.2. A Unified Inner Bound

We here present our unified inner bound for the three categories with RMSI, stated as Theorem 2.

**Theorem 2.** *A rate tuple $(R_0, R_1, R_2, R_3, R_4)$ is achievable for the channel with causal CSIT if it satisfies:*

$$R_0 + R_1 + R_3 < I(U_0, U_1; \tilde{Y}_1) - I(U_0, U_1; S), \tag{11}$$

$$R_0 + R_2 + R_4 < I(U_0, U_2; \tilde{Y}_2) - I(U_0, U_2; S), \tag{12}$$

$$R_0 + R_1 + R_2 + R_3 < I(U_0, U_1; \tilde{Y}_1) + I(U_2; \tilde{Y}_2 \mid U_0) - I(U_1; U_2 \mid U_0) - I(U_0, U_1, U_2; S), \tag{13}$$

$$R_0 + R_1 + R_2 + R_4 < I(U_1; \tilde{Y}_1 \mid U_0) + I(U_0, U_2; \tilde{Y}_2) - I(U_1; U_2 \mid U_0) - I(U_0, U_1, U_2; S), \tag{14}$$

$$2R_0 + R_1 + R_2 + R_3 + R_4 < I(U_0, U_1; \tilde{Y}_1) + I(U_0, U_2; \tilde{Y}_2) - I(U_1; U_2 \mid U_0)$$
$$- I(U_0, U_1, U_2; S) - I(U_0; S), \tag{15}$$

*for some pmf $p(u_0, u_1, u_2)$ and some function $x_0 = \gamma(u_0, u_1, u_2, s)$. It is achievable for the channel with non-causal CSIT if it satisfies (11)–(15) for some pmf $p(u_0, u_1, u_2 \mid s)$ and some function $x_0 = \gamma(u_0, u_1, u_2, s)$. $\tilde{Y}_i = (Y_i, S)$, $i = 1, 2$, if the channel state is available at receiver i and $\tilde{Y}_i = Y_i$ otherwise.*

**Remark 2.** *This inner bound is also used for the discrete memoryless broadcast channel with RMSI, without state, where we assume that the channel state is zero with probability one. Therefore, it is a unified inner bound for all three categories with RMSI.*

**Remark 3.** *Having a unified inner bound for the three categories without RMSI and applying the same pre-coding scheme to them are the two basic reasons why we can also have a unified inner bound for the three categories with RMSI.*

**Proof of Theorem 2.** We prove Theorem 2 by applying our pre-coding scheme to the transmission schemes of the causal and non-causal categories without RMSI.

With causal CSIT: For this category, we apply our pre-coding scheme to the transmission scheme of the discrete memoryless broadcast channel without RMSI, with causal CSIT, described in Section 4.1. Based on our method, Sub-codebook 0 consists of i.i.d. codewords:

$$u_0^n(m_0, m_3, m_4, m_{11}, m_{21}),$$

generated according to $\prod_{j=1}^n p_{U_0}(u_{0,j})$. Sub-codebook $i$, $i = 1, 2$, consists of codewords:

$$u_i^n(m_0, m_3, m_4, m_{11}, m_{21}, m_{i2}, l_i),$$

generated according to:

$$\prod_{j=1}^n p_{U_i|U_0}(u_{i,j} \mid u_{0,j}(m_0, m_3, m_4, m_{11}, m_{21})),$$

where $l_i \in \{1, \ldots, 2^{nR_i'}\}$.

Encoding and decoding are performed similarly to the case without RMSI. For the encoding, given $\{m_i\}_{i=0}^4$, we first find a pair $(l_1, l_2)$, such that:

$$(U_0^n(\cdot), U_1^n(\cdot, l_1), U_2^n(\cdot, l_2)) \in \mathcal{T}_{\epsilon'}^n.$$

If there does not exist one pair, we choose $(l_1, l_2) = (1, 1)$. We then construct the transmitted codeword as $x_{0,j} = \gamma(u_{0,j}(\cdot), u_{1,j}(\cdot), u_{2,j}(\cdot), s_j)$, $j = 1, 2, \ldots, n$.

For the decoding, Receiver 1 decodes $(\hat{m}_{0_1}, \hat{m}_{11}, \hat{m}_{12}, \hat{m}_3)$, if it is the unique tuple that satisfies:

$$(U_0^n(\cdot), U_1^n(\cdot, l_1), \tilde{Y}_1^n) \in \mathcal{T}_\epsilon^n \quad \text{for some } m_{21} \text{ and } l_1;$$

otherwise, an error is declared. Since this receiver knows $M_4$ as side information, it decodes $(u_0^n, u_1^n)$ over a set of $2^{n(R_0 + R_1 + R_3 + R_{21} + R_1')}$ candidates. Receiver 2 decodes $(\hat{m}_{0_2}, \hat{m}_{21}, \hat{m}_{22}, \hat{m}_4)$, if it is the unique tuple that satisfies:

$$(U_0^n(\cdot), U_2^n(\cdot, l_2), \tilde{Y}_2^n) \in \mathcal{T}_\epsilon^n \quad \text{for some } m_{11} \text{ and } l_2;$$

otherwise, an error is declared. Since this receiver knows $M_3$ as side information, it decodes $(u_0^n, u_2^n)$ over a set of $2^{n(R_0 + R_2 + R_4 + R_{11} + R_2')}$ candidates.

Based on error events, written similarly to the case without RMSI and using the packing lemma [4] (p. 45) and the mutual covering lemma [4] (p. 208), sufficient conditions for achievability are:

$$R_1' + R_2' > I(U_1; U_2 \mid U_0),$$
$$R_{12} + R_1' < I(U_1; \tilde{Y}_1 \mid U_0),$$
$$R_0 + R_1 + R_3 + R_{21} + R_1' < I(U_0, U_1; \tilde{Y}_1),$$
$$R_{22} + R_2' < I(U_2; \tilde{Y}_2 \mid U_0),$$
$$R_0 + R_2 + R_4 + R_{11} + R_2' < I(U_0, U_2; \tilde{Y}_2).$$

After performing the Fourier–Motzkin elimination, we obtain Conditions (11)–(15). Note that the terms $I(U_0; S)$, $I(U_0, U_1; S)$, $I(U_0, U_2; S)$ and $I(U_0, U_1, U_2; S)$ are zero for this category. The inner bound is computed over all pmfs $p(u_0, u_1, u_2)$ and all functions $x_0 = \gamma(u_0, u_1, u_2, s)$.

With non-causal CSIT: For this category, we apply our pre-coding scheme to the transmission scheme of the channel without RMSI, with non-causal CSIT, described in Section 4.2. The resulting changes to the codebook construction, encoding and decoding are similar to the ones for the channel

with causal CSIT. Based on error events written similarly to the case without RMSI and using the packing lemma and the multivariate covering lemma [4] (p. 218), sufficient conditions for achievability are:

$$R_0' > I(U_0; S),$$
$$R_0' + R_1' > I(U_0, U_1; S),$$
$$R_0' + R_2' > I(U_0, U_2; S),$$
$$R_0' + R_1' + R_2' > I(U_0, U_1, U_2; S) + I(U_1; U_2 \mid U_0),$$
$$R_{12} + R_1' < I(U_1; \tilde{Y}_1 \mid U_0),$$
$$R_0 + R_1 + R_3 + R_{21} + R_0' + R_1' < I(U_0, U_1; \tilde{Y}_1),$$
$$R_{22} + R_2' < I(U_2; \tilde{Y}_2 \mid U_0),$$
$$R_0 + R_2 + R_4 + R_{11} + R_0' + R_2' < I(U_0, U_2; \tilde{Y}_2).$$

After performing the Fourier–Motzkin elimination, we obtain Conditions (11)–(15). Note that $(U_0, U_1, U_2)$ is not independent of $S$ for this category. The inner bound is computed over all pmfs $p(u_0, u_1, u_2 \mid s)$ and all functions $x_0 = \gamma(u_0, u_1, u_2, s)$.    □

## 6. New Capacity Results

In this section, we present new capacity results for the two-receiver discrete memoryless broadcast channel with RMSI. These results are established using our inner bound in Theorem 2. We also show that our inner bound is tight for all of the cases whose capacity region was known prior to this work.

### 6.1. With RMSI, without State

In this subsection, we first derive a general outer bound for the discrete memoryless broadcast channel with RMSI, without state, stated as Theorem 3. This outer bound is developed based on the Nair-El Gamal outer bound for the discrete memoryless broadcast channel without RMSI, without state [25]. We then establish the capacity region for two new cases in this category: the deterministic channel, stated as Theorem 4, and the more capable channel, stated as Theorem 5.

**Theorem 3.** *If a rate tuple $(R_0, R_1, R_2, R_3, R_4)$ is achievable for the two-receiver discrete memoryless broadcast channel with RMSI, without state, then it must satisfy:*

$$R_0 + R_3 \leq I(U_0; Y_1), \tag{16}$$
$$R_0 + R_4 \leq I(U_0; Y_2), \tag{17}$$
$$R_0 + R_1 + R_3 \leq I(U_0, U_1; Y_1), \tag{18}$$
$$R_0 + R_2 + R_4 \leq I(U_0, U_2; Y_2), \tag{19}$$
$$R_0 + R_1 + R_2 + R_3 \leq I(U_0, U_1; Y_1) + I(U_2; Y_2 \mid U_0, U_1), \tag{20}$$
$$R_0 + R_1 + R_2 + R_4 \leq I(U_1; Y_1 \mid U_0, U_2) + I(U_0, U_2; Y_2), \tag{21}$$

*for some pmf $p(u_1)p(u_2)p(u_0 \mid u_1, u_2)$ and some function $x_0 = \gamma(u_0, u_1, u_2)$.*

**Proof of Theorem 3.** See Appendix B    □

**Theorem 4.** *The capacity region of the two-receiver deterministic broadcast channel with RMSI, without state, is the closure of the set of all rate tuples* $(R_0, R_1, R_2, R_3, R_4)$, *each satisfying:*

$$R_0 + R_1 + R_3 < H(Y_1), \tag{22}$$
$$R_0 + R_2 + R_4 < H(Y_2), \tag{23}$$
$$R_0 + R_1 + R_2 + R_3 < H(Y_1) + H(Y_2 \mid U_0, Y_1), \tag{24}$$
$$R_0 + R_1 + R_2 + R_4 < H(Y_2) + H(Y_1 \mid U_0, Y_2), \tag{25}$$
$$2R_0 + R_1 + R_2 + R_3 + R_4 < I(U_0; Y_1) + H(Y_2) + H(Y_1 \mid U_0, Y_2), \tag{26}$$

*for some pmf* $p(u_0, x_0)$.

**Proof of Theorem 4.** (Achievability) Achievability is proven by setting $(U_0, U_1, U_2, S) = (U_0, Y_1, Y_2, 0)$ in (11)–(15).

(Converse) We start with the outer bound in Theorem 3. By removing Condition (17), adding Conditions (16) and (21) and enlarging the domain of this outer bound, we obtain the following looser outer bound: if a rate tuple is achievable, then it must satisfy:

$$R_0 + R_1 + R_3 \leq I(U_0, U_1; Y_1), \tag{27}$$
$$R_0 + R_2 + R_4 \leq I(U_0, U_2; Y_2), \tag{28}$$
$$R_0 + R_1 + R_2 + R_3 \leq I(U_0, U_1; Y_1) + I(U_2; Y_2 \mid U_0, U_1), \tag{29}$$
$$R_0 + R_1 + R_2 + R_4 \leq I(U_1; Y_1 \mid U_0, U_2) + I(U_0, U_2; Y_2), \tag{30}$$
$$2R_0 + R_1 + R_2 + R_3 + R_4 \leq I(U_0; Y_1) + I(U_1; Y_1 \mid U_0, U_2) + I(U_0, U_2; Y_2), \tag{31}$$

*for some pmf* $p(u_0, u_1, u_2)$ *and some function* $x_0 = \gamma(u_0, u_1, u_2)$.

By relaxing Conditions (27)–(31), we can write them in the form of Conditions (22)–(26). As an example, we show this for (31).

$$
\begin{aligned}
2R_0 + R_1 + R_2 + R_3 + R_4 &\leq I(U_0; Y_1) + I(U_1; Y_1 \mid U_0, U_2) + I(U_0, U_2; Y_2) \\
&\overset{(a)}{=} I(U_0; Y_1) + H(Y_1 \mid U_0, U_2) + H(Y_2) - H(Y_2 \mid U_0, U_2) \\
&\leq I(U_0; Y_1) + H(Y_1, Y_2 \mid U_0, U_2) + H(Y_2) - H(Y_2 \mid U_0, U_2) \\
&= I(U_0; Y_1) + H(Y_1 \mid U_0, U_2, Y_2) + H(Y_2) \\
&\leq I(U_0; Y_1) + H(Y_1 \mid U_0, Y_2) + H(Y_2),
\end{aligned}
$$

where $(a)$ is due to $H(Y_i \mid U_0, U_1, U_2) = 0$, $i = 1, 2$, for the deterministic channel. □

**Theorem 5.** *The capacity region of the two-receiver more capable broadcast channel with RMSI, without state, is the closure of the set of all rate tuples* $(R_0, R_1, R_2, R_3, R_4)$, *each satisfying:*

$$R_0 + R_2 + R_4 < I(U_0; Y_2),$$
$$R_0 + R_1 + R_2 + R_4 < I(U_0; Y_2) + I(X_0; Y_1 \mid U_0),$$
$$R_0 + R_1 + R_2 + R_3 < I(X_0; Y_1),$$

*for some pmf* $p(u_0, x_0)$.

**Proof of Theorem 5.** (Achievability) Achievability is proven by setting $(U_0, U_1, U_2, S) = (U_0, X_0, 0, 0)$ in (11)–(15). Note that $U_2 = 0$ implies that $M_{21} = M_2$ and $R'_1 = 0$.

(Converse) According to Theorem 3, if a rate tuple is achievable, then it must satisfy:

$$R_0 + R_2 + R_4 \leq I(U_0, U_2; Y_2), \tag{32}$$
$$R_0 + R_1 + R_2 + R_3 \leq I(U_0, U_1; Y_1) + I(U_2; Y_2 \mid U_0, U_1), \tag{33}$$
$$R_0 + R_1 + R_2 + R_4 \leq I(U_1; Y_1 \mid U_0, U_2) + I(U_0, U_2; Y_2), \tag{34}$$

for some pmf $p(u_0, u_1, u_2)$ and some function $x_0 = \gamma(u_0, u_1, u_2)$. Since, for the channel without state, we have:

$$I(U_2; Y_2 \mid U_0, U_1) = I(X_0; Y_2 \mid U_0, U_1),$$
$$I(U_1; Y_1 \mid U_0, U_2) = I(X_0; Y_1 \mid U_0, U_2),$$

and for the more capable channel, we have [4] (p. 123):

$$I(X_0; Y_2 \mid U_0, U_1) \leq I(X_0; Y_1 \mid U_0, U_1),$$

we can write (33) and (34) as follows.

$$
\begin{aligned}
R_0 + R_1 + R_2 + R_3 &\leq I(U_0, U_1; Y_1) + I(U_2; Y_2 \mid U_0, U_1) \\
&= I(U_0, U_1; Y_1) + I(X_0; Y_2 \mid U_0, U_1) \\
&\leq I(U_0, U_1; Y_1) + I(X_0; Y_1 \mid U_0, U_1) \\
&= I(X_0; Y_1). 
\end{aligned}
\tag{35}
$$
$$
\begin{aligned}
R_0 + R_1 + R_2 + R_4 &\leq I(U_1; Y_1 \mid U_0, U_2) + I(U_0, U_2; Y_2) \\
&= I(X_0; Y_1 \mid U_0, U_2) + I(U_0, U_2; Y_2). 
\end{aligned}
\tag{36}
$$

From (32), (35), (36) and considering $U_0' = (U_0, U_2)$, if a rate tuple is achievable, then it must satisfy:

$$R_0 + R_2 + R_4 \leq I(U_0'; Y_2),$$
$$R_0 + R_1 + R_2 + R_3 \leq I(X_0; Y_1),$$
$$R_0 + R_1 + R_2 + R_4 \leq I(X_0; Y_1 \mid U_0') + I(U_0'; Y_2),$$

for some pmf $p(u_0', x_0)$. $\square$

*6.2. With RMSI, with Causal CSIT*

In this subsection, we establish the capacity region of the two-receiver (stochastically) degraded broadcast channel with RMSI where the channel state is available causally: (i) at only the transmitter; (ii) at the transmitter and the non-degraded receiver; or (iii) at the transmitter and both receivers.

**Theorem 6.** *The capacity region of the two-receiver degraded broadcast channel with RMSI where the channel state is available causally at only the transmitter is the closure of the set of all rate tuples $(R_0, R_1, R_2, R_3, R_4)$, each satisfying:*

$$R_0 + R_2 + R_4 < I(U_0; Y_2),$$
$$R_0 + R_1 + R_2 + R_3 < I(U_0, U_1; Y_1),$$
$$R_0 + R_1 + R_2 + R_4 < I(U_0; Y_2) + I(U_1; Y_1 \mid U_0),$$

*for some pmf $p(u_0, u_1)$ and some function $x_0 = \gamma(u_0, u_1, s)$.*

**Proof of Theorem 6.** (Achievability) Achievability is proven by setting $U_2 = 0$ in (11)–(15). Note that, for a causal case, $U_2 = 0$ implies that $M_{21} = M_2$ and $R'_1 = 0$. (Converse) See Appendix C. □

**Theorem 7.** *The capacity region of the two-receiver degraded broadcast channel with RMSI where the channel state is available causally either at the transmitter and Receiver 1 or at the transmitter and both receivers is the closure of the set of all rate tuples* $(R_0, R_1, R_2, R_3, R_4)$, *each satisfying:*

$$R_0 + R_2 + R_4 < I(U_0; \tilde{Y}_2),$$
$$R_0 + R_1 + R_2 + R_3 < I(X_0; Y_1 \mid S),$$
$$R_0 + R_1 + R_2 + R_4 < I(U_0; \tilde{Y}_2) + I(X_0; Y_1 \mid U_0, S),$$

*for some pmf* $p(u_0, u_1)$ *and some function* $x_0 = \gamma(u_0, u_1, s)$. $\tilde{Y}_2 = (Y_2, S)$ *if the channel state is available at Receiver 2 and* $\tilde{Y}_2 = Y_2$ *otherwise.*

**Proof of Theorem 7.** (Achievability) Achievability is proven by setting $U_2 = 0$ in (11)–(15). Note that, for a causal case, $U_2 = 0$ implies that $M_{21} = M_2$, $R'_1 = 0$ and:

$$I(U_0, U_1; Y_1, S) = I(U_0, U_1; Y_1 \mid S) = I(X_0; Y_1 \mid S),$$
$$I(U_1; Y_1, S \mid U_0) = I(U_1; Y_1 \mid U_0, S) = I(X_0; Y_1 \mid U_0, S).$$

(Converse) See Appendix D. □

*6.3. With RMSI, with Non-Causal CSIT*

In this subsection, we establish the capacity region of the two-receiver (stochastically) degraded broadcast channel with RMSI where the channel state is available non-causally: (i) at the transmitter and the non-degraded receiver; or (ii) at the transmitter and both receivers.

**Theorem 8.** *The capacity region of the two-receiver degraded broadcast channel with RMSI where the channel state is available non-causally either at the transmitter and Receiver 1 or at the transmitter and both receivers is the closure of the set of all rate tuples* $(R_0, R_1, R_2, R_3, R_4)$, *each satisfying:*

$$R_0 + R_2 + R_4 < I(U_0; \tilde{Y}_2) - I(U_0; S),$$
$$R_0 + R_1 + R_2 + R_3 < I(X_0; Y_1 \mid S),$$
$$R_0 + R_1 + R_2 + R_4 < I(U_0; \tilde{Y}_2) - I(U_0; S) + I(X_0; Y_1 \mid U_0, S),$$

*for some pmf* $p(u_0, u_1 \mid s)$ *and some function* $x_0 = \gamma(u_0, u_1, s)$. $\tilde{Y}_2 = (Y_2, S)$ *if the channel state is available at Receiver 2 and* $\tilde{Y}_2 = Y_2$ *otherwise.*

**Proof of Theorem 8.** (Achievability) Achievability is proven by setting $U_2 = 0$ in (11)–(15). Note that, for a non-causal case, $U_2 = 0$ implies that $M_{21} = M_2$, and:

$$I(U_0, U_1; Y_1, S) - I(U_0, U_1; S) = I(U_0, U_1; Y_1 \mid S) = I(X_0; Y_1 \mid S).$$

(Converse) See Appendix D. □

*6.4. Discussion on Prior Known Results*

We here show that our inner bound in Theorem 2 is also tight for all of the special cases of the two-receiver discrete memoryless channel with RMSI whose capacity region was known prior to this work.

With RMSI, without state: The capacity region of the discrete memoryless channel with complementary RMSI is achieved by multiplexing all of the requested messages into only one

codebook [15,16]. This scheme is a special case of our scheme obtained by setting $(U_0, U_1, U_2, S) = (X_0, 0, 0, 0)$. Note that $M_1$ and $M_2$ are equal to zero in this message setup.

The capacity region of the discrete memoryless channel with degraded message sets (due to RMSI) is achieved by superposition coding [17]. This scheme is a special case of our scheme obtained by setting $(U_0, U_1, U_2, S) = (U_0, X_0, 0, 0)$ or $(U_0, U_1, U_2, S) = (U_0, 0, X_0, 0)$ depending on whether Receiver 1 or Receiver 2 needs to decode the whole set of the source messages, respectively. Note that either $M_1$ or $M_2$ is equal to zero in this message setup.

The less noisy broadcast channel with RMSI is a special case of the more capable broadcast channel with RMSI [4]. Thus, our scheme can also achieve the capacity region of the less noisy channel.

With RMSI, with causal CSIT: By setting $(U_1, U_2) = (0, 0)$, our scheme for the causal category reduces to the one used by Khormuji et al. [19] to establish the capacity region of the two-receiver discrete memoryless broadcast channel with complementary RMSI where the channel state is available causally at only the transmitter.

With RMSI, with non-causal CSIT: By setting $(U_1, U_2) = (0, 0)$, our scheme for the non-causal category reduces to the one used by Oechtering and Skoglund [20] to establish the capacity region of the two-receiver discrete memoryless broadcast channel with complementary RMSI where the channel state is available non-causally at the transmitter and one of the receivers.

## 7. Conclusions

We proposed a pre-coding scheme designed to construct transmission schemes of two-receiver broadcast channels with receiver message side information (RMSI) using the best transmission schemes for the channels without RMSI. This provides a general approach for utilizing RMSI in different two-receiver broadcast channels. Employing our pre-coding scheme, we derived a unified inner bound for three categories of the discrete memoryless broadcast channel with RMSI: (i) channel without state; (ii) channel with causal channel state information at the transmitter (CSIT); and (iii) channel with non-causal CSIT. We showed that our inner bound establishes the capacity region of some new cases in each of the three categories. We also showed that our inner bound is tight for all of the cases whose capacity region was known prior to this work. These results validated our approach for utilizing RMSI in two-receiver broadcast channels.

**Author Contributions:** Behzad Asadi developed this work in discussion with Lawrence Ong and Sarah J. Johnson. Behzad Asadi wrote the article with input from Lawrence Ong and Sarah J. Johnson. All authors have read and approved the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

In this section, we show that the scheme for the causal category, described in Section 4.1, cannot be considered as a special case of the scheme for the non-causal category, described in Section 4.2. However, the rate regions achievable by both schemes have similar expression. This results in a unified inner bound for both causal and non-causal cases from which we can show that the inner bound for a non-causal case is at least as large as the the inner bound for the corresponding causal case. We will explain the reason behind this observation despite them having different schemes.

We consider a scheme as a special case of another scheme when the latter reduces to the former by considering some special cases of its parameters, e.g., superposition coding is a special case of the scheme achieving Marton's inner bound with a common message [4] (p. 212).

Consider the encoding rule of the scheme for the non-causal category where the encoder finds a triple $(U_0^n, U_1^n, U_2^n)$, such that:

$$(U_0^n, U_1^n, U_2^n, S^n) \in \mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2, S).$$

The encoder for the causal category cannot check this rule since this encoder only knows $S^n$ at the end of the transmission. Hence, the scheme for the causal category is not a special case of the scheme for the non-causal category.

By choosing $p(u_0, u_1, u_2 \mid s) = p(u_0, u_1, u_2)$ for all $s$ and setting $R'_0 = 0$, the scheme for the non-causal category has the same codebook construction and decoding approach as the scheme for the causal category. The only difference is that for the channel state realization $s^n$, the encoder for the non-causal category finds a triple $(U_0^n, U_1^n, U_2^n)$, such that:

$$(U_0^n, U_1^n, U_2^n) \in \mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2 \mid s^n),$$

where:

$$\mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2 \mid s^n) = \left\{ (u_0^n, u_1^n, u_2^n) \mid (u_0^n, u_1^n, u_2^n, s^n) \in \mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2, S) \right\},$$

and the encoder for the causal category finds a triple $(U_0^n, U_1^n, U_2^n)$, such that:

$$(U_0^n, U_1^n, U_2^n) \in \mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2).$$

Therefore, the transmitted codewords may be different. However, according to the properties of joint typicality [4] (p. 30), we have:

- $\mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2 \mid s^n) \subseteq \mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2).$
- For sufficiently large $n$,

$$(1 - \epsilon')2^{n(H(U_0, U_1, U_2) \, - \, \epsilon' H(U_0, U_1, U_2))} \leq \left| \mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2) \right| \leq 2^{n(H(U_0, U_1, U_2) \, + \, \epsilon' H(U_0, U_1, U_2))}.$$

- If $s^n \in \mathcal{T}_{\epsilon''}^n$, $\epsilon'' < \epsilon'$, then, for sufficiently large $n$,

$$(1 - \epsilon')2^{n(H(U_0, U_1, U_2) \, - \, \epsilon' H(U_0, U_1, U_2))} \leq \left| \mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2 \mid s^n) \right| \leq 2^{n(H(U_0, U_1, U_2) \, + \, \epsilon' H(U_0, U_1, U_2))}.$$

  Note that in this item, we have also used the fact that $p(u_0, u_1, u_2 \mid s) = p(u_0, u_1, u_2)$, which results in $H(U_0, U_1, U_2 \mid S) = H(U_0, U_1, U_2)$.
- $P\left(S^n \in \mathcal{T}_{\epsilon''}^n(S)\right) \to 1$, as $n$ tends to infinity.

  Consequently,

$$P\left(\mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2 \mid s^n) = \mathcal{T}_{\epsilon'}^n(U_0, U_1, U_2)\right) \to 1,$$

as $n$ tends to infinity. Hence, by choosing $p(u_0, u_1, u_2 \mid s) = p(u_0, u_1, u_2)$, the scheme for the non-causal category asymptomatically almost surely has the same encoding as the scheme for the causal category. This leads the scheme for the non-causal category to achieve the same rate region as the scheme for the causal category.

## Appendix B

In this section, we present the proof of Theorem 3, which is based on the proof of the Nair-El Gamal outer bound for the channel without RMSI [4] (p. 217).

**Proof.** By Fano's inequality [4] (p. 19), we have:

$$H\left(M_0, M_1, M_3 \mid Y_1^n, M_4\right) \leq n\epsilon_{1,n}, \tag{A1}$$

$$H\left(M_0, M_2, M_4 \mid Y_2^n, M_3\right) \leq n\epsilon_{2,n}, \tag{A2}$$

where $\epsilon_{i,n} \to 0$ as $n \to \infty$ for $i = 1, 2$. For the sake of simplicity, we use $\epsilon_n$ instead of $\epsilon_{i,n}$ for the remainder.

Using (A1) and (A2), if a rate tuple $(R_0, R_1, R_2, R_3, R_4)$ is achievable, then it must satisfy:

$$n(R_0 + R_3) \leq I(M_0, M_3; Y_1^n \mid M_4) + n\epsilon_n, \tag{A3}$$

$$n(R_0 + R_4) \leq I(M_0, M_4; Y_2^n \mid M_3) + n\epsilon_n, \tag{A4}$$

$$n(R_0 + R_1 + R_3) \leq I(M_0, M_1, M_3; Y_1^n \mid M_4) + n\epsilon_n, \tag{A5}$$

$$n(R_0 + R_2 + R_4) \leq I(M_0, M_2, M_4; Y_2^n \mid M_3) + n\epsilon_n, \tag{A6}$$

$$n(R_0 + R_1 + R_2 + R_3) \leq I(M_0, M_1, M_3; Y_1^n \mid M_4) + I(M_2; Y_2^n \mid M_0, M_1, M_3, M_4) + 2n\epsilon_n, \tag{A7}$$

$$n(R_0 + R_1 + R_2 + R_4) \leq I(M_1; Y_1^n \mid M_0, M_2, M_3, M_4) + I(M_0, M_2, M_4; Y_2^n \mid M_3) + 2n\epsilon_n. \tag{A8}$$

Inequalities (A3)–(A8) yield Conditions (16)–(21) by using the auxiliary random variables defined as:

$$U_{0,i} = \left( M_0, M_3, M_4, Y_1^{i-1}, Y_{2,i+1}^n \right),$$
$$U_{1,i} = M_1,$$
$$U_{2,i} = M_2,$$

where $Y_1^{i-1} = (Y_{1,1}, Y_{1,2}, \ldots, Y_{1,i-1})$ and $Y_{2,i+1}^n = (Y_{2,i+1}, Y_{2,i+2}, \ldots, Y_{2,n})$.

We here only show how Inequalities (A3) and (A7) yield Conditions (16) and (20), respectively. We just need to follow similar steps for the rest.

In (A3), we expand the mutual information term as follows

$$
\begin{aligned}
I(M_0, M_3; Y_1^n \mid M_4) &= \sum_{i=1}^n I\left( M_0, M_3; Y_{1,i} \mid M_4, Y_1^{i-1} \right) \\
&\leq \sum_{i=1}^n I\left( M_0, M_3, M_4, Y_1^{i-1}; Y_{1,i} \right) \\
&\leq \sum_{i=1}^n I\left( U_{0,i}; Y_{1,i} \right).
\end{aligned}
$$

Then, since $\epsilon_n \to 0$ as $n \to \infty$, by using the standard time-sharing argument [4] (p. 114), we have:

$$R_0 + R_3 \leq I(U_0; Y_1).$$

In (A7), we expand the mutual information terms as follows.

$$
\begin{aligned}
I(M_0, M_1, M_3; Y_1^n \mid M_4) &= \sum_{i=1}^n I\left( M_0, M_1, M_3; Y_{1,i} \mid M_4, Y_1^{i-1} \right) \\
&\leq \sum_{i=1}^n I\left( M_0, M_1, M_3, M_4, Y_1^{i-1}; Y_{1,i} \right) \\
&= \sum_{i=1}^n I\left( M_0, M_1, M_3, M_4, Y_1^{i-1}, Y_{2,i+1}^n; Y_{1,i} \right) \\
&\quad - \sum_{i=1}^n I\left( Y_{2,i+1}^n; Y_{1,i} \mid M_0, M_1, M_3, M_4, Y_1^{i-1} \right), \tag{A9}
\end{aligned}
$$

$$I\left(M_2; Y_2^n \mid M_0, M_1, M_3, M_4\right) = \sum_{i=1}^{n} I\left(M_2; Y_{2,i} \mid M_0, M_1, M_3, M_4, Y_{2,i+1}^n\right)$$

$$\leq \sum_{i=1}^{n} I\left(M_2, Y_1^{i-1}; Y_{2,i} \mid M_0, M_1, M_3, M_4, Y_{2,i+1}^n\right)$$

$$= \sum_{i=1}^{n} I\left(Y_1^{i-1}; Y_{2,i} \mid M_0, M_1, M_3, M_4, Y_{2,i+1}^n\right)$$

$$+ \sum_{i=1}^{n} I\left(M_2; Y_{2,i} \mid M_0, M_1, M_3, M_4, Y_1^{i-1}, Y_{2,i+1}^n\right). \tag{A10}$$

Then, since $\epsilon_n \to 0$ as $n \to \infty$, from (A7), (A9), (A10), the Csiszár sum identity [4] (p. 25) and the time-sharing argument, we have:

$$R_0 + R_1 + R_2 + R_3 \leq I\left(U_0, U_1; Y_1\right) + I\left(U_2; Y_2 \mid U_0, U_1\right). \quad \square$$

**Appendix C**

In this section, we present the converse proof of Theorem 6. In the converse, we assume that the broadcast channel is physically degraded as the capacity region of the stochastically-degraded broadcast channel is equal to its equivalent physically-degraded broadcast channel.

**Proof.** (Converse) By Fano's inequality, we have (A1) and (A2). From (A2) and the physical degradedness of the channel, we have:

$$H\left(M_0, M_2, M_4 \mid Y_1^n, M_3\right) \leq H\left(M_0, M_2, M_4 \mid Y_2^n, M_3\right) \leq n\epsilon_n, \tag{A11}$$

and from (A1) and (A11), we have:

$$H\left(M_0, M_2, M_3 \mid Y_1^n, M_4\right) \leq 2n\epsilon_n. \tag{A12}$$

Using (A1), (A2) and (A12), we obtain the following necessary conditions for achievability:

$$nR_1 \leq I(M_1; Y_1^n \mid M_0, M_2, M_3, M_4) + n\epsilon_n, \tag{A13}$$

$$n(R_0 + R_2 + R_3) \leq I(M_0, M_2, M_3; Y_1^n \mid M_4) + 2n\epsilon_n, \tag{A14}$$

$$n(R_0 + R_2 + R_4) \leq I(M_0, M_2, M_4; Y_2^n \mid M_3) + n\epsilon_n. \tag{A15}$$

We now define the auxiliary random variables $U_{0,i}$ and $U_{1,i}$ as:

$$U_{0,i} = (M_0, M_2, M_3, M_4, Y_1^{i-1}),$$
$$U_{1,i} = (M_1, S^{i-1}),$$

and expand the mutual information terms in (A13)–(A15) respectively as follows.

$$I\left(M_1; Y_1^n \mid M_0, M_2, M_3, M_4\right) = \sum_{i=1}^{n} I\left(M_1; Y_{1,i} \mid M_0, M_2, M_3, M_4, Y_1^{i-1}\right)$$

$$\leq \sum_{i=1}^{n} I\left(M_1, S^{i-1}; Y_{1,i} \mid M_0, M_2, M_3, M_4, Y_1^{i-1}\right)$$

$$= \sum_{i=1}^{n} I\left(U_{1,i}; Y_{1,i} \mid U_{0,i}\right), \tag{A16}$$

$$I\left(M_0, M_2, M_3; Y_1^n \mid M_4\right) = \sum_{i=1}^{n} I\left(M_0, M_2, M_3; Y_{1,i} \mid M_4, Y_1^{i-1}\right)$$

$$\leq \sum_{i=1}^{n} I\left(M_0, M_2, M_3, M_4, Y_1^{i-1}; Y_{1,i}\right)$$

$$= \sum_{i=1}^{n} I\left(U_{0,i}; Y_{1,i}\right), \tag{A17}$$

$$I\left(M_0, M_2, M_4; Y_2^n \mid M_3\right) = \sum_{i=1}^{n} I\left(M_0, M_2, M_4; Y_{2,i} \mid M_3, Y_2^{i-1}\right)$$

$$\leq \sum_{i=1}^{n} I\left(M_0, M_2, M_3, M_4, Y_2^{i-1}; Y_{2,i}\right)$$

$$\leq \sum_{i=1}^{n} I\left(M_0, M_2, M_3, M_4, Y_2^{i-1}, Y_1^{i-1}; Y_{2,i}\right)$$

$$\overset{(a)}{=} \sum_{i=1}^{n} I\left(M_0, M_2, M_3, M_4, Y_1^{i-1}; Y_{2,i}\right)$$

$$= \sum_{i=1}^{n} I\left(U_{0,i}; Y_{2,i}\right), \tag{A18}$$

where $(a)$ follows from the physical degradedness of the channel.

Finally, since $\epsilon_n \to 0$ as $n \to \infty$, substituting (A16)–(A18) into (A13)–(A15) and using the time-sharing argument complete the converse proof. Note that $(U_{0,i}, U_{1,i})$ is independent of $S_i$, and $X_{0,i}$ is a function of $(U_{0,i}, U_{1,i}, S_i)$. $\quad\square$

## Appendix D

In this section, we present the converse proof of Theorems 7 and 8. We here also assume that the broadcast channel is physically degraded.

**Proof.** (Converse) By Fano's inequality, we have:

$$H(M_0, M_1, M_3 \mid Y_1^n, S^n, M_4) \leq n\epsilon_n, \tag{A19}$$

$$H(M_0, M_2, M_4 \mid \tilde{Y}_2^n, M_3) \leq n\epsilon_n, \tag{A20}$$

where $\epsilon_n \to 0$ as $n \to \infty$.

From (A20) and the physical degradedness of the channel, we have:

$$H\left(M_0, M_2, M_4 \mid Y_1^n, S^n, M_3\right) \leq H\left(M_0, M_2, M_4 \mid Y_2^n, S^n, M_3\right) \leq n\epsilon_n, \tag{A21}$$

and from (A19) and (A21), we have:

$$H\left(M_0, M_1, M_2, M_3 \mid Y_1^n, S^n, M_4\right) \leq 2n\epsilon_n. \tag{A22}$$

Using (A19), (A20) and (A22), if a rate tuple $(R_0, R_1, R_2, R_3, R_4)$ is achievable, then it must satisfy:

$$nR_1 \leq I(M_1; Y_1^n \mid M_0, M_2, M_3, M_4, S^n) + n\epsilon_n, \tag{A23}$$

$$n(R_0 + R_1 + R_2 + R_3) \leq I(M_0, M_1, M_2, M_3; Y_1^n \mid M_4, S^n) + 2n\epsilon_n, \tag{A24}$$

$$n(R_0 + R_2 + R_4) \leq I(M_0, M_2, M_4; \tilde{Y}_2^n \mid M_3) + n\epsilon_n. \tag{A25}$$

We now define the auxiliary random variables $U_{0,i}$ and $U_{1,i}$ as:

$$U_{0,i} = (M_0, M_2, M_3, M_4, S^{i-1}, S_{i+1}^n, Y_2^{i-1}),$$

$$U_{1,i} = (M_1, Y_1^{i-1}),$$

and expand the mutual information terms in (A23)–(A25) respectively as follows.

$$
\begin{aligned}
I\left(M_1; Y_1^n \mid M_0, M_2, M_3, M_4, S^n\right) &= \sum_{i=1}^{n} I\left(M_1; Y_{1,i} \mid M_0, M_2, M_3, M_4, S^n, Y_1^{i-1}\right) \\
&\overset{(a)}{=} \sum_{i=1}^{n} I\left(M_1; Y_{1,i} \mid M_0, M_2, M_3, M_4, S^n, Y_1^{i-1}, Y_2^{i-1}\right) \\
&\le \sum_{i=1}^{n} I\left(M_1, Y_1^{i-1}; Y_{1,i} \mid M_0, M_2, M_3, M_4, S^n, Y_2^{i-1}\right) \\
&= \sum_{i=1}^{n} I\left(U_{1,i}; Y_{1,i} \mid U_{0,i}, S_i\right) \\
&= \sum_{i=1}^{n} I\left(X_{0,i}; Y_{1,i} \mid U_{0,i}, S_i\right),
\end{aligned}
\tag{A26}
$$

$$
\begin{aligned}
I\left(M_0, M_1, M_2, M_3; Y_1^n \mid M_4, S^n\right) &= \sum_{i=1}^{n} I\left(M_0, M_1, M_2, M_3; Y_{1,i} \mid M_4, S^n, Y_1^{i-1}\right) \\
&\overset{(b)}{=} \sum_{i=1}^{n} I\left(M_0, M_1, M_2, M_3; Y_{1,i} \mid M_4, S^n, Y_1^{i-1}, Y_2^{i-1}\right) \\
&\le \sum_{i=1}^{n} I\left(U_{0,i}, U_{1,i}; Y_{1,i} \mid S_i\right) \\
&= \sum_{i=1}^{n} I\left(X_{0,i}; Y_{1,i} \mid S_i\right),
\end{aligned}
\tag{A27}
$$

$$
\begin{aligned}
&I(M_0, M_2, M_4; \tilde{Y}_2^n \mid M_3) \\
&= \sum_{i=1}^{n} I(M_0, M_2, M_4; \tilde{Y}_{2,i} \mid M_3, \tilde{Y}_2^{i-1}) \\
&\le \sum_{i=1}^{n} I(M_0, M_2, M_3, M_4, S^{i-1}, \tilde{Y}_2^{i-1}; \tilde{Y}_{2,i}) \\
&= \sum_{i=1}^{n} I(M_0, M_2, M_3, M_4, S^{i-1}, S_{i+1}^n, \tilde{Y}_2^{i-1}; \tilde{Y}_{2,i}) - I(S_{i+1}^n; \tilde{Y}_{2,i} \mid M_0, M_2, M_3, M_4, S^{i-1}, \tilde{Y}_2^{i-1}) \\
&\overset{(c)}{=} \sum_{i=1}^{n} I(M_0, M_2, M_3, M_4, S^{i-1}, S_{i+1}^n, \tilde{Y}_2^{i-1}; \tilde{Y}_{2,i}) - I(\tilde{Y}_2^{i-1}; S_i \mid M_0, M_2, M_3, M_4, S^{i-1}, S_{i+1}^n) \\
&\overset{(d)}{=} \sum_{i=1}^{n} I(M_0, M_2, M_3, M_4, S^{i-1}, S_{i+1}^n, \tilde{Y}_2^{i-1}; \tilde{Y}_{2,i}) - I(M_0, M_2, M_3, M_4, S^{i-1}, S_{i+1}^n, \tilde{Y}_2^{i-1}; S_i) \\
&= \sum_{i=1}^{n} I(U_{0,i}; \tilde{Y}_{2,i}) - I(U_{0,i}; S_i),
\end{aligned}
\tag{A28}
$$

where $(a)$ and $(b)$ follow from the physical degradedness of the channel, $(c)$ from the Csiszár sum identity and $(d)$ from the independence of $(M_0, M_2, M_3, M_4, S^{i-1}, S_{i+1}^n)$ and $S_i$. Note that, for causal cases, $(U_{0,i}, U_{1,i})$ is independent of $S_i$, but for non-causal cases, $(U_{0,i}, U_{1,i})$ and $S_i$ are dependent. For both causal and non-causal cases, $X_{0,i}$ is a function of $(U_{0,i}, U_{1,i}, S_i)$.

Finally, since $\epsilon_n \to 0$ as $n \to \infty$, substituting (A26)–(A28) into (A23)–(A25) and using the time sharing argument complete the converse proof. □

## References

1. Cover, T.M. Broadcast channels. *IEEE Trans. Inf. Theory* **1972**, *18*, 2–14.
2. Ong, L.; Kellett, C.M.; Johnson, S.J. On the equal-rate capacity of the AWGN multiway relay channel. *IEEE Trans. Inf. Theory* **2012**, *58*, 5761–5769.
3. Gallager, R.G. Capacity and coding for degraded broadcast channels. *Probl. Inf. Transm.* **1974**, *10*, 3–14.
4. El Gamal, A.; Kim, Y.H. *Network Information Theory*; Cambridge University Press: Cambridge, UK, 2011.

5.   El Gamal, A. The capacity of a class of broadcast channels. *IEEE Trans. Inf. Theory* **1979**, *25*, 166–169.

6.   Pinsker, M.S. Capacity of noiseless broadcast channels. *Probl. Inf. Transm.* **1978**, *14*, 28–34.

7.   Han, T.S. The capacity region for the deterministic broadcast channel with a common message. *IEEE Trans. Inf. Theory* **1981**, *27*, 122–125.

8.   Marton, K. A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inf. Theory* **1979**, *25*, 306–311.

9.   Körner, J.; Marton, K. General broadcast channels with degraded message sets. *IEEE Trans. Inf. Theory* **1977**, *23*, 60–64.

10.  Steinberg, Y. Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information. *IEEE Trans. Inf. Theory* **2005**, *51*, 2867–2877.

11.  Shannon, C. Channels with side information at the transmitter. *IBM J. Res. Dev.* **1958**, *2*, 289–293.

12.  Lapidoth, A.; Wang, L. The state-dependent semideterministic broadcast channel. *IEEE Trans. Inf. Theory* **2013**, *59*, 2242–2251.

13.  Khosravi-Farsani, R.; Marvasti, F. Capacity bounds for multiuser channels with non-causal channel state information at the transmitters. In Proceedings of the IEEE Information Theory Workshop (ITW), Paraty, Brazil, 16–20 October 2011.

14.  Gel'fand, S.I.; Pinsker, M.S. Coding for channel with random parameters. *Probl. Control Inf. Theory* **1980**, *9*, 19–31.

15.  Oechtering, T.J.; Schnurr, C.; Bjelakovic, I.; Boche, H. Broadcast capacity region of two-phase bidirectional relaying. *IEEE Trans. Inf. Theory* **2008**, *54*, 454–458.

16.  Tuncel, E. Slepian–Wolf coding over broadcast channels. *IEEE Trans. Inf. Theory* **2006**, *52*, 1469–1482.

17.  Kramer, G.; Shamai, S. Capacity for classes of broadcast channels with receiver side information. In Proceedings of the IEEE Information Theory Workshop (ITW), Lake Tahoe, CA, USA, 2–6 September 2007.

18.  Oechtering, T.J.; Wigger, M.; Timo, R. Broadcast capacity regions with three receivers and message cognition. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Cambridge, MA, USA, 1–6 July 2012.

19.  Khormuji, M.N.; Oechtering, T.J.; Skoglund, M. Capacity region of the bidirectional broadcast channel with causal channel state information. In Proceedings of the Tenth International Symposium on Wireless Communication Systems (ISWCS), Ilmenau, Germany, 27–30 August 2013.

20.  Oechtering, T.J.; Skoglund, M. Bidirectional broadcast channel with random states noncausally known at the encoder. *IEEE Trans. Inf. Theory* **2013**, *59*, 64–75.

21.  Kramer, G. Information networks with in-block memory. *IEEE Trans. Inf. Theory* **2014**, *60*, 2105–2120.

22.  Jafar, S. Capacity with causal and noncausal side information: A unified view. *IEEE Trans. Inf. Theory* **2006**, *52*, 5468–5474.

23.  Asadi, B.; Ong, L.; Johnson, S.J. A unified scheme for two-receiver broadcast channels with receiver message side information. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015.

24.  Bracher, A.; Wigger, M. Feedback and partial message side-information on the semideterministic broadcast channel. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015.

25.  Nair, C.; El Gamal, A. An outer bound to the capacity region of the broadcast channel. *IEEE Trans. Inf. Theory* **2007**, *53*, 350–355.