On the Capacity of the Two-User Erasure Broadcast Channel with Mixed CSIT

Zheng Li, Chao He, Sheng Yang

On the Capacity of the Two-User Erasure Broadcast Channel with Mixed CSIT

Zheng Li[‡§], Chao He[‡], and Sheng Yang[‡]
[‡] L2S, CentraleSupélec-CNRS-Université Paris-Sud, 91192, Gif-sur-Yvette, France
[§]Orange Labs Networks, 92326, Chatillon, France
Email: {zheng1.li}@orange.com, {chao.he, sheng.yang}@centralesupelec.fr

***Abstract*—This paper investigates the two-user Erasure Broadcast Channel (EBC), where the channel state information (CSI) is fully known at the destination, while the transmitter is only aware of the strictly causal CSI by state feedback and an estimate of the instantaneous CSI. We propose a novel transmission scheme that exploits both the delayed and the instantaneous CSI. Our scheme includes both the case with full CSI and the case with delayed CSI as special cases. We also derive a new outer bound region for this channel. For the symmetric EBC, we show that our scheme is capacity achieving in some nontrivial cases. Since both the inner and outer bound regions are characterized with linear constraints, numerical evaluation can be done easily.**

I. Introduction

It has been well known that the capacity of a multi-user communication channel depends on the channel state information at the transmitters' side (CSIT). To see this, let us consider a symmetric two-user binary erasure broadcast channel (EBC) in which the signal is erased independently with probability δ at each receiver. The CSI in this case is in fact the knowledge of the erasure event. Without CSIT, this is a degraded broadcast channel [1] for which we know the capacity region [2]. For the current case, we can show that time division multiple access (TDMA) is indeed optimal with the corresponding rate region containing rate pairs satisfying

$$R_1 + R_2 \leq 1 - \delta. \tag{1}$$

If, however, perfect instantaneous CSI can be obtained at the transmitter, then the capacity region becomes

$$\begin{cases} R_1 \leq 1 - \delta, \\ R_2 \leq 1 - \delta, \\ R_1 + R_2 \leq 1 - \delta^2. \end{cases} \tag{2}$$

Intuitively, as long as one of the users is not in erasure (with probability $1 - \delta^2$), the transmitter can still send useful information to this user reliably thanks to the instantaneous CSIT. The above comparison shows that instantaneous CSIT can enhance the sum capacity by a factor of $\frac{1-\delta^2}{1-\delta} = 1 + \delta$. In recent studies [3], [4], it turned out that even delayed CSIT can be very useful. Indeed, an EBC with delayed CSIT can be modeled as an EBC with feedback for which the capacity region has been shown to be

$$\begin{cases} \frac{R_1}{1-\delta} + \frac{R_2}{1-\delta^2} \leq 1, \\ \frac{R_1}{1-\delta^2} + \frac{R_2}{1-\delta} \leq 1. \end{cases} \tag{3}$$

It follows that the capacity gain brought by delayed CSIT is $1 + \frac{\delta}{2+\delta}$. To achieve the capacity region (3), the authors in [3], [4] proposed a two-phase scheme that exploits the multicast opportunities by sending linear combinations of packets that are useful for both users based on the delayed CSIT. Note that the delayed CSIT setting is interesting in practice since feedback is often available for systems with Automatic Repeat reQuest (ARQ).

In this paper, we are interested in a more general CSIT setting for the EBC in which an estimate of the instantaneous CSI is somehow known at the transmitter in addition to the delayed CSI. Such a setting includes both the perfect CSIT and delayed CSIT as special cases by changing the quality of the estimation. The assumption is motivated by the fact that the knowledge of the current state is sometimes available through some unreliable sources. Our goal is to find out how to exploit both the reliable feedback and the unreliable state information simultaneously. It is worth mentioning that our setup is closely related to the one in [5], where a two-user EBC with memory and causal feedback has been studied. In that case, the imperfect instantaneous CSIT is inferred from the past states based on the temporal correlation that is modeled by a stationary Markov process. In our work, we assume that the channel is memoryless (i.e. temporally independent), but the correlation between the estimated and actual instantaneous CSIT can be arbitrary. Therefore, neither setting subsumes the other. The contribution of our work is two-fold. First, we propose a new transmission scheme integrating the block-Markov scheme and the joint source-channel coding (JSC) used in [6]. While the block-Markov structure is not necessary as it has been shown in [6] for the case with only state feedback, it turns out to be important here with instantaneous state information. We then derive an achievable rate region by carefully choosing the input distribution. Second, we derive a new capacity outer bound region that depends on the joint distribution of the states. Remarkably, both inner and outer bounds are subject to linear constraints and can be evaluated numerically with standard tools. Focusing on the symmetric case, our analytical results show that the proposed scheme achieves the capacity region in some nontrivial cases.

The setting of mixed CSIT has been considered for multi-antenna fading broadcast channels in [7], [8], [9] where the optimal degrees of freedom region has been derived. As pointed out in [10], [6], there are close connections between the fading

Gaussian BC and the erasure BC in some special cases. Indeed, both channels fall into the class of state-dependent memoryless channels with feedback for which general achievable rate regions have been investigated, e.g., in [11], [12]. Nevertheless, such results can provide useful insights on particular channels only when the auxiliary random variables are properly chosen in such a way that the region can be evaluated analytically or numerically. Furthermore, finding tractable and tight outer bounds is another obstacle to characterizing the capacity. That being said, this paper derives both inner and outer bounds in a simple form for the EBC with mixed CSIT, which we believe is novel and constructive.

The paper is organized as follows. In Section II, we describe the system model, achievable rate region is derived in Section III while outer bound is proposed in Section IV. Some examples are given in Section V. Finally, Section VI concludes the paper.

II. System Model

Throughout the paper, we use the following notations. Logarithm is to the base 2. The vector of ones is denoted by $\mathbf{1}$. We use $\boldsymbol{u} \preceq \boldsymbol{v}$ to mean that $u_i \le v_i,\ \forall\, i$. $\boldsymbol{u} \circ \boldsymbol{v}$ is the point-wise product between vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. Matrix transpose is denoted by $\cdot^{\mathsf{T}}$.

First, let us consider a general two-user discrete memoryless state-dependent BC $(\mathcal{X} \times \mathcal{S} \times \hat{\mathcal{S}}, p(y|x,s)p(s,\hat{s}), \mathcal{Y}_1 \times \mathcal{Y}_2)$:

$$p\left(\boldsymbol{y}_1,\boldsymbol{y}_2|\boldsymbol{x},\boldsymbol{s}\right)p(\boldsymbol{s},\hat{\boldsymbol{s}}) = \prod_{i=1}^{n} p(y_{1i},y_{2i}|x_i,s_i)p(s_i,\hat{s}_i), \quad (4)$$

where $\boldsymbol{s} \in \mathcal{S}^n, \hat{\boldsymbol{s}} \in \hat{\mathcal{S}}^n, \boldsymbol{x} \in \mathcal{X}^n, \boldsymbol{y}_1 \in \mathcal{Y}_1^n, \boldsymbol{y}_2 \in {\mathcal{Y}_2}^n$ are the sequences of the channel state, estimated channel state, the channel input, and the channel output at receiver 1 and receiver 2, respectively. We assume that at time i, the channel state information $\boldsymbol{s}^{i-1}$ is available to the transmitter via perfect feedback, whereas the estimated channel state information $\hat{\boldsymbol{s}}$ can be somehow obtained non-causally from an unreliable source. At the end of the transmission of n symbols, both $\boldsymbol{s}$ and $\hat{\boldsymbol{s}}$ are known to both receivers for the decoding. In particular, for the transmission of messages (m_1,m_2) to receivers 1 and 2, respectively, with $m_k \in \mathcal{M}_k := [1:2^{nR_k}]$, $k \in \{1,2\}$, the encoding functions are $\{\phi_i : \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{S}^{i-1} \times \hat{\mathcal{S}}^n \to \mathcal{X}\}_{i=1}^n$, whereas the decoding function at receiver k, $k \in \{1,2\}$, is $\varphi_k : \mathcal{Y}_k^n \times \mathcal{S}^n \times \hat{\mathcal{S}}^n \to \mathcal{M}_k$. The rate pair (R_1,R_2) is achievable if the probability of error of each user goes to 0 when $n \to \infty$.

The erasure BC that we consider in this paper is a special case of the above channel where $\mathcal{S} := \{(h_1,h_2) :\ h_1,h_2 \in \{0,1\}\}$ and that the outputs are deterministic functions of the input given the state, namely, at each time i, the output of receiver k is $y_{k,i} = x_i$ when $h_{k,i} = 1$ and $y_{k,i} = ?$ when $h_{k,i} = 0$. Therefore, $\mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{X} \cup \{?\}$. We also assume that $\hat{\mathcal{S}} = \mathcal{S}$ and the states are in short $\{00,01,10,11\}$. The joint distribution of $(\mathrm{S},\hat{\mathrm{S}})$ is $\mathbb{P}\left(\mathrm{S}=s,\hat{\mathrm{S}}=\hat{s}\right)$ for $s,\hat{s} \in \mathcal{S}$. For convenience, we define the following probability vectors, for each $s \in \mathcal{S}$,

$$\boldsymbol{p}_s := \left[\mathbb{P}\left(\mathrm{S}=s,\hat{\mathrm{S}}=\hat{s}\right) :\ \hat{s} \in \hat{\mathcal{S}}\right] \in [0,1]^4 \quad (5)$$

and $\boldsymbol{p}_{\bar{s}} := 1 - \boldsymbol{p}_s$. Further, we define the marginal probability $p_s := \mathbb{P}\left(\mathrm{S}=s\right)$ for each $s \in \mathcal{S}$. Thus we have $p_s = \mathbf{1}^{\mathsf{T}}\boldsymbol{p}_s$. In particular, the erasure probabilities for users 1 and 2 are defined as $\delta_1 := p_{01,00}$ and $\delta_2 := p_{10,00}$, respectively. Finally, let us define the vector of the marginal probabilities of $\hat{S}$

$$\hat{\boldsymbol{p}} := \left[\mathbb{P}(\hat{S}=00),\, \mathbb{P}(\hat{S}=01),\, \mathbb{P}(\hat{S}=10),\, \mathbb{P}(\hat{S}=11)\right]^{\mathsf{T}}. \quad (6)$$

III. Achievable Rate Region

The main result of this paper is stated below.

Theorem 1 (Inner bound)**.** *Let us define*

$$\mathcal{R}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) := \left\{(R_1,R_2) :\ \begin{array}{l} R_1 \le \boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{p}_{\overline{00}} \log|\mathcal{X}|, \\ R_2 \le \boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{p}_{\overline{00}} \log|\mathcal{X}|, \end{array}\right\}. \quad (7)$$

Then an achievable rate region of the two-user EBC with mixed CSIT is the convex hull of the union of $\mathcal{R}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)$ over all $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2 \in [0,1]^4$ such that $\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 \preceq \mathbf{1}$ and

$$\boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{p}_{10,11} + \boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{p}_{\overline{00}} \le 1 - \delta_1, \quad (8)$$

$$\boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{p}_{01,11} + \boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{p}_{\overline{00}} \le 1 - \delta_2. \quad (9)$$

In the rest of the section, we first describe the proposed scheme and derive a general region, then we choose the distribution that leads us to the above result.

A. Proposed scheme

In this part, we describe the proposed scheme in detail. The proposed scheme generalizes the JSC scheme in [6] for the two-user case, by integrating the block-Markov coding and current state information. Hereafter, we refer to our scheme as the BM-JSC scheme.

The BM-JSC scheme works in multiple blocks. In each block, the source encodes and transmits the fresh private messages together with the side information common message. Such side information is generated from the state and private messages of the previous block. The transmission in each block is controlled by the coded time-sharing sequence which also depends on the estimated channel state. At the end of the transmission of all the blocks, a backward decoding is performed by each receiver. To be specific, each receiver decodes first the side information message using the observations of the current and previous blocks. Then, it decodes the private message using the observation of the current block as well as the decoded side information during the "future" block.

In particular, the transmission consists of $B+1$ blocks each of length n. The messages M_k intended for receiver k, $k \in \{1,2\}$, is divided into B sub-messages, i.e., $m_{k,b} \in \mathcal{M}_k$, each one transmitted in block b, $b \in [1:B]$.

Codebook generation: Fix a probability mass function (pmf)

$$p(x|v_0,v_1,v_2,q)p(v_0)p(v_1)p(v_2)p(\hat{y}|v_0,v_1,v_2,s,q)p(q|\hat{s}). \quad (10)$$

1) Before each block, randomly generate the time-sharing sequence according to $\prod_{t=1}^{n} p(q_t \mid \hat{s}_t)$.
2) At the beginning of each block, for each user $k \in \{1,2\}$, randomly and independently generate 2^{nR_k} sequences $\boldsymbol{v}_k(m_k)$, $m_k \in [1:2^{nR_k}]$, each according to $\prod_{t=1}^{n} p(v_{k,t})$.

At the same moment, randomly generate 2^{nR_0} independent sequences $\boldsymbol{v}_0(m_0)$, $m_0 \in [1:2^{nR_0}]$, each according to $\prod_{t=1}^{n} p(v_{0,t})$.

3) At the end of each block, and upon the reception of the state feedback $\boldsymbol{s}$, randomly and independently generate 2^{nR_0} sequences $\hat{\boldsymbol{y}}(m_0)$, $m_0 \in [1:2^{nR_0}]$, each according to $\prod_{t=1}^{n} p(\hat{y}_t|s_t, q_t)$.

Encoding: We assume by convention that $m_{0,0} = 1$ and $m_{1,B+1} = m_{2,B+1} = 1$.

- At the end of block b, $b \in [1:B]$, given the state feedback $\boldsymbol{s}_b$ and messages $m_{0,b-1}$, $m_{1,b}$, and $m_{2,b}$, the encoder looks for a unique message index $m_{0,b}$ such that the corresponding sequence $\hat{\boldsymbol{y}}(m_{0,b})$ satisfies $(\hat{\boldsymbol{y}}(m_{0,b}), \boldsymbol{v}_0(m_{0,b-1}), \boldsymbol{v}_1(m_{1,b}), \boldsymbol{v}_2(m_{2,b}), \boldsymbol{s}_b, \boldsymbol{q}_b, \hat{\boldsymbol{s}}_b) \in \mathcal{T}_{\epsilon_n}^{n}(\hat{Y}V_0V_1V_2SQ\hat{S})$. If there is more than one index, it selects one of them uniformly at random, otherwise an error is declared. According to the covering lemma [13], such an index can be found with high probability if
$$nR_0 \geq nI(\hat{Y}; V_0, V_1, V_2|S, Q, \hat{S}) + n\epsilon_n. \quad (11)$$
- In block b, $b \in [1:B+1]$, the transmitter generates a sequence $\boldsymbol{x}$ from $(\boldsymbol{v}_0(m_{0,b-1}), \boldsymbol{v}_1(m_{1,b}), \boldsymbol{v}_2(m_{2,b}), \boldsymbol{q}_b)$ according to $\prod_{t=1}^{n} p(x_t \,|\, v_{0t}, v_{1t}, v_{2t}, q_t)$.

Decoding:

- After receiving all the signals, a backward decoding is performed at each receiver. For $b \in [1:B]$, with the knowledge of $\hat{m}_{k,b+1}$ and the states information, the receiver k finds a unique index $\hat{m}_{0,b}$ such that both typicalities $(\hat{\boldsymbol{y}}(\hat{m}_{0,b}), \boldsymbol{y}_{k,b}, \boldsymbol{s}_b, \boldsymbol{q}_b, \hat{\boldsymbol{s}}_b) \in \mathcal{T}_{\epsilon_n}^{n}(\hat{Y}Y_kSQ\hat{S})$ and $(\boldsymbol{v}_0(\hat{m}_{0,b}), \boldsymbol{v}_k(\hat{m}_{k,b+1}), \boldsymbol{y}_{k,b+1}, \boldsymbol{s}_{b+1}, \boldsymbol{q}_{b+1}, \hat{\boldsymbol{s}}_{b+1}) \in \mathcal{T}_{\epsilon_n}^{n}(V_0V_kY_kSQ\hat{S})$ hold *simultaneously*. Following the footsteps of error analysis in [14], [6], such an index $\hat{m}_{0,b} = m_{0,b}$ can be found with high probability if
$$nR_0 \leq nI(\hat{Y}; Y_k \,|\, S, Q, \hat{S}) + nI(V_0; Y_k \,|\, V_k, S, Q, \hat{S}) - n\epsilon_n'. \quad (12)$$
- Given that $\hat{m}_{0,b}$ is available for $b \in [1:B]$, the decoder k looks for a unique message index $\hat{m}_{k,b}$ such that the typicality $(\boldsymbol{v}_k(\hat{m}_{k,b}), \boldsymbol{y}_{k,b}, \hat{\boldsymbol{y}}(\hat{m}_{0,b}), \boldsymbol{s}_b, \boldsymbol{q}_b, \hat{\boldsymbol{s}}_b) \in \mathcal{T}_{\epsilon_n}^{n}(V_kY_k\hat{Y}SQ\hat{S})$ holds. Such an index $\hat{m}_{k,b} = m_{k,b}$ can be found with high probability provided that
$$nR_k \leq nI(V_k; Y_k, \hat{Y} \,|\, S, Q, \hat{S}) - n\epsilon_n. \quad (13)$$

From (11)-(13), letting $n \to \infty$, then $B \to \infty$, we obtain the following achievable rate region for the general case.

Proposition 1 (BM-JSC inner bound)**.** *For any pmf* (10)*, a rate pair* (R_1, R_2) *is achievable with the proposed BM-JSC scheme if the following conditions are satisfied*
$$R_1 \leq I(V_1; Y_1, \hat{Y} \,|\, S, Q, \hat{S}), \quad (14)$$
$$R_2 \leq I(V_2; Y_2, \hat{Y} \,|\, S, Q, \hat{S}), \quad (15)$$
$$I(\hat{Y}; V_0, V_1, V_2 \,|\, Y_1, S, Q, \hat{S}) \leq I(V_0; Y_1 \,|\, V_1, S, Q, \hat{S}), \quad (16)$$
$$I(\hat{Y}; V_0, V_1, V_2 \,|\, Y_2, S, Q, \hat{S}) \leq I(V_0; Y_2 \,|\, V_2, S, Q, \hat{S}). \quad (17)$$

In order to derive the rate region in Theorem 1 for the erasure BC, we make the following choices on the pmf (10).

- The coded time-sharing random variable (RV) Q conditional on the estimated state $\hat{S}$ is distributed as
$$\mathbb{P}\left(Q = i \,|\, \hat{S} = \hat{s}\right) = \alpha_{i,\hat{s}}, \quad (18)$$
where $i \in \mathcal{Q} := \{0, 1, 2\}$, $\hat{s} \in \hat{\mathcal{S}}$ with $\alpha_{i,\hat{s}} \geq 0$ and $\sum_{i=0}^{2} \alpha_{i,\hat{s}} = 1$ for $\hat{s} \in \hat{\mathcal{S}}$.
- The signal V's are uniformly distributed over $\mathcal{X}$.
$$\mathbb{P}(V_i = x) = \frac{1}{|\mathcal{X}|}, \quad x \in \mathcal{X},\ i \in \mathcal{Q}.$$
- The input X is a deterministic function of V and Q.
$$X = V_Q. \quad (19)$$
- The side information $\hat{Y}$ is a deterministic function of (V, S, Q).
$$\hat{Y} = \begin{cases} V_1, & \text{if } Q = 1, S = 01, \\ V_2, & \text{if } Q = 2, S = 10, \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

Applying the above choices to (14)-(17), we have, for $k = 1$,
$$\begin{aligned} &I(V_1; Y_1, \hat{Y} \,|\, S, Q, \hat{S}) \\ &= \sum_{s\in\mathcal{S}, \hat{s}\in\mathcal{S}} I(V_1; Y_1\hat{Y} \,|\, Q = 1, S = s, \hat{S} = \hat{s}) \\ &\qquad \cdot \mathbb{P}(Q = 1 \,|\, \hat{S} = \hat{s})\mathbb{P}(S = s, \hat{S} = \hat{s}) \\ &= \sum_{s\neq 00, \hat{s}\in\mathcal{S}} \alpha_{1,\hat{s}}\mathbb{P}(S = s, \hat{S} = \hat{s})H(V_1) \\ &= \boldsymbol{\alpha}_1^{\mathsf{T}} \boldsymbol{p}_{\overline{00}} \log|\mathcal{X}|, \end{aligned} \quad (21)$$
$$\begin{aligned} &I(V_0; Y_1 \,|\, V_1, S, Q, \hat{S}) \\ &= \sum_{s\in\mathcal{S}, \hat{s}\in\mathcal{S}} I(V_0; Y_1 \,|\, Q = 0, S = s, \hat{S} = \hat{s}) \\ &\qquad \cdot \mathbb{P}(Q = 0 \,|\, \hat{S} = \hat{s})\mathbb{P}(S = s, \hat{S} = \hat{s}) \\ &= \sum_{s\in\{10,11\}, \hat{s}\in\mathcal{S}} \alpha_{0,\hat{s}}\mathbb{P}(S = s, \hat{S} = \hat{s})H(V_0) \\ &= \boldsymbol{\alpha}_0^{\mathsf{T}} \boldsymbol{p}_{10,11} \log|\mathcal{X}|, \end{aligned} \quad (22)$$
$$\begin{aligned} &I(\hat{Y}; V_0, V_1, V_2 \,|\, Y_1, S, Q, \hat{S}) \\ &= \sum_{s\in\mathcal{S}, \hat{s}\in\mathcal{S}, i\in\mathcal{Q}} I(\hat{Y}; V_0, V_1, V_2 \,|\, Y_1, Q = i, S = s, \hat{S} = \hat{s}) \\ &\qquad \cdot \mathbb{P}(Q = i \,|\, \hat{S} = \hat{s})\mathbb{P}(S = s, \hat{S} = \hat{s}) \\ &= \sum_{\hat{s}\in\mathcal{S}} \mathbb{P}(Q = 1 \,|\, \hat{S} = \hat{s})\mathbb{P}(S = 01, \hat{S} = \hat{s})H(V_1) \quad (23) \\ &= \boldsymbol{\alpha}_1^{\mathsf{T}} \boldsymbol{p}_{01} \log|\mathcal{X}|, \end{aligned} \quad (24)$$
where we define $\boldsymbol{\alpha}_i := [\alpha_{i,\hat{s}} : \hat{s} \in \mathcal{S}] \in [0,1]^4$, $i \in \mathcal{Q}$, with $\boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 = \mathbf{1}$. Therefore, we have
$$R_1 \leq \boldsymbol{\alpha}_1^{\mathsf{T}} \boldsymbol{p}_{\overline{00}} \log|\mathcal{X}|, \quad (25)$$
$$\boldsymbol{\alpha}_1^{\mathsf{T}} \boldsymbol{p}_{01} \leq \boldsymbol{\alpha}_0^{\mathsf{T}} \boldsymbol{p}_{10,11}. \quad (26)$$

Due to the symmetry, the same calculation applies for $k = 2$. After eliminating $\boldsymbol{\alpha}_0$, we can obtain the result in Theorem 1.

Remark III.1. *Note that the coded-time sharing variable is controlled by the estimated state. If we use the two-phase scheme without the block-Markov encoding, this is still possible, but only in the first phase during which the private information is sent [6]. In this case, the proportion of the phase lengths cannot depend on the state, which yields strictly worse performance than BM-JSC scheme in which coded-time sharing can provide a flexible dependence on the state during the whole transmission.*

IV. Outer bound

The current channel can be seen equivalently as a state-dependent memoryless BC in which $\hat{S}$ is the state known non-causally everywhere and $\tilde{Y}_k := (Y_k, S)$ is the output at receiver $k \in \{1, 2\}$. If a genie provides $\tilde{Y}_2$ to receiver 1, we obtain a physically degraded BC with $X \leftrightarrow \tilde{Y}_1\tilde{Y}_2 \leftrightarrow \tilde{Y}_2$ for which feedback does not enlarge the capacity region [15]. The single-letter capacity region for such a degraded channel is the set of all (R_1, R_2) such that

$$R_1 \le I(X; \tilde{Y}_1, \tilde{Y}_2 \,|\, \hat{S}, U), \tag{27}$$

$$R_2 \le I(U; \tilde{Y}_2 \,|\, \hat{S}), \tag{28}$$

for some pmf $p(x, u \,|\, \hat{s})$. Note that $XU \leftrightarrow \hat{S} \leftrightarrow S$. Thus,

$$\begin{aligned}
R_1 &\le I(X; Y_1, Y_2 \,|\, S, \hat{S}, U) && (29)\\
&= \sum_{s\in\mathcal{S}, \hat{s}\in\mathcal{S}} I(X; Y_1, Y_2 \,|\, S = s, \hat{S} = \hat{s}, U)\mathbb{P}(S = s, \hat{S} = \hat{s})\\
&= \sum_{\hat{s}\in\mathcal{S}} H(X \,|\, \hat{S} = \hat{s}, U)\mathbb{P}(S \neq 00, \hat{S} = \hat{s}) && (30)\\
&= \boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{p}_{\overline{00}} \log|\mathcal{X}|, && (31)\\
R_2 &\le I(U; Y_2 \,|\, S, \hat{S}) && (32)\\
&= \sum_{\hat{s}\in\mathcal{S}} \Big(H(X \,|\, \hat{S} = \hat{s}) - H(X \,|\, \hat{S} = \hat{s}, U)\Big)\\
&\qquad \cdot \mathbb{P}(S \in \{01, 11\}, \hat{S} = \hat{s}) && (33)\\
&\le \sum_{\hat{s}\in\mathcal{S}} (\log|\mathcal{X}| - H(X \,|\, \hat{S} = \hat{s}, U))\\
&\qquad \cdot \mathbb{P}(S \in \{01, 11\}, \hat{S} = \hat{s}) && (34)\\
&= p_{01,11} \log|\mathcal{X}| - \boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{p}_{01,11} \log|\mathcal{X}|, && (35)
\end{aligned}$$

where (34) is from the upper bound of entropy; we define $\log|\mathcal{X}|\boldsymbol{\beta} := \big[H(X \,|\, \hat{S} = \hat{s}, U) : \hat{s} \in \mathcal{S}\big]$ with $\boldsymbol{\beta} \preceq \mathbf{1}$. Obviously, any achievable rate pair (R_1, R_2) of the original channel must satisfy (31) and (35) from the genie-aided argument. Similarly, the genie can provide $\tilde{Y}_1$ to receiver 2 and we obtain the following outer bound region.

Theorem 2. *(Outer bound) Any achievable rate pair (R_1, R_2) for the two-user EBC with feedback and estimated current CSI must satisfy, for all $\lambda \ge 1$,*

$$R_1 + \lambda R_2 \le \big(\lambda(1 - \delta_2) + \mathbf{1}^{\mathsf{T}}(\boldsymbol{p}_{\overline{00}} - \lambda\boldsymbol{p}_{01,11})^+\big) \log|\mathcal{X}|, \tag{36}$$

$$R_2 + \lambda R_1 \le \big(\lambda(1 - \delta_1) + \mathbf{1}^{\mathsf{T}}(\boldsymbol{p}_{\overline{00}} - \lambda\boldsymbol{p}_{10,11})^+\big) \log|\mathcal{X}|. \tag{37}$$

Proof. From (31) and (35), we have

$$R_1 + \lambda R_2 \le \big(\lambda(1 - \delta_2) + \boldsymbol{\beta}^{\mathsf{T}}(\boldsymbol{p}_{\overline{00}} - \lambda\boldsymbol{p}_{01,11})\big) \log|\mathcal{X}|,$$

which can be maximized by letting $\beta_{\hat{s}} = 0$ whenever the corresponding component in $\boldsymbol{p}_{\overline{00}} - \lambda\boldsymbol{p}_{01,11}$ is negative and letting $\beta_{\hat{s}} = 1$ otherwise. This can be written equivalently as in the right hand side of (36). Due to the symmetry, the second bound (37) follows similarly. □

V. Examples

In general, both the inner bound region in Theorem 1 and the outer bound region in Theorem 2 can be evaluated numerically. However, it is harder to compare both regions analytically, due to the generality of the state distribution $p(s, \hat{s})$. In this section, we look at some particular examples and apply the inner and outer bounds. For simplicity, we consider the binary case.

A. Perfect CSIT

In this case, we have $\boldsymbol{p}_{\overline{00}} = [0\ p_{01}\ p_{10}\ p_{11}]^{\mathsf{T}}$, $\boldsymbol{p}_{01,11} = [0\ p_{01}\ 0\ p_{11}]^{\mathsf{T}}$, and $\boldsymbol{p}_{10,11} = [0\ 0\ p_{10}\ p_{11}]^{\mathsf{T}}$. For $\lambda \ge 1$, we have $\mathbf{1}^{\mathsf{T}}(\boldsymbol{p}_{\overline{00}} - \lambda\boldsymbol{p}_{01,11})^+ = p_{10}$ and $\mathbf{1}^{\mathsf{T}}(\boldsymbol{p}_{\overline{00}} - \lambda\boldsymbol{p}_{10,11})^+ = p_{01}$. We can verify that the corner points in the outer bound are $(R_1, R_2) = (p_{10}, 1 - \delta_2)$ and $(R_1, R_2) = (1 - \delta_1, p_{01})$. It can be shown that both corner points are achievable. Indeed, setting $\boldsymbol{\alpha}_1 = [0\ 0\ 1\ 0]^{\mathsf{T}}$ and $\boldsymbol{\alpha}_2 = [0\ 1\ 0\ 1]^{\mathsf{T}}$ in Theorem 1, and we see that $(R_1, R_2) = (p_{10}, 1 - \delta_2)$ is achievable with equality in both (8) and (9). The other corner point follows similarly.

B. Delayed CSIT only

In this case, $\hat{S}$ and S are independent, which implies that $\boldsymbol{p}_{\overline{00}} = p_{\overline{00}}\hat{\boldsymbol{p}}$, $\boldsymbol{p}_{01,11} = (1 - \delta_2)\hat{\boldsymbol{p}}$, and $\boldsymbol{p}_{10,11} = (1 - \delta_1)\hat{\boldsymbol{p}}$, where the vector $\hat{\boldsymbol{p}}$ is defined in (6). It follows that there is only one corner point in the positive quadrant. And the outer bound region is reduced to

$$\frac{R_1}{p_{\overline{00}}} + \frac{R_2}{1 - \delta_2} \le 1, \tag{38}$$

$$\frac{R_2}{p_{\overline{00}}} + \frac{R_1}{1 - \delta_1} \le 1. \tag{39}$$

In fact, this region corresponds to the inner bound region (8) and (9) if we let $R_1 = \boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{p}_{\overline{00}}$ and $R_2 = \boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{p}_{\overline{00}}$ and use the identities $\boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{p}_{\overline{00}} = \boldsymbol{\alpha}_1^{\mathsf{T}}\boldsymbol{p}_{01,11}\frac{p_{\overline{00}}}{1-\delta_2}$ and $\boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{p}_{\overline{00}} = \boldsymbol{\alpha}_2^{\mathsf{T}}\boldsymbol{p}_{10,11}\frac{p_{\overline{00}}}{1-\delta_1}$.

C. Mixed CSIT in symmetric channels

The general joint distribution of $(S, \hat{S})$ involves a large number of parameters. To limit the number of parameters, we consider the symmetric and spatially independent case. Specifically, for $S = (H_1, H_2) \in \{0, 1\}^2$ and $\hat{S} = (\hat{H}_1, \hat{H}_2) \in \{0, 1\}^2$, let us assume $\mathbb{P}(S = (h_1, h_2), \hat{S} = (\hat{h}_1, \hat{h}_2)) = \mathbb{P}(H_1 = h_1, \hat{H}_1 = \hat{h}_1)\mathbb{P}(H_2 = h_2, \hat{H}_2 = \hat{h}_2)$, with $\mathbb{P}(H_k = 0) = \delta$ and $\mathbb{P}(H_k = 1 \,|\, \hat{H}_k = 0) = \mathbb{P}(H_k = 0 \,|\, \hat{H}_k = 1) = \epsilon$, which yields $\mathbb{P}(\hat{H}_k = 0) = \frac{\delta - \epsilon}{1 - 2\epsilon}$, for $k \in \{1, 2\}$. Here ϵ captures the estimation error probability. We can rewrite

$$\boldsymbol{p}_{\overline{00}} = \hat{\boldsymbol{p}} \circ \begin{bmatrix} 1 - \bar{\epsilon}^2 & 1 - \epsilon\bar{\epsilon} & 1 - \epsilon\bar{\epsilon} & 1 - \epsilon^2 \end{bmatrix}, \tag{40}$$

$$\boldsymbol{p}_{01,11} = \hat{\boldsymbol{p}} \circ \begin{bmatrix} \epsilon & \bar{\epsilon} & \epsilon & \bar{\epsilon} \end{bmatrix}, \tag{41}$$

$$\boldsymbol{p}_{10,11} = \hat{\boldsymbol{p}} \circ \begin{bmatrix} \epsilon & \epsilon & \bar{\epsilon} & \bar{\epsilon} \end{bmatrix}. \tag{42}$$

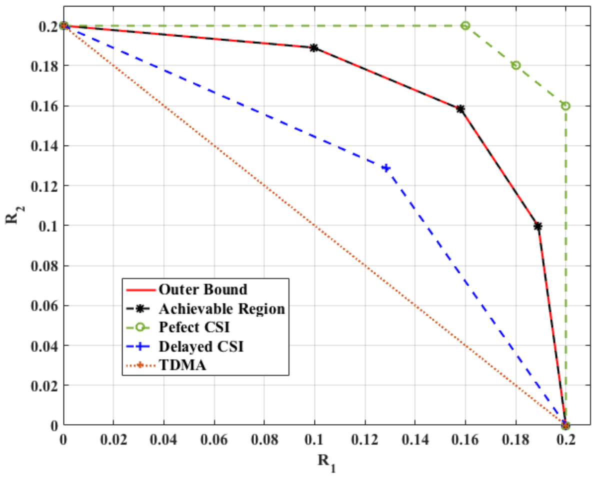

Fig. 1. Rate region for two-user symmetric EBC with $\delta = 0.8, \epsilon = 0.1$.

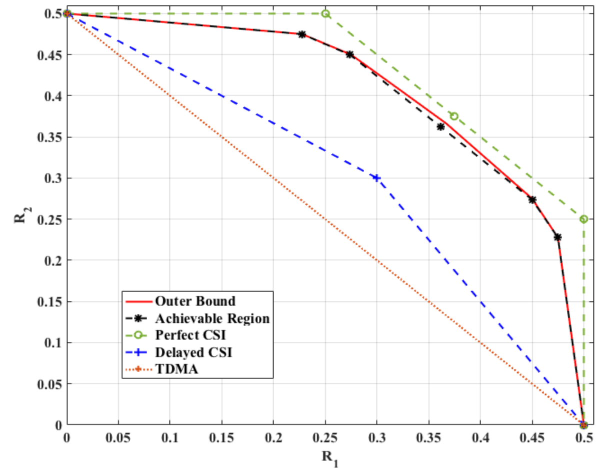

Fig. 2. Rate region for two-user symmetric EBC with $\delta = 0.5, \epsilon = 0.1$.

By inspecting $(\boldsymbol{p}_{\overline{00}} - \lambda \boldsymbol{p}_{01,11})^+$ assuming $\lambda \geq 1$ and $\epsilon \leq 0.5$, we can completely characterize the region from (36) in Theorem 2 over all $\lambda \geq 1$ using the five corner points below:

	R_1	R_2
CP_1	0	$\mathbb{P}(S \in \{01, 11\})$
CP_2	$\mathbb{P}(S \neq 00, \hat{S} = 10)$	$\mathbb{P}(S \in \{01, 11\}, \hat{S} \neq 10)$
CP_3	$\mathbb{P}(S \neq 00, \hat{S} \in \{00, 10\}$	$\mathbb{P}(S \in \{01, 11\}, \hat{S} \in \{01, 11\})$
CP_4	$\mathbb{P}(S \neq 00, \hat{S} \neq 01)$	$\mathbb{P}(S \in \{01, 11\}, \hat{S} = 01)$
CP_5	$\mathbb{P}(S \neq 00)$	0

Specifically, the convex hull $\mathcal{R}_{\text{out}}$ of these corner points is an outer bound. From the symmetry, we know that $\mathcal{R}'_{\text{out}} := \{(R_1, R_2) : (R_2, R_1) \in \mathcal{R}_{\text{out}}\}$ is also an outer bound that can be derived from the second constraint in Theorem 2. Therefore, $\mathcal{R}_{\text{out}} \cap \mathcal{R}'_{\text{out}}$ provides a tighter outer bound with up to seven corner points inside the positive quadrant[1]. Let us take a closer look at two numerical examples.

First, we let $\delta = 0.8$ and $\epsilon = 0.1$. As shown in Figure 1, the outer bound region contains five corner points. It can be verified that there exist $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ for each of these corner points such that the conditions (7)-(9) hold. In particular, to achieve the symmetric point $(0.1582, 0.1582)$, we can let $\boldsymbol{\alpha}_1^\intercal = [\alpha \ 0 \ 1 \ 0]$ and $\boldsymbol{\alpha}_2^\intercal = [\alpha \ 1 \ 0 \ 0]$ with α such that (8) and (9) hold with equality ($\alpha \approx 0.4032$).

Then, we let $\delta = 0.5$ and $\epsilon = 0.1$. In this case, we compute the inner bound numerically with linear programming routines. The outer bound can be obtained with the corner point characterization. As shown in Figure 2, the inner bound is strictly smaller than the outer bound region. It is however unclear whether this is due to the looseness of the outer bound or the sub-optimality of the proposed scheme.

VI. Conclusion

In this work, we proposed a novel transmission scheme that exploits both the delayed CSIT and estimated instantaneous CSIT in a two-user EBC. Our scheme is based on block-Markov coding and joint source-channel coding. We also derived an outer bound region that coincides with the inner bound region for some nontrivial cases. Nevertheless, the capacity region with general state distribution remains unknown.

[1] There can be at most three corner points from $\mathcal{R}_{\text{out}}$ in $\mathcal{R}_{\text{out}} \cap \mathcal{R}'_{\text{out}}$ inside the positive quadrant. and another three from $\mathcal{R}_{\text{out}}$. The intersection $\mathcal{R}_{\text{out}} \cap \mathcal{R}'_{\text{out}}$ creates one more corner point.

References

[1] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, 1972.

[2] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Problemy Peredachi Informatsii*, vol. 10, no. 3, pp. 3–14, 1974.

[3] C.-C. Wang, "On the capacity of 1-to-k broadcast packet erasure channels with channel output feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 931–956, 2012.

[4] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser broadcast erasure channel with feedback-capacity and algorithms," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5779–5804, 2013.

[5] M. Heindlmaier and S. S. Bidokhti, "Capacity regions of two-receiver broadcast erasure channels with feedback and memory," *arXiv preprint arXiv:1612.01487*, 2016.

[6] C. He, S. Yang, and P. Piantanida, "A non-linear transmission scheme for the k-user broadcast channel with state feedback," in *International Symposium on Information Theory and Its Applications (ISITA)*, 2016.

[7] T. Gou and S. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1084–1087, Jul. 2012.

[8] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315–328, Jan. 2013.

[9] J. Chen and P. Elia, "Toward the performance versus feedback tradeoff for the two-user miso broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8336–8356, 2013.

[10] S. Yang and M. Kobayashi, "Secure communication in k-user multi-antenna broadcast channel with state feedback," in *IEEE ISIT 2015*. IEEE, 2015, pp. 1976–1980.

[11] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1329–1345, 2013.

[12] Y. Wu and M. Wigger, "Coding schemes with rate-limited feedback that improve over the no feedback capacity for a large class of broadcast channels," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 2009–2033, 2016.

[13] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.

[14] E. Tuncel, "Slepian-Wolf coding over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, April. 2006.

[15] A. El Gamal, "The feedback capacity of degraded broadcast channels (corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 379–381, 1978.