On Capacity Regions of Two-Receiver Broadcast Packet Erasure Channels with Feedback and Memory

Michael Heindlmaier*, Navid Reyhanian†, Shirin Saeedi Bidokhti*

*Institute for Communications Engineering, Technische Universität München, Munich, Germany
†University of Tehran, Tehran, Iran
Email: michael.heindlmaier@tum.de, n.reyhanian@ut.ac.ir, shirin.saeedi@tum.de

***Abstract*—The two-receiver broadcast packet erasure channel with feedback and memory is studied. Memory is modelled using a finite-state Markov chain representing a channel state. Outer and inner bounds on the capacity region are derived when the channel state is strictly causally known at the transmitter. The bounds are both formulated in terms of feasibility problems and they are matching in all but one of the constraints. The results are extended to feedback with larger delay. Numerical results show that the gains offered through feedback can be quite large.**

I. Introduction

The capacity of broadcast channels (BCs) remains unresolved both without and with feedback. It was shown in [1] that feedback does not increase the capacity of physically degraded BCs. Nevertheless, feedback increases the capacity of general BCs and even partial feedback can help [2], [3]. Feedback also increases the capacity region of AWGN BCs [4], [5].

The capacity region of memoryless broadcast packet erasure channels (BPECs) with feedback (FB) was found in [6] for two receivers. The region is characterized by the closure of all non-negative rate pairs (R_1, R_2) such that

$$\frac{R_1}{1-\epsilon_1} + \frac{R_2}{1-\epsilon_{12}} \leq 1$$
$$\frac{R_1}{1-\epsilon_{12}} + \frac{R_2}{1-\epsilon_2} \leq 1,$$

where ϵ_1 and ϵ_2 are the erasure probabilities at receiver 1 and 2, respectively, and ϵ_{12} is the probability of erasure at both receivers. In particular, feedback increases the capacity and this is of practical interest since the required feedback is only a low-cost ACK/NACK signal that is easy to implement in BPECs.

This result has been extended to certain cases of broadcast channels with more number of receivers in [7], [8], [9]. In all these works, the capacity region is achieved using feedback-based coding algorithms that are based on network coding ideas. The converse theorems are proved by proving genie-aided outer bounds on the capacity region. The trick is that the genie helps the receivers such that the broadcast channel becomes a physically degraded one, for which the capacity region with feedback is known [1], [10], [11].

The capacity region of two-receiver multiple-input BPECs with feedback has been studied in [12] where the capacity region is derived and is shown to be achievable using linear network codes (LNC). The schemes are also applied to partially Markovian and partially controllable broadcast PECs where the linear network coding rate region is characterized by a linear program which exhaustively searches for the LNC scheme(s) with the best possible throughput. During the preparation of this work we were informed that the coding methods developed in [12] are able to achieve the outer bounds derived in Section IV of this work and are thus optimal.

In a recent trend of research, noisy feedback has been studied and achievable schemes are developed in [13], [14].

This paper studies BPECs with memory and feedback. The problem is motivated by the bursty nature of erasures in practical communication systems, e.g., satellite links [15], [16], [17]. We model the memory of a channel by a finite state machine and a set of state-dependent erasure probabilities. For finite state channel models see e.g. [18] and the references therein.

When there is no feedback, one can use erasure correcting codes for memoryless channels in combination with interleavers to decorrelate the erasures. But feedback enables more sophisticated coding methods and several such schemes are discussed in [19]. We remark that [20] studied the general broadcast channel with feedback and memory and considered different cooperation scenarios. The capacity characterizations in [20] are, however, in multi-letter form and not computable.

The main contribution of this paper is to provide lower and upper bounds on the capacity region for two receivers when the channel state is strictly causally known at the transmitter. Both bounds are formulated in terms of feasibility problems, and are similar in all but one set of constraints. Our outer bound is a genie-aided bound. The bound is subtle in the sense that it cannot be derived directly using the results

The authors are in alphabetical order and contributed equally to the work.

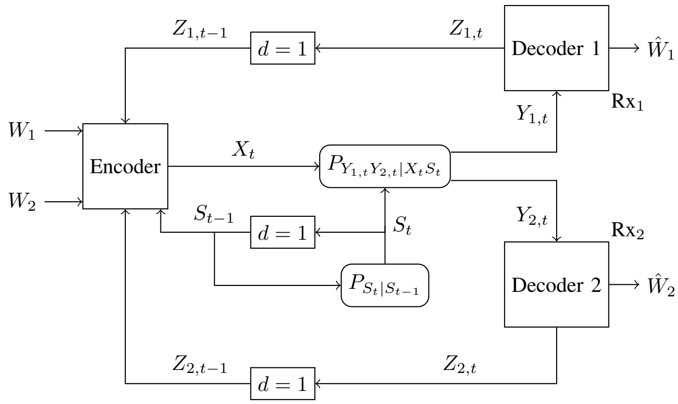

Fig. 1. Block diagram for the broadcast packet erasure channel with visible state. The box marked with $d = 1$ represents a delay of one time unit.

of [1], [10], [11]. The outer bound turns out to have a similar structure as for the memoryless case. Our proposed achievable scheme extends the queue-based algorithms of [6], [8], [21] to incorporate knowledge about the past channel states. The techniques generalize to BPECs with delayed feedback.

This paper is organized as follows: We introduce notation and the system model in Section II, and elaborate on the main result in Section III. The outer bound is presented in Section IV and the inner bound and two achievable schemes are discussed in Section V. In Section VI we discuss implications of our results.

II. Notation and System Model

A. Notation

Random variables are denoted by capital letters. A finite sequence (or string) of random variables $X_1, X_2, \ldots, X_n$ is denoted by X^n. In this context, sequences always refer to sequences in time. Sequences may have subscripts, e.g. X_j^n denotes $X_{j,1}, X_{j,2}, \ldots, X_{j,n}$. It is sometimes convenient to collect random variables that appear at the same time in a vector. Vectors are written with underlined letters, e.g., $\underline{Z}_t = (Z_{1,t}, Z_{2,t})$. Sets are denoted by calligraphic letters, e.g., $\mathcal{X}$. The indicator function $\mathbb{1}\{\cdot\}$ takes on the value 1 if the event inside the brackets is true and 0 otherwise. The probability of a random variable X taking on a realization x given an event $\mathcal{E}$ is written as $\Pr[X = x|\mathcal{E}]$. Often, the conditional event corresponds to another random variable Y taking on some realization y. This conditional probability is written as $\Pr[X = x|Y = y]$ or equivalently $P_{X|Y}(x|y)$. The equivalent expressions $\Pr[X|Y]$ or $P_{X|Y}$ are used to address the conditional probability (distribution) for any outcome of X, Y.

B. System Model

A transmitter wishes to communicate two independent messages W_1 and W_2 (of nR_1, nR_2 packets, respectively) to two receivers Rx_1 and Rx_2 over n channel uses. Communication takes place over a packet erasure broadcast channel with memory and feedback as described below:

The input to the broadcast channel at time t, $t = 1, \ldots, n$, is denoted by $X_t \in \mathcal{X}$. The channel inputs correspond to packets of L bits; we may represent this by choosing $\mathcal{X} = \mathbb{F}_q$ with $q = 2^L$, and $L \gg 1$. Transmission rates are measured in packets per slot and so entropies and mutual information terms are considered with logarithms to the base q.

The channel outputs at time t are written as $Y_{1,t} \in \mathcal{Y}$ and $Y_{2,t} \in \mathcal{Y}$ where $\mathcal{Y} = \mathcal{X} \cup \{E\}$. Each $Y_{j,t}$, $j \in \{1, 2\}$, is either X_t (i.e., received perfectly) or E (i.e., erased).

We define binary random variables $Z_{j,t}$, $j \in \{1, 2\}$, $t = 1, \ldots, n$, to indicate if an erasure occurred at receiver j in time t; i.e., $Z_{j,t} = \mathbb{1}\{Y_{j,t} = E\}$. Clearly, $Y_{j,t}$ can be expressed as a function of X_t and $Z_{j,t}$. Furthermore, $Y_{j,t}$ also determines $Z_{j,t}$. We denote $(Z_{1,t}, Z_{2,t})$ by $\underline{Z}_t$.

The broadcast channel we study has memory that is modeled via a finite state machine with state S_t at time t. The state evolves according to an irreducible aperiodic finite state Markov chain with state space $\mathcal{S}$ and steady-state distribution π_s, $s \in \mathcal{S}$. The initial state S_0 is distributed according to π. Depending on the current random state of the channel, the channel erasure probabilities are specified through the conditional distribution $P_{\underline{Z}_t|S_t}$. Arbitrary correlation between $(Z_{1,t}, Z_{2,t})$ is permitted. The transition probabilities between channel states are known at the transmitter. Note that the sequence $\underline{Z}^n$ is correlated in time in general, hence the channel has memory.

After each transmission, an ACK/NACK feedback is available at the encoder from both receivers. Two possible setups can be considered for the encoding function f_t:

(i) Only ACK/NACK feedback is available at the encoder:

$$X_t = f_t(W_1, W_2, Z_1^{t-1}, Z_2^{t-1}) \tag{1}$$

(ii) ACK/NACK and the previous state feedback is known:

$$X_t = f_t(W_1, W_2, Z_1^{t-1}, Z_2^{t-1}, S^{t-1}) \tag{2}$$

Depending on whether the transmitter knows the previous channel state or not, we call the state *visible* or *hidden*. This paper is focused on the problem with visible states (see Fig. 1). The joint probability mass function of the system then factorizes as

$$P_{W_1 W_2 X^n S^n Y_1^n Y_2^n Z_1^n Z_2^n} = P_{W_1} P_{W_2} P_{S_0} \prod_{t=1}^{n} P_{S_t|S_{t-1}} \cdot$$
$$P_{X_t|S^{t-1} Z_1^{t-1} Z_2^{t-1}} P_{Z_{2,t}|S_t} P_{Z_{1,t}|Z_{2,t} S_t} P_{Y_{1,t}|X_t Z_{1,t}} P_{Y_{2,t}|X_t Z_{2,t}}.$$

The corresponding functional dependency graph (FDG) [3] for the visible case is shown in Fig. 2.

The state can be visible either because it is explicitly available at the transmitter or because it may be determined from the available feedback. The latter is illustrated via the following example.

Example 1: Consider a Gilbert-Elliot model [22], [23] with state space $\mathcal{S} = \{\mathrm{GG}, \mathrm{GB}, \mathrm{BG}, \mathrm{BB}\}$ where G and B respectively refer to a good and bad state at each user.

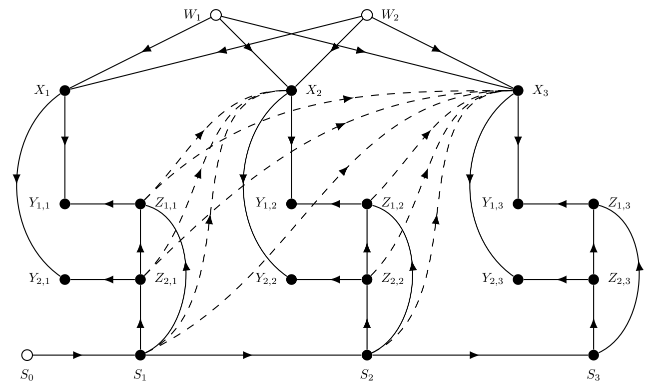

Fig. 2. FDG for the two-receiver broadcast packet erasure channel with memory and ACK/NACK + previous state feedback (visible state), for $n = 3$ and $d = 1$. Dependencies due to feedback are drawn with dashed lines.

One case of interest is when we have erasure in state B and no erasure in state G, i.e.,

$$\begin{aligned} P_{\underline{Z}_t|S_t}(0,0|\text{GG}) = 1, \quad P_{\underline{Z}_t|S_t}(0,1|\text{GB}) = 1 \\ P_{\underline{Z}_t|S_t}(1,0|\text{BG}) = 1, \quad P_{\underline{Z}_t|S_t}(1,1|\text{BB}) = 1. \end{aligned} \tag{3}$$

In such a channel the feedback $\underline{Z}_t$ determines the channel state, and we thus say that the state is visible. We use this channel model for our simulation results in Section VI.

We define the probability of erasure events given the *previous* channel state s as follows:

$$\begin{aligned} \epsilon_{12}(s) &= P_{\underline{Z}_t|S_{t-1}}(1,1|s), & \epsilon_{1\bar{2}}(s) &= P_{\underline{Z}_t|S_{t-1}}(1,0|s), \\ \epsilon_{\bar{1}2}(s) &= P_{\underline{Z}_t|S_{t-1}}(0,1|s), & \epsilon_{\bar{1}\bar{2}}(s) &= P_{\underline{Z}_t|S_{t-1}}(0,0|s), \\ \epsilon_{1}(s) &= \epsilon_{12}(s) + \epsilon_{1\bar{2}}(s), & \epsilon_{2}(s) &= \epsilon_{12}(s) + \epsilon_{\bar{1}2}(s). \end{aligned} \tag{4}$$

Note that these probabilities do not depend on t in our setup.

The goal is to have each decoder Rx_j reliably estimate $\hat{W}_j = h_j(Y_j^n)$ from its received sequence. A rate-pair (R_1, R_2) is said to be achievable if the error probability $\Pr[\hat{W}_1 \neq W_1, \hat{W}_2 \neq W_2]$ can be made arbitrarily small as n gets large. The capacity region $\mathcal{C}_{\text{fb}}^{\text{mem}}$ is the convex closure of the achievable rate pairs.

III. Main Result

The main result of this paper is the following bounds on the capacity region of the two-user packet erasure broadcast channel with memory and ACK/NACK feedback.

Define $\bar{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ to be the closure of rate pairs (R_1, R_2) for which there exist variables x_s, y_s, $s \in \mathcal{S}$ such that

$$0 \le x_s \le 1, \quad 0 \le y_s \le 1 \tag{5}$$

$$R_1 \le \sum_{s \in \mathcal{S}} \pi_s (1 - \epsilon_1(s)) x_s \tag{6}$$

$$R_1 \le \sum_{s \in \mathcal{S}} \pi_s (1 - \epsilon_{12}(s))(1 - y_s) \tag{7}$$

$$R_2 \le \sum_{s \in \mathcal{S}} \pi_s (1 - \epsilon_2(s)) y_s \tag{8}$$

$$R_2 \le \sum_{s \in \mathcal{S}} \pi_s (1 - \epsilon_{12}(s))(1 - x_s). \tag{9}$$

Define, furthermore, $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ to be the closure of rate pairs (R_1, R_2) for which there exist variables x_s, y_s, $s \in \mathcal{S}$ such that

$$0 \le x_s \le 1, \quad 0 \le y_s \le 1 \tag{10}$$

$$x_s + y_s \ge 1, \ \forall s \in \mathcal{S} \tag{11}$$

$$R_1 \le \sum_{s \in \mathcal{S}} \pi_s (1 - \epsilon_1(s)) x_s \tag{12}$$

$$R_1 \le \sum_{s \in \mathcal{S}} \pi_s (1 - \epsilon_{12}(s))(1 - y_s) \tag{13}$$

$$R_2 \le \sum_{s \in \mathcal{S}} \pi_s (1 - \epsilon_2(s)) y_s \tag{14}$$

$$R_2 \le \sum_{s \in \mathcal{S}} \pi_s (1 - \epsilon_{12}(s))(1 - x_s). \tag{15}$$

Note that $\bar{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ and $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ differ in (11).

Theorem 1: The capacity region $\mathcal{C}_{\text{fb}}^{\text{mem}}$ of the two-user broadcast packet erasure channel with feedback and visible state is sandwiched between $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ and $\bar{\mathcal{C}}_{\text{fb}}^{\text{mem}}$; i.e,

$$\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}} \subseteq \mathcal{C}_{\text{fb}}^{\text{mem}} \subseteq \bar{\mathcal{C}}_{\text{fb}}^{\text{mem}} \tag{16}$$

For example, consider Theorem 1 when $\mathcal{S}$ has one state only, say state s, which models a memoryless erasure broadcast channel. One may verify that the two regions $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ and $\bar{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ match, and that eliminating variables x_s, y_s, the well-known result of [6] follows. Let us call this capacity region $\mathcal{C}_{\text{fb}}(s)$. Now consider the case where $|\mathcal{S}|$ is larger: One might guess that the capacity region $\mathcal{C}_{\text{fb}}^{\text{mem}}$ is the average direct sum (set sum) of the capacity regions $\mathcal{C}_{\text{fb}}(s)$ over all states $s \in \mathcal{S}$. However, this is *not* in general the case: the capacity region can be strictly larger than the average direct sum of the $\mathcal{C}_{\text{fb}}(s)$. We expand on this remark in Section VI.

In Section IV, we prove that $\bar{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ forms an outer bound; i.e., for any achievable scheme the problem defined in (5)-(9) is feasible. This is done by bounding the achievable rates R_1, R_2 and expressing them in a manner similar to (5)-(9). Our converse proof is motivated by [10], [1], [24].

In Section V, we introduce two schemes that can achieve any rate-pair in $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$, and thus prove achievability of it. The first scheme is a probabilistic scheme that chooses encoding operations according to a probability distribution. The second scheme uses a deterministic queue-length based algorithm that chooses encoding operations based on the feedback and the current buffer states. This scheme stabilizes all queues in the network for every rate pair in $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$.

While $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ and $\bar{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ match and characterize the capacity in several examples, there are interesting cases where $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ is strictly smaller than $\mathcal{C}_{\text{fb}}^{\text{mem}}$. One such example is given in [25, Sec. II.B]. In this example $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ turns out to be strictly smaller than the capacity region.

In Section VI, we plot our lower and upper bounds on $\mathcal{C}_{\text{fb}}^{\text{mem}}$ for a few examples and address the gain due to feedback and causal knowledge of the channel state. We furthermore discuss how this rate-region is strictly larger that the average direct sum of the $\mathcal{C}_{\text{fb}}(s)$. Finally, we discuss variations of the problem with delayed feedback.

IV. The Converse

In this section, we prove that $\bar{\mathcal{C}}^{\text{mem}}_{\text{fb}}$ is an outer bound on the capacity region. The general idea is to show that for any achievable scheme, there are parameters x_s, y_s, $s \in \mathcal{S}$, as in Theorem 1. We find these parameters by relating them to mutual information terms.

In order to bound R_1 and R_2, for any $\delta > 0$, we write the following multi-letter bounds and single-letterize them properly next. For $j \in \{1,2\}$, we define $\bar{j} \in \{1,2\}$ such that $\bar{j} \neq j$.

$$nR_j \le I(W_j; Y_j^n) + n\delta \tag{17}$$
$$nR_j \le I(W_j; Y_1^n Y_2^n | W_{\bar{j}}) + n\delta \tag{18}$$

In (17)-(18), we have used the independence of the messages and Fano's inequality [26, Chapter 2.10].

For $j = 1$, the single-letterization is done as follows:

$$\begin{aligned}
R_1 - \delta &\le \frac{1}{n} I(W_1; Y_1^n) \\
&\le \frac{1}{n} I(W_1; Y_1^n S^{n-1}) \\
&= \frac{1}{n}\sum_{t=1}^n I(W_1; Y_{1,t} S_{t-1} | Y_1^{t-1} S^{t-2}) \\
&= \frac{1}{n}\sum_{t=1}^n \big[I(W_1; S_{t-1} | Y_1^{t-1} Z_1^{t-1} S^{t-2}) \\
&\qquad + I(W_1; Y_{1,t} | Y_1^{t-1} S^{t-1}) \big] \\
&\overset{(a)}{=} \sum_{t=1}^n \frac{1}{n} I(W_1; Y_{1,t} | Y_1^{t-1} S^{t-1}) \\
&\le \sum_{t=1}^n \frac{1}{n} I(W_1 Y_1^{t-1} S^{t-1}; Y_{1,t} | S_{t-1}) \\
&\overset{(b)}{=} \sum_{t=1}^n \frac{1}{n} I(U_{1,t}; Y_{1,t} | S_{t-1}) \\
&\overset{(c)}{=} I(U_{1,T}; Y_{1,T} | S_{T-1} T) \\
&= \sum_{s\in\mathcal{S}} \pi_s I(U_{1,T}; Y_{1,T} | T, S_{T-1} = s).
\end{aligned} \tag{19}$$

In the above chain of inequalities, (a) follows because Z_1^{t-1} is a function of Y_1^{t-1} and because of the Markov chain

$$W_1 - Y_1^{t-1} Z_1^{t-1} S^{t-2} - S_{t-1},$$

(b) follows by defining $U_{1,t} = (W_1 Y_1^{t-1} S^{t-1})$, and
(c) follows by a standard random time sharing argument with time sharing random variable T. Similarly, one obtains

$$\begin{aligned}
R_1 - \delta &\le \frac{1}{n} I(W_1; Y_1^n Y_2^n | W_2) \\
&\le \sum_{s\in\mathcal{S}} \pi_s I(U_{1,T}; Y_{1,T} Y_{2,T} | U_{2,T} V_T T, S_{T-1} = s),
\end{aligned} \tag{20}$$

where $U_{2,T} = (W_2 Y_2^{T-1} S^{T-1})$ and $V_T = (Y_1^{T-1} Y_2^{T-1} S^{T-1})$.

By symmetry, we also have the following bounds:

$$R_2 - \delta \le \sum_{s\in\mathcal{S}} \pi_s I(U_{2,T}; Y_{2,T} | T, S_{T-1} = s) \tag{21}$$

$$R_2 - \delta \le \sum_{s\in\mathcal{S}} \pi_s I(U_{2,T}; Y_{1,T} Y_{2,T} | U_{1,T} V_T T, S_{T-1} = s) \tag{22}$$

Remark 1: Note that
(i) V_T is a function of $(U_{1,T} U_{2,T})$, and
(ii) $\underline{Z}_T - T S_{T-1} - U_{1,T} U_{2,T} V_T X_T$ forms a Markov chain.

The following lemma extends [24, Lemma 1] and is proven in [27].

Lemma 1: For every $s \in \mathcal{S}$ and $j \in \{1,2\}$, we have:

$$\begin{aligned}
&I(U_{j,T}; Y_{j,T} | T, S_{T-1} = s) \\
&\quad = (1 - \epsilon_j(s)) I(U_{j,T}; X_T | T, S_{T-1} = s),
\end{aligned} \tag{23}$$
$$\begin{aligned}
&I(U_{j,T}; Y_{1,T} Y_{2,T} | U_{\bar{j},T}, V_T T, S_{T-1} = s) \\
&\quad = (1 - \epsilon_{12}(s)) I(U_{j,T}; X_T | U_{\bar{j},T} V_T T, S_{T-1} = s).
\end{aligned} \tag{24}$$

We now replace the mutual information terms in (19) - (22) using Lemma 1 and define the following variables, for $j \in \{1,2\}$ and $s \in \mathcal{S}$.

$$u_s^{(j)} = I(U_{j,T}; X_T | T, S_{T-1} = s) \tag{25}$$
$$z_s^{(j)} = I(U_{j,T}; X_T | U_{\bar{j},T} V_T T, S_{T-1} = s) \tag{26}$$

We have

$$R_j - \delta \le \sum_{s\in\mathcal{S}} \pi_s (1 - \epsilon_j(s)) u_s^{(j)}, \quad j = 1,2, \tag{27}$$
$$R_j - \delta \le \sum_{s\in\mathcal{S}} \pi_s (1 - \epsilon_{12}(s)) z_s^{(j)}, \quad j = 1,2. \tag{28}$$

The following Lemma relates the parameters defined above and is proven in Appendix A.

Lemma 2: For every $j \in \{1,2\}$ and $s \in \mathcal{S}$, we have

$$u_s^{(j)} + z_s^{(\bar{j})} \le 1,$$

Combining the above results and letting δ go to zero, (R_1, R_2) can be achieved only if, for some variables $u_s^{(1)}$, $u_s^{(2)}$, $z_s^{(1)}$, $z_s^{(2)}$, the following inequalities hold:

$$0 \le u_s^{(j)}, z_s^{(j)} \le 1 \qquad \forall j \in \{1,2\},\ \forall s \in \mathcal{S} \tag{29}$$
$$u_s^{(j)} + z_s^{(\bar{j})} \le 1 \qquad \forall j \in \{1,2\},\ \forall s \in \mathcal{S} \tag{30}$$
$$R_j \le \sum_{s\in\mathcal{S}} \pi_s (1 - \epsilon_j(s)) u_s^{(j)} \qquad \forall j \in \{1,2\} \tag{31}$$
$$R_j \le \sum_{s\in\mathcal{S}} \pi_s (1 - \epsilon_{12}(s)) z_s^{(j)} \quad \forall j \in \{1,2\}. \tag{32}$$

The final step is to show that the above outer bound matches $\bar{\mathcal{C}}^{\text{mem}}_{\text{fb}}$ of Theorem 1. This is done by noting that inequality (30) can be made tight without changing the rate region. The equivalence of the two regions then becomes clear by setting $z_s^{(1)} = 1 - y_s$, $z_s^{(2)} = 1 - x_s$, $u_s^{(1)} = x_s$, and $u_s^{(2)} = y_s$.

V. Achievable Schemes

A. Queue and Flow Model

In this section we develop codes that achieve the rate region $\underline{\mathcal{C}}^{\text{mem}}_{\text{fb}}$. For this, we build on the idea of tracking packets that have been received at the wrong destination, as in [6], [7]. The transmitter has two buffers $Q_1^{(1)}$, $Q_1^{(2)}$ to store packets destined for Rx_1, Rx_2, respectively. We consider dynamic arrivals, where packets for Rx_1, Rx_2 arrive in each slot according to a Bernoulli process with probability R_1, R_2, respectively. An analysis for more general arrival processes is possible. The transmitter has two additional buffers $Q_2^{(1)}$ (resp. $Q_2^{(2)}$) for packets that have already been sent, but have been received only by Rx_2 (resp. Rx_1). Hence buffer $Q_2^{(1)}$ contains packets that are destined for Rx_1 and have been received at Rx_2 but not at Rx_1, and vice versa for $Q_2^{(2)}$. These queues are empty before transmission begins. Each user j, $j = 1, 2$, has a buffer $Q_3^{(j)}$ that collects desired packets. These buffers correspond to the system exit and are always empty. The networked queuing system is shown in Fig. 3.

Each receiver has an additional buffer (not depicted in Fig. 3) that collects packets not intended for it, i.e. packets for the other user. Note that packets in this buffer are either also present in $Q_2^{(1)}$, $Q_2^{(2)}$, or have left the system.

A packet for Rx_j will only traverse buffers with superscript j, i.e. $Q_1^{(j)}, Q_2^{(j)}$ or $Q_3^{(j)}$. In the following, slightly abusing notation, we use $Q_{l,t}^{(j)}$ to denote the number of packets stored in buffer $Q_l^{(j)}$ in time slot t. Obviously, $Q_{l,t}^{(j)} \in \mathcal{Q}$ with $\mathcal{Q} = \{0, 1, \ldots, \infty\}$. Define

$$\underline{Q}_t = \left(Q_{1,t}^{(1)}, Q_{2,t}^{(1)}, Q_{1,t}^{(2)}, Q_{2,t}^{(2)} \right) \in \mathcal{Q}^4. \tag{33}$$

Because $Q_3^{(1)} = Q_3^{(2)} = 0$ by definition, the vector $\underline{Q}_t$ determines the queue state at time t.

If both $Q_2^{(1)}$ and $Q_2^{(2)}$ are nonempty, the transmitter can send the XOR combination of these packets. If both users receive this coded packet, both can decode one desired packet and two packets per slot are delivered. In general, the transmitter can select his action A_t in slot t from the set of actions $\mathcal{A} = \{1, 2, 3\}$, where

1 corresponds to sending a packet for Rx_1 from $Q_1^{(1)}$,
2 corresponds to sending a packet for Rx_2 from $Q_1^{(2)}$,
3 corresponds to sending a coded packet.

Actions at time t are based on the *current* queue state $\underline{Q}_t$ and the *previous* channel state S_{t-1}.

Note that we permit actions from the action space $\mathcal{A}$ only. The corresponding stability region consists of all rate tuples (R_1, R_2) for which all queues in the network are strongly stable [28, Definition 3.1], i.e., if

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[Q_t] < \infty. \tag{34}$$

A network is strongly stable if all queues are strongly stable [28, Definition 3.2]. The algorithms developed in the following ensure network stability for rate pairs inside $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$. For the analysis methods from [28], [21] are used, adapted to the setup.

B. Probabilistic Scheme

Consider a strategy that bases decisions for actions only on the previous channel state S_{t-1}, but not on the queue state $\underline{Q}_t$. These strategies are called S-only algorithms in [21]. The decisions are random and independent from previous

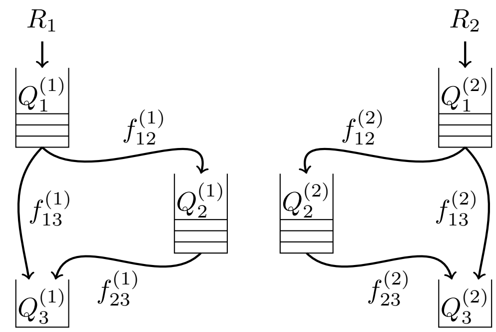

Fig. 3. Networked system of queues.

decisions, according to a probability distribution $P_{A_t|S_{t-1}}$ that does not depend on t.

Let $F_{lm,t}^{(j)}$ denote the number of packets that can travel from buffer $Q_l^{(j)}$ to $Q_m^{(j)}$ in slot t. Clearly, $F_{lm,t}^{(j)}$ depends on the action chosen in slot t. Recall that $Z_{j,t}$ is equal to one if an erasure occurs at time t for Rx_j and is zero otherwise, so that

$$F_{12,t}^{(1)} = \mathbb{1}\{A_t = 1\} Z_{1,t} (1 - Z_{2,t}). \tag{35}$$

The long-term average rate $f_{12}^{(1)}$ is bounded by

$$f_{12}^{(1)} \le \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} F_{12,t}^{(1)} = \mathbb{E}[F_{12,t}^{(1)}], \tag{36}$$

where the expectation in (36) is taken over the random previous channel state S_{t-1}, the random erasure events and the possibly random action A_t. Equality in (36) is achieved if $Q_{1,t}^{(1)} > 0$ whenever $A_t = 1$. Similarly, we have

$$f_{13}^{(1)} \le \mathbb{E}[F_{13,t}^{(1)}], \quad F_{13,t}^{(1)} = \mathbb{1}\{A_t = 1\}(1 - Z_{1,t}) \tag{37}$$
$$f_{23}^{(1)} \le \mathbb{E}[F_{23,t}^{(1)}], \quad F_{23,t}^{(1)} = \mathbb{1}\{A_t = 3\}(1 - Z_{1,t}) \tag{38}$$

and correspondingly for the flows to Rx_2.

Thus, with this scheme, rate tuples (R_1, R_2) can be achieved if there is a distribution $P_{A_t|S_{t-1}}$ such that

$$R_j \le f_{13}^{(j)} + f_{12}^{(j)} \tag{39}$$
$$f_{12}^{(j)} \le f_{23}^{(j)} \tag{40}$$
$$f_{12}^{(j)} \le \sum_{s\in\mathcal{S}} \pi_s P_{A_t|S_{t-1}}(j|s)(\epsilon_j(s) - \epsilon_{12}(s)) \tag{41}$$
$$f_{13}^{(j)} \le \sum_{s\in\mathcal{S}} \pi_s P_{A_t|S_{t-1}}(j|s)(1 - \epsilon_j(s)) \tag{42}$$
$$f_{23}^{(j)} \le \sum_{s\in\mathcal{S}} \pi_s P_{A_t|S_{t-1}}(3|s)(1 - \epsilon_j(s)) \tag{43}$$
$$\forall j \in \{1, 2\}.$$

Note that the region described by (39) - (43) is equivalent to the rate region $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$ described in (12) - (15). This may be seen by setting $P_{A_t|S_{t-1}}(1|s) = 1 - y_s$, $P_{A_t|S_{t-1}}(2|s) = 1 - x_s$, $P_{A_t|S_{t-1}}(3|s) = x_s + y_s - 1$ and eliminating the flow variables $f_{lm}^{(j)}$. Whereas (39) - (43) is a *maximum flow* formulation, (12) - (15) describes the dual *minimum cut* formulation. Note that inequality (11) ensures that $P_{A_t|S_{t-1}}(3|s) \ge 0$. This inequality is implicitly required in this approach but does not appear in the outer bound $\bar{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$.

Action A_t	Weight depending on $\underline{Q}_t$ and $S_{t-1} = s$
1	$[1-\epsilon_1(s)]Q_1^{(1)} + \epsilon_{1\bar{2}}(s)(Q_1^{(1)} - Q_2^{(1)})$
2	$[1-\epsilon_2(s)]Q_1^{(2)} + \epsilon_{\bar{1}2}(s)(Q_1^{(2)} - Q_2^{(2)})$
3	$[1-\epsilon_1(s)]Q_2^{(1)} + [1-\epsilon_2(s)]Q_2^{(2)}$

TABLE I
DETERMINISTIC SCHEME.

C. Deterministic Scheme

In the probabilistic scheme, actions are chosen depending on the channel state, so it can happen that there is no packet to transmit because the corresponding buffer is empty. This can be avoided by a max-weight backpressure-like algorithm [29], [30] basing its actions on both queue and channel states.

In each slot t, weights for each action are computed. These weights depend on the current queue state $\underline{Q}_t$ and the previous channel state $S_{t-1} = s$ and are shown in Table I. The strategy executes the action with highest weight in each slot.

Note that the rule in Table I ensures that actions are chosen only if the corresponding queues contain packets.

Using tools from Lyapunov stability [29], [31], [28], [21] one can show that this rule stabilizes all queues for every rate pair $(R_1+\delta, R_2+\delta) \in \underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$, $\delta > 0$. The detailed proof is omitted due to lack of space and can be found in [27].

The proof uses the T-slot drift similar to [21] but has to be adapted to take into account only previous channel states instead of the current channel. This difference changes the proof and the corresponding max-weight policy. In the model of [32], the authors deal with correlated channels but have the *current* channel state (or an estimate of it) available for the current decision. Similarly, in [31], [33], the current channel state is available at the transmitter. In [34], [35] the authors focus on obtaining channel state information in a scenario that is related to the case of hidden states, however without permitting coding operations. Similarly, [36] investigates the case of delayed channel state information for general networks, without permitting coding operations. During the preparation of this work we were informed that a similar approach was analyzed in [25] in a parallel line of work. More powerful coding actions are permitted in [25] that allow to close the gap to the outer bound.

VI. Discussion

For the discussion and numerical results in this section we use the Gilbert-Elliot model of Example 1. We assume that the individual channels to users 1 and 2 are both Gilbert-Elliot channels with states G and B. The broadcast channel state space is therefore given by $\mathcal{S} = \{\text{GG}, \text{GB}, \text{BG}, \text{BB}\}$ where G and B respectively refer to a good and bad state at each user. Transitions from state B to state G occur with probability g_j for user j, $j = 1, 2$. Similarly, a transition from state G to state B occurs with probability b_j for user j. For simplicity these transitions are assumed to be independent across the two users. The corresponding finite state Markov chain is summarized in Fig. 4. The average (long-term) erasure probability at user

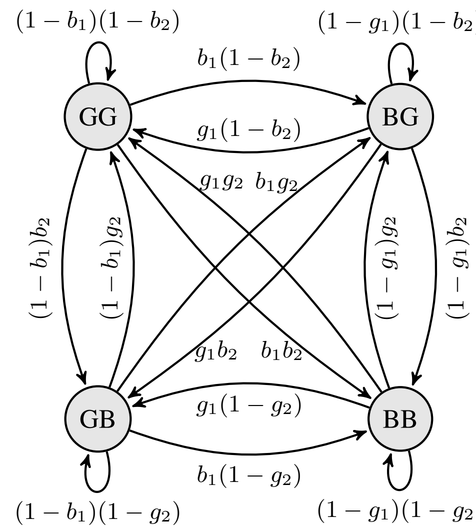

Fig. 4. Markov Chain of channel state space $\mathcal{S}$ with transition probabilities.

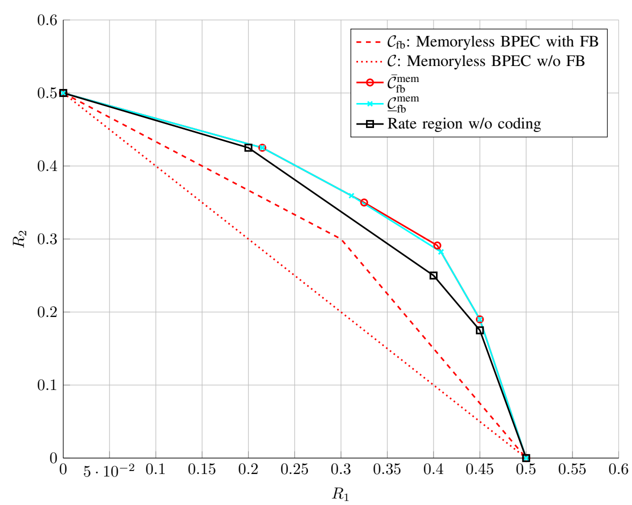

Fig. 5. Bounds on the Capacity region for $\epsilon_1 = 0.5$, $\epsilon_2 = 0.5$, $g_1 = 0.2$, $g_2 = 0.3$. In this case $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ and $\bar{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ almost match.

user j is given by

$$\epsilon_j = \frac{b_j}{g_j + b_j}. \tag{44}$$

Fig. 5 shows the capacity region for a channel with parameters $\epsilon_1 = 0.5$, $\epsilon_2 = 0.5$, $g_1 = 0.2$, $g_2 = 0.3$. In this figure we compare the bounds on the capacity region with that of a memoryless channel with the same average erasure probability (with and without feedback). For comparison we also show the rate region achieved without permitting coding between users as defined in [19].

A. Combination of Memoryless Strategies

Looking at the the characterization of $\underline{\mathcal{C}}_{\text{fb}}^{\text{mem}}$ in Theorem 1, one may wonder if this rate-region can be attained simply by a combination of memoryless capacity achieving schemes. Let $\mathcal{C}_{\text{fb}}(s)$, $s \in \mathcal{S}$, denote the capacity region of a memoryless BPEC with feedback and erasure probabilities $\Pr[\underline{Z}_t | S_{t-1} = s]$. Capacity achieving algorithms for memoryless BPEC with

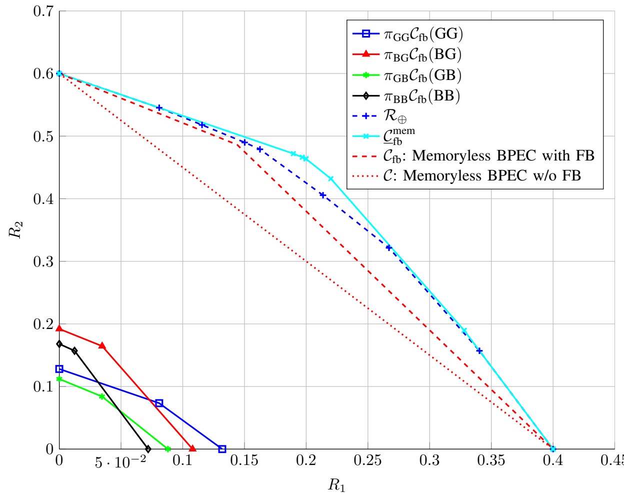

Fig. 6. Individual rate regions and Minkowski sum $\mathcal{R}_{\oplus}$ for $\epsilon_1 = 0.6$, $\epsilon_2 = 0.4$, $g_1 = 0.3$, $g_2 = 0.7$. The region $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$ is strictly larger. For comparison, the corresponding capacity regions for memoryless channels with the same average erasure probability are shown for the cases with and without feedback. The difference between $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$ and $\bar{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$ is not visible in this case, so $\bar{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$ is omitted.

feedback are devised in [6]. A combination of memoryless capacity achieving schemes may be described as follows:

- Choose fractions $\alpha_s \geq 0$ and $\beta_s \geq 0$ such that $\sum_{s\in\mathcal{S}} \alpha_s = \sum_{s\in\mathcal{S}} \beta_s = 1$ and $(\alpha_s R_1, \beta_s R_2) \in \pi_s \mathcal{C}_{\mathrm{fb}}(s)$, for all $s \in \mathcal{S}$.
- Take $n\alpha_s R_1$ packets for Rx_1 and $n\beta_s R_2$ packets for Rx_2 to be transmitted only when the previous channel state is equal to $S_{t-1} = s$, $s \in \mathcal{S}$. For each previous state $s \in \mathcal{S}$, the transmitter chooses an optimal memoryless strategy (e.g., as devised in [6]) corresponding to a memoryless BPEC channel with feedback and erasure probabilities $\Pr[\underline{Z}_t | S_{t-1} = s]$.

Using the above scheme, for large n, one can asymptotically achieve the performance of the memoryless strategy for each state s with the corresponding capacity region $\mathcal{C}_{\mathrm{fb}}(s)$. The overall rate region achievable by this strategy, called $\mathcal{R}_{\oplus}$, is thus a weighted combination of the individual memoryless rate regions (for each state s):

$$\mathcal{R}_{\oplus} = \bigoplus_{s\in\mathcal{S}} \pi_s \mathcal{C}_{\mathrm{fb}}(s), \tag{45}$$

where $\oplus$ denotes the set addition operator[1] (Minkowski sum).

We show in Fig. 6 that $\mathcal{R}_{\oplus}$ can be strictly smaller than $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$.

Remark 2: Note that each memoryless rate region $\mathcal{C}_{\mathrm{fb}}(s)$, $s \in \mathcal{S}$, is a polytope defined by linear inequalities. However, the polytope generated by the Minkowski sum is *not* equal to the one defined by the sum of the individual polytope constraints. That would be the case, for example, if the memoryless rate regions were polymatroids, as pointed out in [26, Chapter 15.3.3],[37]. In that case, $\mathcal{R}_{\oplus}$ would be equal to $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$. However, this is not the case in general.

[1] For example, $\pi_1\mathcal{R}_1 \oplus \pi_2\mathcal{R}_2 = \{\pi_1 \underline{r}_1 + \pi_2 \underline{r}_2 | \underline{r}_1 \in \mathcal{R}_1, r_2 \in \mathcal{R}_2\}$.

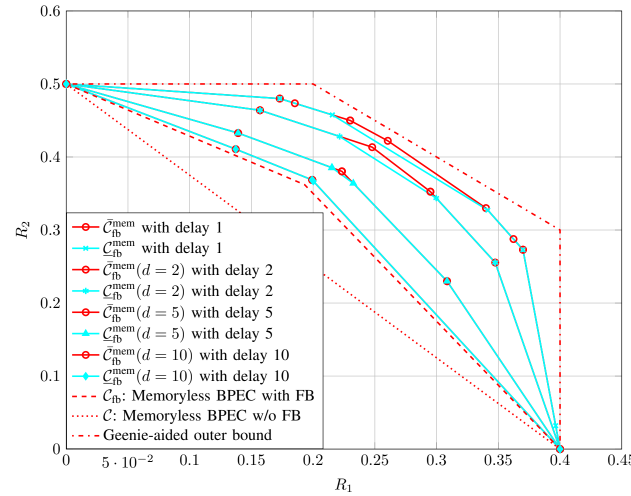

Fig. 7. Capacity regions with delayed feedback, for $\epsilon_1 = 0.6$, $g_1 = 0.1$, $\epsilon_2 = 0.5$, $g_2 = 0.1$. The genie-aided outer bound corresponds to the case when all erasures are known ahead of time.

B. Delayed Feedback

The result in Theorem 1 extends to the scenario where feedback and channel state become available at the encoder with more than a single symbol-time delay. Consider a delay of d time units and call the achievable rate region $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}(d)$. In the converse, one can obtain the corresponding bounds by replacing the sequences S^{T-1}, Y_1^{T-1} Z_1^{T-1}, Y_2^{T-1} and Z_2^{T-1} with S^{T-d}, Y_1^{T-d} Z_1^{T-d}, Y_2^{T-d} and Z_2^{T-d}.

The bounds on the capacity region $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}(d)$ and $\bar{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}(d)$ have thus a characterization as in Theorem 1 by redefining the erasure probabilities in (4) as

$$\begin{aligned}
\epsilon_{12}(s) &= P_{\underline{Z}_t|S_{t-d}}(1,1|s), & \epsilon_{1\bar{2}}(s) &= P_{\underline{Z}_t|S_{t-d}}(1,0|s),\\
\epsilon_{\bar{1}2}(s) &= P_{\underline{Z}_t|S_{t-d}}(0,1|s), & \epsilon_{\bar{1}\bar{2}}(s) &= P_{\underline{Z}_t|S_{t-d}}(0,0|s),\\
\epsilon_1(s) &= \epsilon_{12}(s) + \epsilon_{1\bar{2}}(s), & \epsilon_2(s) &= \epsilon_{12}(s) + \epsilon_{\bar{1}2}(s).
\end{aligned}$$

The corresponding deterministic achievable scheme as in Section V-C uses these redefined conditional erasure probabilities to obtain the same description as in Table I.

Fig. 7 shows the effect of feedback delay for a Gilbert-Elliot channel with parameters $\epsilon_1 = 0.6$, $g_1 = 0.1$, $\epsilon_2 = 0.5$, $g_2 = 0.1$. Two observations can be made: First, obviously, delayed feedback shrinks both the outer and inner bounds, as state information becomes less useful. After a feedback delay of $d = 10$ time units, the region $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}(d = 10)$ is almost the same as for the memoryless case for this example. In general this depends on the convergence speed of the state Markov chain towards its stationary distribution. Second, the more the memoryless capacity region is appraoched, the smaller is the difference between $\underline{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$ and $\bar{\mathcal{C}}_{\mathrm{fb}}^{\mathrm{mem}}$.

VII. Conclusion

We investigated the two-user broadcast packet erasure channel with feedback and memory. We modelled the channel memory by a finite state machine and found outer and inner bounds on the capacity region when the channel state is known strictly causally at the encoder. For the inner bound we proposed a probabilistic achievable scheme and presented a deterministic queue-length based algorithm. The results are extended to feedback with larger delay. Numerical results show that the gains offered through feedback can be quite large and that the difference between the outer and inner bound is small. One future direction is to determine when inner and outer bounds meet.

ACKNOWLEDGMENTS

The authors are supported by the German Ministry of Education and Research in the framework of the Alexander von Humboldt-Professorship and by the grant DLR@Uni of the Helmholtz Allianz. The work of S. Saeedi Bidokhti was partially supported by the Swiss National Science Foundation Fellowship no. 146617. The authors would like to thank Gianluigi Liva and Andrea Munari for motivating the problem and Gerhard Kramer for his helpful comments.

APPENDIX A
PROOF OF LEMMA 2

First note that $I(U_{1,T}U_{2,T}V_T; X_T|T, S_{T-1} = s) \leq 1$ because $H(X_T|T, S_{T-1} = s) \leq 1$. Applying the chain rule, we obtain

$$\begin{aligned}1 \geq& I(U_{1,T}U_{2,T}V_T; X_T|T, S_{T-1} = s)\\ =& I(U_{1,T}; X_T|T, S_{T-1} = s)\\ &+ I(V_T; X_T|U_{1,T}T, S_{T-1} = s)\\ &+ I(U_{2,T}; X_T|U_{1,T}V_TT, S_{T-1} = s)\\ \geq& u_s^{(1)} + I(V_T; X_T|U_{1,T}T, S_{T-1} = s) + z_s^{(2)}\\ \geq& u_s^{(1)} + z_s^{(2)}. \end{aligned} \quad (46)$$

Similarly, $u_s^{(2)} + z_s^{(1)} \leq 1$.

REFERENCES

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

[2] G. Dueck, "Partial feedback for two-way and broadcast channels," *Information and Control*, vol. 46, no. 1, pp. 1–15, 1980.

[3] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 4–21, 2003.

[4] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 667–671, Jul 1984.

[5] S. Bhaskaran, "Gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5252–5257, Nov 2008.

[6] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback-capacity and algorithms," in *Netcod*, 2009.

[7] 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, Sept 2013.

[8] M. Gatzianas, S. Saeedi Bidokhti, and C. Fragouli, "Feedback-based coding algorithms for broadcast erasure channels with degraded message sets," in *Netcod*, 2012.

[9] 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.

[10] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 197–207, 1973.

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

[12] C.-C. Wang and J. Han, "The capacity region of two-receiver multiple-input broadcast packet erasure channels with channel output feedback," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5597–5626, Sept 2014.

[13] 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.

[14] R. Venkataramanan and S. S. Pradhan, "Achievable rates for the broadcast channel with feedback," in *IEEE Int. Symp. Inf. Theory*, 2010.

[15] E. Lutz, D. Cygan, M. Dippold, F. Dolainsky, and W. Papke, "The land mobile satellite communication channel-recording, statistics, and channel model," *IEEE Trans. Vehicular Technology*, vol. 40, no. 2, pp. 375–386, 1991.

[16] F. P. Fontán, M. Vázquez-Castro, C. E. Cabado, J. P. Garcia, and E. Kubista, "Statistical modeling of the LMS channel," *IEEE Trans. Vehicular Technology*, vol. 50, no. 6, pp. 1549–1567, 2001.

[17] M. Ibnkahla, Q. M. Rahman, A. I. Sulyman, H. A. Al-Asady, J. Yuan, and A. Safwat, "High-speed satellite mobile communications: technologies and challenges," *Proc. IEEE*, vol. 92, no. 2, pp. 312–339, 2004.

[18] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state markov modeling of fading channels-a survey of principles and applications," *IEEE Signal Proc. Mag.*, vol. 25, no. 5, pp. 57–80, 2008.

[19] M. Heindlmaier and C. Blöchl, "The two-user broadcast packet erasure channel with feedback and memory," in *Netcod*, 2014.

[20] R. Dabora and A. J. Goldsmith, "Capacity theorems for discrete, finite-state broadcast channels with feedback and unidirectional receiver cooperation," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 5958–5983, 2010.

[21] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.

[22] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Labs Techn. J.*, vol. 39, no. 5, pp. 1253–1265, 1960.

[23] E. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Labs Techn. J.*, vol. 42, no. 5, pp. 1977–1997, 1963.

[24] A. Dana and B. Hassibi, "The capacity region of multiple input erasure broadcast channels," in *IEEE Int. Symp. Inf. Theory*, 2005.

[25] W.-C. Kuo and C.-C. Wang, "Robust and optimal opportunistic scheduling for downlink 2-flow inter-session network coding with varying channel quality," in *IEEE INFOCOM*, April 2014, pp. 655–663.

[26] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition*. Wiley-Interscience, 2006.

[27] M. Heindlmaier, N. Reyhanian, and S. Saeedi Bidokhti, "On capacity regions of two-receiver broadcast packet erasure channels with feedback and memory," *available at arXiv*, 2014.

[28] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource allocation and cross-layer control in wireless networks*. Now Publishers, 2006.

[29] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Automatic Control*, vol. 37, no. 12, pp. 1936–1948, 1992.

[30] ——, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 466–478, 1993.

[31] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *IEEE Journal on Selected Areas in Commun.*, vol. 23, no. 1, pp. 89–103, 2005.

[32] A. Pantelidou, A. Ephremides, and A. L. Tits, "A cross-layer approach for stable throughput maximization under channel state uncertainty," *Wireless Networks*, vol. 15, no. 5, pp. 555–569, 2009.

[33] L. Tassiulas, "Scheduling and performance limits of networks with constantly changing topology," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 1067–1073, 1997.

[34] C.-p. Li and M. J. Neely, "Exploiting channel memory for multi-user wireless scheduling without channel measurement: Capacity regions and algorithms," *Performance Evaluation*, 2011.

[35] ——, "Network utility maximization over partially observable markovian channels," *Performance Evaluation*, vol. 70, no. 7, pp. 528–548, 2013.

[36] L. Ying and S. Shakkottai, "On throughput optimality with delayed network-state information," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5116–5132, 2011.

[37] D. Traskov, M. Heindlmaier, M. Médard, and R. Koetter, "Scheduling for network-coded multicast," *IEEE/ACM Trans. Networking*, vol. 20, no. 5, pp. 1479–1488, 2012.