Coding Schemes for Discrete Memoryless Broadcast Channels with Rate-Limited Feedback

Youlong Wu and Michèle Wigger
Telecom ParisTech,
Paris, France
{youlong.wu, michele.wigger}@telecom-paristech.fr

***Abstract*—We propose two coding schemes for discrete memoryless broadcast channels (DMBCs) with rate-limited feedback. In our first scheme, the encoder does not process the feedback information that it receives, but simply relays it to the other receiver. This first scheme shows that arbitrary small, but positive, feedback rate suffices to improve over the nofeedback capacity for many DMBCs such as: any binary erasure BC (BEBC) with unequal erasure probability at the two receivers, any binary symmetric BC (BSBC) with unequal crossover probability at the receivers, and any binary erasure/binary symmetric BC (BEC/BSC-BC) with nonequal single-user capacity to the receivers. The scheme also improves the entire nofeedback capacity region for any *strictly essentially less-noisy BC*–a new class of BCs introduced in this paper–that is not physically degraded.**

In our second scheme, the encoder decodes all the feedback information and processes it with some local information before sending the result to the receivers. For some setups, this second scheme performs better than our first scheme. In the limit, as the available feedback-rates tend to infinity, our second scheme coincides with a special case of the Shayevitz and Wigger (SW) scheme for DMBCs with generalized feedback. The mentioned special case of the SW-scheme includes several other schemes as further special cases, e.g, the schemes by Dueck and by Maddah-Ali and Tse which achieve capacity or the degrees of freedom on the respectively studied channels.

All our results hold also with noisy feedback when the receivers can code over the feedback links.

I. Introduction

For most discrete memoryless broadcast channels (DMBC), it is not known whether feedback can increase the capacity region, even when the feedback links are noise-free and of infinite rate. There are some exceptions. For all physically degraded DMBCs the capacity regions with and without feedback coincide [1]. The first simple example DMBC where feedback increases capacity was presented by Dueck [2]. His example and coding scheme were generalized by Shayevitz and Wigger [3] who proposed a scheme and achievable region for DMBCs with generalized feedback. In the generalized-feedback model, the feedback to the transmitter is modeled as an additional output of the DMBC that can depend on the input and the receivers' outputs in an arbitrary manner. It has recently been shown [4] that the Shayevitz-Wigger scheme includes as special cases the two-user schemes by Wang [5], by Georgiadis and Tassiulas [6], and by Maddah-Ali and Tse [7], which achieve the capacity region and the degrees of freedom region of their respective channels. Other achievable regions for DMBCs with perfect or noisy feedback, have been proposed by Kramer [8] and Venkataramanan and Pradhan [9]. Kramer's achievable region also shows that feedback improves capacity for some specific *binary symmetric BC* (BSBC).

In this paper, we present two coding schemes for DMBCs with *rate-limited feedback*. Both schemes use a block-Markov strategy with backward decoding, and in each block they apply Marton's coding [10], which to date is the best known coding scheme without feedback. In both schemes, the messages sent over the feedback links are simply compression information that describe the previous channel outputs at the corresponding receiver. In our first scheme, the encoder transmits exactly these compression informations as part of the cloud center of Marton's code in the next block. Thus, here, the encoder only relays the feedback messages from one receiver to the other. Backward decoding is used where in each block each receiver first reconstructs a compressed version of the other receiver's outputs and then applies a modified Marton decoding to these compressed outputs and its own observed outputs. We modify the Marton decoding to account for the fact that each receiver already knows a part of the message sent in the cloud center (namely the compression information it had generated itself after the previous block).

As we will see, in our first scheme, each receiver can decode its intended messages equally well as if the part of the message known to it was not there. Thus in the cloud center we send information that is useful to one of the two receivers without disturbing the other receiver, or in other words, without occupying the other receiver's resources. In this sense, parts of the resources in the cloud center can serve two purposes at the same time.

The proposed scheme is particularly beneficial for the class of *strictly essentially less-noisy* DMBCs, which we define in this paper and which represents a subclass of Nair's essentially less-noisy DMBCs [11]. Our class includes the BSBC and the *binary erasure BC* (BEBC) with unequal crossover probabilities or unequal erasure probabilities at the receivers, and the *binary erasure channel/binary symmetric channel BC* (BEC/BSC-BC) for a large range of parameters. For strictly essentially less-noisy DMBCs Marton's coding achieves capacity [11], and except for physically degraded BCs, our first scheme improves strictly over the nofeedback capacity region no matter how small, but positive, the feedback rates. In fact, for most of these channels our scheme improves over all boundary points $(R_1 > 0, R_2 > 0)$ of the nofeedback

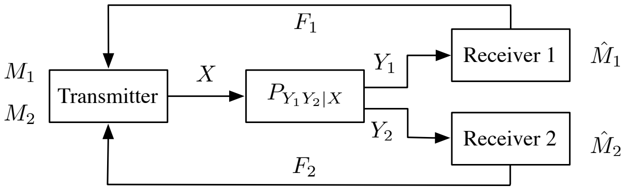

Fig. 1. Broadcast channel with rate-limited feedback

capacity region. The described scheme also improves over the nofeedback capacity region of the BEC/BSC-BC when the DMBC is more capable [12], unless the BEC and BSC have same capacities.

Unlike for previous schemes, with our new schemes we can thus easily show that feedback increases the capacity region for a large set of DMBCs. This holds even when the feedback links are rate-limited to arbitrary small, but positive, rates.

In our second scheme, the encoder first reconstructs the compressed versions of the channel outputs. It then uses these compressed signals together with the previously sent codewords to generate some update information, which it then sends to both receivers as part of the cloud center of Marton's code in the next-following block. For some setups this latter coding scheme performs better than the first scheme. For example, for Dueck's example DMBC, the second scheme can achieve the feedback capacity if the feedback rates are sufficiently large. As the feedback rates increase, the region achieved with this second scheme converges to a special case of the Shayevitz-Wigger region that includes the regions by Wang [5], by Georgiadis and Tassiulas [6], and by Maddah-Ali and Tse [7].

All our results hold also with noisy feedback when the receivers can code over the feedback links.

Notation: $Z \sim \text{Bern}(p)$ denotes that Z is a binary random variable taking values 0 and 1 with probabilities $1-p$ and p. Also, we use the definitions $\bar{a} := (1-a)$ and $a * b := \bar{a}b + a\bar{b}$, for $a, b \in [0,1]$. $H_b(\cdot)$ denotes the binary entropy function.

II. Channel model

Communication takes place over a DMBC with rate-limited feedback, see Figure 1. The setup is characterized by the finite input alphabet $\mathcal{X}$, the finite output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$, the channel law $P_{Y_1Y_2|X}$, and nonnegative feedback rates $R_{\text{Fb},1}$ and $R_{\text{Fb},2}$. Specifically, if at discrete-time t the transmitter sends the channel input $x_t \in \mathcal{X}$, then Receiver $i \in \{1,2\}$ observes the output $Y_{i,t} \in \mathcal{Y}_i$, where the pair $(Y_{1,t}, Y_{2,t}) \sim P_{Y_1Y_2|X}(\cdot,\cdot|x_t)$. Also, after observing $Y_{i,t}$, Receiver i can send a feedback signal $F_{i,t} \in \mathcal{F}_{i,t}$ to the transmitter, where $\mathcal{F}_{i,t}$ denotes the finite alphabet of $F_{i,t}$ and is a design parameter of a scheme. The feedback link between the transmitter and Receiver i is assumed to be instantaneous and noiseless but *rate-limited* to $R_{\text{Fb},i}$ bits on average. Thus, if the transmission takes place over a total blocklength N, then

$$|\mathcal{F}_{i,1}| \times \cdots \times |\mathcal{F}_{i,N}| \leq 2^{NR_{\text{Fb},i}}, \quad i \in \{1,2\}. \tag{1a}$$

The goal of the communication is that the transmitter conveys two independent private messages $M_1 \in \{1, \ldots, \lfloor 2^{NR_1} \rfloor\}$ and $M_2 \in \{1, \ldots, \lfloor 2^{NR_2} \rfloor\}$, to Receiver 1 and 2, respectively. Each M_i, $i \in \{1,2\}$, is uniformly distributed over the set $\mathcal{M}_i := \{1, \ldots, \lfloor 2^{NR_i} \rfloor\}$, where R_i denotes the private rate of transmission of Receiver i.

The transmitter is comprised of a sequence of encoding functions $\{f_t^{(N)}\}_{t=1}^N$ of the form $f_t^{(N)}: \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{F}_{1,1} \times \cdots \times \mathcal{F}_{1,t-1} \times \mathcal{F}_{2,1} \times \cdots \times \mathcal{F}_{2,t-1} \to \mathcal{X}$ that is used to produce the channel inputs as

$$X_t = f_t^{(N)}\big(M_1, M_2, F_1^{t-1}, F_2^{t-1}\big), \qquad t \in \{1, \ldots, N\}. \tag{2}$$

Receiver $i \in \{1,2\}$ is comprised of a sequence of feedback-encoding functions $\{\psi_{i,t}^{(N)}\}_{t=1}^N$ of the form $\psi_{i,t}^{(N)}: \mathcal{Y}_i^t \to \mathcal{F}_{i,t}$ that is used to produce the symbols

$$F_{i,t} = \psi_{i,t}^{(N)}(Y_{i,1}, \ldots, Y_{i,t}), \qquad t \in \{1, \ldots, N\}, \tag{3}$$

sent over the feedback link, and of a decoding function $\Phi_i^{(N)}: \mathcal{Y}_i^N \to \mathcal{M}_i$ used to produce a guess of Message M_i:

$$\hat{M}_i = \Phi_i^{(N)}(Y_i^N). \tag{4}$$

A rate region (R_1, R_2) with averaged feedback rates $R_{\text{Fb},1}$, $R_{\text{Fb},2}$ is called achievable if for every blocklength N, there exists a set encoding functions $\{f_t^{(N)}\}_{t=1}^N$ and for $i = \{1,2\}$ there exists a set of decoding functions $\Phi_i^{(N)}$, feedback alphabets $\{\mathcal{F}_{i,t}\}_{t=1}^N$ satisfying (1), and feedback-encoding functions $\{\psi_{i,t}^{(N)}\}_{t=1}^N$ such that the error probability

$$P_e^{(N)} := \Pr(M_1 \neq \hat{M}_1 \text{ or } M_2 \neq \hat{M}_2) \tag{5}$$

tends to zero as the blocklength N tends to infinity. The closure of the set of achievable rate pairs (R_1, R_2) is called the *feedback capacity region* and is denoted by $\mathcal{C}_{\text{Fb}}$.

In the special case $R_{\text{Fb},1} = R_{\text{Fb},2} = 0$ the feedback signals are constant and the setup is equivalent to a setup without feedback. We denote the capacity region for this setup $\mathcal{C}_{\text{NoFB}}$.

III. Preliminaries

We recall some previous results for DMBCs. The best inner bound *without feedback* is Marton's region [10], $\mathcal{R}_{\text{Marton}}$, which equals the set of all nonnegative rarte-pairs (R_1, R_2) that satisfy

$$R_1 \leq I(U_0, U_1; Y_1) \tag{6a}$$
$$R_2 \leq I(U_0, U_2; Y_2) \tag{6b}$$
$$R_1 + R_2 \leq I(U_0, U_1; Y_1) + I(U_2; Y_2|U_0) - I(U_1; U_2|U_0) \tag{6c}$$
$$R_1 + R_2 \leq I(U_0, U_2; Y_2) + I(U_1; Y_1|U_0) - I(U_1; U_2|U_0) \tag{6d}$$

for some probability mass function (pmf) $P_{U_0U_1U_2}$ and a function $f: \mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \to \mathcal{X}$ such that $X = f(U_0, U_1, U_2)$.

The best known outer bound without feedback [13] is the set of all nonnegative rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(U; Y_1) \tag{7a}$$
$$R_2 \leq I(V; Y_2) \tag{7b}$$
$$R_1 + R_2 \leq I(U; Y_1) + I(X; Y_2|U) \tag{7c}$$
$$R_1 + R_2 \leq I(V; Y_2) + I(X; Y_1|V) \tag{7d}$$

for some pmf P_{UVX}.

We recall the definition of *essentially less-noisy BCs* [11]. A subset $\mathcal{P}_{\mathcal{X}}$ of all pmfs on the input alphabet $\mathcal{X}$ is said to be a *sufficient class* of pmfs for a DMBC if the following holds: Given any joint pmf P_{UVX} there exists a joint pmf P'_{UVX} that satisfies $P'_X(x) \in \mathcal{P}_{\mathcal{X}}$ and

$$\begin{aligned} I_P(U;Y_1) &\le I_{P'}(U;Y_1) \\ I_P(V;Y_2) &\le I_{P'}(V;Y_2) \\ I_P(U;Y_1)+I_P(X;Y_2|U) &\le I_{P'}(U;Y_1)+I_{P'}(X;Y_2|U) \\ I_P(V;Y_2)+I_P(X;Y_1|V) &\le I_{P'}(V;Y_2)+I_{P'}(X;Y_1|V) \end{aligned} \tag{8}$$

where the notations I_P and $I_{P'}$ indicate that the mutual informations are computed assuming that $(U,V,X) \sim P_{UVX}$ and $(U,V,X) \sim P'_{UVX}$, respectively, and $P'_X(x)$ is the marginal obtained from P'_{UVX}. A DMBC is called *essentially less-noisy* if there exists a sufficient class of pmfs $\mathcal{P}_{\mathcal{X}}$ such that whenever $P_X \in \mathcal{P}_{\mathcal{X}}$, then for all conditional pmfs $P_{U|X}$,

$$I(U;Y_1) \le I(U;Y_2). \tag{9}$$

The class of essentially less-noisy DMBCs contains as special cases the BSBC and the BEBC. The capacity region of essentially less-noisy DMBCs equals the *superposition coding* region, which is a special case of Marton's region that results when in (6) we set U_1 =const. and $X = U_2$.

The region $\mathcal{C}^{(1)}_{\text{Enh}}$, given by the set of all nonnegative rate-pairs (R_1,R_2) that for some pmf P_{UX} satisfy

$$R_1 \le I(U;Y_1) \tag{10a}$$
$$R_2 \le I(X;Y_1,Y_2|U), \tag{10b}$$

forms an outer bound to the capacity region with feedback [14]. An analogous outer bound, $\mathcal{C}^{(2)}_{\text{Enh}}$, is obtained by exchanging indices 1 and 2 in the definition above.

IV. Results

The schemes achieving Thoerems 1 and 3 are sketched in Section V. See [15] for details and for proofs of our results.

Theorem 1. *The capacity region $\mathcal{C}_{\text{Fb}}$ includes the set $\mathcal{R}_{\text{relay}}$ of all nonnegative rate-pairs (R_1,R_2) that satisfy*

$$R_1 \le I(U_0,U_1;Y_1,\tilde{Y}_2|Q) - I(\tilde{Y}_2;Y_2|Y_1,Q) \tag{11a}$$
$$R_2 \le I(U_0,U_2;Y_2,\tilde{Y}_1|Q) - I(\tilde{Y}_1;Y_1|Y_2,Q) \tag{11b}$$
$$\begin{aligned} R_1+R_2 \le\ & I(U_0,U_1;Y_1,\tilde{Y}_2|Q) - I(\tilde{Y}_2;Y_2|Y_1,Q) \\ & +I(U_2;Y_2,\tilde{Y}_1|U_0,Q) - I(U_1;U_2|U_0) \end{aligned} \tag{11c}$$
$$\begin{aligned} R_1+R_2 \le\ & I(U_0,U_2;Y_2,\tilde{Y}_1|Q) - I(\tilde{Y}_1;Y_1|Y_2,Q) \\ & +I(U_1;Y_1,\tilde{Y}_2|U_0,Q) - I(U_1;U_2|U_0) \end{aligned} \tag{11d}$$
$$\begin{aligned} R_1+R_2 \le\ & I(U_0,U_1;Y_1,\tilde{Y}_2|Q) - I(\tilde{Y}_2;Y_2|Y_1,Q) \\ & +I(U_0,U_2;Y_2,\tilde{Y}_1|Q) - I(\tilde{Y}_1;Y_1|Y_2,Q) \\ & -I(U_1;U_2|U_0) \end{aligned} \tag{11e}$$

for some pmfs $P_Q, P_{U_0U_1U_2|Q}, P_{\tilde{Y}_1|Y_1Q}, P_{\tilde{Y}_2|Y_2Q}$ and some function $f\colon \mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{Q} \to \mathcal{X}$ that satisfy

$$I(\tilde{Y}_1;Y_1|Y_2,Q) \le R_{\text{Fb},1} \quad \text{and} \quad I(\tilde{Y}_2;Y_2|Y_1,Q) \le R_{\text{Fb},2} \tag{12}$$

where $X = f(U_0,U_1,U_2,Q)$.

For $\tilde{Y}_1 = \tilde{Y}_2$ =const., the region above specializes to $\mathcal{R}_{\text{Marton}}$. For $\tilde{Y}_2$ =const., it specializes to:

Corollary 1. *The capacity region $\mathcal{C}_{\text{Fb}}$ includes the set $\mathcal{R}^{(1)}_{\text{relay,s}}$ of all nonnegative rate pairs (R_1,R_2) that satisfy*

$$R_1 \le I(U_0,U_1;Y_1|Q) \tag{13a}$$
$$R_2 \le I(U_0,U_2;Y_2,\tilde{Y}_1|Q) - I(\tilde{Y}_1;Y_1|Y_2,Q) \tag{13b}$$
$$\begin{aligned} R_1+R_2 \le\ & I(U_0,U_1;Y_1|Q) + I(U_2;Y_2,\tilde{Y}_1|U_0,Q) \\ & -I(U_1;U_2|U_0,Q) \end{aligned} \tag{13c}$$
$$\begin{aligned} R_1+R_2 \le\ & I(U_1;Y_1|U_0,Q) + I(U_0,U_2;Y_2,\tilde{Y}_1|Q) \\ & -I(\tilde{Y}_1;Y_1|Y_2,Q) - I(U_1;U_2|U_0,Q) \end{aligned} \tag{13d}$$

for some pmfs $P_Q, P_{U_0U_1U_2|Q}, P_{\tilde{Y}_1|Y_1Q}$ and some function $f\colon \mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \mathcal{Q} \to \mathcal{X}$ that satisfy

$$I(\tilde{Y}_1;Y_1|Y_2,Q) \le R_{\text{Fb},1}, \tag{14}$$

where $X = f(U_0,U_1,U_2,Q)$.

It also includes the region $\mathcal{R}^{(2)}_{\text{relay,s}}$ which is obtained by exchanging indices 1 and 2 in the above definition of $\mathcal{R}^{(1)}_{\text{relay,s}}$.

In the following example we apply this corollary.

Example 1. *Consider a DMBC where the channel from X to Y_1 is a BSC with crossover probability $p \in (0,1/2)$, and the channel from X to Y_2 is an independent BEC with erasure probability $e \in (0,1)$. We show that our feedback regions $\mathcal{R}^{(1)}_{\text{relay,s}}$ and $\mathcal{R}^{(2)}_{\text{relay,s}}$ improve over a large part of the boundary points of $\mathcal{C}_{\text{NoFB}}$ for all values of e,p unless $H_b(p) = e$, no matter how small $R_{\text{Fb},1}, R_{\text{Fb},2} > 0$.*

We distinguish different parameter ranges of our channel.

- *$0 < e < H_b(p)$: $\mathcal{C}_{\text{NoFB}}$ [11] is formed by the set of rate pairs (R_1,R_2) that for some $s \in [0,1/2]$ satisfy*

$$R_1 \le 1 - H_b(s*p), \tag{15a}$$
$$R_2 \le (1-e)H_b(s), \tag{15b}$$
$$R_1 + R_2 \le 1 - e. \tag{15c}$$

We specialize the region $\mathcal{R}^{(1)}_{\text{relay,s}}$ to the choices $Q = U_1 =$ const., $U_0 \sim$ Bern$(1/2)$, $X = U_2 = U_0 \oplus V$, where $V \sim$ Bern(s) independent of U_0, and $\tilde{Y}_1 = Y_1$ with probability $\gamma \in (0,1)$ and $\tilde{Y}_1 = ?$ with probability $1-\gamma$, where

$$\gamma \le \frac{R_{\text{Fb},1}}{(1-e)H_b(p)+e}. \tag{16}$$

Condition (16) assures that (14) is satisfied. Thus, for any $\gamma \in (0,1)$ satisfying (16), the following region is achievable with feedback when $R_{\text{Fb},1} > 0$:

$$R_1 \le 1 - H_b(s*p) \tag{17a}$$
$$R_2 \le (1-e)H_b(s) + \gamma e(H_b(s*p) - H_b(p)) \tag{17b}$$
$$R_1+R_2 \le 1 - e - \gamma H_b(p). \tag{17c}$$

As shown in [11], the points (R_1,R_2) of the form

$$(1-H_b(s*p), (1-e)H_b(s)), \qquad s \in (0,s_0), \tag{18}$$

are all on the boundary of $\mathcal{C}_{\text{NoFb}}$, where $s_0 \in (0, 1/2)$ is the unique solution to

$$1 - H_b(s_0 * p) + (1-e)H_b(s_0) = 1 - e. \tag{19}$$

Notice that for these boundary points, only the single-rate constraints (15a) *and* (15b) *are active, but not* (15c). *Thus, comparing* (18) *to our feedback region* (17), *we can conclude that by choosing γ sufficiently small, all boundary points* (18) *lie strictly in the interior of our feedback region $\mathcal{R}^{(1)}_{\text{relay,s}}$ when $R_{\text{Fb},1} > 0$.*

- $0 < H_b(p) < e < 1$*: $\mathcal{C}_{\text{NoFb}}$ equals the time-sharing region given by the union of all rate-pairs (R_1, R_2) that for some $\alpha \in [0,1]$ satisfy*

$$R_1 \le \alpha(1 - H_b(p)) \tag{20a}$$
$$R_2 \le (1-\alpha)(1-e). \tag{20b}$$

We specialize the region $\mathcal{R}^{(2)}_{\text{relay},s}$ to the following choices: $Q \sim$ Bern(α); if $Q = 0$ then $U_0 \sim$ Bern$(1/2)$, $X = U_1 = U_0$, and $U_2 = \tilde{Y}_2 = const.$; if $Q = 1$ then $U_0 =$ const., $X = U_1 \sim$ Bern$(1/2)$, U_2 =const, and $\tilde{Y}_2 = Y_2$ with probability $\gamma \in (0,1)$ and $\tilde{Y}_2 = ?$ with probability $1-\gamma$, where to satisfy the feedback rate constraint,

$$\gamma \le \frac{R_{\text{Fb},2}}{\alpha((1-e)H_b(p) + H_b(e))}. \tag{21}$$

For any $\gamma \in (0,1)$ satisfying (21), *the following region is achievable with feedback when $R_{\text{Fb},2} > 0$:*

$$R_1 \le \alpha(1 - H_b(p)) + \alpha(1-e)\gamma H_b(p) \tag{22a}$$
$$R_2 \le (1-\alpha)(1-e) \tag{22b}$$
$$R_1 + R_2 \le 1 - H_b(p) - (1-\alpha)\gamma H_b(e). \tag{22c}$$

Since here $1 - H_b(p) > 1 - e$, for small $\gamma > 0$ the feedback region in (22) *improves over $\mathcal{C}_{\text{NoFB}}$ given in* (20). *In fact,* (22) *improves over all boundary points $(R_1 > 0, R_2 > 0)$ of $\mathcal{C}_{\text{NoFB}}$.*

Remark 1. *The BSC/BEC-BC in Example 1, is particularly interesting, because depending on the values of the parameters e and p, the BC is either degraded, less noisy, more capable, or essentially less-noisy [11]. We conclude that our feedback regions $\mathcal{R}^{(1)}_{\text{relay,s}}$ and $\mathcal{R}^{(2)}_{\text{relay,s}}$ can improve over the nofeedback capacity regions for all these classes of BCs even with only one feedback link that is of arbitrary small, but positive rate.*

We have the following result on the usefulness of feedback.

Theorem 2. *Fix a DMBC. Consider random variables $(U_0^{(\text{M})}, U_1^{(\text{M})}, U_2^{(\text{M})}, X^{(\text{M})})$ such that*

$$\Delta^{(\text{M})} := I(U_0^{(\text{M})}; Y_2^{(\text{M})}) - I(U_0^{(\text{M})}; Y_1^{(\text{M})}) > 0. \tag{23}$$

Let the rate-pair $(R_1^{(\text{M})}, R_2^{(\text{M})})$ satisfy Marton's constraints (6) *when evaluated for $(U_0^{(\text{M})}, U_1^{(\text{M})}, U_2^{(\text{M})}, X^{(\text{M})})$ where Constraint* (6b) *has to hold with strict inequality. Also, let $(R_1^{(\text{Enh})}, R_2^{(\text{Enh})}) \in \mathcal{C}^{(1)}_{\text{Enh}}$.*

If $R_{\text{Fb},1} > 0$, then for all sufficiently small $\gamma \in (0,1)$,

$$\Big((1-\gamma)R_1^{(\text{M})} + \gamma R_1^{(\text{Enh})}, (1-\gamma)R_2^{(\text{M})} + \gamma R_2^{(\text{Enh})}\Big) \in \mathcal{R}^{(1)}_{\text{relay,s}}.$$

An analogous statement holds for exchanged indices 1 & 2.

Remark 2. *For most $U_0^{(\text{M})}, U_1^{(\text{M})}, U_2^{(\text{M})}, X^{(\text{M})}$ satisfying* (23), *the region defined by Marton's constraints* (6) *is a pentagon, and the only point in this region satisfying Constraint* (6b) *with equality is the dominant corner point of maximum R_2-rate.*

If $U_1^{(\text{M})}$ =const. and $U_2^{(\text{M})} = X^{(\text{M})}$ (i.e., when superposition coding is used), then Condition (23) *makes that the region is a quadrilateral and the only active constraints are* (6a) *and* (6c). *In this case, Constraint* (6b) *holds with strict inequality for all points in the region.*

Corollary 2. *Let $R_{\text{Fb},1} > 0$. If there exists a pair $(R_1^{(\text{M})}, R_2^{(\text{M})})$ that satisfies the conditions in Theorem 2 and lies on the boundary of $\mathcal{R}_{\text{Marton}}$ but in the interior of $\mathcal{C}^{(1)}_{\text{Enh}}$, then*

$$\mathcal{R}_{\text{Marton}} \subsetneq \mathcal{C}_{\text{Fb}}. \tag{24}$$

Moreover, if for the considered DMBC $\mathcal{R}_{\text{Marton}} = \mathcal{C}_{\text{NoFB}}$,

$$\mathcal{C}_{\text{NoFB}} \subsetneq \mathcal{C}_{\text{Fb}}. \tag{25}$$

We introduce the term *strictly essentially less-noisy*. The definition of a *strictly essentially less-noisy* DMBC coincides with the definition of an essentially less-noisy DMBC except that Inequality (9) needs to be strict whenever $I(U; Y_1) > 0$.

Corollary 3. *Consider a DMBC where Y_2 is strictly essentially less-noisy than Y_1. Assume $R_{\text{Fb},1} > 0$. We have:*

1) *If a rate-pair (R_1, R_2) lies on the boundary of $\mathcal{C}_{\text{NoFB}}$ but in the interior of $\mathcal{C}^{(1)}_{\text{Enh}}$, then (R_1, R_2) lies in the interior of $\mathcal{C}_{\text{Fb}}$, i.e., it can be improved with feedback.*
2) *If $\mathcal{C}_{\text{NoFB}}$ does not coincide with $\mathcal{C}^{(1)}_{\text{Enh}}$, then $\mathcal{C}_{\text{NoFB}}$ is also a strict subset of $\mathcal{C}_{\text{Fb}}$, i.e., feedback strictly improves capacity.*

Analogous statements hold if indices 1 and 2 are exchanged.

All BSBCs and BEBCs with unequal crossover probabilities or unequal erasure probabilities at the two receivers are strictly essentially less-noisy. Also, for these BCs $\mathcal{C}_{\text{NoFB}}$ has no common boundary points $(R_1 > 0, R_2 > 0)$ with the sets $\mathcal{C}^{(1)}_{\text{Enh}}$ or $\mathcal{C}^{(2)}_{\text{Enh}}$ when the BC is not physically degraded.

Thus, Corollary 3 implies that for not physically degraded BSBCs or BEBCs with unequal crossover or erasure probabilities, rate-limited feedback improves all boundary points $(R_1 > 0, R_2 > 0)$ of $\mathcal{C}_{\text{NoFB}}$ whenever $R_{\text{Fb},1}, R_{\text{Fb},2} > 0$.

The following region sometimes improves over $\mathcal{R}_{\text{relay}}$.

Theorem 3. *The capacity region $\mathcal{C}_{\text{Fb}}$ includes the set $\mathcal{R}_{\text{proc.}}$ of all nonnegative rate pairs (R_1, R_2) that satisfy*

$$\begin{aligned}
R_1 &\le I(U_0, U_1; Y_1, \tilde{Y}_1, V) - I(V; U_0, U_1, U_2, \tilde{Y}_2 | \tilde{Y}_1, Y_1) \\
R_2 &\le I(U_0, U_2; Y_2, \tilde{Y}_2, V) - I(V; U_0, U_1, U_2, \tilde{Y}_1 | \tilde{Y}_2, Y_2) \\
R_1 + R_2 &\le I(U_0, U_1; Y_1, \tilde{Y}_1, V) + I(U_2; Y_2, \tilde{Y}_2, V | U_0) \\
&\quad - I(V; U_0, U_1, U_2, \tilde{Y}_2 | \tilde{Y}_1, Y_1) - I(U_1; U_2 | U_0) \\
R_1 + R_2 &\le I(U_0, U_2; Y_2, \tilde{Y}_2, V) + I(U_1; Y_1, \tilde{Y}_1, V | U_0) \\
&\quad - I(V; U_0, U_1, U_2, \tilde{Y}_1 | \tilde{Y}_2, Y_2) - I(U_1; U_2 | U_0) \\
R_1 + R_2 &\le I(U_0, U_1; Y_1, \tilde{Y}_1, V) - I(V; U_0, U_1, U_2, \tilde{Y}_2 | \tilde{Y}_1, Y_1)
\end{aligned}$$

$$+I(U_0,U_2;Y_2,\tilde{Y}_2,V)$$
$$-I(V;U_0,U_1,U_2,\tilde{Y}_1|\tilde{Y}_2,Y_2)-I(U_1;U_2|U_0)$$

for some pmf $P_{U_0U_1U_2}, P_{\tilde{Y}_1|Y_1}, P_{\tilde{Y}_2|Y_2}, P_{V|U_0U_1U_2\tilde{Y}_1\tilde{Y}_2}$ *and some function* $f\colon \mathcal{X}\to\mathcal{U}_0\times\mathcal{U}_1\times\mathcal{U}_2$ *that satisfy*

$$I(Y_1;\tilde{Y}_1|U_0,U_1,U_2,\tilde{Y}_2)\leq R_{\text{Fb},1} \tag{27a}$$
$$I(Y_2;\tilde{Y}_2|U_0,U_1,U_2,\tilde{Y}_1)\leq R_{\text{Fb},2} \tag{27b}$$
$$I(Y_1,Y_2;\tilde{Y}_1,\tilde{Y}_2|U_0,U_1,U_2)\leq R_{\text{Fb},1}+R_{\text{Fb},2}, \tag{27c}$$

where $X=f(U_0,U_1,U_2)$.

Remark 3. *When the feedback rates* $R_{\text{Fb},1}, R_{\text{Fb},2}$ *are sufficiently large, we can choose* $\tilde{Y}_i = Y_i$ *and our achievable region* $\mathcal{R}_{\text{proc.}}$ *specializes to the Shayevitz-Wigger region for output feedback and for the choice* $V_1 = V_2 = V_0$. *Our region* $\mathcal{R}_{\text{proc.}}$ *can thus recover the two-user capacity results in [2], [5], [6] when* $R_{\text{Fb},1}, R_{\text{Fb},2}$ *are sufficiently large and the degree of freedom achievability result in [7] when* $R_{\text{Fb},1}, R_{\text{Fb},2}\to\infty$.

Remark 4. *All our results remain valid in the related setup where the feedback links are noisy channels of capacities* $R_{\text{Fb},1}$ *and* $R_{\text{Fb},2}$, *when the receivers can code over these channels.*

V. Outline of Coding Schemes

A. Encoder Relays Feedback Messages (Theorem 1)

We use a block-Markov scheme with blocks of length n. In each block $b\in\{1,\dots,B+1\}$ the transmitter uses Marton's nofeedback scheme to send fresh messages $M_{1,b}$ and $M_{2,b}$. In addition, as part of the cloud center it also sends the feedback messages $M_{\text{Fb},1,b-1}$ and $M_{\text{Fb},2,b-1}$ that it received after the previous block. In fact, after each block b, each Receiver $i\in\{1,2\}$ compresses the channel outputs $Y_{i,b}^n$ it observed in this block, and sends the compression index as its feedback message $M_{\text{Fb},i,b}$ to the transmitter. As already mentioned, the encoder only *relays* the information it obtains over the feedback links without any processing. Thus, the compression indices produced by Receiver 1 are really intended for Receiver 2 who already has side-information $Y_{2,b}^n$, and it is therefore constructed using a Wyner-Ziv compression. Similarly for the compression indices produced by Receiver 2.

Notice further that in each block b, some of the messages encoded in Marton's cloud center (namely the feedback messages $M_{\text{Fb},1,b-1}$ and $M_{\text{Fb},2,b-1}$) are already known at one of the two receivers. Simply because the receivers have generated these messages themselves after the previous block $b-1$. In our scheme, the receivers apply a modified Marton decoding rule that is adapted to this.

Decoding is performed backwards. We explain the decoding at Receiver 1; the decoding at Receiver 2 is similar. After the last block $B+1$, Receiver 1 decodes its private message $M_{1,B+1}$ and the feedback message $M_{\text{Fb},2,B}$. To this end, it applies the traditional Marton decoding but in the typicality check it only considers the codewords that correspond to the correct feedback message $M_{\text{Fb},1,B}$, which it had generated in block B. For all other blocks $b=1,\dots,B$, with the aid of the feedback message $M_{\text{Fb},2,b}$ that it had previously decoded in block $b+1$, Receiver 1 first reconstructs a compressed version $\tilde{Y}_{2,b}^n$ of the outputs $Y_{2,b}^n$. Then, it applies the modified Marton decoding explained in the previous paragraph to the enhanced output $(Y_{1,b}^n, \tilde{Y}_{2,b}^n)$.

B. Encoder Processes the Feedback Messages (Theorem 3)

Our second scheme is very similar to the first scheme. It differs in that in each block b, after receiving the feedback messages $M_{\text{Fb},1,b}, M_{\text{Fb},2,b}$, the encoder first reconstructs the compressed versions of the channel outputs, $\tilde{Y}_{1,b}^n$ and $\tilde{Y}_{2,b}^n$, and then newly compresses the quintuple consisting of $\tilde{Y}_{1,b}^n$ and $\tilde{Y}_{2,b}^n$ and the Marton codewords $U_{0,b}^n$, $U_{1,b}^n$, $U_{2,b}^n$ that it had sent during block b. This new compression information is then sent to the two receivers in the next-following block $b+1$ as part of the cloud center of Marton's code for this block.

Backward decoding is applied at the receivers. For each block b, each receiver $i\in\{1,2\}$ uses its observed outputs $Y_{i,b}^n$ to simultaneously reconstruct the encoder's compressed signal and decode its intended messages sent in block b.

Acknowledgment

The authors would like to thank R. Timo for helpful discussions and the city of Paris for supporting their work under the "Emergences" programme.

References

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

[2] G. Dueck, "Partial feedback for two-way and broadcast channels," *Inform. and Control,* vol. 46, pp. 1–15, July 1980.

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

[4] H. Kim, Y.-K. Chia, and Abbas El Gamal, "A note on broadcast channels with stale state information at the transmitter," submitted to *IEEE Trans. on Inf. Theory,*, Sept. 2013, available at `http://arxiv.org/abs/1309.7437`.

[5] C. C. Wang, "Capacity of 1-to-k broadcast erasure channel with output feedback," *in Proc. of Allerton Conf.*, Monticello, IL, Sept. 2010.

[6] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback capacity and algorithms," *in Proc. 5th Workshop on Network Coding, Theory, & Applications,* Lausanne, Switzerland, June 2009.

[7] M. A. Maddah-Ali and D. Tse , "Completely Stale Transmitter Channel State Information is Still Very Useful," *IEEE Trans. on Inf. Theory,* vol. 58, no. 7, pp. 4418–4431, July 2012.

[8] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Trans. on Inf. Theory,* vol. 49, no. 1, pp. 4–20, Jan. 2003.

[9] R. Venkataramanan and S. S. Pradhan, "An achievable rate region for the broadcast channel with feedback," *IEEE Trans. on Inf. Theory,* vol. 59, no. 10, pp. 6175–6191, Oct. 2013.

[10] K. Marton, "A coding theorem for the discrete memoryless broadcast channel, *IEEE Trans. on Inf. Theory*, vol. 25, pp. 306–311, May 1979.

[11] C. Nair, "Capacity regions of two new classes of two-receiver broadcast channels," *IEEE Trans. on Inf. Theory,* vol. 56, no. 9, pp. 4207–4214, Sept. 2010.

[12] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. on Inf. Theory,* vol. 25, no. 2, pp. 166–169, Mar 1979.

[13] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. on Inf. Theory,* vol. 53, no. 1, pp. 350–355, Jan. 2007.

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

[15] Y. Wu and M. Wigger, "Coding schemes for discrete memoryless broadcast channels with rate-limited feedback", extended version of this paper. Available at `http://arxiv.org/pdf/1401.6219.pdf`