An Achievable Region for the Discrete Memoryless Broadcast Channel with Feedback

Ofer Shayevitz
University of California, San Diego
La Jolla, CA, 92093, USA
ofersha@ucsd.edu

Michèle Wigger
Telecom ParisTech
75013 Paris, France
michele.wigger@telecom-paristech.fr

Abstract—**A coding scheme for the discrete memoryless broadcast channel with (possible noisy) feedback is proposed, and the corresponding achievable region derived. The scheme is based on a block-Markov strategy where in each block the transmitter sends fresh data and update information that allows the receivers to improve the channel outputs observed in the previous block. The region is analyzed for two specific broadcast channels: 1) A generalization of Dueck's channel, where it is shown that for noiseless output-feedback the region coincides with the capacity region; 2) A noisy version of Blackwell's channel, where it is shown that for noiseless – and in some cases noisy – output-feedback, the region improves upon the no-feedback capacity region.**

I. INTRODUCTION

We consider a broadcast setup where a single transmitter communicates with two receivers, and where the transmitter has access to a feedback signal. Specifically, we study the capacity region of the discrete memoryless broadcast channel (DMBC) with *generalized feedback*. Generalized feedback refers to a scenario where the feedback signal can be modeled as observing the DMBC's inputs and outputs through a discrete memoryless channel. Special cases of generalized feedback are *noiseless/noisy output-feedback* (noiseless/noisy feedback for short), where the transmitter observes a noiseless/noisy version of the receivers' channel outputs.

Most previous results on the DMBC with feedback focus on noiseless feedback. El Gamal [4] proved that for physically degraded DMBCs the noiseless-feedback capacity coincides with the no-feedback capacity. In contrast, Dueck [7] and Kramer [8] showed by the way of examples that for other DMBCs the noiseless-feedback capacity can exceed the no-feedback capacity.

Determining the feedback capacity of the DMBC is still an open problem. In fact, it is even unknown whether feedback increases the capacity region of a general (non-physically-degraded) DMBC. This is partly because even the no-feedback capacity region is generally unknown, and partly since a computable single-letter achievable region for the DMBC with feedback is missing. Kramer [8] proposed a multi-letter achievable region for the DMBC with noisy or noiseless feedback.

In this paper we propose a coding scheme for the DMBC with generalized feedback, and present a corresponding single-letter achievable region. Our approach is motivated by Dueck's example [7], and is based on the following high-level idea: The transmitter exploits feedback to identify information useful for decoding, and describes this information efficiently on subsequent transmissions, thereby improving the receivers' observed channel outputs. To this end, a block-Markov strategy is adopted, where in each block the transmitter sends a combination of fresh data and compressed update information pertaining to the previous block, using a no-feedback code. Intuitively, this approach may be beneficial if a large part of the update information is common to both receivers.

The region achieved by our feedback scheme is shown to exceed the no-feedback capacity region for some DMBCs with noisy or noiseless feedback: A generalization of Dueck's DMBC is discussed in Subsection VI-A, and a noisy version of Blackwell's DMBC [9] in Subsection VI-B.

II. CHANNEL MODEL

We describe the channel model in more detail. The goal of the communication is that the transmitter conveys a private Message M_1 to a Receiver 1, a private Message M_2 to a Receiver 2, and a common message M_0 to both receivers. The three messages M_0, M_1, and M_2 are assumed to be independent and uniformly distributed over the finite sets $\{1, \ldots, \lfloor 2^{nR_0} \rfloor\}$, $\{1, \ldots, \lfloor 2^{nR_1} \rfloor\}$, and $\{1, \ldots, \lfloor 2^{nR_2} \rfloor\}$ respectively, where n denotes the blocklength and R_0, R_1, R_2 are the corresponding common and private transmission rates.

Communication takes place over a DMBC with generalized feedback. This channel is characterized by a quadruple of finite alphabets $\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2$, and $\widetilde{\mathcal{Y}}$, and a conditional probability law $P_{Y_1 Y_2 \widetilde{Y}|X}(y_1, y_2, \widetilde{y}|x)$ where $x \in \mathcal{X}$, $y_1 \in \mathcal{Y}_1$, $y_2 \in \mathcal{Y}_2$, and $\widetilde{y} \in \widetilde{\mathcal{Y}}$. Given that at time t the transmitter feeds the symbol x_t to the channel, then Receiver 1 and Receiver 2 observe the channel outputs $y_{1,t} \in \mathcal{Y}_1$ and $y_{2,t} \in \mathcal{Y}_2$ respectively and the transmitter observes the generalized feedback $\widetilde{y}_t \in \widetilde{\mathcal{Y}}$, with probability $P_{Y_1 Y_2 \widetilde{Y}|X}(y_{1,t}, y_{2,t}, \widetilde{y}_t|x_t)$.

Thanks to feedback, the transmitter can produce its time-t channel input X_t as a function of the Messages M_0, M_1, M_2 and of the previously observed feedback outputs $\widetilde{Y}^{t-1} \stackrel{\text{def}}{=} (\widetilde{Y}_1, \ldots, \widetilde{Y}_{t-1})$:

$$X_t = \varphi_t^{(n)} \left(M_0, M_1, M_2, \widetilde{Y}^{t-1} \right). \tag{1}$$

The DMBC and its feedback channel are memoryless which is captured by the following Markov relation for $t \in \{1, \ldots, n\}$:

$$(Y_1^{t-1}, Y_2^{t-1}, \widetilde{Y}^{t-1}) \multimap X_t \multimap (Y_{1,t}, Y_{2,t}, \widetilde{Y}_t)$$

where $Y_i^{t-1} \stackrel{\text{def}}{=} (Y_{i,1}, Y_{i,2}, \ldots, Y_{i,t-1})$, for $i \in \{1, 2\}$.

After n channel uses Receiver i decodes its intended messages M_0 and M_i for $i \in \{1, 2\}$. Namely, Receiver i produces the guess:

$$(\hat{M}_{0,i}, \hat{M}_i) = \Phi_i^{(n)}(Y_i^n), \quad i \in \{1, 2\}. \tag{2}$$

A rate triplet (R_0, R_1, R_2) is called achievable if for every blocklength n there exists a set of n encoding functions $\left\{\varphi_t^{(n)}\right\}_{t=1}^n$ and two decoding functions $\Phi_1^{(n)}$ and $\Phi_2^{(n)}$ such that the probability of decoding error, i.e., the probability that

$$(M_0, M_1) \neq (\hat{M}_{0,1}, \hat{M}_1) \quad \text{or} \quad (M_0, M_2) \neq (\hat{M}_{0,2}, \hat{M}_2),$$

tends to 0 as the blocklength n tends to infinity. The closure of the set of achievable rate triplets (R_0, R_1, R_2) is called the *generalized-feedback capacity region* of this setup, and we denote it by $\mathcal{C}_{\text{GenFB}}$. The supremum of the sum $R_1 + R_2$ over all achievable rate triplets is called its *generalized-feedback sum-rate capacity*, and we denote it by $C_{\text{GenFB},\Sigma}$.

The described generalized-feedback setup includes as special cases the *no-feedback* setup where the feedback outputs are deterministic, e.g., $|\widetilde{\mathcal{Y}}| = 1$; the *noiseless-feedback* setup where the feedback output coincides with the pair of channel outputs, i.e., $\widetilde{Y} = (Y_1, Y_2)$; and the *noisy-feedback* setup where the feedback outputs and the channel inputs and outputs satisfy the Markov relation $X_t \multimap (Y_{1,t}, Y_{2,t}) \multimap \widetilde{Y}_t$ for all $t \in \{1, \ldots, n\}$. In these special cases, we denote the sum-rate capacities by $C_{\text{NoFB},\Sigma}$, $C_{\text{NoiselessFB},\Sigma}$, and $C_{\text{NoisyFB},\Sigma}$.

III. PRELIMINARIES

Our generalized-feedback scheme appearing in Section V-B has two main building blocks: The Marton's scheme for the DMBC without feedback [6], [12], and a source-coding scheme for a lossy version of the Gray-Wyner problem with side-information (see [2], [3] for the lossless case). In this section, we review the region achieved by Marton's scheme, and present an achievable region for the lossy Gray-Wyner problem. We outline the coding scheme achieving this region.

We use the notion of strong typicality as defined in [1]. For a distribution P_X over a finite alphabet $\mathcal{X}$, we denote by $\mathcal{T}_\delta^n(P_X)$ the set of all n-length sequences $x^n \in \mathcal{X}^n$ that are δ-strongly-typical with respect to (w.r.t.) the law P_X. Similarly, for a law $P_{X_1 \cdots X_k}$ over a product alphabet $\mathcal{X}_1 \times \cdots \times \mathcal{X}_k$ we denote by $\mathcal{T}_\delta^n(P_{X_1 \cdots X_k})$ the set of all k-tuples of sequences $(x_1^n \in \mathcal{X}_1^n, \ldots, x_k^n \in \mathcal{X}_k^n)$ that are *jointly* δ-strongly typical w.r.t. $P_{X_1 \cdots X_k}$.

A. Marton's Achievable Region for the BC without Feedback

Theorem 1 (From [6], [12]). *A nonnegative rate triplet (R_0, R_1, R_2) is achievable over the DMBC without feedback, if it satisfies*

$$R_0 < \min_i I(U_0; Y_i)$$
$$R_0 + R_1 < I(U_0, U_1; Y_1)$$
$$R_0 + R_2 < I(U_0, U_2; Y_2)$$
$$R_0 + R_1 + R_2 < I(U_1; Y_1 | U_0) + I(U_2; Y_2 | U_0)$$
$$+ \min_i I(U_0; Y_i) - I(U_1; U_2 | U_0)$$

such that $(U_0, U_1, U_2) \multimap X \multimap (Y_1, Y_2)$ forms a Markov chain.

B. Lossy Gray-Wyner Coding with Side Information

Let $\{(X_t, Y_{1,t}, Y_{2,t})\}_{t=1}^n$ be an i.i.d. sequence of triplets of discrete random variables, with marginal distribution $P_{XY_1Y_2}$. We consider a setting where a sender observes the sequence X^n, Receiver 1 observes the side-information Y_1^n, and Receiver 2 observes the side-information Y_2^n. The sender can noiselessly communicate with the receivers by sending a common message from the set $\{1, \ldots, \lfloor 2^{nR_0} \rfloor\}$ to both receivers, a private message from the set $\{1, \ldots, \lfloor 2^{nR_1} \rfloor\}$ to Receiver 1 only, and another private message from the set $\{1, \ldots, \lfloor 2^{nR_2} \rfloor\}$ to Receiver 2 only. Fix two conditional distributions $P_{V_1|X}$ and $P_{V_2|X}$. The goal of the communication is to ensure that for a given $\delta > 0$, Receiver i can construct a sequence V_i^n such that $(X^n, V_i^n) \in \mathcal{T}_\delta^n(P_{XV_i})$ for $i \in \{1, 2\}$, with a probability of failure no larger than δ. A rate triplet (R_0, R_1, R_2) is said to be achievable if that goal can be attained with $\delta \to 0$ as $n \to \infty$.

Theorem 2. *A nonnegative rate triplet (R_0, R_1, R_2) is achievable for the lossy Gray-Wyner problem with side information above, if it satisfies*

$$R_0 > \max_i I(X; V_0 | Y_i),$$
$$R_1 > I(X; V_1 | V_0, Y_1)$$
$$R_2 > I(X; V_2 | V_0, Y_2). \tag{3}$$

where $(V_0, V_1, V_2) \multimap X \multimap (Y_1, Y_2)$ forms a Markov chain.

Proof. We outline the construction, see [10] for details.

1) Codebook Generation: Independently generate codebooks $\mathcal{C}_i$ with $2^{n(R_i' + R_i)}$ codewords by randomly drawing all the entries of the codebook i.i.d. according to P_{V_i}, for $i \in \{0, 1, 2\}$ respectively. Partition each codebook $\mathcal{C}_i$ into 2^{nR_i} bins each containing $2^{nR_i'}$ codewords. All codebooks are revealed to the sender. Codebooks $\{\mathcal{C}_0, \mathcal{C}_i\}$ are revealed to Receiver i, for $i \in \{1, 2\}$.

2) Lossy Gray-Wyner Encoder: The sender observes a sequence X^n, and looks for a triplet of codewords $(V_0^n, V_1^n, V_2^n) \in \mathcal{C}_0 \times \mathcal{C}_1 \times \mathcal{C}_2$, such that $(X^n, V_0^n, V_i^n) \in \mathcal{T}_\delta^n(P_{XV_0V_i})$ for $i \in \{1, 2\}$. It then sends the index J_0 of the bin containing V_0^n over the common link, and the indices J_1, J_2 of the bins containing V_1^n and V_2^n over the respective private links. If no such triplet of codewords exists, the sender transmits arbitrary messages, and an error is declared.

3) Lossy Gray-Wyner Decoder: Receiver i observes the side-information Y_i^n, the common message J_0 and the private message J_i. It seeks a codeword $\widehat{V}_0^n$ in the J_0-th bin of codebook $\mathcal{C}_0$ and a codeword $\widehat{V}_i^n$ in the J_i-th bin of codebook $\mathcal{C}_i$, such that $(\widehat{V}_0^n, \widehat{V}_i^n, Y_i^n) \in \mathcal{T}_\delta^n(P_{V_0V_iY_i})$. If exactly one such pair of codewords exists, Receiver i outputs $\widehat{V}_i^n$. Otherwise, it outputs an arbitrary sequence and an error is declared.

4) Analysis Outline: We write *with high probability (w.h.p.)* to indicate a statement holds with probability $\to 1$ as $n \to \infty$. By the covering Lemma [11], if $R_0' + R_0 > I(X; V_0)$ then w.h.p. the sender can find $V_0^n \in \mathcal{C}_0$ jointly typical with

X^n. Given that, and invoking the covering Lemma again, if $R'_i + R_i > I(V_i; X, V_0)$ then w.h.p. the sender can find $V_i^n \in \mathcal{C}_i$ jointly typical with (X^n, V_0^n). By the conditional typicality Lemma [11] we have that w.h.p. (V_0^n, V_i^n, Y_i^n) are jointly typical. Therefore, by the packing Lemma [11], if $R'_0 < I(V_0; Y_i)$ then w.h.p. V_0^n is the only codeword in the J_0-th bin of codebook $\mathcal{C}_0$ that is jointly typical with Y_i^n. Given that, and invoking the packing Lemma again, if $R'_i < I(V_i; V_0, Y_i)$ then w.h.p. V_i^n is the only codeword in the J_i-th bin of codebook $\mathcal{C}_i$ that is jointly typical with (V_0^n, Y_i^n). Hence, w.h.p. Receiver i will output V_i^n as desired. The result follows by combining the bounds and using Markovity. ∎

IV. MOTIVATION: DUECK'S EXAMPLE

Our feedback scheme in Section V-B is motivated by Dueck's scheme in [7]. Dueck considered the following DMBC. The channel inputs consist of triplets of bits $X = (X_1, X_0, X_2)$ and the outputs of pairs of bits $Y_1 = (Y_{1,1}, Y_{1,0})$ and $Y_2 = (Y_{2,0}, Y_{2,2})$. The channel $P_{Y_1 Y_2|X}$ is described as:

$$Y_{1,0} = Y_{2,0} = X_0,$$
$$Y_{1,1} = X_1 \oplus Z, \quad \text{and} \quad Y_{2,2} = X_2 \oplus Z,$$

where Z is Bern(1/2) and independent of the inputs, and $\oplus$ denotes addition modulo 2.

Obviously, without feedback, the outputs $Y_{1,1}$ and $Y_{2,2}$ are useless. Thus, the no-feedback-capacity is given by the set of all nonnegative rate triplets (R_0, R_1, R_2) satisfying

$$R_0 + R_1 + R_2 \leq 1.$$

With noiseless feedback, the capacity is increased.

Theorem 3 (Dueck [7]). *The noiseless feedback capacity of Dueck's DMBC is given by the set of all nonnegative rate triplets (R_0, R_1, R_2) satisfying*

$$R_0 + R_1 \leq 1 \quad \text{and} \quad R_0 + R_2 \leq 1.$$

The noiseless-feedback capacity is achieved by the following simple scheme of blocklength $(n + 1)$. In the first n channel uses the transmitter uses Inputs $\{X_{1,t}\}$ and $\{X_{2,t}\}$ to send lossless descriptions of the Message pairs (M_0, M_1) and (M_0, M_2), respectively. In channel uses $t = 2, \ldots, (n+1)$, the transmitter additionally sends the previous noise sample $X_{0,t} = Z_{t-1}$. Since each receiver directly observes the inputs $\{X_{0,t}\}_{t=2}^{n+1}$, it can recover the noise sequence $\{Z_t\}_{t=1}^{n}$ and reconstruct its desired channel inputs $\{X_{1,t}\}_{t=1}^{n}$ or $\{X_{2,t}\}_{t=1}^{n}$. Based on these reconstructed channel inputs the receivers then decode their desired pair of messages (M_0, M_1) or (M_0, M_2). Whenever the rates R_1, R_2 do not exceed 1, the described scheme has zero probability of error.

V. MAIN RESULT

A. Achievable Region

Let $\tilde{\mathcal{R}}_{\text{inner}}$ be the set of all nonnegative triplets (R_0, R_1, R_2) that satisfy the inequalities (4) on the top of the next page, for some choice of auxiliary random variables $(U_0, U_1, U_2, V_0, V_1, V_2)$ such that $(U_0, U_1, U_2) \multimap X \multimap (Y_1, Y_2, \widetilde{Y})$ and

$(V_0, V_1, V_2) \multimap (X, \widetilde{Y}) \multimap (Y_1, Y_2, U_0, U_1, U_2)$ are Markov chains. Let $\mathcal{R}_{\text{inner}}$ be the closed convex hull of $\tilde{\mathcal{R}}_{\text{inner}}$.

Theorem 4. *The capacity region of the DMBC with generalized feedback satisfies*

$$\mathcal{R}_{\text{inner}} \subseteq \mathcal{C}_{\text{GenFB}}.$$

Proof. The proof is based on the feedback scheme discussed in the next subsection, which concatenates Marton's nofeedback scheme with the lossy Gray-Wyner scheme of Subsection III-B. When analyzing this concatenated random coding scheme, special attention should be paid to the fact that some of the sequences involved are not i.i.d., but only satisfy a weaker joint typicality condition with high probability. For details, see [10].

∎

B. Feedback-Scheme for the DMBC (Outline)

Inspired by Dueck's scheme, we propose the following scheme for DMBCs with generalized feedback. Our scheme is based on a Block-Markov strategy (as depicted in Figure 1). This means, the blocklength n is divided into $(B + 1)$ blocks of length n' each, and in each block b the transmitter sends to both receivers a combination of fresh data $(M_{0,b}, M_{1,b}, M_{2,b})$ and update information from the previous block $(J_{0,b-1}, J_{1,b-1}, J_{2,b-1})$. Exceptions are the first block where only fresh data is sent, and the last block where only update information is sent (see Figure 1).

The receivers wait until the reception of the last block and then apply the following backward decoding strategy. In a first step, Receiver i, for $i \in \{1, 2\}$, decodes the update information sent to it in the last block $B+1$, i.e., $(J_{0,B}, J_{i,B})$. In the sequel it performs the following iterative decoding steps, starting from the second to last block and ending with the first block. In each block $b \in \{1, \ldots, B\}$, Receiver i first uses the previously decoded update information $(J_{0,b}, J_{i,b})$ – in a way to be explained shortly – to improve its observed channel outputs $Y_{i,b}^{n'} = (Y_{i,(b-1)n'+1}, \ldots, Y_{i,bn'})$, and based on these improved outputs it then decodes the fresh data $(M_{0,b}, M_{i,b})$ and the update information $(J_{0,b-1}, J_{i,b-1})$ sent in this block b.

In what follows we elaborate on the various components of our scheme: 1) Generation of the update information at the transmitter; 2) Computation of the improved channel outputs at the receivers; and 3) Encoding and decoding of the fresh data and the update information in each block.

We first describe the generation of update information. Notice that in our scheme the transmitter can wait until it observed the block-b feedback-outputs $\widetilde{Y}_b^{n'} = (\widetilde{Y}_{(b-1)n+1}, \ldots, \widetilde{Y}_{bn'})$ to generate the update information $J_{0,b}, J_{1,b}, J_{2,b}$ transmitted in block $(b + 1)$. Thus, it can compute the update information by applying the *lossy Gray-Wyner encoder* in Subsection III-B2 to the block-b inputs and the block-b feedback outputs, i.e., to the sequence of pairs $\{(X_{(b-1)n'+k}, \widetilde{Y}_{(b-1)n'+k})\}_{k=1}^{n'}$.

We next describe the computation of the improved channel outputs at Receiver i, for $i \in \{1, 2\}$. Recall that Receiver i

$$R_0 \leq \min_i I(U_0; Y_i, V_i) - \max_i I(V_0; X, \widetilde{Y}|Y_i)$$

$$R_0 + R_1 \leq I(U_0, U_1; Y_1, V_1) - I(X, \widetilde{Y}; V_1|V_0, Y_1) - \max_i I(V_0; X, \widetilde{Y}|Y_i)$$

$$R_0 + R_2 \leq I(U_0, U_2; Y_2, V_2) - I(X, \widetilde{Y}; V_2|V_0, Y_2) - \max_i I(V_0; X, \widetilde{Y}|Y_i)$$

$$R_0 + R_1 + R_2 \leq I(U_1; Y_1, V_1|U_0) + I(U_2; Y_2, V_2|U_0) + \min_i I(U_0; Y_i, V_i)$$

$$-I(U_1; U_2|U_0) - I(X, \widetilde{Y}; V_1|V_0, Y_1) - I(X, \widetilde{Y}; V_2|V_0, Y_2) - \max_i I(V_0; X, \widetilde{Y}|Y_i) \tag{4}$$

Fig. 1. Block-Markov strategy of our feedback-scheme.

has already decoded the update information $(J_{0,b}, J_{i,b})$ before it has to produce its block-b improved outputs. It can thus compute the improved outputs as follows. It first feeds its guess of $(J_{0,b}, J_{i,b})$ and the observed sequence $Y_{i,b}^{n'}$ (as side-information) to the *lossy Gray-Wyner decoder* in Subsection III-B3, and then combines the resulting outputs $V_{i,b}^{n'}$ with the observed sequence $Y_{i,b}^{n'}$. That means the improved outputs are given by $(Y_{i,b}^{n'}, V_{i,b}^{n'})$.

We finally outline the encoding and decoding of the fresh data and the update information in each block b. The transmitter first forms the combined messages $\bar{M}_{0,b} = (M_{0,b}, J_{0,b-1})$, $\bar{M}_{1,b} = (M_{1,b}, J_{1,b-1})$, and $\bar{M}_{2,b} = (M_{2,b}, J_{2,b-1})$, and then encodes these messages using Marton's no-feedback scheme. Receiver i first decodes the combined messages $(\bar{M}_{0,b}, \bar{M}_{i,b})$ by applying Marton's decoding rule to the previously computed improved outputs $(Y_{i,b}^{n'}, V_{i,b}^{n'})$, and then attempts to extract messages $(M_{0,b}, M_{i,b})$ and $(J_{0,b-1}, J_{i,b-1})$ from its guess of $(\bar{M}_{0,b}, \bar{M}_{i,b})$.

VI. EXAMPLES

A. The Generalized Dueck DMBC

Consider a generalized version of Dueck's DMBC where all three binary channels are (possibly) noisy, and the first and third channels are corrupted by (possibly) different noises. As in Dueck's example the channel input consists of three bits, $X = (X_1, X_0, X_2)$, and each output of two bits, $Y_1 = (Y_{1,1}, Y_{1,0})$ and $Y_2 = (Y_{2,0}, Y_{2,2})$. The channel law $P_{Y_1 Y_2|X}$ is described by the relations:

$$Y_{1,0} = Y_{2,0} = X_0 \oplus Z_0,$$
$$Y_{1,1} = X_1 \oplus Z_1, \quad \text{and} \quad Y_{2,2} = X_2 \oplus Z_2,$$

where Z_0, Z_1, Z_2 are binary random variables of a given law $P_{Z_0 Z_1 Z_2}$. We restrict attention to laws $P_{Z_0 Z_1 Z_2}$ such that

$$H(Z_0, Z_1) \leq 1 \qquad \text{and} \qquad H(Z_0, Z_2) \leq 1. \tag{5}$$

Under these conditions and when the transmitter only wishes to send private messages but no common message, our scheme achieves the noiseless-feedback capacity.

Theorem 5. *Under condition* (5) *and for* $R_0 = 0$, *the noiseless-feedback capacity of the Generalized Dueck DMBC is the set of all nonnegative rate pairs* (R_1, R_2) *satisfying*

$$R_1 \leq 2 - H(Z_0, Z_1),$$
$$R_2 \leq 2 - H(Z_0, Z_2),$$
$$R_1 + R_2 \leq 3 - H(Z_0, Z_1, Z_2). \tag{6}$$

Proof. The converse follows from the cutset bound. The direct part follows from Theorem 4 by taking the convex hull of the two achievable regions that result when (4) is evaluated for the following two choices: (U_0, U_1, U_2) i.i.d. $\sim \text{Bern}(1/2)$; $X_0 = U_0$; $X_1 = U_1$; $X_2 = U_2$; $V_1 = (X_0, X_1)$; $V_2 = (X_0, X_2)$; and either $V_0 = (Z_0, Z_1)$ or $V_0 = (Z_0, Z_2)$. ∎

Observation 1. *For zero common rate* $R_0 = 0$, *the no-feedback capacity of the generalized Dueck DMBC is the set of all nonnegative rate pairs* (R_1, R_2) *satisfying*

$$R_1 \leq 2 - H(Z_0, Z_1),$$
$$R_2 \leq 2 - H(Z_0, Z_2),$$
$$R_1 + R_2 \leq 3 - H(Z_0, Z_1, Z_2) - I(Z_1; Z_2|Z_0).$$

Observation 2. *Unless the triplet* $Z_1 - Z_0 - Z_2$ *forms a Markov chain, noiseless feedback strictly increases the capacity of the Generalized Dueck DMBC satisfying Conditions* (5).

B. The Noisy Blackwell DMBC

We consider a noisy version of Blackwell's DMBC with noisy feedback, and evaluate the region in Theorem 4 for a specific choice of the auxiliary random variables. We show that the resulting inner bound on the noisy feedback capacity region exceeds an outer bound on the no-feedback capacity region, for small levels of noise.

The Noisy Blackwell DMBC is described as follows. The input alphabet is ternary $\mathcal{X} = \{0, 1, 2\}$ and both output alphabets are binary $\mathcal{Y}_1 = \mathcal{Y}_2 = \{0, 1\}$. Let $Z \sim \text{Bern}(p)$ for some $p < \frac{1}{2}$ and independent of X. The channel law $P_{Y_1 Y_2|X}$ is described as follows.

$$Y_1 = \begin{cases} Z & X = 0 \\ 1 - Z & X = 1, 2 \end{cases} \qquad Y_2 = \begin{cases} Z & X = 0, 1 \\ 1 - Z & X = 2 \end{cases}$$

For simplicity, we assume a noisy feedback of the form $\widetilde{Y} = (Y_1 + Z', Y_2 + Z')$, where $Z' \sim \text{Bern}(q)$ and (X, Y_1, Y_2) are mutually independent. Let

$$V_1 \stackrel{\text{def}}{=} \begin{cases} 0 & X = 0 \\ 1 & X = 1, 2 \end{cases} \qquad V_2 \stackrel{\text{def}}{=} \begin{cases} 0 & X = 0, 1 \\ 1 & X = 2 \end{cases}$$

$$V_0 \stackrel{\text{def}}{=} V_i \oplus \widetilde{Y}_i = Z \oplus Z'$$

Furthermore, let $U_0 \sim \text{Bern}(\frac{1}{2})$ and

$$P_{X|U_0} = \begin{cases} (\alpha, 1 - \alpha - \beta, \beta) & U_0 = 0 \\ (\beta, 1 - \alpha - \beta, \alpha) & U_0 = 1. \end{cases}$$

for some nonnegative α, β satisfying $\alpha + \beta \le 1$. Finally, set $U_1 \stackrel{\text{def}}{=} V_1, U_2 \stackrel{\text{def}}{=} V_2$. Note that the auxiliary outputs here are simply the deterministic part of the channel actions on its input. Using Theorem 4 with (4) we get the following achievable region (for simplicity we give looser inequalities):

$$R_0 \le h_b\left(\frac{\alpha + \beta}{2}\right) - \frac{1}{2}(h_b(\alpha) + h_b(\beta)) - \lambda(p, q, \alpha, \beta)$$

$$R_0 + R_1 \le h_b\left(\frac{\alpha + \beta}{2}\right) - \lambda(p, q, \alpha, \beta) - h_b(q)$$

$$R_0 + R_2 \le h_b\left(\frac{\alpha + \beta}{2}\right) - \lambda(p, q, \alpha, \beta) - h_b(q)$$

$$R_0 + R_1 + R_2 \le h_b\left(\frac{\alpha + \beta}{2}\right) + \frac{1 - \beta}{2} h_b\left(\frac{\alpha}{1 - \beta}\right)$$
$$+ \frac{1 - \alpha}{2} h_b\left(\frac{\beta}{1 - \alpha}\right) - \lambda(p, q, \alpha, \beta) - 2h_b(q)$$

where

$$\lambda(p, q, \alpha, \beta) \stackrel{\text{def}}{=} h_b(p \star q) + h_b\left(\frac{\alpha + \beta}{2}\right) - h_b\left(\left(\frac{\alpha + \beta}{2}\right) \star p \star q\right)$$

Let us consider the maximal sum-rate $R_1 + R_2$ guaranteed by the region above. To that end, we set $R_0 = 0$ and note it is sufficient to consider only the last inequality. We get:

$$C_{\text{NoisyFB},\Sigma} \ge \sup \left\{ \frac{1}{2} \left(h_b(\alpha) + h_b(\beta) + (1 - \beta)h_b\left(\frac{\alpha}{1 - \beta}\right) \right. \right.$$
$$\left. \left. + (1 - \alpha)h_b\left(\frac{\beta}{1 - \alpha}\right) \right) \right\} - 2h_b(q) \quad (7)$$

where the supremum is taken over all feasible α, β satisfying

$$h_b\left(\frac{\alpha + \beta}{2}\right) - \frac{1}{2}(h_b(\alpha) + h_b(\beta)) - \lambda(p, q, \alpha, \beta) = 0$$

For comparison, let us now upper bound the corresponding no-feedback sum-rate capacity $\mathcal{C}_{\text{NoFB},\Sigma}$. Since the no-feedback capacity of a DMBC depends only on the marginals $P_{Y_1|X}, P_{Y_2|X}$, the capacity region for the Noisy Blackwell channel remains the same if we replace Z with Z_1, Z_2 for Y_1, Y_2 respectively, where $Z_i \sim \text{Bern}(p)$ are independent. Computing the cut-set upper bound for this setting, we get

$$\mathcal{C}_{\text{NoFB},\Sigma} \le \sup_{\alpha \in (0, \frac{1}{2})} \left\{ H[(\alpha(p - \bar{p})^2 + p\bar{p}, \bar{p}^2 + 2\alpha p, \right.$$
$$\left. p^2 + 2\alpha\bar{p}, \alpha(p - \bar{p})^2 + p\bar{p})] - 2h_b(p) \right\}, \quad (8)$$
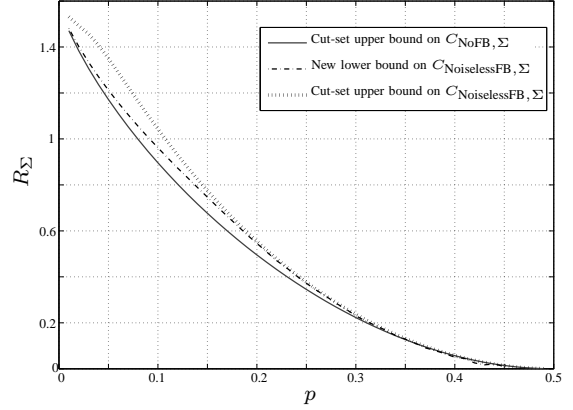

Fig. 2. Bounds on the sum-rate capacity of the noisy Blackwell DMBC

where $\bar{p} \stackrel{\text{def}}{=} 1 - p$. Figure 2 depicts the bounds (7) and (8) together with a cut-set upper bound on $C_{\text{NoiselessFB},\Sigma}$, where (7) is plotted in the noiseless case ($q = 0$). From the continuity of the bound (7) with respect to the noise level q, we conclude the following.

Observation 3. *For any $p \in (0, 1)$ and small enough q, noisy feedback strictly increases the capacity region of the Noisy Blackwell-DMBC. The statement holds also for the feedback $\widetilde{Y} = (Y_1 + Z', Y_2 + Z'')$ where Z', Z'' are independent.*

REFERENCES

[1] T. M. Cover and J. A. Thomas, "Elements of information theory", 2nd edition, Wiley publication.

[2] R. Gray and A. Wyner, "Source Coding for a Simple Network," *Bell System Tech. J.* vol. 48, pp. 1681–1721, Nov. 1974.

[3] R. Timo and A. Grant and T. Chan and G. Kramer, "Source coding for a simple network with receiver side information," in Proc. of ISIT 2008, Toronto, USA, July 6–11, 2008.

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

[5] L. Ozarow and S. Leung, "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, 1984.

[6] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. on Inf. Theory,* vol. 25, pp. 60-64, no. 3, 1979.

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

[8] G. Kramer, "Capacity Results for the Discrete Memoryless Network," *IEEE Trans. on Inf. Theory,* vol. 49, no. 1, pp. 4–21, 2003.

[9] E. C. van der Meulen, "A survey of multi-way channels in information theory: 1961-1976," *IEEE Trans. on Inf. Theory,* vol. 23, no. 1, pp. 1-37, 1977.

[10] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless Broadcast Channel with feedback", *in preparation*.

[11] A. El Gamal and Y-.H Kim, "Lecture Notes on Network Information Theory," *Stanford University and UCSD,* 2009. Available online: http://arxiv.org/abs/1001.3404

[12] S. I. Gelfand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Inf. Transm.,* vol. 16, no. 1, pp. 24-34, 1980.