Capacity Theorems for the Finite-State Broadcast Channel with Feedback

Ron Dabora and Andrea Goldsmith
Dept. of Electrical Engineering, Stanford University

***Abstract*—We consider the discrete, time-varying broadcast channel with memory under the assumption that the channel states belong to a set of finite cardinality. We study the achievable rates in two scenarios where feedback (and cooperation) is available. One scenario is the general finite-state broadcast channel (FSBC) where both receivers send feedback to the transmitter, and in addition one receiver sends his channel outputs to the other receiver through a cooperation link. The second scenario is the degraded FSBC where only the strong receiver sends feedback to the transmitter. We find the capacity regions for both cases. In both scenarios we consider non-indecomposable as well as a class of indecomposable FSBCs.**

I. Introduction

The finite-state channel (FSC) was introduced as early as 1953 to model time-varying channels with memory [1]. In this model memory is captured by the state of the channel at the end of the previous symbol transmission. To make this precise, let X denote the channel input, Y the channel output and S the channel state. The transition function of the FSC at symbol time i satisfies $p(y_i, s_i|x^i, y^{i-1}, s^{i-1}) = p(y_i, s_i|x_i, s_{i-1})$. Capacity analysis of time-varying channels with memory has been the focus of considerable interest, especially in the Gaussian setup. This has been motivated by the proliferation of mobile communications in which the channel is subject to multipath and correlated fading. The correlation of the fading process introduces memory into the time-varying channel. The capacity of non-indecomposable FSCs without feedback was originally studied by Gallager [2] and has been applied to specific channels, see [3] and references therein. Recently, the capacity of FSCs with feedback has been studied. Specifically, it was shown in [4] that feedback can increase the capacity of some FSCs with memory. The capacity of the FS-MAC with and without feedback was also considered recently [5].

In this work we focus on the capacity of finite-state broadcast channels with feedback. In this scenario a single transmitter communicates with two receivers over a finite-state channel characterized by $p(y, z, s|x, s')$ (see Figure 1 for notations), where s' denotes the state of the channel at the end of the previous symbol transmission. In a previous work we considered the degraded FSBC without feedback [6]. Specifically, we first defined the notions of physical and stochastic degradedness for channels with memory and then showed that capacity is achieved with a superposition codebook with memory. The capacity region was obtained as a limiting expression by taking the blocklength to infinity. In this work we study the effect of feedback and cooperation on the capacity region of the FSBC. The degraded FSBC with feedback is of special interest, as it was shown in [7] that feedback does not increase the capacity region of the physically degraded *memoryless* broadcast channel (BC), and it would be interesting to contrast this with the case with memory (see discussion in Section III).

The authors are with the Wireless Systems Lab, Department of Electrical Engineering, Stanford University, Stanford, CA 94305. Email: {ron,andrea}@wsl.stanford.edu. This work was supported in part by the DARPA ITMANET program under grant 1105741-1-TFIND.

In the context of discrete multi-user channels with states, we note that the capacity of the degraded arbitrarily varying BC (DAVBC) has been recently investigated in [8] and [9]. This channel is characterized by the transition function $p(y^n, z^n|x^n, s^n) = \prod_{i=1}^{n} p(y_i, z_i|x_i, s_i)$. This models a memoryless channel whose parameters vary with time in an arbitrary manner. In [8] DAVBCs with causal and non-causal side information at the transmitter were considered, and in [9] the capacity for DAVBCs with causal side information at the transmitter and non-causal side information at the good receiver was derived.

The most general scenario for the FSBC with feedback and cooperation is depicted in Figure 1.

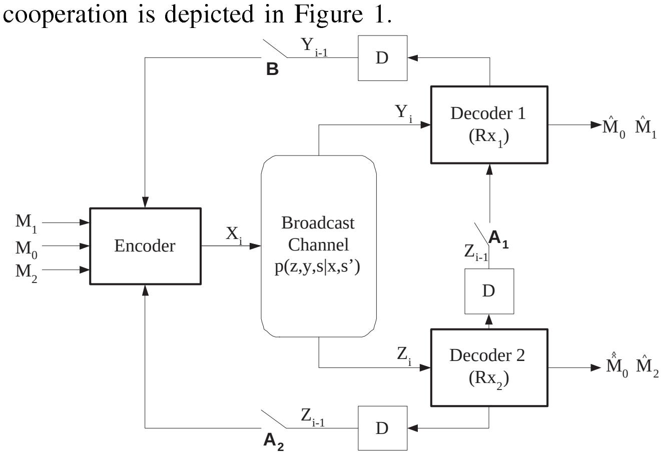

Fig. 1. The general FSBC with feedback and cooperation. D indicates a single symbol-time delay. Each of the three switches A_1, A_2 or B can be either open or closed. Switch A_1 enables cooperation from Rx_2 to Rx_1, switch A_2 enables feedback from Rx_2 to the transmitter, and switch B enables feedback from Rx_1 to the transmitter.

From Figure 1 we see that there are eight possible configurations. In this work we consider two of the eight scenarios (the scenario with all switches open was considered in [6]). We denote them by SC1 and SC2:

SC1: The *general* FSBC with all switches closed. Therefore, Rx_2 sends its channel output to both Rx_1 and the transmitter. This results in a physically degraded channel $X^n|s_0 \leftrightarrow (Y^n, Z^n)|s_0 \leftrightarrow Z^n|s_0$, but it is *not* required that $X^n|s_0 \leftrightarrow Y^n|s_0 \leftrightarrow Z^n|s_0$ form a Markov chain. The scenario in which the channel output at one receiver is available to the other receiver is also called the augmented BC [10]. Note that the delay from Rx_2 to Rx_1 is of no significance as the receivers can wait until the end of the entire block before beginning to decode.

SC2: The physically degraded BC with only switch B closed. Therefore, Rx_2 does not send feedback, and the codebook is generated without knowledge of the channel output at Rx_2. For SC2 we shall assume that the degradedness condition $X^n|s_0 \leftrightarrow Y^n|s_0 \leftrightarrow Z^n|s_0$ holds.

The scenario SC1 can be viewed as representing an uplink scenario that combines receiver cooperation and feedback: the weak (e.g. far) receiver sends its channel output to the strong receiver, and the feedback from the strong receiver to the transmitter contains both its own feedback and the feedback of the weak receiver, obtained through the cooperation link. The scenario SC2 can represent a cellular situation without cooperation in which the mobile Rx_2 is too far from the base station to reliably send feedback to it, but the mobile Rx_1 is close to the base station and thus has a reliable feedback path to it. In Section III we explain why, despite the fact that only one receiver sends feedback, the rates to both receivers increase. This is an important benefit of feedback in multi-user scenarios.

II. Channel Model

First, a word about notation. In the following we denote random variables with upper case letters, e.g. X, Y, and their realizations with lower case letters x, y. A random variable (RV) X takes values in a set $\mathcal{X}$. We use $||\mathcal{X}||$ to denote the cardinality of a finite, discrete set $\mathcal{X}$, $\mathcal{X}^n$ to denote the n-fold Cartesian product of $\mathcal{X}$, and $p_X(x)$ to denote the probability mass function (p.m.f.) of a discrete RV X on $\mathcal{X}$. For brevity we may omit the subscript X when it is obvious from the context. We use $p_{X|Y}(x|y)$ to denote the conditional p.m.f. of X given Y. We denote vectors with boldface letters, e.g. $\mathbf{x}$, $\mathbf{y}$; the i'th element of a vector $\mathbf{x}$ is denoted with x_i and we use x_i^j where $i < j$ to denote the vector $(x_i, x_{i+1}, ..., x_{j-1}, x_j)$; x^j is short form notation for x_1^j, and $\mathbf{x} \equiv x^n$. A vector of random variables is denoted by $\mathbf{X} \equiv X^n$, and similarly we define $X_i^j \triangleq (X_i, X_{i+1}, ..., X_{j-1}, X_j)$ for $i < j$. We use $H(\cdot)$ to denote the entropy of a discrete random variable and $I(\cdot;\cdot)$ to denote the mutual information between two random variables, as defined in [12, Chapter 2]. $I(\cdot;\cdot)_q$ denotes the mutual information evaluated with a p.m.f. q on the random variable and *co* $\mathcal{R}$ denotes the convex hull of the set $\mathcal{R}$. Finally, we recall the definitions of directed mutual information and causal conditioning (see also [5]):

$$I(X^n \to Y^n|Z^n) = \sum_{i=1}^n I(X^i; Y_i|Y^{i-1}, Z^n)$$

$$I(X^n \to Y^n||Z^n) = \sum_{i=1}^n I(X^i; Y_i|Y^{i-1}, Z^i)$$

$$Q(x^n||y^n, z^{n-1}) = \prod_{i=1}^n p(x_i|x^{i-1}, y^i, z^{i-1}).$$

Definition 1: The *finite-state broadcast channel* is defined by the triplet $\{\mathcal{X} \times \mathcal{S}, p(y, z, s|x, s'), \mathcal{Y} \times \mathcal{Z} \times \mathcal{S}\}$ where X is the input symbol, Y and Z are the output symbols, S' is the channel state at the end of the previous symbol transmission and S is the channel state at the end of the current symbol transmission. $\mathcal{S}$, $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ are discrete alphabets of finite cardinalities. The p.m.f of a block of n transmissions is

$$\begin{aligned}
&p(y^n, z^n, s^n, x^n|s_0) \\
&= \prod_{i=1}^n p(y_i, z_i, s_i, x_i|y^{i-1}, z^{i-1}, s^{i-1}, x^{i-1}, s_0) \\
&\overset{(a)}{=} \prod_{i=1}^n p(x_i|y^{i-1}, z^{i-1}, x^{i-1}) p(y_i, z_i, s_i|y^{i-1}, z^{i-1}, s^{i-1}, x^i, s_0) \\
&\overset{(b)}{=} \prod_{i=1}^n p(x_i|y^{i-1}, z^{i-1}, x^{i-1}) \prod_{i=1}^n p(y_i, z_i, s_i|x_i, s_{i-1}),
\end{aligned}$$

where s_0 is the initial state. Here, (a) is because the transmitter is oblivious of the channel states and (b) captures the fact that given S_{i-1}, the symbols at time i are independent of the past.

Definition 2: The FSBC is called *physically degraded* if its p.m.f. satisfies

$$p(y_i|x^i, y^{i-1}, z^{i-1}, s_0) = p(y_i|x^i, y^{i-1}, s_0), \quad \text{(1a)}$$

$$p(z_i|x^i, y^i, z^{i-1}, s_0) = p(z_i|y^i, z^{i-1}, s_0). \quad \text{(1b)}$$

Condition (1a) captures the intuitive notion of degradedness, namely that Z^{i-1} is a degraded version of Y^{i-1}, thus it does not add information when Y^{i-1} is given. Note that in the memoryless case this condition is not necessary as, given X_i, Y_i is independent of the history. Condition (1b) follows from the standard notion of degradedness, namely that Y^i makes Z^i independent of X^i. Note that condition (1b) does not eliminate memory. Using conditions (1a) and (1b) *in scenario SC2* we obtain (when $p(y^n, x^n|s_0) > 0$)

$$\begin{aligned}
&p(z^n|y^n, x^n, s_0) = \frac{p(z^n, y^n, x^n|s_0)}{p(y^n, x^n|s_0)} \\
&= \frac{\prod_{i=1}^n p(z_i, y_i, x_i|z^{i-1}, y^{i-1}, x^{i-1}, s_0)}{p(y^n, x^n|s_0)} \\
&= \frac{\prod_{i=1}^n p(x_i|z^{i-1}, y^{i-1}, x^{i-1}) \prod_{i=1}^n p(z_i, y_i|z^{i-1}, y^{i-1}, x^i, s_0)}{p(y^n, x^n|s_0)} \\
&\overset{(a)}{=} \frac{\prod_{i=1}^n p(x_i|y^{i-1}, x^{i-1}) \prod_{i=1}^n p(z_i, y_i|z^{i-1}, y^{i-1}, x^i, s_0)}{\prod_{i=1}^n p(x_i|y^{i-1}, x^{i-1}) \prod_{i=1}^n p(y_i|y^{i-1}, x^i, s_0)} \\
&\overset{(b)}{=} \frac{\prod_{i=1}^n p(y_i|y^{i-1}, x^i, s_0) \prod_{i=1}^n p(z_i|z^{i-1}, y^i, x^i, s_0)}{\prod_{i=1}^n p(y_i|y^{i-1}, x^i, s_0)} \\
&\overset{(c)}{=} \prod_{i=1}^n p(z_i|z^{i-1}, y^i, s_0)
\end{aligned}$$

where (a) is because A_2 is open in SC2, (b) follows from (1a) and (c) follows from (1b). We therefore conclude that when (1) holds, $p(z^n|y^n, x^n, s_0) = p(z^n|y^n, s_0)$. Hence,

$$p(y^n, z^n|x^n, s_0) = p(y^n|x^n, s_0) p(z^n|y^n, s_0). \quad (2)$$

Note that Z^n is a degraded version of Y^n but still depends on the state sequence (i.e. degradedness does not eliminate the memory). A special case of the physically degraded FSBC occurs when in (1b) we have $p(z_i|x^i, y^i, z^{i-1}, s_0) = p(z_i|y_i)$. Hence,

$$p(z^n|y^n, s_0) = p(z^n|y^n) = \prod_{i=1}^n p(z_i|y_i). \quad (3)$$

Equation (3) is similar to the definition of degradedness for the DAVBC used in [8]. Condition (3) does not constitute only a mathematical convenience, but represents a physical scenario, as shown in [6].

Definition 3: The FSBC is called *stochastically degraded* if there exists a p.m.f. $\tilde{p}(z|y)$ such that

$$p(z^n|x^n, s_0) = \sum_{\mathcal{Y}^n} p(z^n, y^n|x^n, s_0) = \sum_{\mathcal{Y}^n} p(y^n|x^n, s_0) \prod_{i=1}^{n} \tilde{p}(z_i|y_i). \quad (4)$$

Definition 4: (see [2, Section 4.6]) The FSBC is called *indecomposable* if for every $\epsilon > 0$ there exists $N_0(\epsilon)$ such that for all $n > N_0(\epsilon)$, $|p(s_n|\mathbf{x}, s_0) - p(s_n|\mathbf{x}, s_0')| < \epsilon$, for all s_n, $\mathbf{x}$, and initial states s_0 and s_0'.

Definition 5: An (R_0, R_1, R_2, n) *deterministic code* for the FSBC with feedback consists of three message sets, $\mathcal{M}_0 = \{1, 2, ..., 2^{nR_0}\}$, $\mathcal{M}_1 = \{1, 2, ..., 2^{nR_1}\}$ and $\mathcal{M}_2 = \{1, 2, ..., 2^{nR_2}\}$, and a collection of mappings $(\{f_i\}_{i=1}^n, g_{\mathbf{y}}, g_{\mathbf{z}})$ such that

$$f_i : \mathcal{M}_0 \times \mathcal{M}_1 \times \mathcal{M}_2 \times \mathcal{Z}^{i-1} \times \mathcal{Y}^{i-1} \mapsto \mathcal{X} \quad (5)$$

is the encoder, and

$$g_{\mathbf{y}} : \mathcal{Y}^n \mapsto \mathcal{M}_0 \times \mathcal{M}_1, \quad g_{\mathbf{z}} : \mathcal{Z}^n \mapsto \mathcal{M}_0 \times \mathcal{M}_2,$$

are the decoders. Here, $\mathcal{M}_0$ is the set of common messages and $\mathcal{M}_1$ and $\mathcal{M}_2$ are the sets of private messages to Rx_1 and Rx_2, respectively. Note that we assume no knowledge of the states at the transmitter and receivers.

Definition 6: The *average probability of error* of a code of blocklength n is given by $\max_{s_0 \in \mathcal{S}} P_e^{(n)}(s_0)$ where

$$P_e^{(n)}(s_0) = \Pr\left(g_{\mathbf{y}}(Y^n) \neq (M_0, M_1) \text{ or } g_{\mathbf{z}}(Z^n) \neq (M_0, M_2)|s_0\right),$$

and the messages $M_0 \in \mathcal{M}_0$, $M_1 \in \mathcal{M}_1$ and $M_2 \in \mathcal{M}_2$ are selected independently and uniformly.

Remark: In Definitions 5 and 6 we assume no receiver cooperation. However, since in SC1 Z^n is also available at Rx_1, then for this scenario we modify Definitions 5 and 6 to include cooperation by replacing $g_{\mathbf{y}}(Y^n)$ with $g_{\mathbf{y}}(Y^n, Z^n)$.

Definition 7: A rate triplet (R_0, R_1, R_2) is called *achievable* for the FSBC if for every $\epsilon > 0$ and $\delta > 0$ there exists an $n(\epsilon, \delta) \in \mathbb{N}$ such that $\forall n > n(\epsilon, \delta)$ an $(R_0 - \delta, R_1 - \delta, R_2 - \delta, n)$ code with $P_e^{(n)}(s_0) \leq \epsilon$, $\forall s_0 \in \mathcal{S}$ can be constructed.

Definition 8: The *capacity region* of the FSBC is the convex hull of all achievable rate triplets.

III. Main Results and Discussion

The main results are stated in the following theorems. The proof of Theorem 1 is outlined in Section IV. The proof of Theorem 2 follows from similar arguments and is thus omitted.

Theorem 1 (capacity region for SC1): For the FSBC $p(y, z, s|x, s')$, let $\mathcal{Q}_n^{SC1}$ be the set of all distributions on $(\times_{i=1}^n \mathcal{U}_i \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n \times \mathcal{S}^n)$ that satisfy

$$\mathcal{Q}_n^{SC1} = \Big\{ q(u^n, x^n, y^n, z^n, s^n|s_0) : q(u^n, x^n, y^n, z^n, s^n|s_0) = \prod_{i=1}^{n} p(u_i|u^{i-1}, z^{i-1}) \times p(x_i|x^{i-1}, z^{i-1}, y^{i-1}, u^i) p(y_i, z_i, s_i|x_i, s_{i-1}) \text{ for all } s_0 \in \mathcal{S} \Big\}, \quad \forall n \in \mathbb{N}.$$

Define the region $\mathcal{R}_n^{SC1}$ as

$$\mathcal{R}_n^{SC1} = co \bigcup_{q_n \in \mathcal{Q}_n^{SC1}} \Big\{ (R_0, R_1, R_2) : R_0 \geq 0, R_1 \geq 0, R_2 \geq 0, R_1 \leq \min_{s_0' \in \mathcal{S}} \frac{1}{n} I(X^n \to Z^n, Y^n || U^n, s_0')_{q_n}, R_0 + R_2 \leq \min_{s_0 \in \mathcal{S}} \frac{1}{n} I(U^n \to Z^n | s_0)_{q_n} \Big\}.$$

For the general FSBC with feedback such that switches A_1, A_2 and B are closed, the capacity region is given by

$$\mathcal{C}_{fb}^{SC1} = \lim_{n \to \infty} \mathcal{R}_n^{SC1}, \quad (6)$$

and the limit exists.

The definition of the limit of regions can be found, e.g., in [5]. Here, the auxiliary RV U^n represents the information transmitted to Rx_2. When feedback and cooperation are not available from Rx_2 and the channel is physically degraded we obtain the following capacity region:

Theorem 2 (capacity region for SC2): Let $\mathcal{Q}_n^{SC2}$ be the set of all distributions on $(\mathcal{U}^n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n \times \mathcal{S}^n)$ that satisfy

$$\mathcal{Q}_n^{SC2} = \Big\{ q(u^n, x^n, y^n, z^n, s^n|s_0) : q(u^n, x^n, y^n, z^n, s^n|s_0) = p(u^n) \prod_{i=1}^{n} p(x_i|x^{i-1}, y^{i-1}, u^n) p(y_i, z_i, s_i|x_i, s_{i-1}) \text{ for all } s_0 \in \mathcal{S} \text{ and } ||\mathcal{U}^n|| \leq \min\{||\mathcal{X}|| \cdot ||\mathcal{Y}||, ||\mathcal{Z}||\}^n + 1 \Big\}, \forall n \in \mathbb{N},$$

where $p(y, z, s|x, s')$ is the physically degraded FSBC satisfying Equation (2)*. Define the region $\mathcal{R}_n^{SC2}$ as*

$$\mathcal{R}_n^{SC2} = co \bigcup_{q_n \in \mathcal{Q}_n^{SC2}} \Big\{ (R_0, R_1, R_2) : R_0 \geq 0, R_1 \geq 0, R_2 \geq 0, R_1 \leq \min_{s_0' \in \mathcal{S}} \frac{1}{n} I(X^n \to Y^n | U^n, s_0')_{q_n}, R_0 + R_2 \leq \min_{s_0 \in \mathcal{S}} \frac{1}{n} I(U^n; Z^n | s_0)_{q_n} \Big\}.$$

For the general FSBC with feedback such that switch B is closed, the capacity region is given by

$$\mathcal{C}_{fb}^{SC2} = \lim_{n \to \infty} \mathcal{R}_n^{SC2}, \quad (7)$$

and the limit exists.

Since the capacity of the broadcast channel depends only on the conditional marginals $p(y^n|x^n, s_0)$ and $p(z^n|x^n, s_0)$ (see [12, Chapter 14.6]), the capacity region of the stochastically degraded FSBC for scenario SC2 is the same as the corresponding physically degraded FSBC:

Corollary 1: For the stochastically degraded FSBC of Definition 3, the capacity region is given by Theorem 2 with the appropriate $\tilde{p}(z|y)$ satisfying (4)*.*

The finite-state Markov BC (FSMBC) is defined by

$$p(y, z, s|x, s') = p(s|s') p(y, z|x, s').$$

When the Markov chain is homogeneous, irreducible and aperiodic then the FSMBC is an indecomposable channel, denoted H-FSMBC. For the H-FSMBC we have the following corollary:

Corollary 2:

1) *For the H-FSMBC of scenario SC1, the capacity region is given by Theorem 1 where, in the definition of $\mathcal{R}_n^{SC1}$, the conditioning on s_0 and s_0' are omitted from the mutual information expressions.*
2) *For the physically degraded H-FSMBC of scenario SC2, the capacity region is obtained from Theorem 2 where, in the definition of $\mathcal{R}_n^{SC2}$, s_0 and s_0' are omitted from the mutual information expressions.*
3) *For the stochastically degraded H-FSMBC of scenario SC2, the capacity region is obtained from Corollary 1 where, in the definition of $\mathcal{R}_n^{SC2}$, s_0 and s_0' are omitted from the mutual information expressions.*

Proof outline: Loosely speaking, the corollary is true since for n large enough the effect of the initial state fades away. Therefore, the maximum over all $s_0 \in \mathcal{S}$ equals the minimum. Hence, for all initial states the limits for $n \to \infty$ are the same.

Discussion

We make the following observations:

1) We note that in both scenarios we actually have physically degraded situations. In SC1 the physical degradedness is due to the cooperation link (switch A_1) and the Markov chain $X^n \leftrightarrow (Y^n, Z^n) \leftrightarrow Z^n$. In SC2 the channel is physically degraded by definition of the scenario. Therefore, in the derivation we need to consider only the two private messages case as the common message can be incorporated by splitting the rate to Rx_2 into private and common rates, as in [12, Theorem 14.6.4].

2) Since in SC1 the channel is effectively a physically degraded broadcast channel, a superposition *codetree* achieves capacity. Interpreting superposition coding in terms of "cloud centers" (U^n) and "cloud elements" ($X^n|U^n$) we note that in SC1 the cloud centers are in fact codetrees on which cloud elements, which are also codetrees, are superimposed.

3) It was shown in [7] that feedback does not increase the capacity of the physically degraded memoryless BC. The intuition is as follows: in the physically degraded memoryless BC, Y is a memoryless transformation of X and Z is a memoryless transformation of Y. Now, as feedback does not help the memoryless point-to-point channel it cannot help the cascade of two such channels. As explained earlier, when the channel has memory, feedback can increase its capacity. Feedback also helps the stochastically degraded BC [10]. Therefore, feedback can increase the capacity of the FSBC.

4) In SC1, both receivers decode the cloud center based on Z^n, as Z^n is available also at Rx_1 through the cooperation link. Rx_1 now proceeds to decode the cloud element. However, since the channel is not physically degraded, after decoding the cloud center using Z^n, Rx_1 uses both (Z^n, Y^n) rather than only Y^n to decode the cloud element. This is because $p(z^n|x^n, y^n, s_0) \neq p(z^n|y^n, s_0)$.

5) A superposition codebook structure achieves capacity for both SC1 and SC2. This introduces a structural constraint when optimizing the codebook for achieving the maximal rate pairs. Note, however, that for SC1 we are not able to establish a cardinality bound for the auxiliary RV.

6) In SC2 feedback is available only from Rx_1. However, due to the physically degraded structure, this helps both receivers: degradedness constrains the sum-rate to be bounded by the sum-rate at Rx_1. When this sum-rate is increased, then for the same rate R_1 it is possible to obtain a higher rate R_2. The maximum rate to Rx_2, though, remains the same.

IV. Proof Outline for Theorem 1 (SC1)

A. Achievability

The achievability proof consists of the following steps:

- Fix a collection of probability distributions $\{p(u_i|u^{i-1}, z^{i-1}), p(x_i|u^i, y^{i-1}, z^{i-1}, x^{i-1})\}_{i=1}^n \triangleq \{Q_1, Q_2\}$.
- Using maximum-likelihood decoding at Rx_2 according to

$$\arg\max_{m_2} \sum_{s_0 \in \mathcal{S}} \frac{1}{||\mathcal{S}||} \prod_{i=1}^n p(z_i|z^{i-1}, u^i, s_0)$$

we conclude that a positive error exponent for decoding M_2 can be obtained as long as $R_2 \le \min_{s_0 \in \mathcal{S}} \frac{1}{n} I(U^n \to Z^n|s_0) - \frac{\log_2 ||\mathcal{S}||}{n} \triangleq I_2^n(Q_1, Q_2)$. As Rx_1 receives Z^n through the cooperation link, then also for decoding M_2 at Rx_1 the error exponent is positive.

- Using maximum-likelihood decoding at Rx_1 according to

$$\arg\max_{m_1} \sum_{s_0 \in \mathcal{S}} \frac{1}{||\mathcal{S}||} \prod_{i=1}^n p(y_i, z_i|x^i, y^{i-1}, z^{i-1}, s_0)$$

we conclude that a positive error exponent for decoding M_1 at Rx_1, given that M_2 was correctly decoded, can be achieved as long as $R_1 \le \min_{s_0 \in \mathcal{S}} I(X^n \to Y^n, Z^n||U^n, s_0) - \frac{\log_2 ||\mathcal{S}||}{n} \triangleq I_1^n(Q_1, Q_2)$.

- Next, for a fixed n and some integer b let $n_0 = nb$. Construct distributions by taking the product of the basic distribution for a block of n symbols b times:

$$\begin{aligned}
&\tilde{Q}_2(x^{n_0}||u^{n_0}, y^{n_0-1}, z^{n_0-1}) \\
&\quad = \prod_{b'=1}^{b} \prod_{i=1}^{n} p\Big(x_{(b'-1)n+i}\Big|x_{(b'-1)n+1}^{(b'-1)n+i-1}, \\
&\qquad\qquad y_{(b'-1)n+1}^{(b'-1)n+i-1}, z_{(b'-1)n+1}^{(b'-1)n+i-1}, u_{(b'-1)n+1}^{(b'-1)n+i}\Big) \\
&\tilde{Q}_1(u^{n_0}||z^{n_0-1}) \\
&\quad = \prod_{b'=1}^{b} \prod_{i=1}^{n} p\Big(u_{(b'-1)n+i}\Big|z_{(b'-1)n+1}^{(b'-1)n+i-1}, u_{(b'-1)n+1}^{(b'-1)n+i-1}\Big).
\end{aligned}$$

Now, taking b large enough results in an average probability of error that is arbitrarily small, hence $(I_1^n(Q_1, Q_2), I_2^n(Q_1, Q_2))$ is achievable.

- Finally, for $\lambda > 0$ define $C_{fb,SC1}^n(\lambda)$:

$$C_{fb,SC1}^n(\lambda) = \max_{Q_1(u^n||z^{n-1})Q_2(x^n||u^n, z^{n-1}, y^{n-1})} F_n(\lambda, Q_1, Q_2), \quad (8)$$

$$\begin{aligned}
F_n(\lambda, Q_1, Q_2) = \Big\{ &\min_{s_0 \in \mathcal{S}} \frac{1}{n} I(U^n \to Z^n|s_0) \\
&+\lambda \min_{s_0' \in \mathcal{S}} \frac{1}{n} I(X^n \to Z^n, Y^n||U^n, s_0') \Big\} - (1+\lambda)\frac{\log ||\mathcal{S}||}{n}.
\end{aligned}$$

We show that $C_{fb,SC1}^n(\lambda)$ is sup-additive: $C_{fb,SC1}^\infty(\lambda) \triangleq \lim_{n\to\infty} C_{fb,SC1}^n(\lambda) = \sup_n C_{fb,SC1}^n(\lambda)$. Therefore, the boundary of the achievable region can be written as

$$R_2(R_1) = \inf_{0 \le \lambda \le 1} \left(C_{fb,SC1}^\infty(\lambda) - \lambda R_1 \right). \quad (9)$$

B. Converse

Theorem 3: If for some $\lambda > 0$,

$$R_2 + \lambda R_1 > C^{\infty}_{fb,SC1}(\lambda) + \epsilon,$$

then there exist initial states $s'_0, s_0 \in \mathcal{S}$ *for which*

$$P^{(n)}_{e2}(s_0)R_2 + \lambda P^{(n)}_{e1}(s'_0)R_1 > \epsilon - (1+\lambda)\frac{1 + \log ||\mathcal{S}||}{n}. \quad (10)$$

The implication of (10), as explained in [11], is that for n large enough the probability of error cannot be made arbitrarily small, outside the region whose boundary is given by (9).

Proof: Recall that $P^{(n)}_{e2}(s_0)$ and $P^{(n)}_{e1}(s_0)$ denote probabilities of error for initial state s_0, when the decoders are ignorant of the initial state. From Fano's inequality we have that for initial state s_0

$$H(M_2|Z^n, s_0) \le P^{(n)}_{e2}(s_0)nR_2 + 1 \quad (11a)$$

$$H(M_1|Z^n, Y^n, s'_0) \le P^{(n)}_{e1}(s'_0)nR_1 + 1. \quad (11b)$$

Denote with $s_{0,n}$ the initial state that maximizes $H(M_2|Z^n, s_0)$ and with $s'_{0,n}$ the initial state that maximizes $H(M_1|Z^n, Y^n, s'_0)$. Now, note that

$$\begin{aligned}\min_{s_0\in\mathcal{S}} I(M_2; Z^n|s_0) &= \min_{s_0\in\mathcal{S}} \{H(M_2|s_0) - H(M_2|Z^n, s_0)\} \\ &= nR_2 - \max_{s_0\in\mathcal{S}} H(M_2|Z^n, s_0), \quad (12)\end{aligned}$$

$$\begin{aligned}&\min_{s'_0\in\mathcal{S}} I(M_1; Z^n, Y^n|M_2, s'_0) \\ &\quad = nR_1 - \max_{s'_0\in\mathcal{S}} H(M_1|Y^n, Z^n, M_2, s'_0) \\ &\quad \ge nR_1 - \max_{s_0\in\mathcal{S}} H(M_1|Y^n, Z^n, s'_0). \quad (13)\end{aligned}$$

We next have

$$\begin{aligned}I(M_2; Z^n|s_0) &= \sum_{i=1}^{n}\big(H(Z_i|Z^{i-1}, s_0) - H(Z_i|Z^{i-1}, M_2, s_0)\big) \\ &= \sum_{i=1}^{n}\big(H(Z_i|Z^{i-1}, s_0) - H(Z_i|Z^{i-1}, U^i, s_0)\big) \\ &= I(U^n \to Z^n|s_0)\end{aligned}$$

where $U_i = (M_2, Z^{i-1})$, $i = 1, 2, ..., n$. Note that $U_i - (U^{i-1}, Z^{i-1}) - Y^{i-1}$ and also $U^i|Z^{i-1}, Y^{i-1}, s_0 - X^i|Z^{i-1}, Y^{i-1}, s_0 - Y_i, Z_i|Z^{i-1}, Y^{i-1}, s_0$.

We also have that

$$\begin{aligned}&I(M_1; Y^n, Z^n|M_2, s'_0) \\ &= \sum_{i=1}^{n}\big(H(Z_i, Y_i|Z^{i-1}, Y^{i-1}, M_2, s'_0) \\ &\qquad - H(Z_i, Y_i|Z^{i-1}, Y^{i-1}, M_1, M_2, s'_0)\big) \\ &\le \sum_{i=1}^{n}\big(H(Z_i, Y_i|Z^{i-1}, Y^{i-1}, U^i, s'_0) \\ &\qquad - H(Z_i, Y_i|Z^{i-1}, Y^{i-1}, X^i, U^i, s'_0)\big) \\ &= I(X^n \to Z^n, Y^n||U^n, s'_0).\end{aligned}$$

Combining the above we have that for our choice of U^n:

$$\begin{aligned}&\min_{s_0\in\mathcal{S}} I(M_2; Z^n|s_0) + \lambda \min_{s'_0\in\mathcal{S}} I(M_1; Z^n, Y^n|M_2, s'_0) \\ &\le \min_{s_0\in\mathcal{S}} I(U^n \to Z^n|s_0) + \lambda \min_{s'_0\in\mathcal{S}} I(X^n \to Z^n, Y^n||U^n, s'_0) \\ &\le nC^n_{fb,SC1}(\lambda) + (1+\lambda)\log||\mathcal{S}|| \\ &\qquad \le nC^{\infty}_{fb,SC1}(\lambda) + (1+\lambda)\log||\mathcal{S}||, \quad (14)\end{aligned}$$

since $C^n_{fb,SC1}(\lambda)$ is obtained by maximizing over all joint distributions $Q_1(u^n||z^{n-1})Q_2(x^n||u^n, z^{n-1}, y^{n-1})$ and also because $C^n_{fb,SC1}(\lambda)$ is sup-additive.

Plugging (12) and (13) into (14) yields

$$\begin{aligned}&nR_2 - H(M_2|Z^n, s_{0,n}) + \lambda(nR_1 - H(M_1|Z^n, Y^n, s'_{0,n})) \\ &\qquad \le nC^{\infty}_{fb,SC1}(\lambda) + (1+\lambda)\log||\mathcal{S}|| \\ &\Rightarrow H(M_2|Z^n, s_{0,n}) + \lambda H(M_1|Z^n, Y^n, s'_{0,n}) + (1+\lambda)\log||\mathcal{S}|| \\ &\qquad \ge n\left(R_2 + \lambda R_1 - C^{\infty}_{fb,SC1}(\lambda)\right) > n\epsilon.\end{aligned}$$

Combined with Fano's inequalities (11), we obtain (10), which means that at least one of the states $s_{0,n}$, $s'_{0,n}$ results in a probability of error (at the respective receiver) that is bounded away from zero, completing the proof of the converse. ∎

V. Conclusions

We have derived the capacity region of a two-user FSBC with feedback under two different assumptions about the nature of the feedback and user cooperation. An important property of the first scenario is that the channel output at one receiver is also available to the other receiver through cooperation. When this cooperation is not possible, then the first user's receiver cannot use information about received signal at the second receiver to decode its own message. This implies that a superposition codebook is not necessarily optimal, even if the channel is physically degraded. In our second model, as decoder 2's channel output is not available at the transmitter and the channel is physically degraded, a superposition codebook in which the cloud centers are generated without feedback is optimal. An important property of feedback in multi-user scenarios is that feedback from one user can help other users as well.

References

[1] B. McMillan. "The Basic Theorems of Information Theory". *The Annals of Mathematical Statistics*, Vol-24(2):196–219, 1953.

[2] R. G. Gallager. *Information Theory and Reliable Communication.* John Wiley and Sons Inc., 1968.

[3] T. Holliday, A. Goldsmith, and P. Glynn. "On Entropy and Lyapunov Exponents for Finite State Channels". *IEEE Tr. on IT*, IT-52(8), 2006.

[4] H. Permuter, P. Cuff, B. Van Roy, and T. Weissman. "Capacity of the Trapdoor Channel with Feedback". Submitted to *IEEE Trans. on IT.*

[5] H. Permuter and T. Weissman. "Capacity Region of the Finite-State Multiple Access Channel with and without Feedback". Submitted to the *IEEE Trans. Inform. Theory*, 2007.

[6] R. Dabora and A. Goldsmith. "The Capacity of the Degraded Finite-State Broadcast Channel". Accepted to the *Information Theory Workshop (ITW)*, 2008.

[7] A. El-Gamal. "The Feedback Capacity of Degraded Broadcast Channels". *IEEE Trans. Inform. Theory*, IT-24(3):379–381, 1978.

[8] Y. Steinberg. "Coding for the Degraded Broadcast Channel With Random Parameters, With Causal and Noncausal Side Information". *IEEE Trans. Inform. Theory*, IT-51(8):2867–2877, 2005.

[9] A. Winshtok and Y. Steinberg. "The Arbitrarily Varying Degraded Broadcast Channel with States Known at the Encoder". *International Symp. on Inform. Theory (ISIT)* 2006, Seattle, WA, pp. 2156–2160.

[10] L. H. Ozarow. *Coding and Capacity for Additive White Gaussian Noise Multi-User Channels with Feedback* Ph.D. dissertation, Mass. Inst. Tech., Cambridge, MA, May 1979.

[11] R. G. Gallager. "Capacity and Coding for Degraded Broadcast Channels". *Problemy Peredachi Informatsii*, vol. 10(3):3–14, 1974.

[12] T. M. Cover and J. Thomas. *Elements of Information Theory.* John Wiley and Sons Inc., 1991.