# Multiple Access Channels with Intermittent Feedback and Side Information

Ashish Khisti
Dept. of Electrical and Computer Engineering
University of Toronto
Toronto, ON, M5S 3G4, Canada
Email: akhisti@comm.utoronto.ca

Amos Lapidoth
Dept. of Information Technology & Electrical Engineering
ETH Zurich,
Zurich, 8092, Switzerland
Email: lapidoth@isi.ee.ethz.ch

*Abstract*—We study two multiple-access scenarios with encoders that are informed only intermittently. The first is the Gaussian multiple-access channel with an intermittent feedback link. Here we assume that, depending on the current binary state which evolves in a memoryless fashion, the previous channel output is either revealed to the two encoders or not. For this scenario we obtain an outer bound on the capacity region that approaches the capacity region without feedback when the probability that the channel output will be fed back approaches zero. We also propose an inner bound that converges to the capacity region with ideal feedback and the capacity region with no feedback in the associated extreme cases.

In the second scenario the encoders always observe ideal feedback, and in addition they can crib intermittently. For this scenario we establish the capacity region for the special class of semi-deterministic multiple-access channels. The capacity is achieved using the Superposition Block Markov Coding technique of Cover and Leung.

For both scenarios the outer bounds are tighter than those obtained by revealing the underlying state sequence non-causally to the encoders.

## I. INTRODUCTION

It is well known that the capacity region of the multiple-access channel (MAC) is achieved using independent input symbols at the encoders. However when some side information is made available to the encoders, the input symbols can be dependent and the capacity region is in general unknown. For the MAC with feedback, Cover and Leung [1] propose the Superposition Block Markov Coding technique, which introduces dependence between the input symbols, and Willems establishes its optimality for a special class of channels [2]. For the two-user Gaussian MAC, Ozarow [3] proposes an iterative refinement scheme that can attain any point on the cut-set outer bound. Some extensions to the case of imperfect feedback have appeared recently [4]–[6]. A different form of encoder side information is considered in the cribbing MAC model introduced by Willems and van der Mulen [7]. In this setup each encoder is revealed the input symbols generated by the other encoder. Both causal and strictly causal side information are considered and the capacity region has been characterized for each case. For some recent extensions of the cribbing setup see, e.g., [8] and references therein.

In this paper we study the case when the encoders are revealed side information only intermittently. In section II, we study an AWGN MAC with intermittent feedback. At each
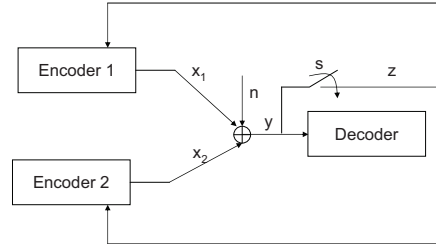


Fig. 1: AWGN MAC with Intermittent Feedback. The switch is closed when $s = 0$ and the encoders observe $z = y$. The switch remains open when $s = 1$ and the encoders observe $z = \star$. The state sequence $s^n$ is i.i.d.

time instant — depending on a binary valued state variable, which evolves in a memoryless fashion — the encoders either observe the previous channel output or an erasure symbol $\star$. A practical motivation for this model is that the destination node may need to harvest energy for providing feedback. When sufficient energy cannot be harvested, the destination will not be able to provide such feedback and an erasure is declared. In section III we study another MAC where the encoders are always revealed perfect feedback, but crib intermittently. At each time instant — depending on the state variable — each encoder either observes the previous input symbol of the other encoder or both observe an erasure. A motivation for this setup is that while the encoders may be located close to one another to overhear their transmissions, they could also be interfered at certain times by other nodes in the network which can lead to erasures. We also briefly discuss an extension when only one encoder can crib.

## II. AWGN MAC WITH INTERMITTENT FEEDBACK

We consider a two-user memoryless AWGN MAC with input symbols $x_1$ and $x_2$, output symbol $y$, and a feedback symbol $z$. The channel output symbol at time $i$ is $y(i) = x_1(i) + x_2(i) + n(i)$ where $n(i) \sim \mathcal{N}(0, 1)$ is independently sampled at each time, and we impose average power constraints $E[\frac{1}{N} \sum_{i=1}^{N} x_k^2(i)] \leq P_k$, for $k = 1, 2$. We associate a binary valued state sequence $s^N$ with the feedback link. Let $\{s(i) = 1\}$ denote the erasure event i.e., $z(i) = \star$, and $\{s(i) = 0\}$ denote the feedback event i.e., $z(i) = y(i)$. We

assume throughout that the state sequence is sampled i.i.d. with

$$\Pr(\mathsf{s}(i) = 1) = \beta, \quad \Pr(\mathsf{s}(i) = 0) = 1 - \beta \triangleq \bar{\beta}. \quad (1)$$

We assume that the state sequence is revealed in a strictly causal manner to all the three terminals. The input symbols at each time can depend on all the past observations at the encoder i.e., $\mathsf{x}_k(i) = f_{k,i}(\mathsf{w}_k, \mathsf{s}^{i-1}, \mathsf{z}^{i-1})$, for $k = 1, 2$. After $N$ channel uses the decoder generates an estimate $(\hat{\mathsf{w}}_1, \hat{\mathsf{w}}_2) = g(\mathsf{y}^N, \mathsf{s}^N)$. We say that a rate pair $(R_1, R_2)$ is achievable if there exists a sequence of encoding and decoding functions such that the average error probability approaches zero as $N \to \infty$. The capacity region is the union of all achievable rate pairs.

*A. Outer Bound*

**Theorem 1.** *Any achievable rate pair $(R_1, R_2)$ satisfies the following:*

$$R_1 \leq \frac{\bar{\beta}}{2} \log\left(1 + \frac{P_1}{\sigma_1^2}\right) + \frac{\beta}{2} \log(1 + P_1) \quad (2)$$

$$R_2 \leq \frac{\bar{\beta}}{2} \log\left(1 + \frac{P_2}{\sigma_2^2}\right) + \frac{\beta}{2} \log(1 + P_2) \quad (3)$$

$$R_1 + R_2 \leq \bar{\beta}\left\{\frac{1}{2} \log\left(1 + \frac{P_1}{\sigma_1^2}\right) + \frac{1}{2} \log\left(1 + \frac{P_2}{\sigma_2^2}\right)\right\}$$
$$+ \beta\frac{1}{2} \log(P_1 + P_2 + 1). \quad (4)$$

*The above inequalities hold for every pair $(\sigma_1^2, \sigma_2^2)$ with $\sigma_1^2 + \sigma_2^2 = 1$.*

**Remark 1.** *As the erasure probability on the feedback link approaches 1, the capacity region in Theorem 1 approaches the MAC capacity region without feedback.*

**Remark 2.** *The outer bound in Theorem 1 precludes power adaptation based on the state variable. An approach that reveals the sequence $\mathsf{s}^n$ non-causally to the encoders needs to permit such adaptation and in general yields a weaker outer bound.*

We establish Theorem 1 in the rest of this sub-section. We consider an enhanced MAC with the following property. When $\mathsf{s} = 1$, as before we let $\mathsf{z} = \star$ and the receiver output is $\mathsf{y} = \bar{\mathsf{y}} \triangleq \mathsf{x}_1 + \mathsf{x}_2 + \mathsf{n}$. When $\mathsf{s} = 0$, we let $\mathsf{y} = \mathsf{z} = (\mathsf{u}_1, \mathsf{u}_2)$ where

$$\mathsf{u}_1 = \mathsf{x}_1 + \mathsf{n}_1, \quad \mathsf{u}_2 = \mathsf{x}_2 + \mathsf{n}_2 \quad (5)$$

and the noise variables $\mathsf{n}_1$ and $\mathsf{n}_2$ are independent Gaussian random variables of variances $\sigma_1^2$ and $\sigma_2^2$ respectively that satisfy $\mathsf{n} = \mathsf{n}_1 + \mathsf{n}_2$. Note that the output symbol for the original MAC satisfies $\mathsf{y} = \mathsf{u}_1 + \mathsf{u}_2$ and thus the proposed channel indeed enhances the original channel. It has the following additional property which is useful in establishing the outer bound.

**Lemma 1.** *For each $i = 1, \ldots, N$ we have that $\mathsf{x}_{1,i} \to (\mathsf{s}^{i-1}, \mathsf{z}^{i-1}) \to \mathsf{x}_{2,i}$.*

The proof of Lemma 1 appears in Appendix A

We begin by upper bounding the sum-rate as follows.

$$n(R_1 + R_2) = H(\mathsf{w}_1, \mathsf{w}_2) \quad (6)$$

$$= H(\mathsf{w}_1, \mathsf{w}_2 | \mathsf{y}^N, \mathsf{s}^N) + I(\mathsf{w}_1, \mathsf{w}_2; \mathsf{y}^N, \mathsf{s}^N) \quad (7)$$

$$\leq N\epsilon_N + I(\mathsf{w}_1, \mathsf{w}_2; \mathsf{y}^N, \mathsf{s}^N) \quad (8)$$

where the last step follows from Fano's inequality since $(\mathsf{w}_1, \mathsf{w}_2)$ is decoded from $(\mathsf{y}^N, \mathsf{s}^N)$ with high probability. Now consider

$$I(\mathsf{w}_1, \mathsf{w}_2; \mathsf{y}^N, \mathsf{s}^N) = \sum_{i=1}^{N} I(\mathsf{w}_1, \mathsf{w}_2; \mathsf{y}(i), \mathsf{s}(i) | \mathsf{y}^{i-1}, \mathsf{s}^{i-1}) \quad (9)$$

$$= \sum_{i=1}^{N} I(\mathsf{w}_1, \mathsf{w}_2; \mathsf{y}(i) | \mathsf{y}^{i-1}, \mathsf{s}^{i-1}, \mathsf{s}(i)) \quad (10)$$

$$\leq \sum_{i=1}^{N} I(\mathsf{x}_1(i), \mathsf{x}_2(i), \mathsf{w}_1, \mathsf{w}_2; \mathsf{y}(i) | \mathsf{y}^{i-1}, \mathsf{s}^{i-1}, \mathsf{s}(i))$$

$$= \sum_{i=1}^{N} I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{y}(i) | \mathsf{y}^{i-1}, \mathsf{s}^{i-1}, \mathsf{s}(i)) \quad (11)$$

$$= \sum_{i=1}^{N} I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{y}(i) | \mathsf{y}^{i-1}, \mathsf{s}^{i-1}, \mathsf{z}^{i-1}, \mathsf{s}(i)) \quad (12)$$

$$\leq \sum_{i=1}^{N} I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{y}(i) | \mathsf{s}^{i-1}, \mathsf{z}^{i-1}, \mathsf{s}(i)) \quad (13)$$

where (10) follows from the fact that $\mathsf{s}(i)$ is independent of $(\mathsf{w}_1, \mathsf{w}_2, \mathsf{y}^{i-1}, \mathsf{s}^{i-1})$, so $I(\mathsf{w}_1, \mathsf{w}_2; \mathsf{s}_i | \mathsf{y}^{i-1}, \mathsf{s}^{i-1}) = 0$; (11) follows from the fact that since the channel is memoryless the Markov condition: $\mathsf{y}(i) \to (\mathsf{x}_1(i), \mathsf{x}_2(i)) \to (\mathsf{y}^{i-1}, \mathsf{s}^i, \mathsf{w}_1, \mathsf{w}_2)$ is satisfied; (12) follows from the fact that by construction $\mathsf{z}^{i-1}$ is a deterministic function of $(\mathsf{y}^{i-1}, \mathsf{s}^{i-1})$; (13) follows from the fact that $(\mathsf{s}^{i-1}, \mathsf{z}^{i-1}, \mathsf{y}^{i-1}) \to (\mathsf{x}_1(i), \mathsf{x}_2(i), \mathsf{s}(i)) \to \mathsf{y}(i)$ and conditioning reduces entropy. For the remainder of the steps we define $\Omega_i \triangleq (\mathsf{s}(i), \mathsf{z}(i))$ and let $\Omega^i = (\Omega_1, \ldots, \Omega_i)$. Observe that $(\mathsf{x}_1(i), \mathsf{x}_2(i), \Omega^{i-1})$ is independent of $\mathsf{s}(i)$. Furthermore, when $\mathsf{s}(i) = 0$ we have that $\mathsf{y}(i) = (\mathsf{u}_1(i), \mathsf{u}_2(i))$ whereas when $\mathsf{s}(i) = 1$ we have that $\mathsf{y}(i) = \bar{\mathsf{y}}(i) = \mathsf{x}_1(i) + \mathsf{x}_2(i) + \mathsf{z}(i)$. Therefore we have:

$$I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{y}(i) | \Omega^{i-1}, \mathsf{s}(i))$$
$$= \bar{\beta} I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{u}_1(i), \mathsf{u}_2(i) | \Omega^{i-1}, \mathsf{s}(i) = 0)$$
$$+ \beta I(\mathsf{x}_1(i), \mathsf{x}_2(i); \bar{\mathsf{y}}(i) | \Omega^{i-1}, \mathsf{s}(i) = 1) \quad (14)$$
$$= \bar{\beta} I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{u}_1(i), \mathsf{u}_2(i) | \Omega^{i-1})$$
$$+ \beta I(\mathsf{x}_1(i), \mathsf{x}_2(i); \bar{\mathsf{y}}(i) | \Omega^{i-1}). \quad (15)$$

Note that the noise variables $(\mathsf{n}_1(i), \mathsf{n}_2(i))$ in (5) are mutually independent and each is also i.i.d. Furthermore let $E[\mathsf{x}_k^2(i)] = P_k(i)$ denote the average power in symbol $\mathsf{x}_k(i)$ then

$$I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{u}_1(i), \mathsf{u}_2(i) | \Omega^{i-1}) \leq \sum_{k=1}^{2} I(\mathsf{x}_k(i); \mathsf{u}_k(i))$$

$$\leq \sum_{k=1}^{2} \frac{1}{2} \log\left(1 + \frac{P_k(i)}{\sigma_k^2}\right).$$

Thus applying Jensen's inequality we have that

$$\sum_{i=1}^{N} I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{u}_1(i), \mathsf{u}_2(i)|\Omega^{i-1}) \leq \frac{N}{2} \sum_{k=1}^{2} \log\left(1 + \frac{P_k}{\sigma_k^2}\right) \tag{16}$$

where $P_k$ denotes the average power constraint for encoder $k \in \{1, 2\}$. We next consider the second term in (15). From Lemma 1 we have that $\mathsf{x}_1(i) \rightarrow \Omega^{i-1} \rightarrow \mathsf{x}_2(i)$. Also let $E[\mathsf{x}_k^2(i)|\Omega^{i-1} = \omega^{i-1}] = P_k(i; \omega^{i-1})$ and let the associated covariance matrix be

$$\mathrm{Cov}(\mathsf{x}_1(i), \mathsf{x}_2(i)|\Omega^{i-1} = \omega^{i-1}) = \begin{bmatrix} P_1(i; \omega^{i-1}) & 0 \\ 0 & P_2(i; \omega^{i-1}) \end{bmatrix}. \tag{17}$$

By replacing the distribution $p(\mathsf{x}_1(i), \mathsf{x}_2(i)|\omega^{i-1})$, with a Gaussian distribution of the same covariance matrix we have that

$$I(\mathsf{x}_1(i), \mathsf{x}_2(i); \bar{\mathsf{y}}(i)|\Omega^{i-1} = \omega^{i-1}) \leq$$
$$\frac{1}{2} \log\left(1 + P_1(i; \omega^{i-1}) + P_2(i; \omega^{i-1})\right). \tag{18}$$

Thus we have that

$$I(\mathsf{x}_1(i), \mathsf{x}_2(i); \bar{\mathsf{y}}(i)|\Omega^{i-1})$$
$$\leq E_{\Omega^{i-1}}\left[\frac{1}{2} \log\left(1 + P_1(i; \omega^{i-1}) + P_2(i; \omega^{i-1})\right)\right] \tag{19}$$
$$\leq \frac{1}{2} \log\left(1 + E_{\Omega^{i-1}}[P_1(i; \omega^{i-1})] + E_{\Omega^{i-1}}[P_2(i; \omega^{i-1})]\right) \tag{20}$$
$$= \frac{1}{2} \log\left(1 + P_1(i) + P_2(i)\right) \tag{21}$$

where (20) follows from the Jensen's inequality and (21) follows by letting $P_k(i) = E[P_k(i; \Omega^{i-1})]$ to be the average power used by encoder $k$ at time $i$. Therefore we have that

$$\sum_{i=1}^{N} I(\mathsf{x}_1(i), \mathsf{x}_2(i); \mathsf{y}(i)|\Omega^{i-1}) \leq \frac{1}{2} \sum_{i=1}^{N} \log\left(1 + P_1(i) + P_2(i)\right) \tag{22}$$
$$\leq \frac{N}{2} \log\left(1 + \frac{1}{N} \sum_{i=1}^{N}(P_1(i) + P_2(i))\right) \tag{23}$$
$$\leq \frac{N}{2} \log(1 + P_1 + P_2) \tag{24}$$

where (23) follows from the Jensen's inequality. Combining (13), (15), (16) and (24) and using the fact that $\epsilon_N \rightarrow 0$ as $N \rightarrow \infty$, we have shown that

$$R_1 + R_2 \leq \bar{\beta}\left\{\frac{1}{2} \log\left(1 + \frac{P_1}{\sigma_1^2}\right) + \frac{1}{2} \log\left(1 + \frac{P_2}{\sigma_2^2}\right)\right\}$$
$$+ \beta \frac{1}{2} \log\left(P_1 + P_2 + 1\right). \tag{25}$$

We next establish an upper bound on the rate $R_1$.

$$NR_1 = H(\mathsf{w}_1) = H(\mathsf{w}_1|\mathsf{w}_2) \tag{26}$$
$$= H(\mathsf{w}_1|\mathsf{w}_2, \mathsf{y}^N, \mathsf{s}^N) + I(\mathsf{w}_1; \mathsf{y}^N, \mathsf{s}^N|\mathsf{w}_2). \tag{27}$$

From Fano's inequality we have that $H(\mathsf{w}_1|\mathsf{w}_2, \mathsf{y}^N, \mathsf{s}^N) \leq N\epsilon_N$. By following the steps similar to those leading to (15):
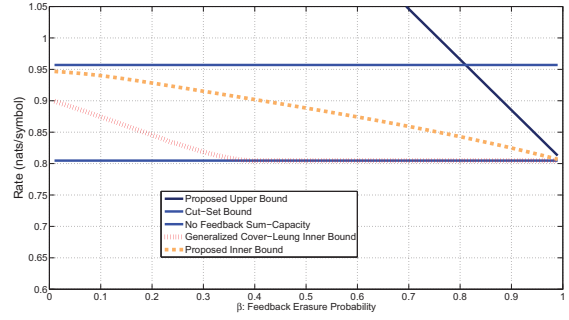


Fig. 2: Upper and Lower Bounds on the sum-rate of AWGN MAC with Intermittent Feedback. We let $P_1 = P_2 = 2$.

$$NR_1 \leq \sum_{i=1}^{N}\left\{\beta I(\mathsf{x}_1(i); \mathsf{y}(i)|\mathsf{x}_2(i))\right.$$
$$\left. + \bar{\beta} I(\mathsf{x}_1(i); \mathsf{u}_1(i), \mathsf{u}_2(i)|\mathsf{x}_2(i))\right\}. \tag{28}$$

Using Jensen's inequality and the fact that conditioning reduces entropy:

$$\sum_{i=1}^{N} I(\mathsf{x}_1(i); \mathsf{y}(i)|\mathsf{x}_2(i)) \leq \sum_{i=1}^{N} \frac{1}{2} \log(1 + P_1(i)) \tag{29}$$
$$\leq \frac{N}{2} \log(1 + P_1). \tag{30}$$

Furthermore since the random variables $(\mathsf{n}_1(i), \mathsf{n}_2(i))$ are mutually indepenent, we can show that

$$I(\mathsf{x}_1(i); \mathsf{u}_1(i), \mathsf{u}_2(i)|\mathsf{x}_2(i)) \leq I(\mathsf{x}_1(i); \mathsf{u}_1(i)) \tag{31}$$
$$\leq \frac{1}{2} \log\left(1 + \frac{P_1(i)}{\sigma_1^2}\right). \tag{32}$$

Therefore using Jensen's inequality we have

$$\sum_{i=1}^{N} I(\mathsf{x}_1(i); \mathsf{u}_1(i), \mathsf{u}_2(i)|\mathsf{x}_2(i)) \leq \frac{N}{2} \log\left(1 + \frac{P_1}{\sigma_1^2}\right). \tag{33}$$

Thus from (28), (30) and (33) we have that

$$R_1 \leq \frac{\bar{\beta}}{2} \log\left(1 + \frac{P_1}{\sigma_1^2}\right) + \frac{\beta}{2} \log(1 + P_1). \tag{34}$$

In a similar way we can show (3).

### B. Inner Bound

Our coding scheme is a variant of the concatenated coding scheme [4] for AWGN MAC with noisy feedback. Each encoder samples an i.i.d. Gaussian $\mathcal{N}(0, 1)$ codebook of length $N$. Thus associated with message $\mathsf{w}_1 \in [1, 2^{NR_1}]$, we have a codeword $(\psi_{1,1}(\mathsf{w}_1), \ldots, \psi_{1,N}(\mathsf{w}_1))$ and likewise associated with message $\mathsf{w}_2 \in [1, 2^{NR_2}]$ we have a codeword $(\psi_{2,1}(\mathsf{w}_2), \ldots, \psi_{2,N}(\mathsf{w}_2))$. The inner encoders iteratively refine each codeword symbol until an erasure is encountered. At that point they switch to the next fresh symbol. More specifically for $k = 1$ the encoders transmit

$$\mathsf{x}_1(1) = \sqrt{P_1}\psi_{1,1}, \quad \mathsf{x}_2(1) = \sqrt{P_2}\psi_{2,1}. \tag{35}$$

Assume that the first erasure on the feedback link happens at time $t = \Theta_1$. Then for $t = 2, 3, \ldots, \Theta_1$ the encoders transmit

$$x_1(k) = \sqrt{\frac{P_1}{\beta_{1,k}}}\left(\psi_{1,1} - E[\psi_{1,1}|y^{k-1}]\right), \quad (36)$$

$$x_2(k) = (-1)^{k-1}\sqrt{\frac{P_2}{\beta_{2,k}}}\left(\psi_{2,1} - E[\psi_{2,1}|y^{k-1}]\right), \quad (37)$$

where we select: $\beta_{1,k} = \mathrm{Var}\left(\psi_{1,1} - E[\psi_{1,1}|y^{k-1}]\right)$ and $\beta_{2,k} = \mathrm{Var}\left(\psi_{2,1} - E[\psi_{2,1}|y^{k-1}]\right)$. At time $k = \Theta_1 + 1$ the two encoders transmit

$$x_1(k) = \sqrt{P_1}\psi_{1,2}, \quad x_2(k) = \sqrt{P_2}\psi_{2,2}, \quad (38)$$

and continue refining these symbols until a feedback erasure happens at time $k = \Theta_2 + 1$.

Thus our concatenated coding scheme induces a channel $(x_1, x_2) \rightarrow (y_1, \ldots, y_\Theta, \Theta)$ between the encoders and the decoder. An achievable rate region for this MAC is:

$$R_j \leq \frac{1}{E[\Theta]}I(x_j; y_1, \ldots, y_\Theta|\Theta, x_{\bar{j}}), \quad j \in \{1, 2\} \quad (39)$$

$$R_1 + R_2 \leq \frac{1}{E[\Theta]}I(x_1, x_2; y_1, \ldots, y_\Theta|\Theta), \quad (40)$$

where we select $x_j \sim \mathcal{N}(0, P_j)$ and $x_1$ and $x_2$ are mutually independent and independent of $\Theta$. Also note that $\Pr(\Theta = t) = (1-\beta)^{t-1}\beta$ where $\beta$ is the probability of erasure on the feedback channel.

Furthermore from [4, Appendix E] we can show that when $\Theta = t$ we have:

$$R_j(t) \triangleq I(x_j; y_1, \ldots, y_t|x_{\bar{j}})$$
$$= \frac{1}{2}\sum_{k=1}^{t}\log\left(1 + P_j(1 - \rho_{k-1}^2)\right), \quad j = 1, 2 \quad (41)$$

$$R_1(t) + R_2(t) \triangleq I(x_1, x_2; y_1, \ldots, y_t)$$
$$= \frac{1}{2}\sum_{k=1}^{t}\log(1 + P_1 + P_2 + 2(-1)^{k-1}\rho_{k-1}\sqrt{P_1 P_2}), \quad (42)$$

where the correlation coefficient $\rho_k$ is recursively defined using

$$\rho_k = \frac{\rho_{k-1} - (-1)^{k-1}\sqrt{P_1 P_2}(1 - \rho_{k-1}^2)}{\sqrt{1 + P_1(1 - \rho_{k-1}^2)}\sqrt{1 + P_2(1 - \rho_{k-1}^2)}}, \quad k \geq 1$$
$$(43)$$

and $\rho_0 = 0$.

**Remark 3.** *The proposed inner bound approaches the capacity region with perfect feedback and no feedback in the limits of $\beta \rightarrow 0$ and $\beta \rightarrow 1$ respectively.*

Fig. 2 illustrates the upper and lower bounds on the sum-rate of the AWGN MAC when $P_1 = P_2 = 2$. The x-axis denotes the erasure probability in feedback, whereas the y-axis denotes the achievable rate. The lowermost horizontal line is the sum-capacity achieved without feedback. The upper horizontal line is the cut-set bound. The other solid (black) line denotes the

new upper bound on the sum-rate in Theorem 1. It is only tighter than the cut-set bound for large $\beta$. The dashed (orange) line corresponds to the inner bound in (41)-(42) while the dotted (red) line is the rate achieved by an extension of the Cover-Leung scheme [9] for MAC with generalized feedback. We note that it degenerates to the no-feedback rate for $\beta > 0.4$ and does not achieve the cut-set bound as $\beta \rightarrow 0$. In general the proposed iterative refinement scheme appears to be superior.

## III. MAC WITH FEEDBACK AND INTERMITTENT CRIBBING

We study discrete memoryless multiple-access channel $p(y|x_1, x_2)$ where the side information symbol is specified by:

$$z(i) = \begin{cases} (y(i), x_1(i), x_2(i)), & s(i) = 1 \\ y(i), & s(i) = 0, \end{cases} \quad (44)$$

and the input symbols satisfy $x_k(i) = f_{k,i}(w_k, z^{i-1}, s^{i-1})$ for $k \in \{1, 2\}$. The associated state sequence is i.i.d. with $\Pr(s(i) = 1) = \beta$ and $\Pr(s(i) = 0) = 1 - \beta = \bar{\beta}$. Note that in the proposed model, the encoders are always revealed ideal feedback, and in addition can crib intermittently. We will restrict our discussion to the case when one of the input symbols can be computed given the output symbol and the other input symbol i.e.,

$$x_1(i) = h(y(i), x_2(i)). \quad (45)$$

**Theorem 2.** *The capacity region under the condition* (45) *is given by:*

$$R_1 \leq \beta I(x_1; y|u, x_2) + \bar{\beta}H(x_1|u) \quad (46)$$
$$R_2 \leq \beta I(x_2; y|u, x_2) + \bar{\beta}H(x_2|u) \quad (47)$$
$$R_1 + R_2 \leq I(x_1, x_2; y) \quad (48)$$

*for some joint distribution $p(x_1, x_2, u, y)$ that satisfies $x_1 \rightarrow u \rightarrow x_2$, $u \rightarrow (x_1, x_2) \rightarrow y$, and the cardinality of $u$ is no greater than $|\mathcal{X}_1||\mathcal{X}_2| + 3$. The capacity region is achieved using a Superposition Block Markov Coding scheme.*

**Remark 4.** *When $\beta = 1$ i.e., only ideal feedback is available to both encoders the capacity region reduces to the MAC with perfect feedback. This region was obtained by Willems [2]. When $\beta = 0$, the capacity region reduces to the the case of strictly causal cribbing encoders [7]. In particular the additional feedback symbol does not increase the capacity region, as recently noted in [8].*

**Remark 5.** *The capacity achieving scheme in Theorem 2 does not require the knowledge of $s^N$ at the destination. The feedback symbol is also not used by the encoders whenever cribbing is available.*

We only provide details for the converse in Theorem 2. Using Fano's inequality we can show that $NR_1 \leq N\epsilon_N + I(w_1; z^N, s^N|w_2)$. Now consider

$$I(w_1; z^N, s^N|w_2) = \sum_{i=1}^{N}I(w_1; z_i, s_i|w_2, z^{i-1}, s^{i-1}) \quad (49)$$

$$= \sum_{i=1}^{N} I(\mathsf{w}_1; \mathsf{z}_i, \mathsf{s}_i | \mathsf{w}_2, \mathsf{z}^{i-1}, \mathsf{s}^{i-1}, \mathsf{x}_2^{i-1}, \mathsf{x}_{2,i}) \qquad (50)$$

$$= \sum_{i=1}^{N} I(\mathsf{w}_1; \mathsf{z}_i, \mathsf{s}_i | \mathsf{w}_2, \mathsf{z}^{i-1}, \mathsf{s}^{i-1}, \mathsf{x}_2^{i-1}, \mathsf{x}_1^{i-1}, \mathsf{x}_{2,i}) \qquad (51)$$

$$\leq \sum_{i=1}^{N} I(\mathsf{x}_{1,i}; \mathsf{z}_i, \mathsf{s}_i | \mathsf{w}_2, \mathsf{z}^{i-1}, \mathsf{s}^{i-1}, \mathsf{x}_2^{i-1}, \mathsf{x}_1^{i-1}, \mathsf{x}_{2,i}) \qquad (52)$$

$$\leq \sum_{i=1}^{N} I(\mathsf{x}_{1,i}; \mathsf{z}_i, \mathsf{s}_i | \mathsf{z}^{i-1}, \mathsf{s}^{i-1}, \mathsf{x}_1^{i-1}, \mathsf{x}_{2,i}) \qquad (53)$$

$$= \sum_{i=1}^{N} I(\mathsf{x}_{1,i}; \mathsf{z}_i, \mathsf{s}_i | \mathsf{u}_i, \mathsf{x}_{2,i}) \qquad (54)$$

$$= \sum_{i=1}^{N} I(\mathsf{x}_{1,i}; \mathsf{z}_i, | \mathsf{u}_i, \mathsf{x}_{2,i}, \mathsf{s}_i) \qquad (55)$$

$$= n \cdot I(\mathsf{x}_{1,Q}; \mathsf{z}_Q, | \mathsf{u}_Q, Q, \mathsf{x}_{2,Q}, \mathsf{s}_Q) \qquad (56)$$

where (50) follows from the fact that $\mathsf{x}_{2,i}$ is a function of $(\mathsf{w}_2, \mathsf{z}^{i-1}, \mathsf{s}^{i-1})$; (51) follows directly from (45); (54) follows by using $\mathsf{u}_i = (\mathsf{z}^{i-1}, \mathsf{s}^{i-1}, \mathsf{x}_1^{i-1})$; The justification of the remaining steps is omitted due to space constraint. It can be shown along the lines of Lemma 1, that for each $i$ the relation $\mathsf{x}_{1,i} \to \mathsf{u}_i \to \mathsf{x}_{2,i}$ is satisfied. Therefore we have $\mathsf{x}_{1,Q} \to (\mathsf{u}_Q, Q) \to \mathsf{x}_{2,Q}$, $(\mathsf{u}_Q, Q) \to (\mathsf{x}_{1,Q}, \mathsf{x}_{2,Q}) \to \mathsf{y}_Q$ and furthermore $(\mathsf{y}_Q, \mathsf{x}_{1,Q}, \mathsf{x}_{2,Q}, \mathsf{u}_Q, Q)$ are independent of $\mathsf{s}_Q$. Substituting $\mathsf{x}_{1,Q} = \mathsf{x}_1$, $\mathsf{u} = (\mathsf{u}_Q, Q)$ etc., we have

$$R_1 \leq I(\mathsf{x}_1; \mathsf{z} | \mathsf{s}, \mathsf{x}_2, \mathsf{u})$$
$$= \Pr(\mathsf{s} = 0) I(\mathsf{x}_1; \mathsf{z} | \mathsf{s} = 0, \mathsf{x}_2, \mathsf{u})$$
$$\quad + \Pr(\mathsf{s} = 1) I(\mathsf{x}_1; \mathsf{z} | \mathsf{s} = 1, \mathsf{x}_2, \mathsf{u}) \qquad (57)$$
$$= \bar{\beta} I(\mathsf{x}_1; \mathsf{x}_1, \mathsf{x}_2, \mathsf{y} | \mathsf{x}_2, \mathsf{u}) + \beta I(\mathsf{x}_1; \mathsf{y} | \mathsf{x}_2, \mathsf{u}) \qquad (58)$$
$$= \bar{\beta} H(\mathsf{x}_1 | \mathsf{x}_2, \mathsf{u}) + \beta I(\mathsf{x}_1; \mathsf{y} | \mathsf{x}_2, \mathsf{u}) \qquad (59)$$
$$\leq \bar{\beta} H(\mathsf{x}_1 | \mathsf{u}) + \beta I(\mathsf{x}_1; \mathsf{y} | \mathsf{x}_2, \mathsf{u}) \qquad (60)$$

where (58) follows from using (44); (59) follows from the fact that $\mathsf{s}$ is independent of $(\mathsf{x}_1, \mathsf{x}_2, \mathsf{y}, \mathsf{u})$; (60) follows from the fact that conditioning reduces entropy. The upper bonds on $R_2$ and $R_1 + R_2$ can be established in a similar fashion and will be omitted.

The above result can be naturally extended to other cribbing scenarios. For example if only encoder 1 cribs intermittently and both observe ideal feedback the rate constraints in Theorem 2 are modified as follows:

$$R_1 \leq I(\mathsf{x}_1; \mathsf{y} | \mathsf{x}_2, \mathsf{u}), \qquad (61)$$
$$R_2 \leq \bar{\beta} H(\mathsf{x}_2 | \mathsf{u}) + \beta I(\mathsf{x}_2; \mathsf{y} | \mathsf{x}_1, \mathsf{u}) \qquad (62)$$
$$R_1 + R_2 \leq I(\mathsf{x}_1, \mathsf{x}_2; \mathsf{y}) \qquad (63)$$

## APPENDIX A
### PROOF OF LEMMA 1

Our proof uses the following fact.

**Lemma 2.** *Let* $(\mathsf{x}, \mathsf{y}, \mathsf{z})$ *be random variables with a joint distribution of* $p(\mathsf{x}, \mathsf{y}, \mathsf{z})$. *The condition* $\mathsf{x} \to \mathsf{z} \to \mathsf{y}$ *holds if and only if we can factorize* $p(\mathsf{x}, \mathsf{y}, \mathsf{z}) = g(\mathsf{x}, \mathsf{z}) h(\mathsf{y}, \mathsf{z})$.

Lemma 1 follows from the following stronger statement:

**Proposition 1.** *For each* $i = 1, 2, \ldots, N$ *we have that* $\mathsf{w}_1 \to (\mathsf{z}^{i-1}, \mathsf{s}^{i-1}) \to \mathsf{w}_2$. *and* $\mathsf{x}_{1,i} \to (\mathsf{z}^{i-1}, \mathsf{s}^{i-1}) \to \mathsf{x}_{2,i}$.

For convenience we let $\Omega_i = (\mathsf{z}(i), \mathsf{s}(i))$ and $\Omega^i = (\mathsf{z}^i, \mathsf{s}^i)$. To establish $\mathsf{w}_1 \to \Omega^i \to \mathsf{w}_2$, it suffices to show that we can factorize $p(\mathsf{w}_1, \mathsf{w}_2, \Omega^i) = g(\mathsf{w}_1, \Omega^i) h(\mathsf{w}_2, \Omega^i)$. Consider

$$p(\mathsf{w}_1, \mathsf{w}_2, \Omega^j)$$
$$= p(\mathsf{w}_1) p(\mathsf{w}_2) \prod_{j=1}^{i-1} p(\mathsf{z}(j), \mathsf{s}(j) | \mathsf{w}_1, \mathsf{w}_2, \Omega^{j-1}) \qquad (64)$$
$$= p(\mathsf{w}_1) p(\mathsf{w}_2) \prod_{j=1}^{i-1} p(\mathsf{s}(j)) p(\mathsf{z}(j) | \mathsf{w}_1, \mathsf{w}_2, \Omega^{j-1}, \mathsf{s}(j)) \qquad (65)$$

Thus it is sufficient to show that the term $p(\mathsf{z}(j) | \mathsf{w}_1, \mathsf{w}_2, \Omega^{j-1}, \mathsf{s}(j))$ factorizes as $\{g_j(\mathsf{w}_1, \Omega^j) \cdot h_j(\mathsf{w}_2, \Omega^j)\}$. When $\mathsf{s}(j) = 0$ it follows that $\mathsf{z}(j) = \star$ and the factorization trivially follows. When $\mathsf{s}(j) = 1$ it follows that $\mathsf{z}(j) = (\mathsf{u}_1(j), \mathsf{u}_2(j))$. Therefore

$$p(\mathsf{z}(j) | \mathsf{w}_1, \mathsf{w}_2, \Omega^{j-1}, \mathsf{s}(j) = 0)$$
$$= p(\mathsf{u}_1(j), \mathsf{u}_2(j) | \mathsf{w}_1, \mathsf{w}_2, \Omega^{j-1}) \qquad (66)$$
$$= p(\mathsf{u}_1(j) | \mathsf{w}_1, \Omega^{j-1}) p(\mathsf{u}_2(j) | \mathsf{w}_2, \Omega^{j-1}) \qquad (67)$$
$$\triangleq g_j(\mathsf{w}_1, \Omega^j) \cdot h_j(\mathsf{w}_2, \Omega^j) \qquad (68)$$

where the last step follows from the fact that $\mathsf{x}_k(j) = f_{k,j}(\mathsf{w}_k, \Omega^{j-1})$ for $k = 1, 2$. Thus the desired factorization follows. Finally the relation $\mathsf{x}_{1,i} \to \Omega^{i-1} \to \mathsf{x}_{2,i}$ directly follows from the relation $\mathsf{w}_1 \to \Omega^{i-1} \to \mathsf{w}_2$ and $\mathsf{x}_{1,i} = f_{1,i}(\mathsf{w}_1, \Omega^{i-1})$ and $\mathsf{x}_{2,i} = f_{2,i}(\mathsf{w}_2, \Omega^{i-1})$.

## REFERENCES

[1] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inform. Theory*, vol. 27, pp. 292–298, May 1981.

[2] F. M. J. Willems, "The feedback capacity region of a class of discrete memoryless multiple access channels," *IEEE Trans. Inform. Theory*, vol. 28, pp. 93–95, Jan. 1982.

[3] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. Inform. Theory*, vol. 30, pp. 623–629, Jul. 1984.

[4] A. Lapidoth and M. A. Wigger, "On the AWGN MAC with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5432–5476, 2010.

[5] R. Tandon and S. Ulukus, "Dependence balance based outer bounds for Gaussian networks with cooperation and feedback," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4063–4086, 2011.

[6] D. Shaviv and Y. Steinberg, "On the multiple-access channel with common rate-limited feedback," in *International Zurich Seminar*, march 2008, pp. 108 –111.

[7] F. M. J. Willems and E. C. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Transactions on Information Theory*, vol. 31, no. 3, pp. 313–327, 1985.

[8] A. Bracher, A. Lapidoth, and Y. Steinberg, "On feedback, cribbing, and causal state-information on the multiple-access channel," in *Information Theory Workshop (ITW)*, sept. 2012, pp. 134 –138.

[9] F. M. J. Willems, E. van der Mulen, and J. Schalkwijk, "Generalized feedback for the discrete memoryless multiple-access channel," in *Proc. Allerton Conf. Commun., Contr., Computing*, 1983, pp. 284–292.