

A New Achievable Rate Region for the Discrete Memoryless Multiple-Access Channel with Feedback

Ramji Venkataramanan
Dept. of Electrical Engineering
Stanford University, USA
Email: vramji@stanford.edu

S. Sandeep Pradhan
Dept. of Electrical Engineering & Computer Science
University of Michigan, Ann Arbor, USA
Email: pradhanv@eecs.umich.edu

Abstract—A single-letter achievable rate region for the two-user discrete memoryless multiple-access channel is proposed. The rate region includes the Cover-Leung region [1], and it is shown that the inclusion is strict. The proof uses a block-Markov superposition strategy based on the observation that the messages of the two users are correlated given the feedback. The rates of transmission are too high for each encoder to decode the other's message directly using the feedback, so they transmit correlated information in the next block in order to learn the message of one another. They then cooperate in the following block to resolve the residual uncertainty of the decoder. Our scheme may be viewed as a natural generalization of the Cover-Leung scheme with a delay of one extra block and a pair of additional auxiliary random variables. The scheme can also be extended to obtain larger rate-regions with more auxiliary random variables.

I. INTRODUCTION

The two-user discrete memoryless multiple-access channel (MAC) is shown in Figure 1. The channel has inputs X, Y , output Z and is characterized by the conditional law $P_{Z|XY}$. The capacity region for this channel without feedback (S_1 and S_2 open in Figure 1) was determined by Ahlswede [2] and Liao [3]. Gaarder and Wolf [4] demonstrated that feedback can enlarge the capacity region of this channel using the example of a binary erasure MAC. Cover and Leung [1] then established a single-letter achievable rate region for a discrete memoryless MAC with feedback. The Cover-Leung (C-L) region was shown to be the feedback capacity region for a class of discrete memoryless MACs [5]. However the C-L region is smaller than the feedback capacity in general, the white Gaussian MAC being a notable example [6], [7]. The capacity region of the MAC with feedback was characterized by Kramer [8], [9] in terms of directed information. This is a 'multi-letter' characterization and the existence of a single-letter capacity characterization for the discrete memoryless MAC with feedback remains an open question. Recently, a single-letter achievable region that includes the C-L region was reported in [10].

The basic idea behind the Gaarder-Wolf and the C-L schemes is the following. Each message pair is conveyed to the decoder over two successive blocks of transmission. In

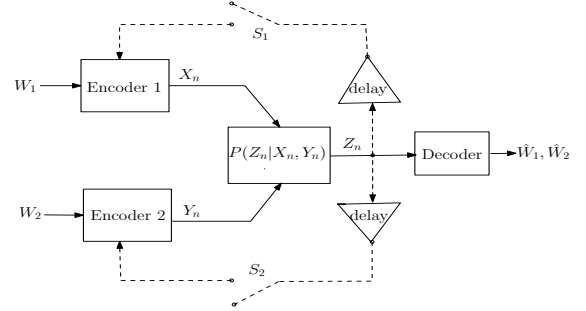


Fig. 1. The multiple-access channel. When S_1, S_2 are closed there is feedback to both encoders.

the first block, the two encoders transmit messages at rates outside the no-feedback capacity region. At the end of this block, the decoder cannot decode the message pair; however, each encoder can decode the message of the other using the feedback (the rates must be low enough to ensure this). The two encoders can now cooperate and send a common message to resolve the decoder's uncertainty in the next block. This procedure is repeated over several blocks with fresh information superimposed over resolution information in every block. This block-Markov superposition scheme yields the following rate region for the MAC with feedback.

Cover-Leung Region: Consider a joint distribution of the form $P_{UXYZ} = P_U P_{X|U} P_{Y|U} P_{Z|XY}$, where $P_{Z|XY}$ is fixed by the channel and U is a discrete random variable with cardinality $\min\{|\mathcal{X}| \cdot |\mathcal{Y}| + 1, |\mathcal{Z}| + 2\}$. Then the following rate pairs (R_1, R_2) are achievable.

$$\begin{aligned} R_1 &< I(X; Z|YU), \quad R_2 < I(Y; Z|XU), \\ R_1 + R_2 &< I(XY; Z) \end{aligned} \quad (1)$$

In the C-L strategy described above, suppose that the rates are too high for the encoders to decode the message of one another using the feedback. In this case, encoder 1 has its own message plus a list of possible messages for encoder 2. This list, created from all the Y sequences that are jointly typical with the known (X, Z) sequence pair, is clearly smaller than the original message set of encoder 2. Similarly, encoder 2 has

a list of all possible messages for encoder 1. Thus at the end of the first block, the encoders have correlated information they wish to transmit over the next block. Note that this is more general than the C-L scheme which imposes *perfect* correlation between the encoders' messages given a block of channel output. Transmission of correlated sources and correlated message graphs over the MAC have been studied in [11] and [12], respectively. We use these ideas to develop a new block-Markov superposition coding scheme. This scheme yields a single-letter rate region that includes the C-L region. We show that the inclusion is strict using the example of the AWGN MAC.

II. PRELIMINARIES AND MAIN RESULT

We use uppercase letters to denote random variables, lowercase for their realizations and bold-face notation for random vectors. Unless otherwise stated, all vectors have length N . Thus $\mathbf{A} \triangleq A^N \triangleq (A_1, \dots, A_N)$. For any α such that $0 < \alpha < 1$, $\bar{\alpha} \triangleq 1 - \alpha$. The typical set of a random variable with pmf P is denoted $T(P)$. Logarithms are with base 2, and entropy and mutual information are measured in bits.

Consider a two-user discrete memoryless multiple-access channel with input alphabets $\mathcal{X}$ and $\mathcal{Y}$, output alphabet $\mathcal{Z}$ defined by the probability mass functions $P_{Z|XY}(\cdot|x, y)$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$. The channel is memoryless, i.e., it satisfies

$$P(Z_n|X^n, Y^n, Z^{n-1}) = P(Z_n|X_n, Y_n), \quad n = 1, 2, \dots$$

There is feedback to both encoders (S_1 and S_2 are both closed in Figure 1).

Definition 1: An $(N, 2^{NR_1}, 2^{NR_2})$ code with block length N and rates (R_1, R_2) for a MAC with feedback consists of

- 1) A sequence of mappings for each encoder:

$$\begin{aligned} e_{1n} &: \{1, \dots, 2^{NR_1}\} \times \mathcal{Z}^{n-1} \rightarrow \mathcal{X}, \quad n = 1, \dots, N \\ e_{2n} &: \{1, \dots, 2^{NR_2}\} \times \mathcal{Z}^{n-1} \rightarrow \mathcal{Y}, \quad n = 1, \dots, N \end{aligned}$$

- 2) A decoder mapping given by

$$g: \mathcal{Z}^N \rightarrow \{1, \dots, 2^{NR_1}\} \times \{1, \dots, 2^{NR_2}\}$$

Assuming the messages (W_1, W_2) are drawn uniformly from the set $\{1, \dots, 2^{NR_1}\} \times \{1, \dots, 2^{NR_2}\}$, the error probability is

$$P_{e,N} = \frac{\sum_{m_1=1, m_2=1}^{2^{NR_1}, 2^{NR_2}} \Pr(g(\mathbf{Z}) \neq (m_1, m_2) | W_1, W_2 = m_1, m_2)}{2^{NR_1} 2^{NR_2}}$$

A rate pair (R_1, R_2) is achievable for the discrete memoryless MAC with feedback if there exists a sequence of $(N, 2^{NR_1}, 2^{NR_2})$ codes such that $P_{e,N} \rightarrow 0$ as $N \rightarrow \infty$. The closure of all achievable rate pairs is the capacity region with feedback. The following theorem is the main result of this paper.

Theorem 1: For the MAC described by $P_{Z|XY}$, consider any joint distribution of the form

$$P_{UWABXYZ} = P_U P_{WAB} P_{X|UWA} P_{Y|UWB} P_{Z|XY} \quad (2)$$

with U, W, A, B are discrete random variables such that:

- 1) There exists a function $f: \mathcal{Z} \times \mathcal{A} \times \mathcal{B} \times \mathcal{U} \times \mathcal{W} \rightarrow \mathcal{W}$ such that

$$P_W(k) = \sum_{\substack{(z, a, b, u, w): \\ f(z, a, b, u, w) = k}} P_{ZABUW}(z, a, b, u, w) \quad (3)$$

where P_{ZABUW} is the marginal derived from (2).

- 2) There exist conditional probability distributions $Q_{\bar{A}|ZXABUW}$ (denoted Q^1) and $Q_{\bar{B}|ZYABUW}$ (denoted Q^2) such that for all $(\bar{a}, \bar{b}) \in \mathcal{A} \times \mathcal{B}$ and all $(z, a, b, u, w) \in \mathcal{Z} \times \mathcal{A} \times \mathcal{B} \times \mathcal{U} \times \mathcal{W}$

$$\begin{aligned} &\sum_{x, y} Q^1(\bar{a}|z, x, a, b, u, w) Q^2(\bar{b}|z, y, a, b, u, w) \\ &\quad \cdot P_{XY|ZABUW}(x, y|z, a, b, u, w) \\ &= P_{AB|W}(\bar{a}, \bar{b}|f(w, u, a, b, z)), \end{aligned} \quad (4)$$

where $P_{XY|ZABUW}$ and $P_{AB|W}$ are marginals derived from (2).

For such a joint distribution, the following rate-region is achievable.

$$\begin{aligned} R_1 &< \min\{I(X; Z|YBWU), \\ &\quad I(X; Z|YABWU) + I(A; Z|BWU) + I(U; Z)\} \\ R_2 &< \min\{I(Y; Z|XAWU), \\ &\quad I(Y; Z|XABWU) + I(B; Z|AWU) + I(U; Z)\} \\ R_1 + R_2 &< I(XY; Z|UW) + I(U; Z) \end{aligned}$$

A. Overview of Coding Scheme

Consider B blocks of transmission, with a fresh pair of messages in each block. Let (m_{1b}, m_{2b}) denote the messages, and $\mathbf{X}_b, \mathbf{Y}_b$ the input sequences for block b . The input $\mathbf{X}_b$ has three components: $\mathbf{U}_b, \mathbf{A}_b$ and a fresh message from user 1. $\mathbf{Y}_b$ comprises $\mathbf{U}_b, \mathbf{B}_b$ and a fresh message from user 2. The fresh messages transmitted by the two users in any block are independent; however, conditioned on the output of that block, the messages are correlated. Consider the block $(b-1)$, at the end of which the decoder as well as both encoders receive $\mathbf{Z}_{b-1}$. At this point, encoder 1 makes a list of possible $m_{2(b-1)}$. This list, exponentially smaller than 2^{NR_2} in size, is resolved by the sequence $\mathbf{B}_b$ that is decoded by encoder 1 at the end of block b . Similarly, encoder 2 determines $m_{1(b-1)}$ at the end of block b by decoding $\mathbf{A}_b$. At this point, the decoder cannot decode the pair $(\mathbf{A}_b, \mathbf{B}_b)$ since it is disadvantaged compared to the encoders. Hence, at the end of block b , it is left with a list of possible message pairs $(m_{1(b-1)}, m_{2(b-1)})$. This list is resolved by both encoders together in block $(b+1)$ using a common sequence $\mathbf{U}_{b+1}$.

Note that $\mathbf{A}_b$ and $\mathbf{B}_b$ are correlated random sequences since $(m_{1(b-1)}, m_{2(b-1)})$ is a correlated message pair [12] given $\mathbf{Z}_{b-1}$. The correlation between $\mathbf{A}_b$ and $\mathbf{B}_b$ cannot be arbitrary—it is generated using the information available at each encoder at the end of block $(b-1)$. At this time, both encoders know $(\mathbf{z}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1}$. In addition, encoder 1 also knows $\mathbf{x}_{b-1}$ and generates $\mathbf{A}_b \sim Q^1(\cdot | (\mathbf{z}, \mathbf{x}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1})$. Similarly, encoder 2 generates $\mathbf{B}_b \sim Q^2(\cdot | (\mathbf{z}, \mathbf{y}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1})$.

In order to obtain a single-letter characterization, we need to make sure that the $(\mathbf{A}_b, \mathbf{B}_b)$ pair has the same joint distribution in each block b . This is achieved using $\mathbf{W}_b$, which summarizes the information known to both encoders just before they transmit block b . $\mathbf{w}_b$ is defined as $f((\mathbf{z}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1})$, where (3) ensures that $\mathbf{w}_b \sim P_W$. Then (4) ensures that given $\mathbf{w}_b$, the $(\mathbf{A}_b, \mathbf{B}_b)$ pairs—each corresponding to one possible $(m_{1(b-1)}, m_{2(b-1)})$ —are as if picked independently of each other from the typical set $T(P_{AB|W})$. Conditioned on $\mathbf{w}_b$, (4) guarantees that all the sequences corresponding to block b will be independent of all the sequences corresponding to block $b-1$, $\forall b$. This ensures stationarity of the coding scheme.

The coding scheme described above is similar in spirit to that of Han for two-way channels [13]. If we set $W = A = B = \phi$, the conditions are trivially satisfied and we obtain the Cover-Leung region (1). In the C-L coding scheme, the rates (R_1, R_2) are low enough so that each encoder can decode the message of the other at the end of the *same* block. In our strategy, the rates are too high to do this, so the encoders wait one more block to decode their messages (using $\mathbf{A}$ and $\mathbf{B}$). In both schemes, the encoders cooperate to resolve the decoders list one block after they decode each other's message.

It is easy to see that our scheme can be extended as follows. Suppose that the rates are high enough so that the correlated information $\mathbf{A}$ and $\mathbf{B}$ cannot be decoded, i.e., the encoders are not able to decode the messages even with a delay of one block. We can then introduce another pair of (correlated) auxiliary random variables that the encoders decode with a delay of two blocks. In other words, we can extend our rate region by introducing an extra pair of auxiliary random variables. Clearly, this can be repeated to obtain potentially larger rate-regions, albeit with more auxiliary random variables.

III. PROOF SKETCH OF THEOREM 1

Fix a joint distribution $P_{UWABXYZ}$ as in (2) with corresponding f, Q^1, Q^2 such that (3) and (4) are satisfied.

A. Random Codebook generation

Choose 2^{NR_0} sequences $\mathbf{u}[m]$ independently from the typical set $T(P_U)$, $1 \leq m \leq 2^{NR_0}$. For each $(\mathbf{w}, \mathbf{u}, \mathbf{a}, \mathbf{b}) \in T(P_{WUAB})$, generate 2^{NR_1} sequences $\mathbf{x}[i]$ with $x_n \sim P_{X|UWA}$, $1 \leq i \leq 2^{NR_1}$ and 2^{NR_2} sequences $\mathbf{y}[j]$ with $y_n \sim P_{Y|UWB}$, $1 \leq j \leq 2^{NR_2}$.

For each $(\mathbf{z}, \mathbf{x}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w}) \in T(P_{ZXABUW})$, generate one sequence $\bar{\mathbf{a}}$ with $\bar{a}_n \sim Q^1(\cdot | z_n, x_n, a_n, b_n, u_n, w_n)$. Similarly, for each $(\mathbf{z}, \mathbf{y}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w}) \in T(P_{ZYABUW})$, generate one sequence $\bar{\mathbf{b}}$ with $\bar{b}_n \sim Q^2(\cdot | z_n, y_n, a_n, b_n, u_n, w_n)$.

B. Encoding

For block 1, $(\mathbf{u}, \mathbf{w}, \mathbf{a}, \mathbf{b})_1$ are chosen a priori at random from $T(P_{UWAB})$, and $\mathbf{x}_1[i], \mathbf{y}_1[j]$ are transmitted from the respective codebooks, if (i, j) are the messages in block 1. For $2 \leq b \leq B$, the encoders generate the inputs $\mathbf{X}_b, \mathbf{Y}_b$ for block b as follows.

- 1) Both encoders agree on the index in $\{1, \dots, 2^{NR_0}\}$ to resolve the decoder's list of message pairs for block $(b-2)$, and choose $\mathbf{u}_b$ corresponding to that index.

- 2) At the end of block $(b-1)$, both encoders know $(\mathbf{z}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1}$, and thus $\mathbf{w}_b = f((\mathbf{z}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1})$. Encoder 1 chooses $\mathbf{A}_b = \bar{\mathbf{a}}$, where $\bar{\mathbf{a}}$ is the sequence corresponding to $(\mathbf{z}, \mathbf{x}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1}$ (cf. Section III-A). Similarly, encoder 2 chooses $\mathbf{B}_b = \bar{\mathbf{b}}$, with $\bar{\mathbf{b}}$ corresponding to $(\mathbf{z}, \mathbf{y}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1}$.
- 3) Finally, $(\mathbf{u}, \mathbf{w}, \mathbf{a}, \mathbf{b})_b$ determines the codebooks of encoder 1 and 2, from which $\mathbf{x}_b[i]$ and $\mathbf{y}_b[j]$ are transmitted if the messages are i and j , respectively. No new messages are transmitted in blocks $(B-1)$ and B — $(\mathbf{x}_{B-1}[1], \mathbf{y}_{B-1}[1])$ and $(\mathbf{x}_B[1], \mathbf{y}_B[1])$ are taken to be the codewords for these two blocks. The reduction in rate due to this is insignificant if B is very large.

C. Decoding

Upon receiving $\mathbf{z}_b$, the decoder decodes the message pair corresponding to block $(b-2)$, while the encoders decode the messages (of one another) corresponding to block $(b-1)$. This is explained below.

At the Decoder ($3 \leq b \leq B$): At the end of block $(b-1)$,

- the decoder has correctly decoded $\mathbf{u}_{b-1}$ from $\mathbf{z}_{b-1}$,
- it knows the tuple $(\mathbf{z}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2}$ (thus also $\mathbf{w}_{b-1}$).

At this point, the decoder has the following list of all possible message pairs for block $(b-2)$:

$$\begin{aligned} \mathcal{L}_{b-2} \triangleq \{ & (i, j) : (\mathbf{z}, \mathbf{x}[i], \mathbf{y}[j], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2} \text{ jointly typ.} \\ & \text{AND } (\mathbf{z}_{b-1}, \mathbf{a}_i, \mathbf{b}_j, \mathbf{u}_{b-1}, \mathbf{w}_{b-1}) \text{ jointly typ.} \}, \end{aligned} \quad (5)$$

where $\mathbf{a}_i$ denotes the $\bar{\mathbf{a}}$ -sequence corresponding to $(\mathbf{z}, \mathbf{x}[i], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2}$, and $\mathbf{b}_j$ denotes the $\bar{\mathbf{b}}$ -sequence corresponding to $(\mathbf{z}, \mathbf{y}[j], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2}$ (cf. Section III-A).

Upon receiving $\mathbf{z}_b$, the decoder determines the unique $m^* \in \{1, \dots, 2^{NR_0}\}$ such that $(\mathbf{z}_b, \mathbf{u}[m^*])$ is jointly typical. This is done correctly with high probability if

$$R_0 < I(U; Z) - \epsilon \quad (6)$$

If the list-size $|\mathcal{L}_{b-2}|$ is smaller than 2^{NR_0} , the decoder can resolve it using m^* . Assuming (i^*, j^*) is the message pair transmitted, we have $\mathcal{L}_{b-2} = \{(i^*, j^*)\} \cup \mathcal{L}_Y \cup \mathcal{L}_X \cup \mathcal{L}_{XY}$, where

$$\begin{aligned} \mathcal{L}_Y \triangleq \{ & (i^*, j), j \neq j^* : (\mathbf{z}, \mathbf{x}[i^*], \mathbf{y}[j], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2} \text{ jointly typ.} \\ & \text{AND } (\mathbf{z}_{b-1}, \mathbf{a}_{i^*}, \mathbf{b}_j, \mathbf{u}_{b-1}, \mathbf{w}_{b-1}) \text{ are jointly typ.} \}, \\ \mathcal{L}_X \triangleq \{ & (i, j^*), i \neq i^* : (\mathbf{z}, \mathbf{x}[i], \mathbf{y}[j^*], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2} \text{ jointly typ.} \\ & \text{AND } (\mathbf{z}_{b-1}, \mathbf{a}_i, \mathbf{b}_{j^*}, \mathbf{u}_{b-1}, \mathbf{w}_{b-1}) \text{ are jointly typ.} \}, \\ \mathcal{L}_{XY} \triangleq \{ & (i, j), i \neq i^*, j \neq j^* : (\mathbf{z}, \mathbf{x}[i], \mathbf{y}[j], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2} \text{ jointly} \\ & \text{typ. AND } (\mathbf{z}_{b-1}, \mathbf{a}_i, \mathbf{b}_j, \mathbf{u}_{b-1}, \mathbf{w}_{b-1}) \text{ are jointly typ.} \}, \end{aligned}$$

Consider $\mathcal{L}_Y$. The probability of $(\mathbf{z}, \mathbf{x}[i^*], \mathbf{y}[j], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2}$ being jointly typical can be upper-bounded by $2^{-N(I(Y; Z|XABWU) - \epsilon)}$. Conditioned on $\mathbf{w}_{b-1}$, note that all the sequences in block $b-1$ are independent of all the sequences in block $b-2$. Since $\mathbf{w}_{b-1}$ is known at both encoders at the beginning of block $b-1$, the probability of $(\mathbf{z}_{b-1}, \mathbf{a}_{i^*}, \mathbf{b}_j, \mathbf{u}_{b-1}, \mathbf{w}_{b-1})$ being jointly typical, given

$(\mathbf{z}, \mathbf{x}[i^*], \mathbf{y}[j], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-2}$ is jointly typical, can be upper-bounded by $2^{-N(I(B;Z|ABUW)-\epsilon)}$. We therefore have

$$E|\mathcal{L}_Y| < 2^{NR_2} 2^{-N(I(Y;Z|XABUW)-\epsilon)} 2^{-N(I(B;Z|AUW)-\epsilon)}$$

Using similar upper bounds for $E|\mathcal{L}_X|$ and $E|\mathcal{L}_{XY}|$, we obtain

$$\begin{aligned} E|\mathcal{L}_{b-2}| &< 1 + 2^{N(R_2-I(Y;Z|XABUW)-I(B;Z|AUW)+2\epsilon)} \\ &\quad + 2^{N(R_1-I(X;Z|YABUW)-I(A;Z|BUW)+2\epsilon)} \\ &\quad + 2^{N(R_1+R_2-I(XY;Z|UW)+2\epsilon)} \end{aligned} \quad (7)$$

Using (7) in the Markov inequality, one can show that with high probability, $|\mathcal{L}_{b-2}| < 2^{N(A+4\epsilon)}$, where

$$\begin{aligned} A &\triangleq \max\{R_1 + R_2 - I(XY;Z|UW), \\ &\quad R_1 - I(X;Z|YABUW) - I(A;Z|BUW), \\ &\quad R_2 - I(Y;Z|XABUW) - I(B;Z|AUW)\}. \end{aligned}$$

Thus $\mathbf{U}_b = \mathbf{u}[m^*]$ can resolve the list $\mathcal{L}_{b-2}$ if $2^{N(A+4\epsilon)} < 2^{NR_0}$. Combining this with (6), we conclude that m^* can uniquely index the correct message pair (i^*, j^*) in $\mathcal{L}_{b-2}$ if

$$\begin{aligned} R_1 + R_2 - I(XY;Z|UW) &< I(U;Z) - 5\epsilon, \\ R_1 - I(X;Z|YABUW) - I(A;Z|BUW) &< I(U;Z) - 5\epsilon, \\ R_2 - I(Y;Z|XABUW) - I(B;Z|AUW) &< I(U;Z) - 5\epsilon. \end{aligned} \quad (8)$$

The decoder now also knows the tuple $(\mathbf{z}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1}$ since $\mathbf{a}_{b-1} = \mathbf{a}_{i^*}$ and $\mathbf{b}_{b-1} = \mathbf{b}_{j^*}$. This is the information required for decoding at the end of the next block $b+1$.

At the Encoders ($2 \leq b \leq B$): At the end of block $(b-1)$, the encoders agree on a common $\mathbf{u}_b$ and also know $(\mathbf{z}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1}$ and $\mathbf{w}_b = f((\mathbf{z}, \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1})$. In the next block, $\mathbf{z}_b$ is received. Let $\mathbf{b}_j$ denote the $\mathbf{b}$ -sequence corresponding to $(\mathbf{z}, \mathbf{y}[j], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1}$, $1 \leq j \leq 2^{NR_2}$. If there is a unique j^* such that $(\mathbf{z}_b, \mathbf{x}_b, \mathbf{a}_b, \mathbf{b}_{j^*}, \mathbf{u}_b, \mathbf{w}_b)$ is jointly typical, encoder 1 decodes j^* to be the message sent by encoder 2 in block $(b-1)$. An error is made only if there exists $j \neq j^*$ such that $(\mathbf{z}, \mathbf{y}[j], \mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{w})_{b-1}$ and $(\mathbf{z}_b, \mathbf{x}_b, \mathbf{a}_b, \mathbf{b}_j, \mathbf{u}_b, \mathbf{w}_b)$ are each jointly typical. The probability of this event can be upper bounded by

$$\sum_{j=1}^{2^{NR_2}} 2^{-N(I(Y;Z|XABUW)-\epsilon)} 2^{-N(I(B;Z|XAUW)-\epsilon)}.$$

A similar upper bound can be obtained for the probability of encoder 2 incorrectly decoding i^* , the message sent by encoder 1 in block $(b-1)$. Hence the probability of each encoder decoding the other's message incorrectly can be made arbitrarily small if

$$\begin{aligned} R_2 + 2\epsilon &< I(Y;Z|XABUW) + I(B;Z|XAUW) \\ &= I(YB;Z|XAUW) = I(Y;Z|XAUW), \\ R_1 + 2\epsilon &< I(X;Z|YABUW) + I(A;Z|YBWU) \\ &= I(XA;Z|XBWU) = I(X;Z|YBWU). \end{aligned} \quad (9)$$

At this point, the decoder has a list $\mathcal{L}_{b-1}$ of possible message pairs for block $(b-1)$ (cf. (5)). The encoders both set $\mathbf{U}_{b+1} = \mathbf{u}[m^*]$, where $m^* \in \{1, \dots, 2^{NR_0}\}$ uniquely

determines $(i^*, j^*) \in \mathcal{L}_{b-1}$. Note that if no error is made by the encoders in decoding i^*, j^* at the end of block b , both encoders know $(\mathbf{a}_b, \mathbf{b}_b)$ to be equal to $(\mathbf{a}_{i^*}, \mathbf{b}_{j^*})$, as required for the next block.

Probability of Error: Using (4), one can show that $(\mathbf{u}, \mathbf{w}, \mathbf{a}, \mathbf{b}, \mathbf{x}, \mathbf{y}, \mathbf{z})_b$ is jointly typical according to (2) with high probability $\forall b$. Further, we have shown that the probability of error at the decoder and each encoder can be made arbitrarily small if (8) and (9) are satisfied.

IV. EXAMPLE

Consider the AWGN MAC with power constraint P on each of the inputs. This channel, with $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \mathbb{R}$, is defined by

$$Z = X + Y + N \quad (10)$$

where N is a Gaussian noise random variable with mean 0 and variance σ^2 that is independent of X and Y . The inputs $\mathbf{x}$ and $\mathbf{y}$ for each block satisfy $\frac{1}{N} \sum_{n=1}^N x_n^2 \leq P$, $\frac{1}{N} \sum_{n=1}^N y_n^2 \leq P$. For this channel, the equal-rate point on the boundary of the C-L region [1] is (R_{CL}, R_{CL}) where

$$R_{CL} = \frac{1}{2} \log \left(2\sqrt{1 + \frac{P}{\sigma^2}} - 1 \right) \quad (11)$$

The achievable rate region of Theorem 1 for the discrete memoryless case can be extended to the AWGN MAC using a similar proof, recognizing that in the Gaussian case superposition is equivalent to addition.

For the joint distribution $P_{UWABXYZ}$ in (2), define $U \sim \mathcal{N}(0, 1)$ and (W, A, B) jointly Gaussian with mean zero and covariance matrix

$$K_{WAB} = \begin{bmatrix} 1 & \gamma & -\gamma \\ \gamma & 1 & \lambda \\ -\gamma & \lambda & 1 \end{bmatrix}. \quad (12)$$

The input distributions $P_{X|UWA}$ and $P_{Y|UWB}$ are defined by

$$\begin{aligned} X &= \sqrt{\alpha P} I_X + \sqrt{\beta P} A + \sqrt{\alpha + \beta P} U, \\ Y &= \sqrt{\alpha P} I_Y + \sqrt{\beta P} B + \sqrt{\alpha + \beta P} U \end{aligned} \quad (13)$$

where I_X, I_Y are independent $\mathcal{N}(0, 1)$ random variables, $\alpha, \beta > 0$ and $\alpha + \beta \leq 1$. I_X and I_Y represent the fresh information and U is the resolution information for the decoder that the encoders cooperate to send. A is the information that encoder 2 decodes using feedback, and B the information that encoder 1 decodes. We need to verify that the two conditions of Theorem 1 are satisfied by this joint distribution and then evaluate the rates.

A. Verifying the conditions

At the end of block b , both encoders know $(\mathbf{Z}, \mathbf{A}, \mathbf{B}, \mathbf{U}, \mathbf{W})_b$. This tuple is represented as $\mathbf{W}_{b+1}$ in the next block $b+1$ using a function $f((\mathbf{Z}, \mathbf{A}, \mathbf{B}, \mathbf{U}, \mathbf{W})_b)$. Define the function as

$$\begin{aligned} \bar{W} &= f(Z, A, B, U) \\ &\triangleq \frac{Z - \sqrt{\beta P} A - \sqrt{\beta P} B - 2\sqrt{\alpha + \beta P} U}{\sqrt{2\alpha P + \sigma^2}}. \end{aligned} \quad (14)$$

TABLE I
COMPARISON OF EQUAL-RATE BOUNDARY POINTS

	P/σ^2				
	0.5	1	5	10	100
R_{CL}	0.2678	0.4353	0.9815	1.2470	2.1277
R^*	0.2754	0.4499	1.0067	1.2709	2.1400
R_{FBcap}	0.2834	0.4642	1.0241	1.2847	2.1439

From (13) and (10), we see that f removes the effect of $(\mathbf{U}, \mathbf{A}, \mathbf{B})_b$ from $\mathbf{Z}_b$, resulting in a $\mathbf{W}_{b+1}$ that is the sum of the fresh information in block b and the channel noise.

Sequences $\mathbf{A}_{b+1}$ and $\mathbf{B}_{b+1}$ are generated using $\mathbf{X}_b$ and $\mathbf{Y}_b$, respectively, along with the common tuple $(\mathbf{Z}, \mathbf{A}, \mathbf{B}, \mathbf{U}, \mathbf{W})_b$. Define the following functions to generate $(\mathbf{A}, \mathbf{B})_{b+1}$:

$$\begin{aligned} q_1 : \bar{A} &= k_1 \frac{X - \sqrt{\alpha + \beta P} U - \sqrt{\beta P} A}{\sqrt{\alpha P}} + k_2 f(U, A, B, Z), \\ q_2 : \bar{B} &= -k_1 \frac{Y - \sqrt{\alpha + \beta P} U - \sqrt{\beta P} B}{\sqrt{\alpha P}} - k_2 f(U, A, B, Z) \end{aligned} \quad (15)$$

where $k_1, k_2 \in \mathbb{R}$. With respect to (4), this is equivalent to

$$\begin{aligned} Q^1(\bar{a}|z, x, a, b, u, w) &= \mathbf{1}(q_1(z, x, a, b, u, w) = \bar{a}), \\ Q^2(\bar{b}|z, y, a, b, u, w) &= \mathbf{1}(q_2(z, y, a, b, u, w) = \bar{b}). \end{aligned}$$

(15) ensures that $(\mathbf{A}, \mathbf{B})_{b+1}$ depend on $(\mathbf{Z}, \mathbf{A}, \mathbf{B}, \mathbf{U})_b$ only through ' $\mathbf{Z}_b$ with the effect of $(\mathbf{A}, \mathbf{B}, \mathbf{U})_b$ removed'. In other words, conditioned on $\mathbf{W}_{b+1} = f((\mathbf{Z}, \mathbf{A}, \mathbf{B}, \mathbf{U})_b)$, $(\mathbf{A}, \mathbf{B})_{b+1}$ is independent of all the sequences in block b as required by (4). Indeed, it can be formally verified that (4) holds if the following equalities are satisfied.

$$E[\bar{A}^2] = E[\bar{B}^2] = 1, \quad E[\bar{A}\bar{B}] = \lambda, \quad E[\bar{A}\bar{W}] = -E[\bar{B}\bar{W}] = \gamma. \quad (16)$$

Evaluating the conditions in (16) using (14) and (15), we have

$$1 = E[\bar{A}^2] = k_1^2 + k_2^2 + 2k_1k_2\sqrt{\frac{\alpha P}{2\alpha P + \sigma^2}}, \quad (17)$$

$$\lambda = E[\bar{A}\bar{B}] = -k_2^2 - 2k_1k_2\sqrt{\frac{\alpha P}{2\alpha P + \sigma^2}}, \quad (18)$$

$$\gamma = E[\bar{A}\bar{W}] = k_2 + k_1\sqrt{\frac{\alpha P}{2\alpha P + \sigma^2}}. \quad (19)$$

Once we choose (α, β, λ) , (17) and (18) determine k_1 and k_2 . Using these in (19), we obtain the value of γ in order for the joint distribution to be valid for Theorem 1.

B. Evaluating the rates

For a valid $(\alpha, \beta, \lambda, \gamma)$, the achievable rates can be evaluated from Theorem 1 to be

$$\begin{aligned} R_1, R_2 &< \min\{G, H\}, \\ R_1 + R_2 &< \frac{1}{2} \left(1 + \frac{2P}{\sigma^2} + \frac{2P}{\sigma^2} (\overline{\alpha + \beta} + \lambda\beta) \right), \end{aligned} \quad (20)$$

where

$$\begin{aligned} G &= \frac{1}{2} \log \left(1 + (\alpha + \beta) \frac{P}{\sigma^2} \right. \\ &\quad \left. - \beta \frac{P}{\sigma^2} \cdot \frac{\lambda^2 + \gamma^2(1 - \lambda^2 + 2\lambda) - \gamma^4(1 + 2\lambda)}{(1 - \gamma^2)^2} \right), \\ H &= \frac{1}{2} \log \left(1 + \frac{4\overline{\alpha + \beta}P/\sigma^2}{2(\alpha + \beta + \beta\lambda)P/\sigma^2 + 1} \right) + \frac{1}{2} \log \left(1 + \alpha \frac{P}{\sigma^2} \right) \\ &\quad + \frac{1}{2} \log \left(1 + \frac{\beta(1 + \lambda)P/\sigma^2}{1 + 2\alpha P/\sigma^2} \left(2 - \frac{1 + \lambda}{1 - \gamma^2} \right) \right). \end{aligned}$$

For different values of the signal-to-noise ratio P/σ^2 , we (numerically) compute the equal-rate point (R^*, R^*) on the boundary of (20). For various values of P/σ^2 , Table I compares R^* with R_{CL} , the equal-rate point of the C-L region given by (11), and with the equal rate-point R_{FBcap} on the boundary of the feedback capacity region [6]. We observe that our equal-rate points represent a significant improvement over the C-L region, and are close to the feedback capacity for large SNR. In [8, pp. 82-84], a two-letter generalization of the Cover-Leung region is given. The equal-rate point on the boundary of that region for $P/\sigma^2 = 10$ is 1.2566, which is smaller than the corresponding R^* given in Table I.

It is part of ongoing work to compare our rate region with that of [10]. Though direct comparison of the two regions appears difficult in general, it may be possible for specific channels. Moreover, as explained at the end of Section II, our strategy can be extended to yield larger rate-regions by adding more auxiliary random variables.

REFERENCES

- [1] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 27, no. 3, pp. 292-298, 1981.
- [2] R. Ahlswede, "Multi-way communication channels," in *Proc. Second Int. Symp. Inform. Transmission*, Tsahkadsor, Armenia, USSR, Hungarian Press, 1971.
- [3] H. D. Liao, *Multiple-Access Channels*. PhD thesis, Univ. Hawaii, 1972.
- [4] N. T. Gaarder and J. K. Wolf, "The capacity region of a multiple-access discrete memoryless channel can increase with feedback," *IEEE Trans. on Inf. Theory*, vol. 21, no. 1, pp. 100-102, 1975.
- [5] F. M. J. Willems, "The feedback capacity region of a class of discrete memoryless multiple access channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 93-95, 1982.
- [6] L. H. Ozarow, "The capacity of the white gaussian multiple access channel with feedback," *IEEE Trans. Inf. Theory*, vol. 30, no. 4, pp. 623-628, 1984.
- [7] M. Wigger, *Cooperation on the Multiple-Access Channel*. PhD thesis, Swiss Federal Institute of Technology, Zurich, 2008.
- [8] G. Kramer, *Directed Information for channels with Feedback*. PhD thesis, Swiss Federal Institute of Technology, Zurich, 1998.
- [9] G. Kramer, "Capacity Results for the Discrete Memoryless Network," *IEEE Trans Inf Theory*, vol. 49, pp. 4-20, January 2003.
- [10] S. Bross and A. Lapidoth, "An improved achievable region for the discrete memoryless two-user multiple-access channel with noiseless feedback," *IEEE Trans Inf Theory*, vol. IT-51, pp. 811-833, March 2005.
- [11] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. 26, no. 6, pp. 648-657, 1980.
- [12] S. S. Pradhan, S. Choi, and K. Ramchandran, "A graph-based framework for transmission of correlated sources over multiple-access channels," *IEEE Trans Inf Theory*, vol. 53, no. 12, pp. 4583-4604, 2007.
- [13] T. S. Han, "A general coding scheme for the two-way channel," *IEEE Trans Inf Theory*, vol. 30, no. 1, pp. 35-43, 1984.