

On the Capacity of Multiple Access Channels with State Information and Feedback

Wei Wu, Sriram Vishwanath and Ari Arapostathis

Abstract—In this paper, the multiple access channel (MAC) with channel state is analyzed in a scenario where a) the channel state is known non-causally to the transmitters and b) there is perfect causal feedback from the receiver to the transmitters. An achievable region and an outer bound are found for a discrete memoryless MAC that extend existing results, bringing together ideas from the two separate domains of MAC with state and MAC with feedback. Although this achievable region does not match the outer bound in general, special cases where they meet are identified.

In the case of a Gaussian MAC, a specialized achievable region is found by using a combination of dirty paper coding and a generalization of the Schalkwijk-Kailath [1], Ozarow [2] and Merhav-Weissman [3] schemes, and this region is found to be capacity achieving. Specifically, it is shown that additive Gaussian interference that is known non-causally to the transmitter causes no loss in capacity for the Gaussian MAC with feedback.

Index Terms—Network information theory, multiple access channel, dirty-paper coding, feedback capacity, Gel'fand-Pinsker coding

I. INTRODUCTION

The capacity of channels with perfect feedback [4]–[7] is of great interest, as it provides us with an outer limit on the performance of any feedback-based scheme. Although the capacity of single user memoryless channels is unaffected by feedback [8], that of multiple access channels (MAC) is known to be enhanced by feedback from the receiver to both transmitters [2]. In the two-user Gaussian MAC case, the entire capacity region can be found in closed form employing Ozarow's ingenious extension [2] of the Schalkwijk-Kailath (SK) coding scheme [1] that enables limited cooperation between transmitters by using the feedback information.

On a parallel track, the capacity of channels with state has been studied under a varied set of assumptions on state knowledge [9]–[11]. An important result in this family is when the state in a discrete memoryless (DM) single user channel is known non-causally to the transmitter [10], [11]. In the all-Gaussian (additive noise and interference) case, this result, combined with a clever choice for the auxiliary variable translates into the Costa-coding result [12]. The capacity of discrete memoryless multiple-access channels is still an open

problem, but in the Gaussian case a result that is similar the Costa result is shown in [13].

Recently, these two classes of problems were combined in the study of memoryless single user channels with non-causal state knowledge and feedback [3]. It was shown in this work that there is no capacity gain from feedback, and in the all-Gaussian case, interference causes no loss in capacity.

A. Our contributions

In this work, from one point of view, we are interested in understanding the impact of feedback in a MAC channel with state, and from the other, the impact of state on a MAC channel with feedback. As a concrete research challenge, we focus on a MAC channel with state and feedback where the state variable is known non-causally to both the transmitters, and find a general achievable region and outer bound on the capacity region. We find this achievable region and outer bound to meet for a non-trivial class of channels which includes the binary symmetric and binary erasure channels.

Next, we specialize our study to the case of a Gaussian MAC with additive interference and feedback. The achievable region we obtain for this scenario coincides with a simple outer bound on the system, and thus results in the capacity region. From this capacity characterization, we find that:

- Feedback enhances the capacity of the MAC channel with state.
- Interference when non-causally known at the transmitter has no impact on the capacity region of the Gaussian 2 user MAC channel.

Thus, our results are analogous to the Merhav-Weissman [3] and Costa [12] results for the single-user case.

B. Organization

The rest of the paper is organized as follows. In Section II, basic definitions and notation used are introduced. The main results of the paper for the discrete memoryless channel case are presented in Section III. The Gaussian case is handled in Section IV. Detailed proofs for Sections III are presented in Section V and the correspondence concludes with Sectionsec:conclude.

II. NOTATIONS AND PRELIMINARIES

A. Notation

We adopt the following notation throughout the correspondence. For matrix A , A^T , A^{-1} denote the transpose and inverse of A respectively. Random variables (RVs) will be

This research was supported in part by NSF grants CCF-0448181, CCF-0552741, ECS-0218207 and ECS-0225448, THECB ARP 010115-0013-2006, the Office of Naval Research through the Electric Ship Research and Development Consortium and a grant from Freescale Semiconductor Corporation. Wei Wu was also supported by the Hemphill-Gilmore Student Endowed Fellowship through the University of Texas at Austin.

The authors are with Wireless Networking and Communications Group, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712, USA (e-mail: {wwu,sriram,ari}@ece.utexas.edu).

noted by capital letters, while their realizations will be denoted by the respective lower case letters. X_m^n denotes the random vector (X_m, \dots, X_n) , and X_i^n denotes the random vector $(X_{i,1}, \dots, X_{i,n})$. Both $X_{i,j}$ and $X_i(j)$ is used to denote the j -th random variable of a random vector X_i . $\mathbb{E}[X]$ denotes the expectation of random variable X and the correlation coefficient of two scalar random variable X_1, X_2 is defined as

$$\rho_{X_1 X_2} = \frac{\mathbb{E}[X_1 X_2]}{\sqrt{\mathbb{E}[X_1 X_1^T] \mathbb{E}[X_2 X_2^T]}}.$$

The alphabet of a random variable X will be designated by a calligraphic letter \mathcal{X} , and that of the n -fold Cartesian power of \mathcal{X} will be denoted as \mathcal{X}^n .

$$X \Rightarrow Y \Rightarrow Z$$

will be used to denote the conditional independence of X and Z given Y .

B. Models and definitions

A two-user *multiple access channel with random parameters* $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{S}, \mathcal{Y}, P(y|x_1, x_2, s))$ is a channel with two input alphabets $\mathcal{X}_1, \mathcal{X}_2$; state space \mathcal{S} , output alphabet \mathcal{Y} , and transition probability $P(y|x_1, x_2, s)$. The states s take values in \mathcal{S} according to the probability mass function (PMF) $P(s)$. It is assumed that the channel and the state are both memoryless, namely,

$$P(y^n|x_1^n, x_2^n, s^n) = \prod_{i=1}^n P(y_i|x_{1,i}, x_{2,i}, s_i)$$

and

$$P(s^n) = \prod_{i=1}^n P(s_i).$$

Here we assume the state variable is noncausally known. Both *noncausal* state information and *feedback* are incorporated into the channel model via the definition of a noncausal feedback code (R_1, R_2, n) as follows.

Definition 2.1: An (R_1, R_2, n) code for the MAC $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{S}, \mathcal{Y}, P(y|x_1, x_2, s))$ with feedback and noncausal state information is defined by encoding functions and decoding functions:

- 1) The encoding functions for user i are the mappings

$$f_{i,k} : \{1, \dots, 2^{nR_i}\} \times \mathcal{Y}^{k-1} \rightarrow \mathcal{X}_i, \\ i = 1, 2, k = 1, 2, \dots, n,$$

or in other words, for the message of user i , $w_i \in \{1, 2, \dots, 2^{nR_i}\}$, $i = 1, 2$, the channel input is expressed as

$$x_{i,k} = f_{i,k}(w_i, y^{k-1}, s^n), \quad (1)$$

where s^n is the noncausal state information of the whole block and y^{k-1} is the perfect feedback of channel output up to time $t - 1$.

- 2) The decoding functions for the receiver are the mappings

$$g : \mathcal{Y}^n \rightarrow \{1, \dots, 2^{nR_1}\} \times \{1, \dots, 2^{nR_2}\},$$

or in other words, the decoder gives the estimates of the two messages w_1, w_2 , $\hat{w} = (\hat{w}_1, \hat{w}_2)^T$,

$$\hat{w} = g(y^n). \quad (2)$$

We shall use the average probability of error criterion P_e assuming that the messages (w_1, w_2) are drawn according to uniform distribution over $\{1, \dots, 2^{nR_1}\} \times \{1, \dots, 2^{nR_2}\}$.

Definition 2.2: A rate pair (R_1, R_2) per channel use is achievable for the MAC with feedback and noncausal state information if there exists a sequence of (R_1, R_2, n) codes such that $P_e \rightarrow 0$ as $n \rightarrow \infty$.

The capacity region of MAC with feedback and noncausal state information, $\mathcal{C}_{fb,nc}^{MAC}$, is the closure of the set of all achievable rates. $\mathcal{C}_{fb,nc}^{MAC}$ is known to be convex by time multiplexing of achievable rates.

III. MAIN RESULTS

Let \mathcal{P} stand for the collection of all RVs $(S, U, X_1, X_2, V_1, V_2, Y)$ while U, V_1, V_2 are auxiliary variables introduced to form a Markov chain

$$S \Rightarrow U \Rightarrow ((X_1, V_1), (X_2, V_2)) \Rightarrow Y.$$

Define \mathcal{R}_i^{MAC} to be the set of all rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq I(X_1; Y | X_2, U, S) \\ R_2 &\leq I(X_2; Y | X_1, U, S) \\ R_1 + R_2 &\leq I(V_1, V_2; Y) - I(V_1, V_2; S) \end{aligned} \quad (3)$$

with joint distribution $P(u|s)P(v_1, x_1|u, s)P(v_2, x_2|u, s)$. As stated in the next theorem, the set \mathcal{R}_i^{MAC} is an inner bound on the capacity region $\mathcal{C}_{fb,nc}^{MAC}$.

Theorem 3.1: The capacity region of the MAC channel $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{S}, \mathcal{Y}, P(y|x_1, x_2, s))$, with feedback and noncausal state information at both encoders, satisfies

$$\mathcal{R}_i^{MAC} \subseteq \mathcal{C}_{fb,nc}^{MAC}.$$

In this achievable region, the auxiliary variable “U” reflects the amount of common information shared between the two transmitters, while “ V_i ” is the auxiliary variable associated with the message from Transmitter i . This expression is highly intuitive - the sum rate expression in (3) resembles a generalized Gel’fand-Pinsker expression where there is noncausal side information at the transmitters and no information at the receivers; while the two individual constraints reflect the scenario where both the transmitter and the receiver know the channel state.

A proof of this is given in Section V-A. This proof builds on the Cover-Leung [14] and Gel’fand-Pinsker arguments [10]. It differs from them in the following ways:

- i) binning is used to determine common transmission (U),
- ii) backward decoding is employed at the receiver; and
- iii) sequences v_1^n and v_2^n are not placed in separate bins at each transmitter.

All three changes are introduced both to facilitate the result and as simplifying mechanisms to make the proof tractable.

The outer bound is stated next. Define \mathcal{R}_o^{MAC} to be the set of all rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq I(V_1; Y|V_2) - I(V_1; S|V_2) \\ R_2 &\leq I(V_2; Y|V_1) - I(V_2; S|V_1) \\ R_1 + R_2 &\leq I(V_1, V_2; Y) - I(V_1, V_2; S) \end{aligned} \quad (4)$$

for all joint distribution $P(v_1, v_2, x_1, x_2|s)$.

Theorem 3.2: The capacity region of the MAC channel $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{S}, \mathcal{Y}, P(y|x_1, x_2, s))$, with feedback and noncausal state information at both encoders, satisfies

$$\mathcal{C}_{fb,nc}^{MAC} \subseteq \mathcal{R}_o^{MAC}.$$

This outer bound expression is quiet intuitive. As V_1, V_2 represent the messages from Transmitters 1 and 2, the region resembles a combination of the multiple access capacity region combined with a generalization of the Gel'fand-Pinsker expression. In that spirit, the proof for this outer bound is an extension of the Gel'fand-Pinsker arguments [10] with one important modification. For the proof and the modification to the Gel'fand-Pinsker argument, see Section V-B. with one important modification. For the proof and the modification to the Gel'fand-Pinsker argument, see Section V-B.

Although the achievable region and outer bounds do not meet in general, one can determine non-trivial sufficient conditions for them to do so. Consider the class of MAC channels (called class Γ) that satisfy either

$$H(X_1|S, X_2, Y) = 0 \quad \text{or} \quad H(X_2|S, X_1, Y) = 0. \quad (5)$$

Theorem 3.3: The capacity region of a MAC channel in class Γ , $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{S}, \mathcal{Y}, P(y|x_1, x_2, s))$, with feedback and noncausal state information at both encoders, is given by

$$\mathcal{C}_{fb,nc}^{MAC} = \mathcal{R}_i^{MAC}.$$

The proof of Theorem 3.3 is provided in Section V-C. The key idea is to use the condition in (5) to prove a tighter outer bound for the class channels belonging to Γ .

The condition in (5) covers a wide class of discrete memoryless MAC channels. For the input sets and state set $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{S} = \{0, 1\}$, both the binary MAC adder channel,

$$Y = (X_1 + X_2 + S) \bmod 2, \quad \mathcal{Y} = \{0, 1\},$$

and the MAC "erasure-type" channel,

$$Y = (X_1 + X_2 + S) \bmod 3, \quad \mathcal{Y} = \{0, 1, 2\},$$

satisfy (5), thus the capacity region can be characterized using Theorem 3.3.

On the other hand, the Gaussian MAC channel does not belong to class Γ thus Theorem 3.3 does not apply. In the next section, we will develop a coding strategy tailored to the Gaussian MAC channel. This coding scheme builds on Ozarow's feedback coding scheme [2] and the dirty paper coding strategy by Costa [12] to obtain the full capacity region.

IV. THE CAPACITY REGION OF THE GAUSSIAN MAC CHANNEL WITH RANDOM PARAMETERS

Consider the Gaussian multiple access channel

$$Y(k) = X_1(k) + X_2(k) + S(k) + Z(k),$$

where $\{S(k)\}$ denotes the interfering signal, which are i.i.d. Gaussian random variables with $\mathbb{E}S(k) = 0$ and $\sigma_S^2 = \mathbb{E}S^2(k) \leq \infty$. It is assumed to be known to the encoder non-causally. $Z(k)$ in this model is zero-mean, white Gaussian noise with variance σ_Z^2 . The average powers for the two transmitters are assumed to be P_1, P_2 respectively.

One of the first results for this class of channels is in [3], where Merhav and Weissman show feedback does not increase the capacity of the point-to-point version of this channel, but one can significantly reduce the coding complexity and achieve *doubly exponential* error exponent by using an extension of the Schalkwijk-Kailath (SK) [1] coding scheme.

In this section, we will show for the Gaussian MAC channel with feedback and noncausal state information at the transmitters, the capacity is as if the interference were absent, with a double exponential decay in probability of error.

The coding scheme we employ to show our results is a mixture of two different schemes - the Ozarow coding scheme [2] and the Merhav-Weissman scheme [3]; each one of which is a generalization of the original Schalkwijk-Kailath coding scheme.

A. Coding Scheme that achieves sum capacity

We use the following notation in this subsection: k is the time index, i is the transmitter index (denoted by T_i). The main idea of this scheme is to set an initial condition, and over time send corrections (innovations) to the receiver to allow the receiver to converge to this initial condition. Here is the coding scheme:

Before transmission. Given a message for transmitter T_i , m_i , $m_i \in \{0, 1, \dots, M_i - 1\}$, $M_i = 2^{nR_i}$, map m_i to a point on the real line as follows: $\theta_i^0 = (m_i + 1/2)/M_i$. Define $a_i \triangleq (-1)^{i-1}2^{R_i}$. Given the noncausal state $S^n = (S(1), \dots, S(n))$, compute a *precancelling* message $\theta_i(k)$ for $k = 2, \dots, n$, as

$$\theta_i(p) = \theta_i^0 + \sum_{j=p+1}^n \frac{l_i S(j)}{a_i^{p-2}},$$

where $p \in \{2, \dots, n\}$ and l_i is a scaling factor that will be presented later in this section.

Initialization. This is what we call the first two transmissions ($k = 1, 2$). At time $k = 1$, T_2 sends nothing while T_1 sends

$$X_1(1) = \theta_1(2) - S(1).$$

The receiver obtains $Y_1 = X_1(1) + S(1) + Z(1) = \theta_1(2) + Z(1)$ and receiver finds an initial estimate $(\hat{\theta}_1)$ for θ_1^0 to be $\hat{\theta}_1(1) = Y(1)$. At time $k = 2$, T_1 sends nothing while T_2 sends

$$X_2(2) = \theta_2(2) - S(2).$$

Then receiver sets its initial estimate for θ_2^0 to be

$$\hat{\theta}_2(2) = Y(2) = \theta_2(2) + Z(2).$$

Estimation recursion. This is the remainder of the transmissions $k = 3, \dots, n$. Defining $\epsilon(k) \triangleq \hat{\theta}_i(k) - \theta_i(k)$. At time k , T_i transmits a scaled version of ϵ_k :

$$X_i(k) = a_i^{k-2} \epsilon_i(k-1). \quad (6)$$

At the end of time k , the receiver updates its estimate of message θ_i^0 as:

$$\hat{\theta}_i(k) = \hat{\theta}_i(k-1) - a_i^{-(k-2)} l_i Y(k), \quad (7)$$

while Transmitter T_i updates ϵ_i as

$$\begin{aligned} \epsilon_i(k) &= \hat{\theta}_i(k) - \theta_i(k) \\ &= (\hat{\theta}_i(k-1) - \theta_i(k-1)) + \frac{l_i S(k)}{a_i^{k-2}} \\ &\quad - \frac{l_i (X_1(k) + X_2(k) + S(k) + Z(k))}{a_i^{k-2}} \\ &= \epsilon_i(k-1) - \frac{l_i (X_1(k) + X_2(k) + Z(k))}{a_i^{k-2}}. \end{aligned}$$

Analysis

The estimation error at the receiver for T_i at the end of the k th transmission, denoted $\tilde{\epsilon}_i(k)$ is

$$\tilde{\epsilon}_i(k) = \hat{\theta}_i(k) - \theta_i^0 = \epsilon_i(k) + (\theta_i(k) - \theta_i^0).$$

At the end of the block $k = n$, $\theta_i(n) = \theta_i^0$ thus $\tilde{\epsilon}_i(n) = \epsilon_i(n)$. As long as $\epsilon_i(n)$ goes to zero as $n \rightarrow \infty$, the estimation error at the receiver goes to zero also.

We denote $Y' = X_1 + X_2 + Z$. Notice that all the steps taken above were *linear*. We can combine them into one system equation as:

$$X(k+1) = AX(k) - LY'(k), \quad (8)$$

where

$$X = [X_1, X_2]^T, \quad A = \text{diag}(a_1, a_2), \quad L = [a_1 l_1, a_2 l_2]^T.$$

Along the lines of [2] and [6], we chose the optimal L that minimizes the mean squared error as:

$$L(k) = \frac{\mathbb{E}[Y'(k)X(k)]}{\mathbb{E}[Y'(k)^2]}.$$

Denoting $Q = \mathbb{E}[X(k)X^T(k)]$, we obtain

$$Q(k+1) = A[Q(k) - Q(k)H^T(HQ(k)H^T + \sigma_Z^2)^{-1}HQ(k)]A,$$

from (8) where $H = [1 \ 1]$ and $Y' = HX + Z$. Since (H, A) is detectable, the matrix recursion above converges and $Q(k) \rightarrow Q$ as k goes to ∞ , satisfying

$$Q = A[Q - QH^T(HQH^T + \sigma_Z^2)^{-1}HQ]A. \quad (9)$$

From (9), we solve for Q in terms of a_i and σ_Z^2 . Note that the diagonal elements of Q are the individual power constraints P_i , and the off-diagonal element is the correlation between the transmit signals ρ . These turn out to be: (9),

$$P_1 = \frac{(a_1^2 - 1)(|a_1 a_2| + 1)^2}{(|a_1| + |a_2|)^2} \sigma_Z^2 \quad (10)$$

$$P_2 = \frac{(a_2^2 - 1)(|a_1 a_2| + 1)^2}{(|a_1| + |a_2|)^2} \sigma_Z^2, \quad (11)$$

and the correlation coefficient ρ between X_1 and X_2 satisfies,

$$\rho = \sqrt{\frac{(a_1^2 - 1)(a_2^2 - 1)}{(|a_1| + |a_2| + 1)^2}}. \quad (12)$$

One can find the achievable rate (R_1, R_2) in terms of the power constraints (P_1, P_2) and ρ by rewriting the equations above as:

$$\begin{aligned} R_1 &= \frac{1}{2} \log \left(1 + \frac{P_1(1 - \rho^2)}{\sigma_Z^2} \right) \\ R_2 &= \frac{1}{2} \log \left(1 + \frac{P_2(1 - \rho^2)}{\sigma_Z^2} \right) \\ R_1 + R_2 &= \frac{1}{2} \log \left(1 + \frac{P_1 + P_2 + 2\rho\sqrt{P_1 P_2}}{\sigma_Z^2} \right). \end{aligned} \quad (13)$$

It is well-known that (13) is the sum-capacity point for the normal two-user Gaussian MAC with feedback, which is certainly the outer bound of the counterpart with state variable. Thus (13) is the also the sum-capacity for Gaussian MAC with feedback and noncausal state.

Remark 4.1: Note that this precancellation strategy requires non-causal knowledge of the interference at both transmitters and causal may not be sufficient. Also note that the receiver updates the estimates *as if* there is no interference and by the end of the block, and due to the pre-cancellation, the estimation errors are not affected by the state.

Remark 4.2: Note $\mathbb{E}[X_1(1)^2] < \infty$ and $\mathbb{E}[X_2(2)^2] < \infty$, due to the average power constraint, as long as the power assumption at the *initialization stage* is finite, the average power is asymptotically close to P_i as $n \rightarrow \infty$. Moreover, $\{S(k)\}$ is not restricted to be Gaussian to achieve sum-capacity (13). All that is required is $\sigma_S^2 < \infty$.

Remark 4.3: Using the same type of pre-canceling, one can extend the coding scheme to Gaussian broadcast (BC) channel and Gaussian interference channel [4] and [6].

B. Hybrid coding that achieve other points in the capacity region

In the previous section, we focus on sum-capacity (13). In this section, we combine the coding scheme in Subsection IV-A with dirty-paper coding (Costa coding) to obtain the full capacity region. The strategy here mimics the approach used in [2] with an important difference - instead of superposition coding, we employ dirty paper coding.

Let one transmitter, say T_1 , have two messages which we call $m_1^{(1)}$ and $m_1^{(2)}$ at rates $R_1^{(1)}$ and $R_1^{(2)}$ respectively. Given the noncausal state S^n , let T_1 use power αP_1 , ($0 \leq \alpha \leq 1$) to transmit $m_1^{(2)}$ and use the remainder $\bar{\alpha} P_1$ ($\bar{\alpha} = 1 - \alpha$) to transmit $m_1^{(1)}$. $m_1^{(1)}$ is transmitted using the feedback coding scheme described in Subsection IV-A, while $m_1^{(2)}$ is sent independently. Meanwhile, T_2 uses all his power to transmit m_2 using the feedback coding scheme in Subsection IV-A.

Decoding: at the receiver, messages $m_1^{(1)}$ and m_2 are first decoded by treating the code letters of $m_1^{(2)}$ as noise. Finally, $m_1^{(2)}$ is decoded.

According to (13), $m_1^{(1)}$ and m_2 will be transmitted reliably iff

$$R_1^{(1)} \leq \frac{1}{2} \log \left(1 + \frac{\bar{\alpha} P_1(1 - \rho^2)}{\sigma_Z^2 + \alpha P_1} \right) \quad (14)$$

$$R_2 \leq \frac{1}{2} \log \left(1 + \frac{P_2(1 - \rho^2)}{\sigma_Z^2 + \alpha P_1} \right), \quad (15)$$

where ρ satisfies

$$\left(1 + \frac{\bar{\alpha}P_1(1-\rho^2)}{\sigma_Z^2 + \alpha P_1}\right) \left(1 + \frac{P_2(1-\rho^2)}{\sigma_Z^2 + \alpha P_1}\right) = \left(1 + \frac{\bar{\alpha}P_1 + P_2 + 2\rho\sqrt{\bar{\alpha}P_1P_2}}{\sigma_Z^2 + \alpha P_1}\right). \quad (16)$$

At the end of the block, after decoding $m_1^{(1)}$ and m_2 , one can, with high probability, subtract the feedback coding from the channel output and obtain \tilde{Y}^n , where

$$\tilde{Y}^n = X_1^n(m_1^{(2)}) + S^n + Z^n,$$

Here $X_1^n(m_1^{(2)})$ denotes the dirty paper coded transmit vector corresponding to message $m_1^{(2)}$. Then $m_1^{(2)}$ can be decoded reliably at rate

$$R_1^{(2)} \leq \frac{1}{2} \log \left(1 + \frac{\alpha P_1}{\sigma_Z^2}\right). \quad (17)$$

By adding (14) and (17), we obtain the total rate for T_1 as:

$$\begin{aligned} R_1 &= R_1^{(1)} + R_1^{(2)} \\ &\leq \frac{1}{2} \log \left[1 + \frac{(1 - \bar{\alpha}\rho^2)P_1}{\sigma_Z^2}\right]. \end{aligned}$$

Using (16) and (15), we can obtain

$$\begin{aligned} R_2 &\leq \frac{1}{2} \log \left[1 + \frac{\bar{\alpha}P_1 + P_2 + 2\rho\sqrt{\bar{\alpha}P_1P_2}}{\sigma_Z^2 + \alpha P_1}\right] \\ &\quad - \frac{1}{2} \log \left(1 + \frac{\bar{\alpha}P_1(1-\rho^2)}{\sigma_Z^2 + \alpha P_1}\right) \\ &= \frac{1}{2} \log \left[\frac{\sigma_Z^2 + P_1 + P_2 + 2\rho\sqrt{\bar{\alpha}P_1P_2}}{\sigma_Z^2 + (1 - \bar{\alpha}\rho^2)P_1}\right] \end{aligned}$$

Note by defining $\rho' = \sqrt{\bar{\alpha}}\rho$, $|\rho'| \leq 1$ and for any $0 \leq \rho' \leq \rho$,

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log \left[1 + \frac{(1 - \rho'^2)P_1}{\sigma_Z^2}\right] \\ R_2 &\leq \frac{1}{2} \log \left[\frac{\sigma_Z^2 + P_1 + P_2 + 2\rho'\sqrt{P_1P_2}}{\sigma_Z^2 + (1 - \rho'^2)P_1}\right] \end{aligned}$$

is achievable. A similar region can be obtained when T_2 splits the power to combine feedback coding and dirty paper coding. Thus the achievable region can be written as

$$\begin{aligned} &\bigcup_{0 \leq \rho \leq 1} \left\{ (R_1, R_2) : 0 \leq R_1 \leq \frac{1}{2} \log \left[1 + \frac{(1 - \rho^2)P_1}{\sigma_Z^2}\right] \right. \\ &\quad \left. 0 \leq R_2 \leq \frac{1}{2} \log \left[1 + \frac{(1 - \rho^2)P_2}{\sigma_Z^2}\right] \right. \\ &\quad \left. 0 \leq R_1 + R_2 \leq \frac{1}{2} \log \left[1 + \frac{P_1 + P_2 + 2\rho\sqrt{P_1P_2}}{\sigma_Z^2}\right] \right\}. \quad (18) \end{aligned}$$

Note (18) is the same as the capacity region of Gaussian MAC with feedback but without interference [2], which is clearly the outer bound of the channel with interference, thus (18) is also the capacity region of Gaussian MAC with feedback and noncausal interference, $C_{MAC}^{fb,S}$.

V. PROOFS

A. Proof of Theorem 3.1

Here we provide a proof for Theorem 3.1, the achievability of (3).

Code Generation: 2^{nR_0} strongly typical sequences $\sim p(u^n|s^n)$ are first generated. At transmitter i , for each sequence $u^n(j)$, $j \in \{1, \dots, 2^{nR_0}\}$, 2^{nR_i} V_i^n sequences are generated using $p(v_i^n|u^n, s^n)$ and are indexed using $k \in \{1, \dots, 2^{nR_1}\}$ and $l \in \{1, \dots, 2^{nR_2}\}$ respectively.

The index pairs (k, l) are thrown uniformly into 2^{nR_0} bins such that each bin receives $2^{n(R_1+R_2-R_0)}$ of them. These bins are indexed by j , $j \in \{1, \dots, 2^{nR_0}\}$.

A total of B (B large) messages are sent over blocks of length n each. At the beginning of block b , each transmitter successfully decodes the other transmitter's message sent in block $b-1$, while the receiver waits till the end of transmission to decode all messages. Using shared information from block $b-1$, the transmitters cooperatively transmit "cloud centers" $u^n(j)$ in block b . This process is diagrammatically illustrated in Figure 1.

Encoding: Suppose $j^*(b) \in \{1, \dots, 2^{nR_0}\}$ is the common index to be sent by the two transmitters in block b . Let $k^*(b)$ be the message to be transmitted by Transmitter 1 and $l^*(b)$ that by Transmitter 2. Transmitter 1 locates $v_1^n(k^*(b))$ given the particular $u^n(j^*(b))$ and $s^n(b)$, then determines a sequence x_1^n that is jointly typical with the pair v_1^n, s^n and transmits it.

Similarly, Transmitter 2 locates $v_2^n(l^*(b))$ that is jointly typical with $u^n(k^*(b)), s^n(b)$, then generates a sequence x_2^n that is jointly typical with this particular v_2^n and s^n . x_1^n and x_2^n are transmitted in block b .

Decoding: We employ backward decoding at the receiver, i.e., we wait till all $B+1$ transmissions are complete before we begin decoding. Each transmitter, however, decodes the other transmitter's message at the end of each block.

Decoding at Transmitter 1: Transmitter 1 looks for a unique index $l(b)$ such that the set $(x_1^n(k^*(b)), v_2^n(l), u^n(j^*(b)), s^n(b), y^n(b))$ are jointly typical. Note here that $x_1^n(k^*(b)), u^n(j^*(b)), s^n(b)$ are all known at transmitter 1, and so $l(b)$ can be determined uniquely iff

$$R_2 < I(X_2; Y|X_1, U, S) \quad (19)$$

Now, $j_1^*(b+1)$ at Transmitter 1 is determined as the bin index in which the pair $k^*(b), l(b)$ lie.

Decoding at Transmitter 2: Similarly, Transmitter 2 can determine the unique bin index corresponding to Transmitter 1's message if $R_1 < I(X_1; Y|X_2, U, S)$. $j_2^*(b+1)$ is determined as the bin index to which the pair $k^*(b), l(b)$ belong.

For n large, with high probability we have $j_1^*(b+1) = j_2^*(b+1) \triangleq j^*(b+1)$.

Backward decoding at receiver: In block $B+1$ only the common "cloud center" $u^n(j^*(B+1))$ is communicated at a rate of $R_0 < I(U; Y)$. Using $j^*(B+1)$ as the bin index, the subset of all possible choices for $k(B), l(B)$ are determined. We call this subset S_B . Note that the cardinality of S_B is $2^{n(R_1+R_2-R_0 \pm \epsilon)}$.

In block B , the cloud center $u^n(j(B))$ is first determined, which can be performed at a rate $R_0 < I(U; Y)$. Next, the

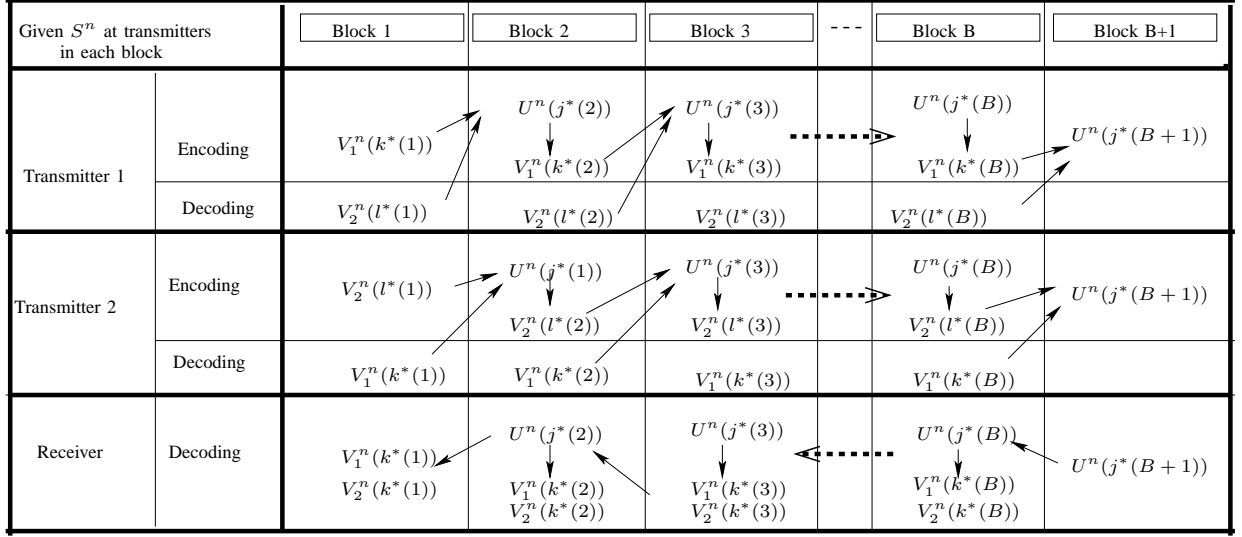


Fig. 1. The Modified Cover-Leung Coding Procedure

unique pair of indices $k(B), l(B)$ in the set S_B are located, if they exist, such that $(u^n(j(B)), v_1^n(k(B)), v_2^n(l(B)), y^n)$ are all jointly typical.

The set of possible errors determines the bound on the pair R_1, R_2 . Here we ignore trivial error cases and concentrate on those that provide bounds on rates. Specifically, the event

$$E_{k'l'} = \{(u^n(j(B)), v_1^n(k'), v_2^n(l'), y^n) \in T_\epsilon^n | (v_1^n(k'), v_2^n(l'), s^n) \in T_\epsilon^n\}$$

is of concern when either $k' \neq k^*(B)$ or $l' \neq l^*(B)$.

So

$$\begin{aligned} P_e &= P\left(\bigcup_{k' \neq k^*(B), l' \neq l^*(B), (k', l') \in S(B)} E_{k'l'}\right) \\ &\leq \sum_{i=2}^{2^n(R_1+R_2-R_0)} P(E_{k'l'}) \\ &\leq 2^{n(R_1+R_2-R_0)} 2^{-n(I(V_1, V_2; Y|U) - I(V_1, V_2; S) - 6\epsilon)} \end{aligned}$$

Thus it is sufficient if

$$R_1 + R_2 - R_0 \leq I(V_1, V_2; Y|U) - I(V_1, V_2; S).$$

Proof of (20). By Bayes' rule

$$\begin{aligned} P(E_{k'l'}) &= \frac{P(\{(u^n(j(B)), v_1^n(k'), v_2^n(l'), y^n) \in T_\epsilon^n, (v_1^n(k'), v_2^n(l'), s^n) \in T_\epsilon^n\})}{P(\{(v_1^n(k'), v_2^n(l'), s^n) \in T_\epsilon^n\})} \\ &\leq 2^{-n(I(V_1, V_2; Y|U) - 3\epsilon)} / 2^{-n(I(V_1, V_2; S) + 3\epsilon)} \end{aligned}$$

Thus we have the result.

B. Proof of the outer bound of (4)

The outer bound given in (4) is proved here.

$$nR_1 = H(W_1) \leq I(W_1; Y^n) \quad (20)$$

$$\leq I(W_1; Y^n | W_2) \quad (21)$$

$$\begin{aligned} &= \sum_i I(W_1; Y_i | W_2, Y^{i-1}) \\ &= \sum_i I(W_1, S_{i+1}^n; Y_i | W_2, Y^{i-1}) \\ &\quad - \sum_i I(S_{i+1}^n; Y_i | W_2, Y^{i-1}, W_1) \end{aligned} \quad (22)$$

$$\begin{aligned} &= \sum_i (I(W_1, Y_i | W_2, Y^{i-1}, S_{i+1}^n) \\ &\quad + I(S_{i+1}^n; Y_i | W_2, Y^{i-1}) \\ &\quad - I(S_{i+1}^n; Y_i | W_2, Y^{i-1}, W_1)) \end{aligned} \quad (23)$$

Here, (20) results from Fano's inequality ($n\epsilon_n$ is dropped as a convenience) [15], (21) from the fact that conditioning reduces entropy, (22) and (23) from the chain rule [15].

Note by the property of mutual information in [16] and the fact that S_i is independent of W_1, W_2 and S_{i+1}^n , we have

$$\begin{aligned} &\sum_i I(S_{i+1}^n; Y_i | W_2, Y^{i-1}) \\ &= \sum_i I(S_i; Y^{i-1} | W_2, S_{i+1}^n) \\ &= \sum_i I(W_2, S_{i+1}^n, Y^{i-1}; S_i) \\ &= \sum_i I(S_{i+1}^n; Y_i | W_1, W_2, Y^{i-1}) \\ &= \sum_i I(S_i; Y^{i-1} | W_1, W_2, S_{i+1}^n) \\ &= \sum_i I(W_1, W_2, S_{i+1}^n, Y^{i-1}; S_i). \end{aligned}$$

By defining the auxiliary variables V_1, V_2 as

$$\begin{aligned} V_{1,i} &= (W_1, Y^{i-1}, S_{i+1}^n) \\ V_{2,i} &= (W_2, Y^{i-1}, S_{i+1}^n), \end{aligned} \quad (24)$$

we have that

$$\begin{aligned} nR_1 &\leq \sum_i (I(W_1, Y_i | W_2, Y^{i-1}, S_{i+1}^n) + I(W_2, S_{i+1}^n, Y^{i-1}; S_i)) \\ &\quad - I(W_1, W_2, S_{i+1}^n, Y^{i-1}; S_i)) \\ &= \sum_i (I(V_{1,i}; Y_i | V_{2,i}) - I(W_1; S_i | Y^{i-1}, S_{i+1}^n, W_2)) \\ &= \sum_i (I(V_{1,i}; Y_i | V_{2,i}) - I(V_{1,i}; S_i | V_{2,i})). \end{aligned}$$

In the same way, one can obtain the inequality for R_2 . For the sum rate, we follow the Gel'fand-Pinsker arguments [10] almost exactly:

$$\begin{aligned} n(R_1 + R_2) &= H(W_1, W_2) \\ &\leq I(W_1, W_2; Y^n) \\ &\leq \sum_i I(W_1, W_2, Y^{i-1}; Y_i) \\ &= \sum_i (I(W_1, W_2, Y^{i-1}, S_{i+1}^n; Y_i) \\ &\quad - I(S_{i+1}^n; Y_i | Y^{i-1}, W_1, W_2)) \\ &= \sum_i (I(V_{1,i}, V_{2,i}; Y_i) \\ &\quad - I(Y^{i-1}; S_i | W_1, W_2, S_{i+1}^n)) \\ &= \sum_i (I(V_{1,i}, V_{2,i}; Y_i) - I(V_{1,i}, V_{2,i}; S_i)). \end{aligned}$$

C. Proof of Theorem 3.3

The achievability is the same as the proof of the general achievable region in Theorem 3.1.

Now we prove (3) is also the outer bound. Without any loss of generality, we may assume $H(X_1 | X_2, S, Y) = 0$, or there exists deterministic functions F_k , such that,

$$X_{1,k} = F_k(X_{2,k}, S^n, Y^k). \quad (25)$$

Then for R_1 , we have

$$\begin{aligned} nR_1 &= H(W_1) = H(W_1 | W_2, S^n) \\ &\leq I(W_1; Y^n | W_2, S^n) \end{aligned} \quad (26)$$

$$\begin{aligned} &= \sum_{k=1}^n I(W_1; Y_k | W_2, S^n, Y^{k-1}) \\ &= \sum_{k=1}^n (H(Y_k | W_2, S^n, Y^{k-1}, X_1^{k-1}, X_2^k) \\ &\quad - H(Y_k | W_1, W_2, S^n, Y^{k-1}, X_1^k, X_2^k)) \end{aligned} \quad (27)$$

$$\begin{aligned} &\leq \sum_{k=1}^n (H(Y_k | S^n, Y^{k-1}, X_1^{k-1}, X_{2,k}) \\ &\quad - H(Y_k | X_1^{k-1}, Y^{k-1}, S^n, X_{1,k}, X_{2,k})) \quad (28) \\ &= \sum_{k=1}^n I(X_{1,k}; Y_k | S^n, Y^{k-1}, X_1^{k-1}, X_{2,k}) \\ &= \sum_{k=1}^n I(X_{1,k}; Y_k | S_k, X_{2,k}, U_k), \end{aligned}$$

where the auxiliary variables U_k are defined as

$$U_k = (Y^{k-1}, X_1^{k-1}, S^{k-1}, S_{k+1}^n).$$

Thus for R_2 ,

$$\begin{aligned} nR_2 &= H(W_2) = H(W_2 | W_1, S^n) \\ &\leq I(W_2; Y^n | W_1, S^n) \end{aligned} \quad (29)$$

$$\begin{aligned} &= \sum_{k=1}^n I(W_2; Y_k | W_1, S^n, Y^{k-1}) \\ &= \sum_{k=1}^n (H(Y_k | W_1, S^n, Y^{k-1}, X_1^k) \\ &\quad - H(Y_k | W_1, W_2, S^n, Y^{k-1}, X_1^k, X_2^k)) \end{aligned} \quad (30)$$

$$\begin{aligned} &\leq \sum_{k=1}^n (H(Y_k | S^n, Y^{k-1}, X_1^{k-1}, X_{1,k}) \\ &\quad - H(Y_k | S^n, Y^{k-1}, X_{1,k}, X_{2,k})) \quad (31) \\ &= \sum_{k=1}^n I(Y_k; X_{2,k} | U_k, S_k, X_{1,k}). \end{aligned}$$

Note (26) and (29) are due to Fano's inequality [15]. (27) and (30) are from (25) and the fact that the channel input $X_i = G_i(W_i, S^n, Y)$ for feedback coding. (28) and (31) are because

$$(W_1, W_2, Y^{k-1}, X_1^{k-1}, X_2^{k-1}) \Rightarrow (X_{1,k}, X_{2,k}, S_k) \Rightarrow Y_k$$

forms a Markov chain. Now define the auxiliary variables $V_{1,k}, V_{2,k}$ as (24) and note given $U_k, V_{1,k}$ and $V_{2,k}$ are independent, thus the proof for the sum-rate in theorem 3.2 can be applied directly to establish (3).

VI. CONCLUSIONS

In this paper, both the discrete memoryless and the Gaussian two-user multiple access channel with state and feedback are analyzed, where the state is non-causally known at the transmitters. For the discrete memoryless case, both an outer bound

and an achievable region are derived, and sufficient conditions under which they meet are obtained. For the all-Gaussian case, the entire capacity region can be found. This capacity region is the same as that of a Gaussian MAC with feedback with channel state known to the transmitters and the receiver. The proofs in this paper are obtained as generalizations of the Merhav-Weissman, Ozarow, Costa, Gel'fand-Pinsker and Cover-Leung coding schemes.

ACKNOWLEDGMENT

The authors would like to thank Dr. Gerhard Kramer for helpful discussions and comments.

REFERENCES

- [1] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback - part I: No bandwidth constraint," *IEEE Transactions on Information Theory*, vol. 12, no. 4, pp. 172–182, April 1966.
- [2] L. H. Ozarow, "An capacity of the white Gaussian multiple access channel with feedback," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 623–629, April 1984.
- [3] N. Merhav and T. Weissman, "Coding for the feedback Gel'fand-Pinsker channel and the feedforward Wyner-Ziv source," in *Proceedings of IEEE Int. Symp. Info. Theory, Adelaide, Australia*, 2005, p. 1.
- [4] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE transactions on Information theory*, vol. 48, no. 6, pp. 1423–1438, June 2002.
- [5] A. Sahai, S. Draper, and M. Gastpar, "Boosting reliability over AWGN networks with average power constraints and noiseless feedback," in *Proceedings of IEEE Int. Symp. Info. Theory, Adelaide, Australia*, September 2005.
- [6] W. Wu, S. Vishwanath, and A. Arapostathis, "Feedback strategies of Gaussian interference networks: duality, sum capacity and dynamic team problems," in *Proceedings of 43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois*, September 2005.
- [7] Y.-H. Kim, "On the feedback capacity of stationary Gaussian channels," in *Proceedings of 43rd Annual Allerton Conference on Communication, Control, and Computing, Monticello, Illinois*, September 2005.
- [8] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Inform. Theory*, vol. 2, no. 9, pp. 8–19, Sept. 1956.
- [9] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Dev.*, vol. 2, pp. 289–293, Oct. 1958.
- [10] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Peredachi Inform. (Probl. Inform. Trans.)*, vol. 9, no. 1, pp. 19–31, 1980.
- [11] C. Heegard and A. E. Gamal, "On the capacity of computer memory with defects," *IEEE transactions on Information theory*, vol. 29, pp. 731–739, Sep. 1983.
- [12] M. H. Costa, "Writing on dirty paper," *IEEE transactions on information theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [13] Y.-H. Kim, A. Sutivong, and S. Sigurjonsson, "Multiple user writing on dirty paper," in *Proceedings of IEEE Int. Symp. Info Theory, Chicago, Illinois*, June 2004.
- [14] T. Cover and S. Leung, "An achievable rate region for the multiple access channel with feedback," *IEEE transactions on Information theory*, vol. 27, pp. 292–298, May 1981.
- [15] T. M. Cover and J. A. Thomas, *Elements of information theory*, ser. Wiley Series in Telecommunications. New York: John Wiley & Sons Inc., 1991, a Wiley-Interscience Publication.
- [16] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest: Akademiai Kiado, 1997.