EE 87005 Summary of Multiple Access Channels with Generalized Feedback

Mei Chen

I. Introduction

A multiple access channel (MAC) with generalized feedback (GF) is a communication network where two or more sources transmit information to a single destination, and each source observes a different feedback from the channel output. Such a channel with two users is shown in Fig. 1, where W_1, W_2 are the messages to be transmitted to the receiver, X_1, X_2 are signals transmitted by the encoders, Y is the channel output, Y_1 and Y_2 are feedback to encoder 1 and 2 respectively, and $\hat{W}_1$ and $\hat{W}_2$ are the decoded messages.

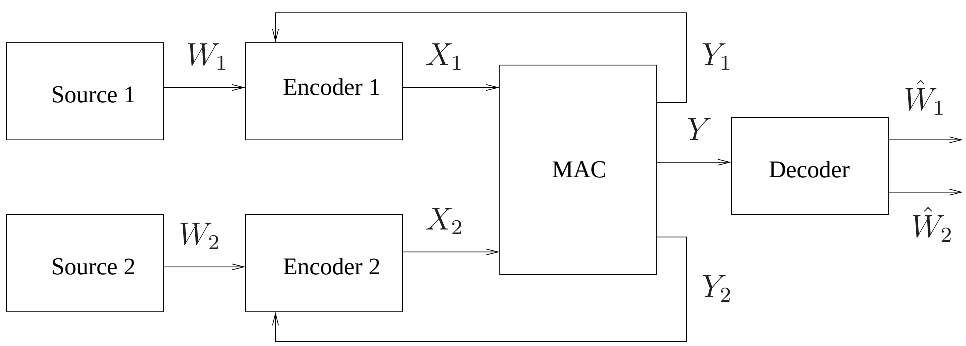

Fig. 1. Two-user MACGF.

The capacity region for MACGF is not yet known. On one hand, Carleial [1] and Willems [2] have independently found the achievable rate region via superposition and block Markov coding. Though it is difficult to compare these two regions for general cases, Zeng and Kuhlmann [3] have shown that for some special cases, Willem's region is strictly larger than Carleial's. Another approach for obtaining the achievable rate region is developed by Gastpar [4], which is an extension of the linear feedback strategies of Ozarow [5] for the AWGN MAC with noiseless feedback. On the other hand, Gastpar and Kramer [6] have established outer bounds for a special case of MACGF, where the noisy feedback is a degraded version of the channel output.

The rest of the summary is organized as follows. Section II discusses Willems' and Carleial's achievable rate regions for discrete memoryless (d.m.) MACGF. Carleial's and Gastpar's regions for AWGN MACGF are briefly introduced in Section III. The outer bounds developed by Gastpar and Kramer are then illustrated in Section IV.

II. D.M. MACGF

A. Channel Model

The two-user d.m. MACGF is depicted in Fig. 1. Denote the channel by $(\mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1, y_2|x_1, x_2), \mathcal{Y} \times \mathcal{Y}_1 \times \mathcal{Y}_2)$, where $\mathcal{X}_1$ and $\mathcal{X}_2$ are the input alphabets, $\mathcal{Y}$ is the output alphabet, $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are the feedback output alphabets, and $p(y, y_1, y_2|x_1, x_2)$ is the conditional probability law of $(y, y_1, y_2) \in \mathcal{Y} \times \mathcal{Y}_1 \times \mathcal{Y}_2$ given $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$. All the alphabets are finite. Since the channel is memoryless, we have

$$\begin{aligned} p(y^N, y_1^N, y_2^N|x_1^N, x_2^N) &= \Pi_{n=1}^N p(y_n, y_{1n}, y_{2n}|x_{1n}, x_{2n}, y^{n-1}, y_1^{n-1}, y_2^{n-1}) \\ &= \Pi_{n=1}^N p(y_n, y_{1n}, y_{2n}|x_{1n}, x_{2n}), \end{aligned} \tag{1}$$

where $y^N \in \mathcal{Y}^N$, $y_1^N \in \mathcal{Y}_1^N$, $y_2^N \in \mathcal{Y}_2^N$, $x_1^N \in \mathcal{X}_1^N$, and $x_2^N \in \mathcal{X}_2^N$.

An (M_1, M_2, N) code for the d.m. MACGF consists of:

- Two independent message sources 1 and 2, which produce random integers $W_1 \in \{1, 2, \cdots, M_1\}$ and $W_2 \in \{1, 2, \cdots, M_2\}$.
- Two sets of N encoder functions

$$x_{1n} = f_{1n}(W_1, Y_1^{n-1}), \tag{2}$$

$$x_{2n} = f_{2n}(W_2, Y_2^{n-1}), n = 1, 2, \cdots, N. \tag{3}$$

- One decoder function

$$(\hat{w}_1, \hat{w}_2) = g(Y^N). \tag{4}$$

The error probability P_e, for W_1, W_2 uniform and independent, is defined as

$$P_e = \Pr\{(\hat{W}_1, \hat{W}_2) \neq (W_1, W_2)\}. \tag{5}$$

A rate pair (R_1, R_2) is said to be achievable for a d.m. MACGF, if for any $\epsilon > 0$ and N sufficiently large, there exists an (M_1, M_2, N) code such that $\log(M_i)/N > R_i - \epsilon, i = 1, 2$, and $P_e < \epsilon$.

$\mathcal{C}_{gf}$, the capacity region for a d.m. MACGF is the closure of the set of all achievable rate pairs (R_1, R_2).

B. Willems' Achievable Rate Region

Theorem 1: For a d.m. MACGF $(\mathcal{X}_1 \times \mathcal{X}_2, p(y, y_1, y_2|x_1, x_2), \mathcal{Y} \times \mathcal{Y}_1 \times \mathcal{Y}_2)$, the capacity region $\mathcal{C}_{gf} \supseteq \mathcal{R}_{gf}$, where the achievable rate region $\mathcal{R}_{gf}$ is defined as:

$$\begin{aligned} \mathcal{R}_{gf} \triangleq \big\{(R_1, R_2) :& R_1 = R_{12} + R_{11}, R_2 = R_{21} + R_{22}, \\ & 0 \leq R_{12} \leq I(V_1; Y_2|X_2, U), && (6) \\ & 0 \leq R_{21} \leq I(V_2; Y_1|X_1, U), && (7) \\ & 0 \leq R_{11} \leq I(X_1; Y|X_2, V_1, U), && (8) \\ & 0 \leq R_{22} \leq I(X_2; Y|X_1, V_2, U), && (9) \\ & 0 \leq R_{11} + R_{22} \leq I(X_1, X_2; Y|V_1, V_2, U), && (10) \\ & 0 \leq R_{12} + R_{21} + R_{11} + R_{22} \leq I(X_1, X_2; Y) \text{ for} && (11) \\ & p(u, v_1, v_2, x_1, x_2, y) = p(u)p(v_1|u)p(v_2|u)p(x_1|v_1, u) \\ & \qquad \cdot p(x_2|v_2, u)p(y, y_1, y_2|x_1, x_2)\big\}. \end{aligned}$$

Note that $\mathcal{R}_{gf}$ is convex so that the convex hull is not needed.

Proof: To prove the achievability of $\mathcal{R}_{gf}$, Willems constructs a random coding scheme with block Markov encoding and backward decoding that yields arbitrarily small probability of error.

First fix the distribution $p(u, v_1, v_2, x_1, x_2) = p(u)p(v_1|u)p(v_2|u)p(x_1|v_1, u)p(x_2|v_2, u)$. The message $W_{1b} = (W_{12b}, W_{11b})$, $W_{12b} \in \{1, 2, \cdots, M_{12} = \lceil \exp(NR_{12}) \rceil\}$, $W_{11b} \in \{1, 2, \cdots, M_{11} = \lceil \exp(NR_{11}) \rceil\}$, $W_{2b} = (W_{21b}, W_{22b})$, $W_{21b} \in \{1, 2, \cdots, M_{21} = \lceil \exp(NR_{21}) \rceil\}$, $W_{22b} \in \{1, 2, \cdots, M_{22} = \lceil \exp(NR_{22}) \rceil\}$, $b = 1, 2, \cdots, B-1$, will be sent over the MACGF in B blocks each of N channel uses. W_{12b}, W_{11b}, W_{21b}, W_{22b} are uniformly distributed and independent of each other. The same codebook is used for each of the B blocks

The encoding process is outlined as follows:

1) Generate $M_{12}M_{21}$ sequences $\mathbf{u} = (u_1, u_2, \cdots, u_N)$ with $\Pr(\mathbf{u}) = \Pi_{n=1}^{N} p(u_n)$. Label them $\mathbf{u}(w_0)$, $w_0 = (w_{01}, w_{02})$, $w_{01} \in \{1, 2, \cdots, M_{12}\}$, and $w_{02} \in \{1, 2, \cdots, M_{21}\}$.
2) For every $\mathbf{u}(w_0)$ generate M_{12} sequences $\mathbf{v}_1 = (v_{11}, v_{12}, \cdots, v_{1N})$ and M_{21} sequences $\mathbf{v}_2 = (v_{21}, v_{22}, \cdots, v_{2N})$, each with probability $\Pr\{\mathbf{v}_m|\mathbf{u}(w_0)\} = \Pi_{n=1}^{N} p(v_{mn}|u_n(w_0))$, $m = 1, 2$. Label them as $\mathbf{v}_1(w_0, w_{12}), w_{12} \in \{1, 2, \cdots, M_{12}\}$, and $\mathbf{v}_2(w_0, w_{21}), w_{21} \in \{1, 2, \cdots, M_{21}\}$, respectively.
3) For every pair $(\mathbf{u}(w_0), \mathbf{v}_1, \mathbf{v}_2)$, generate M_{11} sequences $\mathbf{x}_1 = (x_{11}, x_{12}, \cdots, x_{1N})$ with probability $\Pr\{\mathbf{x}_1|\mathbf{v}_1(w_0, w_{12}), \mathbf{u}(w_0)\} = \Pi_{n=1}^{N} p(x_{1n}|v_{1n}(w_0, w_{12}), u_n(w_0))$, and M_{22} sequences $\mathbf{x}_2 = (x_{21}, x_{22}, \cdots, x_{2N})$ with probability $\Pr\{\mathbf{x}_2|\mathbf{v}_2(w_0, w_{21}), \mathbf{u}(w_0)\} = \Pi_{n=1}^{N} p(x_{2n}|v_{2n}(w_0, w_{21}), u_n(w_0))$. Label them as $\mathbf{x}_1(w_0, w_{12}, w_{11}), w_{11} \in \{1, 2, \cdots, M_{11}\}$ and $\mathbf{x}_2(w_0, w_{21}, w_{22}), w_{22} \in \{1, 2, \cdots, M_{22}\}$, respectively.
4) The messages $W_{12b}, W_{11b}, W_{21b}$ and W_{22b}, $b = 1, 2, \cdots, B-1$ are block Markov encoded in the following way:
   a) In block 1,

$$\begin{aligned} \mathbf{x}_{11} &= \mathbf{x}_1((1,1), w_{121}, w_{111}), \\ \mathbf{x}_{21} &= \mathbf{x}_2((1,1), w_{211}, w_{221}). \end{aligned} \tag{12}$$

   b) Before block b, $b = 2, 3, \cdots, B$, because of the feedback links, encoder 1 has the estimate $\hat{w}'_{21(b-1)}$ for message $w_{21(b-1)}$ and encoder 2 has the estimate $\hat{w}''_{12(b-1)}$ for message $w_{12(b-1)}$. Then in block b, $b = 2, 3, \cdots, B-1$

$$\begin{aligned} \mathbf{x}_{1b} &= \mathbf{x}_1(\hat{w}'_{0b}, w_{12b}, w_{11b}), \\ \mathbf{x}_{2b} &= \mathbf{x}_2(\hat{w}''_{0b}, w_{21b}, w_{22b}), \end{aligned} \tag{13}$$

   and in block B

$$\begin{aligned} \mathbf{x}_{1B} &= \mathbf{x}_1(\hat{w}'_{0B}, 1, 1), \\ \mathbf{x}_{2B} &= \mathbf{x}_2(\hat{w}''_{0B}, 1, 1), \end{aligned} \tag{14}$$

   where $\hat{w}'_{0b} = (w_{12(b-1)}, \hat{w}'_{21(b-1)})$ and $\hat{w}''_{0b} = (\hat{w}''_{12(b-1)}, w_{21(b-1)})$, $b = 2, 3, \cdots, B$.

The decoding process uses the principle of backward decoding.

1) First, in the last block B, the decoder looks for $\hat{w}_{0B} = (\hat{w}_{01B}, \hat{w}_{02B})$ such that

$$\begin{aligned} &(\mathbf{u}(\underline{\hat{w}_{0B}}), \mathbf{v}_1(\underline{\hat{w}_{0B}}, 1), \mathbf{v}_2(\underline{\hat{w}_{0B}}, 1), \mathbf{x}_1(\underline{\hat{w}_{0B}}, 1, 1), \mathbf{x}_2(\underline{\hat{w}_{0B}}, 1, 1), \mathbf{y}_B) \\ &\in \mathcal{A}_\epsilon(U, V_1, V_2, X_1, X_2, Y), \end{aligned} \tag{15}$$

   where $\mathcal{A}_\epsilon(\cdot)$ denotes jointly typical set [7]. The unknown messages are underlined.
2) In block b, $b = B-1, B-2, \cdots, 1$, the decoder sets

$$\begin{aligned} \hat{w}_{12b} &= \hat{w}_{01(b+1)}, \\ \hat{w}_{21b} &= \hat{w}_{02(b+1)}. \end{aligned} \tag{16}$$

   Then in block b the decoder looks for the triple $(\hat{w}_{0b}, \hat{w}_{11b}, \hat{w}_{22b})$ such that

$$\begin{aligned} &(\mathbf{u}(\underline{\hat{w}_{0b}}), \mathbf{v}_1(\underline{\hat{w}_{0b}}, \hat{w}_{12b}), \mathbf{v}_2(\underline{\hat{w}_{0b}}, \hat{w}_{21b}), \mathbf{x}_1(\underline{\hat{w}_{0b}}, \hat{w}_{12b}, \underline{\hat{w}_{11b}}), \mathbf{x}_2(\hat{w}_{0b}, \hat{w}_{21b}, \underline{\hat{w}_{22b}}), \mathbf{y}_b) \\ &\in \mathcal{A}_\epsilon(U, V_1, V_2, X_1, X_2, Y). \end{aligned} \tag{17}$$

   Note in block 1, since $w_{01} = (1, 1)$, the decoder looks for $(\hat{w}_{111}, \hat{w}_{221})$ satisfying (17) with $b = 1$ instead.
3) Recall that the encoders need to obtain the estimates $\hat{w}'_{21b}$ and $\hat{w}''_{12b}$ from the feedback link at the end of block b. Therefore, the encoders also perform decoding operations. Encoder 1 chooses $\hat{w}'_{21b}$

such that

$$(\mathbf{u}(\hat{w}'_{0b}), \mathbf{v}_1(\hat{w}'_{0b}, w_{12b}), \mathbf{v}_2(\hat{w}'_{0b}, \underline{\hat{w}'_{21b}}), \mathbf{x}_1(\hat{w}'_{0b}, w_{12b}, w_{11b}), \mathbf{y}_{1b})$$
$$\in \mathcal{A}_\epsilon(U, V_1, V_2, X_1, Y_1), \tag{18}$$

and encoder 2 chooses $\hat{w}''_{12b}$ such that

$$(\mathbf{u}(\hat{w}''_{0b}), \mathbf{v}_1(\hat{w}''_{0b}, \underline{\hat{w}''_{12b}}), \mathbf{v}_2(\hat{w}''_{0b}, w_{21b}), \mathbf{x}_2(\hat{w}''_{0b}, w_{21b}, w_{22b}), \mathbf{y}_{2b})$$
$$\in \mathcal{A}_\epsilon(U, V_1, V_2, X_2, Y_2), \tag{19}$$

where $\hat{w}'_{0b} = (w_{12(b-1)}, \hat{w}'_{21(b-1)})$ and $\hat{w}''_{0b} = (\hat{w}''_{12(b-1)}, w_{21(b-1)})$ are from the results of the previous decoding operations performed by the encoders after block $b-1$.

From the encoding and the decoding process, it is clear that w_{11} and w_{22} contain information to be sent directly to the receiver. w_{12} and w_{21} are new information with high rate to be sent to the receiver via feedback, while w_0 is the resolution information the encoders cooperatively send to resolve the receiver's residual uncertainty.

The error probability analysis for this random coding scheme is similar to that of other multiuser channels, such as MAC without feedback and the broadcast channel, which uses union bounds and properties of jointly typical sequences. Details can be found in [2]. It can be shown that given $\epsilon > 0$ and N sufficiently large, if $(R_{12}, R_{21}, R_{11}, R_{22})$ satisfies (6) to (11), then $P_e < \epsilon$, which completes the proof. ■

C. Carleial's Achievable Rate Region

Carleial [1] develops another achievable rate region for the channel depicted in Fig. 1, also via superposition and block Markov coding, but windowed decoding is used rather than backward decoding. This superposition coding scheme is graphically represented in Fig. 2.

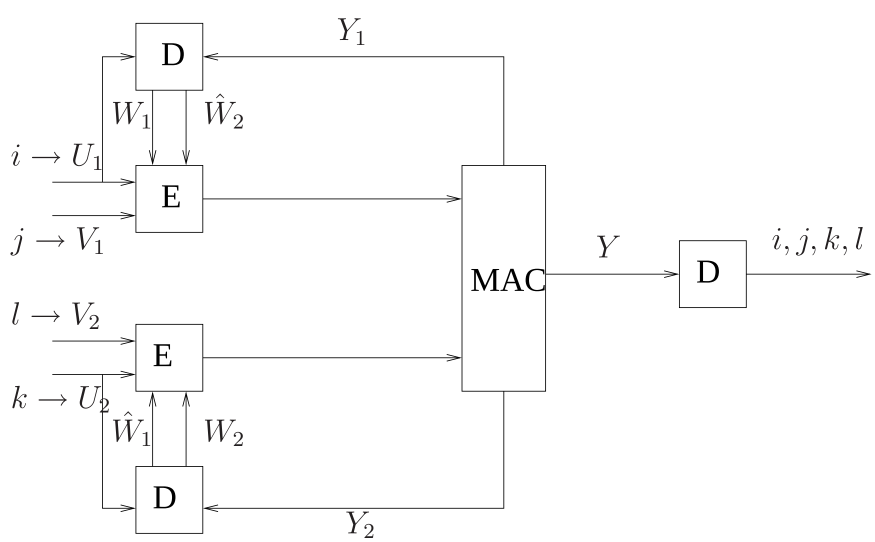

Fig. 2. Representation of Carleial's superposition coding scheme for MACGF.

The achievable rate region for d.m. MACGF is given in [1], and its proof is similar to that of Willems' region. The details will not be discussed here. We only summarize Carleial's idea as follows:

1) Time-sharing random variable Q is used to make the region convex such that the convex hull is not needed.
2) Messages are sent in B blocks of N channel uses.
3) The messages of source m, $m = 1, 2$ are decomposed into two components (i, j for source 1 and k, l for source 2), represented by random sequences U_m and V_m. U_m represents the high rate new

message that is sent to the receiver via feedback links, while V_m is the message that is sent directly to the receiver.

4) At the end of each block, encoder 1 makes an estimate $\hat{W}_2$ of U_2 via feedback Y_1, and encoder 2 makes an estimate $\hat{W}_1$ of U_1 via feedback Y_2.
5) The encoding functions $\mathbf{x}_1 = \phi_1(i, j, i', \hat{k}')$ and $\mathbf{x}_2 = \phi_1(k, l, \hat{i}', k')$ are denoted by

$$x_1^t = f_1(u_{1i}^t, v_{1j}^t, w_{1i'}^t, \hat{w}_{2\hat{k}'}^t, q)$$
$$x_2^t = f_2(u_{2k}^t, v_{2l}^t, \hat{w}_{1\hat{i}'}^t, w_{2k'}^t, q), \tag{20}$$

for $t = 1, 2, \cdots, N$. $'$ denotes that the message is pertaining to the last preceding block, and $\mathbf{q} = (q_1, q_2, \cdots, q_n)$ is the time-sharing sequence.
6) The decoder uses channel outputs in two blocks simultaneously for decoding. The decisions are made by means of joint typicality. Moreover, the encoders have to decode U_2 and U_1 from the previous block. The decisions are also made by means of joint typicality.
7) Probability of error analysis is performed in the same way as the achievability proofs of other channels.

Remarks:

1) Compared to Willems' approach, Carleial uses two random variables W_1 and W_2 for the cooperative information while Willems only uses one random variable U.
2) The encoder outputs X_1 and X_2 are deterministic given U_1, V_1, W_1 and U_2, V_2, W_2. Therefore, they do not appear in the conditions of the achievable rate region.

D. Comparison of Willems' and Carleial's Achievable Rate Region

For general cases, it is not clear whether Willems' region is larger than Carleial's or not. However, Zeng and Kuhlmann [3] studied two special cases of d.m. MACGF, and proved that for these two special cases, the region obtained by Willems' method is strictly larger than that by Carleial's. These two special cases are:

1) Channel with Y degraded from Y_1 and Y_2;
2) Channel with one-sided feedback Y_1 and Y is also degraded from Y_1.

Here Y is degraded from Y_1 or Y_2 in the following sense [1]:

Definition 1: In the MACGF of Fig. 1, Y is a degraded version of Y_1 with respect to the channel inputs if there exists a "test channel" from Y_1 to a random variable Y' such that

1) $(X_1, X_2), Y_1$ and Y' constitute a Markov chain $X_1, X_2 \to Y_1 \to Y'$, i.e., $p(y'|x_1, x_2, y_1) = p(y'|y_1)$, and
2) Y' is statistically equivalent to Y with respect to the inputs, i.e., the sample spaces of Y' and Y are isomorphic, and for all corresponding values $y' \sim y, p(y'|x_1, x_2) = p(y|x_1, x_2)$.

III. AWGN MACGF

A. Channel Model

A two-user AWGN MACGF, which is also depicted by Fig. 1, has real input and output signals X_1, X_2, Y, Y_1, Y_2. The two input signals subject to average-power constraints P_1 and P_2:

$$\frac{1}{N} \sum_{t=1}^{N} \left(x_i^t\right)^2 \le P_i, i = 1, 2. \tag{21}$$

The output signals are

$$Y = X_1 + X_2 + Z,$$
$$Y_1 = X_1 + X_2 + Z_1,$$
$$Y_2 = X_1 + X_2 + Z_2, \tag{22}$$

where Z, Z_1, Z_2 are zero-mean Gaussian random variables and constitute the channel noise. Those noise powers are $N = \mathrm{E}[Z^2]$, $N_1 = \mathrm{E}[Z_1^2]$, and $N_2 = \mathrm{E}[Z_2^2]$. Without loss of generality, we assume $N_1 \leq N_2$ and $N = 1$ by appropriate normalization.

While Carleial develops an achievable rate region for the general AWGN MACGF, Gastpar [4] considers a special case of the AWGN MACGF, where the noisy feedback is the channel output plus additional white Gaussian noise,

$$\begin{aligned} Y_1 &= Y + V_1, \\ Y_2 &= Y + V_2. \end{aligned} \tag{23}$$

Here V_1 and V_2 are i.i.d. Gaussian with zero mean and variance $\sigma_{V_1}^2$ and $\sigma_{V_2}^2$ respectively. The block diagram of the channel model is shown in Fig. 3.

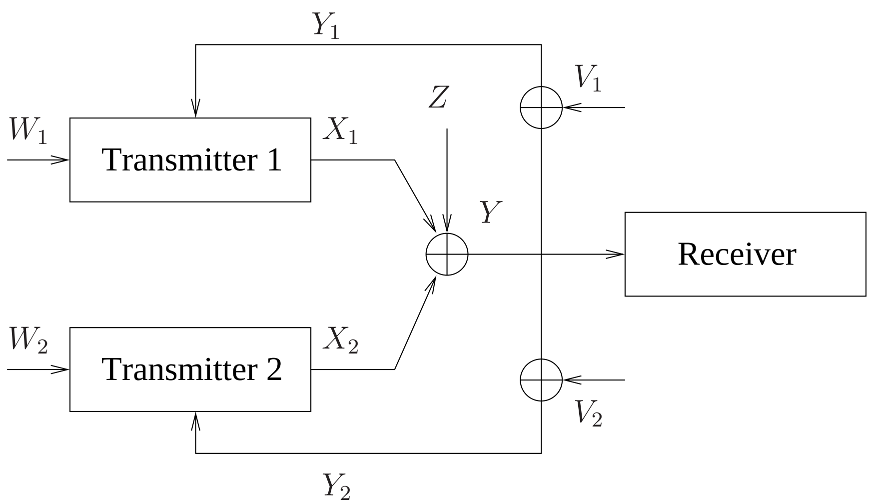

Fig. 3. The AWGN MAC with degraded noisy feedback.

This channel model is further simplified by adding the same noise to the feedback links, and thus

$$Y_1 = Y_2 = Y + V_1 = Y + V_2. \tag{24}$$

Gastpar's achievable rate region is obtained for this simplified case.

B. Carleial's Achievable Region

The achievable rate region is given in [1]. It is formally similar to the region for the d.m. MACGF with the following differences:

1) The time-sharing random variable Q is not used.
2) The encoding functions f_1 and f_2 are linear combinations of the information-bearing auxiliary random variables, which have the zero-mean unit-variance Gaussian distribution.

$$\begin{aligned} x_1^t &= \left[\left(\alpha_1\bar{\beta}_1\right)^{1/2} u_{1i}^t + \left(\alpha_1\beta_1\right)^{1/2} v_{1j}^t + \left(\bar{\alpha}_1\lambda_1\right)^{1/2} w_{1i'}^t + \left(\bar{\alpha}_1\lambda_2\right)^{1/2} w_{2\hat{k}'}^t\right] P_1^{1/2}, \\ x_2^t &= \left[\left(\alpha_2\bar{\beta}_2\right)^{1/2} u_{2k}^t + \left(\alpha_2\beta_2\right)^{1/2} v_{2l}^t + \left(\bar{\alpha}_2\lambda_1\right)^{1/2} w_{1\hat{i}'}^t + \left(\bar{\alpha}_2\lambda_2\right)^{1/2} w_{2k'}^t\right] P_2^{1/2}, \end{aligned} \tag{25}$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2, \lambda_1$ are power-allocation parameters each freely chosen between zero and unity, and $\lambda_2 = 1 - \lambda_1$, $\bar{\alpha}_1 = 1 - \alpha_1$ and $\bar{\alpha}_2 = 1 - \alpha_2$.

More details about the decoding process and the probability of error analysis are provided in [1].

Remarks:

1) Computation of the region is very complicated by the large number of free parameters involved.

2) Fortunately, the region can be much simplified if Y is at least as noisy as Y_1, the least noisy of the feedback signals. In other words, $N_1 \leq 1$.
3) When $1 < N_1 \leq N_2$, it turns out that not much simplification can be made, and thus the region is hard to calculate.

C. Gastpar's Achievable Rate Region

Gastpar [4] considers the special case where the feedback is at least as noisy as the channel output, and the same noise is added to the feedback links. As is discussed before, this is the scenario where Carleial's region is hard to compute.

Unlike Willems and Carleial, who use the random coding strategy, Gastpar [4] employs the linear deterministic feedback strategy of Ozarow [5] to obtain the achievable rate region via joint source-channel coding.

First, a statement analogous to the converse to the separation theorem for a point-to-point channel can be formalized for the two-user AWGN MAC described in Section III-A.

Proposition 1: Denote the capacity region of the network by $\mathcal{C}(P_1, P_2)$, where P_1, P_2 are the average channel input cost constraints. Consider a set of sources that need to be transmitted accross this channel network, and suppose the rate-distortion region for the sources is given by $\mathcal{R}(D_1, D_2)$. If a distortion pair (D_1, D_2) is achievable, then $\mathcal{R}(D_1, D_2) \cap \mathcal{C}(P_1, P_2) \neq \varnothing$.

The interpretation to Proposition 1 is

1) For arbitrary sources with respect to an arbitrary distortion measure, we devise a joint source-channel coding strategy that observe the channel input cost constraint (P_1, P_2) and incurs a certain distortion (D_1, D_2).
2) Then achievable rates can be derived according to Proposition 1 that $\mathcal{R}(D_1, D_2) \cap \mathcal{C}(P_1, P_2) \neq \varnothing$.

Gastpar's strategy to obtain the achievable region follows Ozarow's approach for AWGN MAC with noiseless feedback [5], but Ozarow's approach is slightly changed in order to use Proposition 1.

1) First, consider the transmission of two independent Gaussian sources S_1 and S_2 of mean zero and variance σ_s^2 with respect to mean-squared error.
2) In the first two channel uses, the two transmitters take turns and transmit their source without further coding, using a power of P_1 and P_2, respectively.
3) In subsequent channel uses, the sources send maximally informative updates according to the feedback links $Y_1(Y_2)$.
4) The receiver produces the MMSE estimates of S_1 and S_2 after N channel uses. Denote the vector channel outputs by $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_N)^T$. The MMSE estimate can be written as

$$\hat{S}_m^{(N)} = \mathrm{E}[S_m \mathbf{Y}^T](\mathrm{E}[\mathbf{Y}\mathbf{Y}^T])^{-1}\mathbf{Y}, \text{ for } m = 1, 2. \tag{26}$$

5) The distortion is

$$D_m^{(N)} = \sigma_S^2 - \mathrm{E}[S_m \mathbf{Y}^T](\mathrm{E}[\mathbf{Y}\mathbf{Y}^T])^{-1}\,\mathrm{E}[\mathbf{Y} S_m], \text{ for } m = 1, 2. \tag{27}$$

6) The achievable rate is then given by

$$R_{nF} = \max_N \left(\frac{1}{2}\log_2 \frac{\sigma_S^2}{D_1^{(N)}} + \frac{1}{2}\log_2 \frac{\sigma_S^2}{D_2^{(N)}} \right). \tag{28}$$

Remarks:

1) The maximum of (28) is not achieved as $N \to \infty$.
2) R_{nF} does not have a closed-form solution, but numerical results can be obtained through simulation.
3) Strictly speaking, what Gastpar has derived is not the achievable rate region but the achievable sum-rate.

IV. Outer Bounds

Gastpar and Kramer [6] introduce an outer bound on the capacity region of a special case of AWGN MACGF where the noisy feedback is the channel output plus additional white Gaussian noise (Fig. 3), and generalize the result to all MAC with degraded noisy feedback. Here the noisy feedback is said to be degraded in the sense that (X_1, X_2), Y, (Y_1, Y_2) constitute a Markov chain: $X_1, X_2 \to Y \to Y_1, Y_2$.

A. General Outer Bound

Definition 2: For arbitrary random vectors A and B, define the following quantity:

$$V_{A|B} \triangleq \frac{\det \sum_{A,B}}{\det \sum_{B}}, \tag{29}$$

where $\sum_{A,B}$ denotes the covariance matrix of the random vector (A, B), and $\sum_B$ the covariance matrix of the random vector B.

Theorem 2: The capacity region of the MAC with degraded noisy feedback is contained within the union of the regions

$$R_1 \leq I(X_1; Y|X_2, T) \tag{30}$$

$$R_2 \leq I(X_2; Y|X_1, T) \tag{31}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y, T) \tag{32}$$

over all $p(t, x_1, x_2, y, y_1, y_2)$ that satisfy the following conditions:

$$(i) \quad T \to (X_1, X_2) \to Y \to (Y_1, Y_2), \tag{33}$$

$$(ii) \quad I(X_1; X_2|T) \leq I(X_1; X_2|Y_1, Y_2, T), \tag{34}$$

$$(iii) \quad 0 \leq \sum_t p(t) \log \left(\frac{V_{Y_1,Y_2|X_1,T=t} V_{Y_1,Y_2|X_2,T=t}}{V_{Y_1,Y_2|T=t} V_{Y_1,Y_2|X_1,X_2}} \right), \tag{35}$$

where T has cardinality at most 3.

Remarks:

1) Auxiliary random variable $T = (Q, Y_1^{Q-1}, Y_2^{Q-1})$, where Q is the time-sharing random variable. Therefore, T satisfies (33).
2) Starting from the seemingly trivial inequality

$$0 \leq I(W_1; W_2|Y_1^N, Y_2^N), \tag{36}$$

we can arrive at condition (34).

3) Condition (35) is proved with the aid of the Hadamard-Fischer inequality.
4) Weaker outer bounds are obtained by omitting any or all of the conditions (33)-(35).

B. Degraded AWGN MAC Outer Bound

For the AWGN MACGF, consider the scenario where the noisy feedback is the channel output plus independent white Gaussian noise, given in (23).

Theorem 3: The capacity region of the AWGN MAC with noisy feedback as in (23) is contained within

$$R_1 \leq I(X_1; Y|X_2) \tag{37}$$

$$R_2 \leq I(X_2; Y|X_1) \tag{38}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y) \tag{39}$$

for joint Gaussian X_1 and X_2 satisfying

$$I(X_1; X_2) \leq I(X_1; X_2|Y_1, Y_2). \tag{40}$$

Remark:

1) For the case $Y_1 = Y_2 = Y$, this theorem coincides exactly with the cut-set bound.
2) This outer bound can be evaluated explicitly and is compared to lower bounds appearing previously in the literature in [6].

REFERENCES

[1] A. B. Carleial, "Multiple-access channels with different generalized feedback signals," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 6, pp. 841–850, Nov. 1982.

[2] F. M. J. Willems, *Informationtheoretical results for the discrete memoryless multiple access channel*, Ph.D. thesis, Katholieke Universiteit Leuven, Leuven, Belgium, 1982.

[3] C.-M. Zeng and F. Kuhlmann, "Achievability proof of some multiuser channel coding theorems using backward decoding," *IEEE Trans. Inform. Theory*, vol. 35, no. 6, pp. 1160–1165, Nov. 1989.

[4] M. Gastpar, "On noisy feedback in Gaussian networks," in *Proc. Annu. Allerton Conf. Communication, Control, Computing*, Monticello, IL, Sept. 2005.

[5] L. H. Ozarow, "The capacity of the white gaussian multiple access channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-30, no. 4, pp. 623–629, July 1984.

[6] M. Gastpar and G. Kramer, "On cooperation via noisy feedback," in *Int. Zurich Seminar on Commun. (IZS)*, Feb. 2006.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, 1991.