EE8510 Multi-user Information Theory Project

Feedback Capacity of Multiple Access and Broadcast Channels

Shahrokh Farahmand *

May 6, 2005

Abstract

This report is a literature survey that investigates the effect of feedback on the capacity region of multiple access channels (MAC) and broadcast channels (BC).

1 Introduction

Effect of feedback on capacity was first considered by Shannon, where he proved that the capacity of a single user discrete memoryless channel does not increase with feedback. This is shown to be the case for white Gaussian channels as well. However, if the Gaussian noise is correlated, some capacity increase can be achieved. There are two well-known results [3]:

1. 1 Feedback at most doubles the capacity of a channel with correlated Gaussian noise.
2. 2 Feedback at most increases the capacity of a channel with correlated Gaussian noise by half a bit.

In the next two sections, the major results on the effect of feedback on MAC and BC are summarized.

2 Multiple Access Channel

It is expected that capacity region for a discrete memoryless MAC increases by feedback. This is due to the fact that there are several independent transmitters, so feedback helps each transmitter learn about the messages that other transmitters are sending. As a result, transmitters can align themselves with each other and transmit cooperatively leading to improved capacity. An

*Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street SE, Minneapolis, Minnesota 55455. Email: shahrokh@ece.umn.edu.

achievable region for MAC with feedback was given by [1] using block Markov coding and a random coding argument. An outer bound for the capacity of MAC with feedback is given by [7]. There is a nonzero gap between the outer bound and the achievable region so capacity is unknown in general. However, for the special case of white Gaussian MAC with two transmitters Ozarow [7] derived an achievable region using a deterministic coding argument proposed by [10] which he proved to be the optimal coding strategy, thus yielding capacity region.

Capacity region for a discrete memoryless MAC without feedback is given as [3]:

$$C_{nf} = \bigcup_{p(x_1)p(x_2)} \Big\{ (R_1, R_2) : \begin{aligned} R_1 &\leq I(X_1; Y|X_2) \\ R_2 &\leq I(X_2; Y|X_1) \\ R_1 + R_2 &\leq I(X_1, X_2; Y) \end{aligned} \Big\}, \tag{1}$$

where $\bigcup$ denotes convex hull operation over all achievable rate pairs. The achievable region of Cover and Leung is described in the following theorem:

Theorem 1 (Cover and Leung) *Let U be a discrete random variable which takes on values in the set $||U|| = \min\{||X_1|| \times ||X_2||, ||Y||\}$, where $||X_i||$ denotes the alphabet size of user i. Consider the set $\mathcal{P}$ of all joint distributions of the form $p(u, x_1, x_2, y) = p(u)p(x_1|u)p(x_2|u)p(y|x_1, x_2)$ where $p(y|x_1, x_2)$ is fixed by the channel. An achievable region for MAC with feedback is given by [1]:*

$$C_{fb} = \bigcup_{\mathcal{P}} \Big\{ (R_1, R_2) : \begin{aligned} R_1 &\leq I(X_1; Y|X_2, U) \\ R_2 &\leq I(X_2; Y|X_1, U) \\ R_1 + R_2 &\leq I(X_1, X_2; Y) \end{aligned} \Big\}. \tag{2}$$

This achievable region is strictly larger than 1. Sum-rate is also increased in 2 with respect to 1.[1] To prove 2, a block Markov coding scheme with random coding is used [1]: Assume there are B blocks. In each block we transmit at a rate of $(2^{nR_1}, 2^{nR_2}, n)$, and the messages to be sent are $w_1 \in \{1, \cdots, 2^{nR_1}\}$, $w_2 \in \{1, \cdots, 2^{nR_2}\}$ and an index $j \in \{1, \cdots, J := 2^{nR_0}\}$. Index j is transmitted to help the receiver remove any uncertainty it has in decoding the previous block. For each j produce a random code $u^n(j) \sim \prod_{i=1}^n p(u_i(j))$. Suppose that at block b, j^* will be transmitted. Transmitter 1 maps w_1 to the codeword $x_1^n(w_1, j) \sim \prod_{i=1}^n p(x_{1i}|u_i(j))$, and Transmitter 2 maps w_2 to the codeword $x_2^n(w_2, j) \sim \prod_{i=1}^n p(x_{2i}|u_i(j))$. At block b, transmitter 1 sends $x_1^n(w_1, j)$, and transmitter 2 sends $x_2^n(w_2, j)$. The receiver gets y^n and performs the decoding procedure in two steps. In step1 receiver finds the unique $\hat{j}$ such that $(u^n(\hat{j}), y)$ are jointly typical and uses this index to remove any uncertainty it has from the previous block. In step 2, it finds the set of all the pairs $(\hat{w}_1, \hat{w}_2)$ such that $(x_1^n(\hat{w}_1, j), x_2^n(\hat{w}_2, j), y^n)$ are jointly typical. Let us denote this set by S_y. At the end of block b, each transmitter gets a feedback from the receiver. Using this feedback, each transmitter tries to find the message of the other

[1]During the presentation I gave a wrong answer. In fact sum-rate is increased. Although 2 and 1 are similar in their third equation, the pdf over which convex hull is taken is different.

transmitter. So, transmitter 1 tries to estimate unique $\hat{w}_2$ such that $(x_1^n(\hat{w}_1, j), x_2^n(\hat{w}_2, j), y^n)$ are jointly typical. Other transmitter does the same thing. The probability that each transmitter decode the message of the other transmitter with error should be arbitrarily low. This is how the first two inequalities are derived. Third inequality appears due to constraints on $J := 2^{nR_0}$. Index j is used to remove the residual uncertainty at the receiver, so receiver should be able to decode it with arbitrarily small probability of error, yielding $R_0 \leq I(U;Y)$. On the other hand j is the index of the true transmitted pair (w_1, w_2) in the previous block inside the set of all candidate pairs (S_y) which is kept by the receiver, thus in order for the receiver to be able to determine the correct pair without ambiguity, we should have $J \geq ||S_y||$. Combining these two constraints on J gives the third inequality in 2. Note that at the first block there is no previous uncertainty to be removed, so only new data are sent. At the last block, there is no new data so the actual rate is $B-1/B$, but as B becomes large rate loss becomes negligible. A complete proof can be found in [1].

There exists no converse to show that this achievable region is capacity. However, an outer bound exists. To derive the outer bound, we assume that transmitters can fully cooperate in their transmission. Converse for this outer bound is proved in [7]. It is given by:

$$C_{out} = \bigcup_{p(x_1,x_2)} \Big\{(R_1, R_2) : R_1 \leq I(X_1, Y|X_2) \quad R_2 \leq I(X_2, Y|X_1) \quad R_1 + R_2 \leq I(X_1, X_2; Y)\Big\}. \tag{3}$$

For a white Gaussian MAC scenario with two transmitters, we know capacity is larger than 2. This result, which is summarized in the following theorem, is due to Ozarow.

Theorem 2 (Capacity for white Gaussian MAC) *For a white Gaussian MAC with two transmitters with average power constraints* $\mathrm{E}[X_1^2] \leq P1$, *and* $\mathrm{E}[X_1^2] \leq P1$, *capacity region is given as [7]:*

$$C_{fb} = \bigcup_{0\leq\rho\leq1} \Big\{R_1 \leq \frac{1}{2}\log(1 + \frac{P_1}{N}(1-\rho^2)) \quad R_2 \leq \frac{1}{2}\log(1 + \frac{P_2}{N}(1-\rho^2)) \quad R_1 + R_2 \leq \frac{1}{2}\log(1 + \frac{P_1 + P_2 + 2\rho\sqrt{P_1P_2}}{N})\Big\} \tag{4}$$

Ozarow used the deterministic coding proposed in [10] to prove his result. A sketch of this coding strategy is given below. A complete proof can be found in [7]. Suppose transmitter i has a message set $w_i \in \{0, \cdots, M_i - 1\}$. Transmitters map each message into a number on the real line and transmit that number instead of the message itself. Transmitter i maps message m_i to $\theta_i = \frac{m_i}{M_{i-1}} - \frac{1}{2}$, so each message is mapped to a number in $[-1/2, 1/2]$. If the total number of messages M_{i-1} is large, θ_i will have a uniform distribution with variance 1/12. At time $k = 1$, transmitter 1 sends $\sqrt{12P_1}\theta_1$, where the square root expression is used to satisfy the average power constraint, while transmitter 2 is quiet. Receiver gets $Y_1 = \sqrt{12P_1}\theta_1 + Z_1$, where $Z_1 \sim \mathcal{N}(0, N)$ is the channel noise, and forms an estimate of θ_1 as $\hat{\theta}_1^1 = Y_1/\sqrt{12P_1}$. At time $k = 2$, transmitter 2 sends θ_2 while transmitter 1 is quiet. Receiver gets $Y_2 = \sqrt{12P_2}\theta_2 + Z_2$, and forms an estimate of θ_2 as $\hat{\theta}_2^2 = Y_2/\sqrt{12P_2}$. For $k \geq 2$, the receiver feeds back its previous estimates $\hat{\theta}_1^k, \hat{\theta}_1^k$ to the transmitters. Transmitter i computes the error $\epsilon_{i,k} = \hat{\theta}_i^k - \theta_i$, and at time $k+1$ sends only the error to the receiver. Note that for time indices $k \geq 3$ both transmitters send simultaneously. Transmitted signals are given by:

$$
\begin{aligned}
X_{1,k+1} &= \sqrt{\frac{P_1}{\alpha_{1,k}}}\epsilon_{1,k} \\
X_{2,k+1} &= \sqrt{\frac{P_2}{\alpha_{2,k}}}\epsilon_{2,k}\mathrm{sign}(\rho_\mathrm{k}),
\end{aligned}
$$

where $\alpha_{i,k}$ is the variance of $\epsilon_{i,k}$ which is the error of receivers estimate of the transmitter i message at time k, and ρ_k is the correlation coefficient between $\epsilon_{1,k}$ and $\epsilon_{2,k}$. Note that by using a sign equal to ρ_k, second transmitter tries to somehow align itself with the first transmitter. After each transmission receiver update its estimates using the recursive LMMSE given by:

$$
\hat{\theta}_i^{k+1} = \hat{\theta}_i^k - \frac{\overline{Y_{k+1}\epsilon_{i,k}}}{\overline{Y_{k+1}^2}} Y_{k+1},
$$

and feeds back its new estimates to the transmitters. This procedure is repeated until arbitrarily low error variance is achieved.

In Fig. 1, the capacity region with feedback is plotted for a 2 transmitter white Gaussian MAC scenario where $P_1 = P_2 = 10$ and noise power $N = 1$. The capacity without feedback is also plotted and is the pentagonal region depicted by small squares. From this figure, it is obvious that a positive gain is achieved by using feedback.

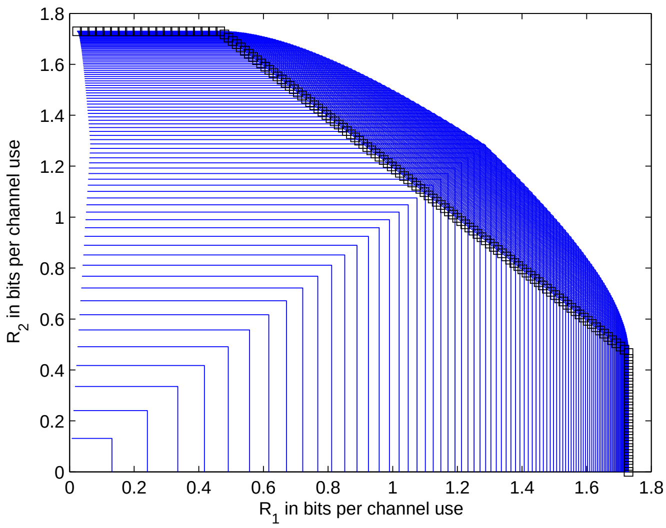

Figure 1: MAC capacity: feedback versus no feedback

3 Broadcast Channel

For the physically degraded discrete memoryless BC, El Gamal proved that capacity does not increase with feedback [4]. Later, he proved the same result for physically degraded white Gaussian BC. In his proof he used a modified version of entropy power inequality (EPI) given as [5]:

$$e^{\frac{2H(Y)}{n}} \geq e^{[\frac{2}{n}\sum_{i=1}^{n} H(X_i|Y^{i-1})]} + e^{\frac{2H(Z)}{n}},$$

where $Y = X + Z$, and X is dependent on Y. However, it is conditionally independent of Y given $(Y_1, \cdots, Y_{i-1})$.

For stochastically degraded white Gaussian BC, Ozarow and Leung [8] developed a deterministic coding scheme (Similar to one we described in MAC section) that achieved a region strictly larger than the no feedback capacity. However, they could not prove a converse to show that their achievable region is in fact the capacity, but they showed their scheme comes very close to the established outer bound.

Before proceeding further, we define the system model. The model of a white Gaussian BC is depicted in Fig. 2. If either N_1 or N_2 are zero, we have a physically degraded BC. So, we assume both N_1 and N_2 are nonzero. The dotted arrow from one receiver to the other is used to establish the upper bound on capacity and can be neglected for now. The following theorem describes the achievable region derived by Ozarow and Leung [8].

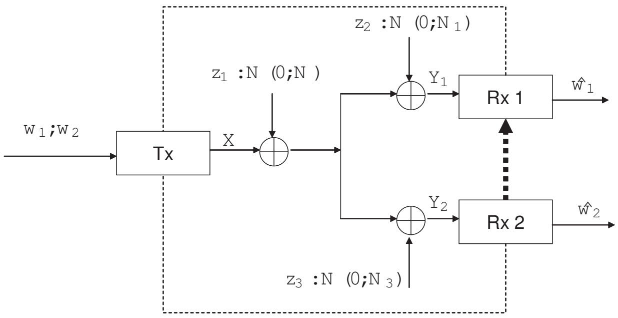

Figure 2: System model for Gaussian BC

Theorem 3 (Ozarow and Leung) *For a white Gaussian BC channel as in Fig. 2 with an av-*

erage power constraint $\mathrm{E}[X^2] \leq P$, *there exists a code that achieves the rate region defined by:*

$$C_{bc} = \bigcup_{g\geq 0} \left\{ (R_1, R_2) : R_1 \leq \frac{1}{2}\log\left[\frac{N+N_1+P}{N+N_2+\frac{P}{D^*}g^2(1-{\rho^*}^2)}\right] \right. \tag{5}$$

$$\left. R_2 \leq \frac{1}{2}\log\left[\frac{N+N_2+P}{N+N_2+\frac{P}{D^*}(1-{\rho^*}^2)}\right] \right\},$$

where $D^* = 1 + g^2 + 2g\rho^*$, *and* $g \geq 0$ *allows trade-off between transmission rates to receivers 1 and 2, and* ρ^* *is given by the following equation:*

$$-\rho^* = \frac{(N\Sigma + N_1N_2)\rho^* - \frac{P\Sigma}{D^*}g(1-{\rho^*}^2)}{\sqrt{\Pi}\sqrt{N+N_1+\frac{P}{D^*}g^2(1-{\rho^*}^2)}\sqrt{N+N_2+\frac{P}{D^*}(1-{\rho^*}^2)}},$$

and

$$\begin{aligned} \Sigma &= P + N + N_1 + N_2 \\ \Pi &= (P+N+N_1)(P+N+N_2). \end{aligned}$$

Ozarow also derives a simple upper bound on capacity with feedback. To establish this upper bound, suppose that receiver 1 somehow knows anything that receiver 2 receives. This is depicted in Fig. 2 as the dotted arrow. The capacity region is enlarged by this process as receiver 1 has more information available to decode its intended message. Now $(X, (Y_1, Y_2), Y_2)$ form a Markov chain resulting in a physically degraded BC, for which we know the capacity region. The same argument can be made with the rule of receivers 1 and 2 reversed yielding another region. The feedback capacity lies in the intersection of these two regions.

Theorem 4 (Outer Bound for BC) *Let us denote feedback capacity of the white Gaussian channel by* C_{fb}. *Then* $C_{fb} \subseteq C_1 \cap C_2$, *where:*

$$C_1 = \bigcup_{0\leq\alpha\leq 1} \left\{ (R_1, R_2) : R_1 \leq \frac{1}{2}\log\left[1 + \frac{\alpha P}{N + \frac{N_1N_2}{N_1+N_2}}\right] \right.$$

$$\left. R_2 \leq \frac{1}{2}\log\left[1 + \frac{(1-\alpha)P}{N+N_2+\alpha P}\right] \right\},$$

and

$$C_2 = \bigcup_{0\leq\alpha\leq 1} \left\{ (R_1, R_2) : R_2 \leq \frac{1}{2}\log\left[1 + \frac{\alpha P}{N + \frac{N_1N_2}{N_1+N_2}}\right] \right.$$

$$\left. R_1 \leq \frac{1}{2}\log\left[1 + \frac{(1-\alpha)P}{N+N_1+\alpha P}\right] \right\}.$$

Fig. 3 shows the outer bound versus the inner bound (no feedback) for the capacity of the white Gaussian BC. Outer bound is the intersection of the two regions confined by the two solid curves. Inner bound which is the no feedback capacity region is confined by the dashed curve. In this figure we have assumed $P = 10, N_1 = N_2 = 1$, and $N = 0$. For this setting, Ozarow showed that the rate pair of $R_1 = R_2 = 0.70468$ nats is achievable with his coding scheme. This is very close to the rate given by the outer bound which is $R_1 = R_2 = 0.71956$, while it is a great improvement to the no feedback rate of $R_1 = R_2 = 0.59947$. However, his code falls short of reaching the outer bound and capacity for white Gaussian BC is not yet known.

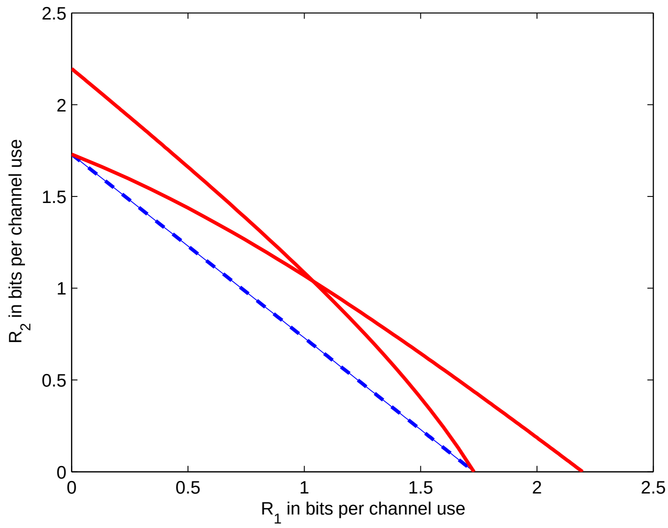

Figure 3: Inner and outer bound for BC

4 A few more results...

Pombra and Cover [9] proved that capacity of a white Gaussian MAC with correlated noise is at most doubled by feedback. Kramer [6] extended the feedback coding method of [10] to an interference white Gaussian channel with arbitrary number of transmitters and receivers. His results are not in closed form, and are given as a recursive equation. Given the channel gain matrix, the iterations can be carried out by a computer. His results provide an achievable sum-rate, which for MAC he showed can achieve the capacity found by Ozarow. One of his conclusions concerning a strong interference channel is mentioned here. Consider an interference channel of the form:

$$\begin{aligned} Y_1 &= X_1 + X_2 + \cdots + X_K + Z_1 \\ Y_2 &= X_1 + X_2 + \cdots + X_K + Z_2 \\ \vdots & \quad \vdots \\ Y_K &= X_1 + X_2 + \cdots + X_K + Z_K, \end{aligned} \tag{6}$$

where Y_i is the intended receiver of transmitter X_i, and Z_i is the channel noise. Each Y_i receives interference from all other transmitters. Kramer's coding strategy achieves a sum-rate of $(\log K)/2 + \log\log K$ as K becomes large. This is a factor of $\log\log K$ larger than the no feedback sum-rate capacity which is $[\log(1 + PK)]/2 \approx (\log K)/2$. Here P is the average power constraint on each transmitter.

5 Concluding Remarks

Feedback can strictly enlarge the capacity for discrete memoryless MAC and white Gaussian MAC. For discrete memoryless MAC with feedback, an achievable region and an outer bound exists but capacity is not known. For white Gaussian MAC with two transmitters capacity is known and it is larger than the no feedback capacity region. Capacity of physically degraded discrete memoryless BC and physically degraded white Gaussian BC can not be enlarged by feedback. This is not the case for general white Gaussian BC where capacity is in fact increased by feedback. For Gaussian BC, an outer bound is established. Also, codes were derived which achieve rates higher than the no feedback capacity, and very close to the outer bound, but fall short of reaching it. So, the capacity is still unknown.

References

[1] T. M. Cover, and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. on Information Theory*, vol. IT-27, no. 3, pp. 292-298, May 1981.

[2] T. M. Cover, and S. Pombra, "Gaussian feedback capacity," *IEEE Trans. on Information Theory*, vol. 35, no. 1, pp. 37-43, Jan. 1989.

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

[4] A. El Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. on Information Theory*, vol. IT-24, no. 4, pp. 379-381, May 1978.

[5] A. El Gamal, "The capacity of the physically degraded Gaussian broadcast channel with feedback," *IEEE Trans. on Information Theory*, vol. IT-27, no. 4, pp. 508-511, July 1981.

[6] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. on Information Theory*, vol. 48, no. 6, pp. 1423-1438, June 2002.

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

[8] L. H. Ozarow, and S. K. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback," *IEEE Trans. on Information Theory*, vol. IT-30, no. 4, pp. 667-671, July 1984.

[9] S. Pombra, and T. M. Cover, "Non white Gaussian multiple access channel with feedback," *IEEE Trans. on Information Theory*, vol. 40, no. 3, pp. 885-892, May 1994.

[10] J. P. M. Schalkwijk, and T. Kailath, "A coding scheme for additive noise channels with feedback - Part I: No bandwidth constraint," *IEEE Trans. on Information Theory*, vol. IT-12, no. 2, pp. 172-182, April 1966.