A Survey of Literature Regarding the Multiple Access Channel and Broadcast Channel with Noiseless Feedback

Michael L Dickens
Department of Electrical Engineering
University of Notre Dame
mdickens@nd.edu
DRAFT of November 16, 2006

1 Introduction

This report is a survey of currently known literature with respect to noiseless ("perfect") feedback (FB) for the multi–user access channel (MAC, "MISO") and broadcast channel (BC, "SIMO"), for both the discrete memoryless (DM) case as well as for the additive white gaussian noise (AWGN) case. A motivating example is the FB uplink from a controlling ground terminal to a satellite, where for all practical purposes the signal power can completely dominate the noise and hence the link can be viewed as noiseless.

A review of literature on noiseless FB in information theory is provided in the next section. The MAC achievable rates as found in [1] and [2] will be discussed in Section 2.1. The BC achievable rates as found in [3] will be discussed in Section 2.2. Complete determination of the capacity region of some MACs and BCs with FB remain an open problem, e.g. the inner and outer bounds do not coincide.

1.1 Literature Review

The effects of FB on capacity was initiated by Shannon [4], where he proved that for a DM single–user ("SISO") channel the channel capacity is not increased through the use of FB. This was subsequently shown to be true for additive white Gaussian noise (AWGN) FB channels (see, e.g. [5][6][7] and more recently [8]). In particular, Schalkwijk and Kailath published a set of papers in 1966 [9][10] providing coding schemes for additive noise channels for both bandlimited and non–bandlimited channels. They showed that, although the capacity does not increase with FB for white ("time independent") additive noise channels, the complexity of encoding and decoding as well as the transmission latency can be greatly reduced with the added FB. Many of the achievability arguments for additive noise channels with FB are based upon the "Kailath–Schalkwijk" coding scheme.

If the added noise is not white (i.e. "correlated" or "colored": the channel is not memoryless), then the channel capacity can be increased [11] since the FB provides the encoders with better knowledge of the channel and hence allows them to combat the noise more effectively. Cover and Pombra in 1989 [12] summarize and characterize the AWGN FB capacity, with the results that $C_{NFB} \leq C_{FB} \leq 2C_{NFB}$[1] and $C_{NFB} \leq C_{FB} \leq C_{NFB} + \frac{1}{2}$, where C_{NFB} and C_{FB} are the capacity without and with FB respectively. Cover's celebrated Feedback Capacity Conjecture ("Feedback Doesn't Help Much" [14]) has been discussed by Ordentlich in 1996 [15], Chen and Yanagi in 1997–9 [16], and most recently by Kim in 2006 [17].

Gaardner and Wolf [18] are credited with publishing in 1975 the first results showing that capacity can be increased by moving from a DM SISO channel with (or without) FB to a DM 2–user MAC with FB. In essence, the senders can decode each other's transmissions as the receiver does (before, at the same time, or

[1] This was proved by Ebert in 1972 [13] and repeated by Cover and Pombra in 1989 [12].

after – depending on the model). The encoders can then cooperate to resolve the uncertainty at the receiver, sending information at the higher cooperative rate rather than the lower non–cooperative rates. Cover and Leung [1] in 1981 determined an achievable region for the 2–user MAC with FB using superposition, block Markov encoding, random coding, and list decoding. This same result was found independently in 1977 by Carleial [19], who went on to consider the case of two users observing different (generalized) FB in 1982 [20]. Ozarow's 1979 dissertation [21] showed that FB can increase capacity of the AWGN MAC, and provided the capacity region for the 2–user DM MAC with FB; Ozarow's findings were reported in 1979 [22][2] and 1984 [2]. Thomas in 1987 [23] proved that the total capacity (sum of the rates of all the senders) of a k–user AWGN MAC with FB is less than twice the total capacity without FB. In 1989, Zeng, Kuhlmann, and Buzo [24] offered a different[3] proof of the achievability for the 2–user DM MAC with FB. In 1994, Pombra and Cover [25] show that this "factor of 2 bound" on capacity holds even when cooperation and prediction are combined in a k–user MAC with FB with non–white additive Gaussian noise. Iacobucci and Benedetto in 1997–8 [26][27] developed a k–user coding scheme for the MAC with FB using the constraint of equal power per user. They found that their scheme provided varying capacity gains depending on the number of users, with the greatest capacity gains for $k \in \{2, 3\}$. Kramer in 1999 [28] discusses FB strategies for some 2–user MACs.

Cover [29] published the initial research into the broadcast channel (BC) in 1972, and Bergmans provided a simple converse to show the complete capacity region of the k–user AWGN BC in 1974 [30]. In 1977 Leighton and Tan [31] showed that the addition of noiseless FB in a 2–user strongly degraded DM BC does not increase the system capacity. El Gamal in 1977 [32] discussed the 2–user BC with and without FB, and later showed that the capacity of the 2–user physically degraded[4] DM BC[33] and physically degraded AWGN BC [34] do not change with FB. Ozarow's 1979 dissertation [21] showed that FB can increase capacity of the AWGN BC, and provided the an achievable region and outer bound for the 2–user DM BC with FB; Ozarow's findings were reported in 1979 [22]. Dueck in 1980 [35] is credited with the first *peer–reviewed* publication showing by example that the capacity of a 2–user BC can be improved through partial FB. In 1984 Ozarow and Leung [3] showed a new way to achieve the capacity region for the 2–user Gaussian BC[5] with FB. In 1998, Cover [36] provided a general overview of the BC, though he only touched on FB by primarily citing the references provided here to that date.

Kramer in 2002 [37] discussed FB strategies for AWGN interference networks, including the MAC and BC as special cases. He extends the Kailath–Schalkwijk coding scheme from 2 users to k users via the Fourier modulated estimated correction ("Fourier MEC"), which assigned each user's message in the network a distinct modulation frequency from the rows of a DFT matrix. The Fourier MEC achieves the sum–rate capacity for MACs as long as the transmitter powers are all the same and beyond some threshold. In 2005, Wu, Vishwanath, and Arapostathis [38] conjectured[6] that the k–user AWGN MAC and BC with FB are duals of each other under the constraint that the total transmit power remain constant. Kramer and Gastpar in 2006 [39] generalize and refine the Hekstra and Willems [40] dependence balance bounds, and apply their updated dependence balance to the AWGN MAC with FB. They find the sum–rate capacity when all users have the same per–symbol power constraints.

For "a blast from the past" make sure to review Van der Meulen's 1977 survey of multi–way chan-

[2]This document was a NASA internal report which was apparently not further disseminated.

[3]The authors claim: "and simpler".

[4]A physically degraded, or cascaded, channel, is one for which $p(y_1, y_2|x) = p(y_1|x)p(y_2|y_1)$.

[5]By definition, the AWGN BC with $N_a \neq N_b \; \forall \; a \neq b$ is stochastically degraded, i.e. $p(y_2|x) = \sum_{y_1} p(y_1|x)p(y_2|y_1)$, but not necessarily physically degraded.

[6]The conjecture is in their accompanying presentation.

nels [41] and El Gamal and Cover's 1980 summary of multiple user information theory [42]. Both of these articles have sections on FB, though their use is marginal since they came before much of the work on MAC and BC with FB was published.

1.2 Acronyms and Notation

The following notation is used throughout this report.

A_b	refers to the "time" index "b" for the variable "A".
$A_{a,b}$	refers to the "time" index "b" for the variable "A_a".
$\mathbf{A}^b$	generally refers to a b–length vector holding the results $A_0, A_1, \ldots, A_b$.

For quicker reference, here is a list of acronyms used throughout this report, provided in alphabetical order.

AWGN	Additive White Gaussian Noise
BC	Broadcast Channel (a type of SIMO channel)
DM	Discrete and Memoryless
FB	Feedback
LMMSE	Linear Minimum Mean Squared Error
MAC	Multi–Access Channel (a type of MISO channel); also Multi–user Access Channel
MISO	Multiple Input Single Output
SIMO	Single Input Multiple Output
SISO	Single Input Single Output

2 Results

In this Section results from the achievability and converse (where available) for the MAC and BC with FB will be presented. For the purposes of this discussion, feedback is causal in that the received values (Y_i) are available to the encoders before the next values (X_{i+1}) are to be encoded. Note that all of the theorems discussed here are for $k = 2$ users, and in general are difficult to expand to more users without significant changes.

2.1 MAC with FB

The capacity region for a MAC with FB would intuitively be no smaller than that without due to the encoders' learning about each other's transmissions via feedback and hence "cooperating" with each other indirectly. In the worst case, encoders can learn nothing about each other and the situation reverts back to the MAC without FB. In the best case, encoders have full knowledge each other's messages, will fully cooperate, and hence can realize maximum system capacity.

Figure 1 shows a generic MAC with FB. Regarding this figure, unless otherwise stated for specific derivations: (1) the messages $m_1, \ldots, m_k$ are generated (a) independently of each other, and (b) uniformly over each source's alphabet; (2) the message encoders $f_1, \ldots, f_k$ have knowledge of each other through the feedback only, not through any direct information exchange, and thus, conditioned upon the feedback, generate output codewords independently of each other; and (3) the feedback can be information obtained before, during, or after processing by the decoder.

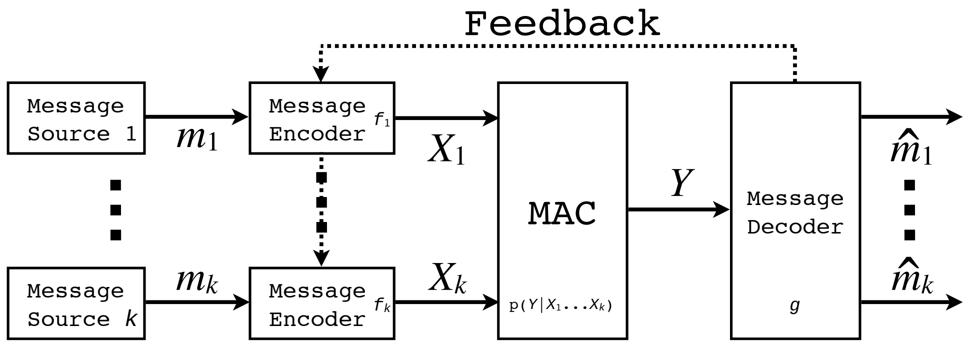
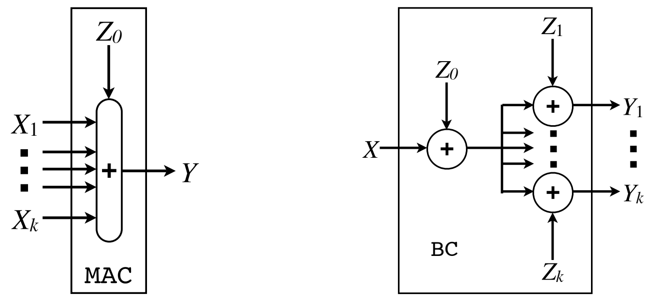


Figure 1: Generic k–User Multiple Access Channel with Feedback



(a) (b)

Figure 2: k–User AWGN Replacements for the Similarly–Labeled Boxes from (a) Figure 1 and (b) Figure 4.

Theorem 1 (Cover and Leung [1]) *Let U be a discrete random variable which takes on values in the set $\mathcal{U} = \{1, 2, \ldots, m\}$, where $m = \min\{\|\mathcal{X}_1\| \cdot \|\mathcal{X}_2\|, \|\mathcal{Y}\|\}$ and $\|\mathcal{X}_k\|$ denotes the alphabet size of user $k \in \{1, 2\}$. Consider the set $\mathcal{P}$ of all joint distributions of the form*

$$P_{UX_1X_2Y}(u, x_1, x_2, y) = P_U(u)P_{X_1|U}(x_1|u)P_{X_2|U}(x_2|u)P_{Y|X_1X_2}(y|x_1, x_2) \ ,$$

where $P_{Y|X_1X_2}(y|x_1, x_2)$ is fixed for the channel, and the channel is DM. Then an achievable rate region for the 2–user DM MAC with FB as shown in Figure 1 is $\mathcal{R}_{DM,MAC,FB}$, defined[7] as the convex hull of all rate pairs (R_1, R_2) satisfying

$$\begin{aligned} 0 \le R_1 &< I(X_1; Y|X_2, U) \ , \\ 0 \le R_2 &< I(X_2; Y|X_1, U) \ , \textit{ and} \\ R_1 + R_2 &< I(X_1, X_2; Y) \ . \end{aligned}$$

An outline of the achievable coding scheme is as follows: B blocks each of length n will be sent, and the messages to be sent at each block are $m_k \in \{1, 2, \ldots, \lceil 2^{nR_k} \rceil\}$ and $k \in \{1, 2\}$; this range for k will be used throughout the discussion. For the messages to be transferred in block $b \in \{1, 2, \ldots, B\}$, $(j_b, i_{1,b}, i_{2,b})$ denotes the indices of interest to encode, with j_b being the index sent to resolve uncertainty from the previous block and $i_{k,b} \in \{1, 2, \ldots, \lceil 2^{nR_k} \rceil\}$ being indices representing the new messages to be sent for the current block. The encoders actually generate $\mathbf{x}_k^n(i_{k,b}, j_b)$ for transmission, which are found as follows: Fix a choice of $P_U(u)$ and $P_{X_k|U}(x_k|u)$. Generate a sequence of i.i.d. random vectors $\mathbf{u}^n = (u_1, \ldots, u_n)$ according to $P_U(u_1, \ldots, u_n) = \prod_{i=1}^n P_U(u_i)$, indexed as $\mathbf{u}^n(j_b)$, for $j_b \in \{1, 2, \ldots, \lceil 2^{nR_0} \rceil\}$; it turns out that $R_0 < I(Y; U)$. Now for each $\mathbf{u}^n(j_b)$, generate the conditionally–independent sequences $\mathbf{x}_k^n(i_{k,b}, j_b)$ according to $\prod_{i=1}^n P_{X_k|U}(x_{k,i}|u_i(j_b))$ where $x_{k,i}$ is the i^{th} entry of this $\mathbf{x}_k^n(i_{k,b}, j_b)$.

[7] We disagree with Cover and Leung that the rate region has closure.

Encoding is done as follows: In the first block, all encoders use the same predetermined ("initialization") index $\hat{j}_1$ and look up $\mathbf{x}_k^n(\hat{i}_{k,1}, \hat{j}_1)$ for transmission from the indices set $(\hat{j}_1, \hat{i}_{1,1}, \hat{i}_{2,1})$. For any block, each encoder sends its n values of $x_k^n(i_{k,b}, j_b)$. In block $b : 1 < b < B$ the transmitters split the available rate into two parts: one part (the j_b) provides enough information so that the receiver can resolve any uncertainly left over from the previous block; the other part (the $i_{k,b}$) is new information. The rate of this new information must be small enough so that, through the FB link, each transmitter can reliably determine each other's message. In the last block (B), each encoder sends only the information needed to resolve the uncertainly of the previous block; no new information is sent (i.e. $i_{k,B}$ are set to predetermined "termination" indices).

Decoding is done as follows, for any block b: (1) The receiver declares $\tilde{j}_b$ was sent iff it is the unique value of j_b such that $(\mathbf{u}^n(j_b), \mathbf{y}^n)$ are jointly typical, where $\mathbf{y}^n$ is the set of n current received values. (2) Given the $\tilde{j}_b$ from part (1), the receiver declares $\tilde{i}_{k,b-1}$ as the intended message indices iff they are the unique values of $i_{k,b-1}$ such that $(x_k^n(i_{k,b-1}, \tilde{j}_b), y_{b-1})$ are jointly typical. (3) Both encoders estimate the transmitted message of the other encoder. This is done by declaring that (for encoder 1) $\tilde{i}_{2,b}$ was sent iff it is the unique value of $i_{2,b}$ such that $(\mathbf{x}_1^n(i_{1,b}, j_b), \mathbf{x}_2^n(i_{2,b}, j_b), \mathbf{y}^n)$ are jointly typical; encoder 1 knows the values for $i_{1,b}$, j_b, $\mathbf{x}_1^n(i_{1,b}, j_b)$, and $\mathbf{y}^n$ (through FB).

The bounds on each encoder's transmission rate are found by examining the probability of error at both the receiver and the encoders required for successful decoding. If the number of blocks B is large, then the effective transmission rate will be negligibly affected by the lack of new information in the last block. The bound on the cardinality of $\mathcal{U}$ was determined by Salehi [43].

There exists no converse to show that the achievable rate region in Theorem 1 is the capacity region, and this still remains an open problem. However, Ozarow [2] provides a converse for an outer bound assuming that the encoders fully cooperate to determine what to transmit. This bound, $\mathcal{C}_{DM,MAC,FB,Outer}$ is the union of all rate pairs (R_1, R_2) satisfying

$$\begin{aligned} 0 \le R_1 &\le I(X_1; Y|X_2) \ , \\ 0 \le R_2 &\le I(X_2; Y|X_1) \ , \text{and} \\ R_1 + R_2 &\le I(X_1, X_2; Y) \end{aligned}$$

for some joint $p(x_1, x_2)$. This result is the same as provided by the cut–set bound [11, Section 14.10] for 3 nodes (2 transmitting and 1 receiving). Interestingly, this result can also be reverted back to the known capacity region for the DM MAC with *no* FB by constraining $p(x_1, x_2) = p(x_1)p(x_2)$ and applying the closure of the convex hull of all rate pairs.

In the same article, Ozarow also found the capacity region for the AWGN MAC with FB and 2 users:

Theorem 2 (Ozarow [2]) *For the AWGN MAC with FB and $k = 2$ users as shown in Figures 1 and 2(a) under the constraints*

$$\begin{aligned} k &\in \{1, 2\} \\ i &\in \{1, 2, \ldots, N\} \ \textit{is the time index for each block} \\ Y_i &= X_{1,i} + X_{2,i} + Z_i \\ Z_i &\sim \mathcal{N}(0, N_0) \\ \mathbf{E}[X_{k,i}^2] &\le P_k \end{aligned}$$

where the expectation $\mathbf{E}$ is over all messages m_k and feedback $\mathbf{Y}^{i-1}$, then the capacity region is

$$\mathcal{C}_{AWGN,MAC,FB} = \bigcup_{0 \le \rho \le 1} \Bigg\{ (R_1, R_2) : 0 \le R_1 \le \frac{1}{2} \log\left(1 + \frac{P_1}{N_0}\left(1-\rho^2\right)\right) \ ,$$
$$0 \le R_2 \le \frac{1}{2} \log\left(1 + \frac{P_2}{N_0}\left(1-\rho^2\right)\right) \ , \textit{and}$$
$$R_1 + R_2 \le \frac{1}{2} \log\left(1 + \frac{P_1 + P_2 + 2\rho\sqrt{P_1 P_2}}{N_0}\right) \Bigg\} \ .$$

Note that when $\rho = 0$, then there is no cooperation and the capacity region reverts back to that for the AWGN MAC with no FB. Here is an outline of the achievable coding scheme: Transmissions are in blocks of length N; one incoming message $m_{k,j} \in \{0, 1, \ldots, M_k - 1\}$ is transmitted during each block. Assume that the incoming messages $m_{k,j}$ are equally–likely and independent for each source. Each encoder maps the incoming message via the function

$$\theta_k(m_{k,j}) = \frac{m_{k,j}}{M_k - 1} - \frac{1}{2} \tag{1}$$

such that θ_k is uniformly distributed over the M_k equally–spaced values in $[-\frac{1}{2}, \frac{1}{2}]$. When M_k are both large, then $\mathbf{E}(\theta_k^2) \simeq \frac{1}{12}$. After transmission $i \in \{1, 2, \ldots, N\}$, the receiver computes a linear minimum mean squared error (LMMSE) estimate $\hat{\theta}_{k,i}$ of θ_k. Define the estimation error $\epsilon_{k,i} = \hat{\theta}_{k,i} - \theta_k$, and its variance $\alpha_{k,i} = \text{var}(\epsilon_{k,i})$. At time $i = 1$, only encoder 1 sends information, and it sends $X_{1,1} = \theta_1\sqrt{12P_1}$. The receiver gets $Y_1 = X_{1,1} + Z_1$, and forms an estimate of θ_1 via $\hat{\theta}_{1,1} = \theta_1 + \epsilon_{1,1}$ where $\epsilon_{1,1} \sim \mathcal{N}(0, \frac{N_0}{12P_1})$; thus $\alpha_{1,1} = \frac{N_0}{12P_1}$. At time $i = 2$, only the second encoder sends information, and it sends $X_{2,2} = \theta_2\sqrt{12P_2}$. The receiver gets $Y_2 = X_{2,2} + Z_2$, and similarly forms $\hat{\theta}_{2,2} = \theta_2 + \epsilon_{2,2}$, where $\epsilon_{2,2} \sim \mathcal{N}(0, \frac{N_0}{12P_2})$; thus $\alpha_{2,2} = \frac{N_0}{12P_2}$. Also at time 2, the receiver sets $\hat{\theta}_{1,2} = \hat{\theta}_{1,1}$ so–as to be prepared for the rest of the transmission. At each time i, the decoder feeds back the received value Y_{i-1}, with $Y_0 = 0$. Since the encoder knows θ_k, it can compute $\epsilon_{k,i}$ at each step.

At time $3 \le i \le N$, the encoders send

$$X_{1,i} = \epsilon_{1,i-1}\sqrt{\frac{P_1}{\alpha_{1,i-1}}} \qquad \text{and} \qquad X_{2,i} = \text{sgn}(\rho_{i-1})\epsilon_{2,i-1}\sqrt{\frac{P_2}{\alpha_{2,i-1}}}$$

with

$$\text{sgn}(x) = \begin{cases} 1 & x \ge 0 \\ -1 & x < 0 \end{cases} \ , \tag{2}$$

and the correlation coefficient between the $\epsilon_{k,i}$

$$\rho_i = \frac{\mathbf{E}[\prod_k \epsilon_{k,i}]}{\sqrt{\prod_k \alpha_{k,i}}} \ . \tag{3}$$

The "sgn" in $X_{2,i}$ is used to keep both encoders' transmissions the same sign. The receiver gets $Y_i = Z_i + \sum_k X_{k,i}$, and forms the LMMSE estimate of θ_k given the previous estimates ($\hat{\theta}_{k,i-1}$), the previous estimation errors ($\epsilon_{k,i-1}$), and the current received information (Y_i) as[8]

$$\hat{\theta}_{k,i} = \hat{\theta}_{k,i-1} - \frac{\mathbf{E}[Y_i \epsilon_{k,i-1}]}{\mathbf{E}[Y_i^2]} Y_i \ .$$

This process is continued until time $i = N$, or when a predetermined variance threshold as been met, at which point the receiver will use the closest available true value of θ_k as the final estimate and reverse

[8] Ozarow's paper has the typo "1" instead of "i" for the receiver index in Equations (5) and (7).

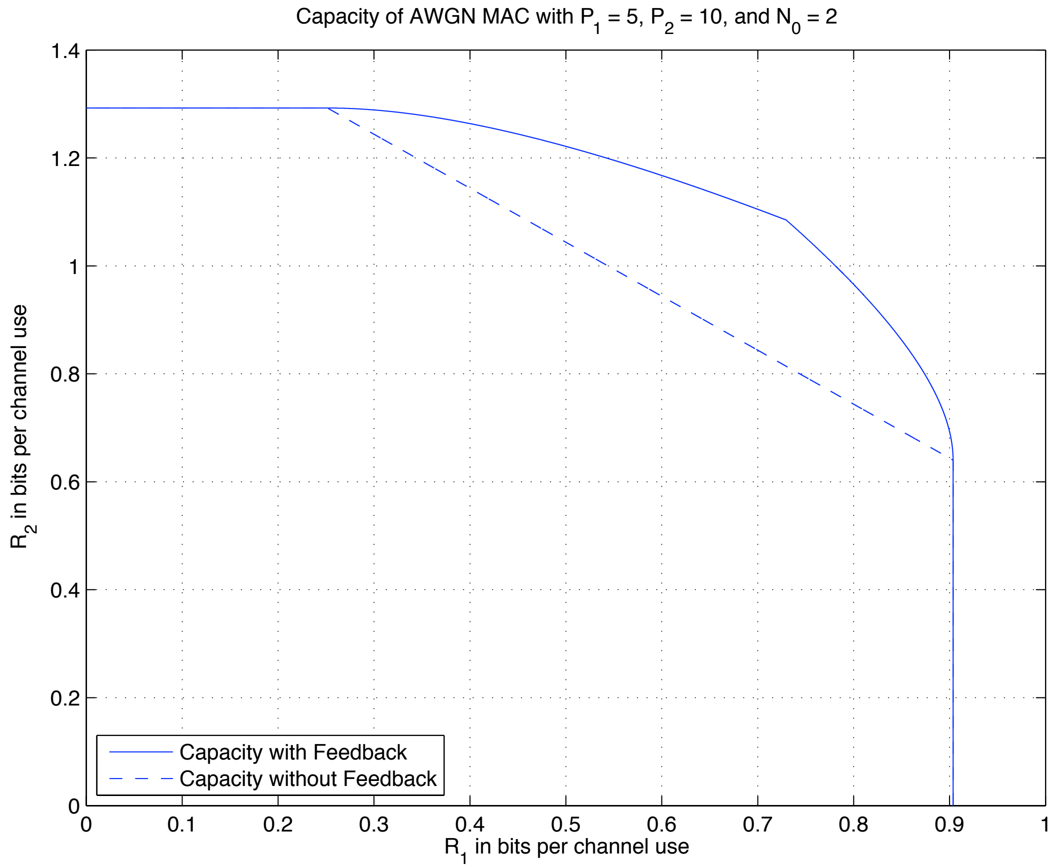


Figure 3: Capacity Regions for the AWGN MAC with and without FB

computes $m_{k,j}$ from Equation 1. The capacity bounds are found by examining the probability of error at the receiver required for successful decoding.

The converse to Theorem 2 starts by proving an outer bound for the general DM MAC with FB, and noting that the capacity of the AWGN MAC with FB must be included within this region. The capacity region is then (1) specialized for the particular $p(y|x_1x_2)$ being jointly Gaussian and thus $H(Y|X_1X_2) = \frac{1}{2}\log 2\pi e N_0$, and (2) generalized noting that for any random variable W with variance σ_W^2, $H(W) \leq \frac{1}{2}\log 2\pi e \sigma_W^2$. Jensen's inequality is used to swap expectation and $\log$ functions multiple times, with the final result that the achievable rate region as found in Theorem 2 represents the actual capacity region.

Figure 3 shows the capacity region for the AWGN MAC with and without FB. Note that the point at around $(R_1, R_2) = (0.73, 1.08)$ (where the AWGN MAC with FB capacity curve would not have a continuous derivative) is most likely an artifact of the computation, and this curve should ideally be continuously differentiable. This point is where $R_1 + R_2$ equals the sum-rate constraint. Note that at least for the 2 user system FB provides capacity gains relative to no FB.

2.2 BC with FB

The generic BC with FB is shown in Figure 4. Regarding this figure, unless otherwise stated for specific derivations: (1) the messages $m_1, \ldots, m_k$ are generated (a) independently of each other, and (b) uniformly over each source's alphabet; (2) the message decoders $g_1, \ldots, g_k$ have no knowledge of any other decoder's information, and thus generate both the feedback and output message independently of other decoders; and

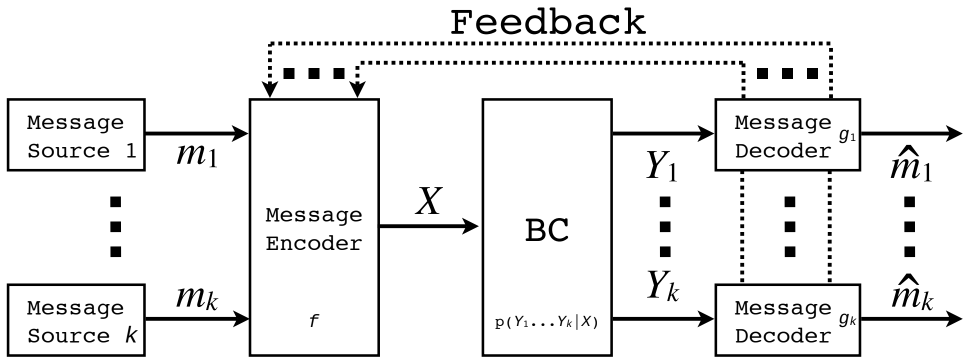


Figure 4: Generic k–User Broadcast Channel with Feedback

(3) the feedback can be information obtained before, during, or after processing by the decoders, though generally all feedback is taken at the same processing location.

All of the theorems presented here are for $k = 2$ users, and in general cannot easily be expanded to higher numbers of users without significant changes. All derivations summarized in this report are found under the following conditions:

$$\begin{aligned} k &\in \{1, 2\} \\ i &\in \{1, 2, \dots, N\} \text{ is the time index for each block} \\ Y_{k,i} &= X_i + Z_{0,i} + Z_{k,i} \\ Z_{l,i} &\sim \mathcal{N}(0, N_l) \text{ for } l \in \{0, 1, 2\} \\ N_l &\neq 0 \text{ for } l \in \{0, 1, 2\} \\ \mathbf{E}[Z_{m,i} Z_{l,i}] &= 0 \;\forall\; m \neq l \;,\; m, l \in \{0, 1, 2\} \\ \mathbf{E}[X_i^2] &\leq P \end{aligned}$$

where the last expectation $\mathbf{E}$ is over all messages m_k and feedback $\mathbf{Y}_1^{i-1} \dots \mathbf{Y}_k^{i-1}$, and P is the average power.

Theorem 3 (Ozarow and Leung [3]) *For the AWGN BC with FB and $k = 2$ encoders as shown in Figures 4 and 2(b), there exists a code that achieves the rate region defined by*

$$\mathcal{R}_{AWGN,BC,FB} = \bigcup_{0 \leq g} \left\{ (R_1, R_2) \;:\; 0 \leq R_1 < \frac{1}{2} \log \left[\frac{N_0 + N_1 + P}{N_0 + N_1 + \frac{P}{D^*} g^2 \left(1 - \rho^{*2}\right)} \right] , \right.$$
$$\left. 0 \leq R_2 < \frac{1}{2} \log \left[\frac{N_0 + N_2 + P}{N_0 + N_2 + \frac{P}{D^*} \left(1 - \rho^{*2}\right)} \right] \right\}$$

where $D^* = 1 + g^2 + 2g\rho^*$, g is a non–negative number which allows trading off of achievable rates between the receivers, and ρ^* is the largest solution in $(0, 1)$ of[9]

$$-\rho^* = \frac{(N_0 \Sigma + N_1 N_2)\rho^* - \frac{P\Sigma}{D^*} g \left(1 - \rho^{*2}\right)}{\sqrt{\Pi}\sqrt{N_0 + N_1 + \frac{P}{D^*} g^2 \left(1 - \rho^{*2}\right)}\sqrt{N_0 + N_2 + \frac{P}{D^*}\left(1 - \rho^{*2}\right)}}$$
$$\Sigma = P + N_0 + N_1 + N_2$$
$$\Pi = (P + N_0 + N_1)(P + N_0 + N_2)$$

[9] There is a typo in [3, Equation 8] of the denominator's second square root: "σ_2^2" should read "σ_1^2".

Note that this region is not necessarily convex, and nor is the closure or a convex hull operation applied to the union. Here is an outline of the achievable coding scheme, somewhat following that provided for Theorem 2: Transmissions are in blocks of length N, during which time the encoder will transmit two messages, call them m_{1j} and m_{2j}, one each from the alphabets $\mathcal{M}_k = \{0, 1, \ldots, M_k - 1\}$ for $k \in \{1, 2\}$. Each message is uniformly generated independently of the other, and is mapped into a real number in $[-\frac{1}{2}, \frac{1}{2}]$ via Equation 1. When M_k are both large, then $\mathbf{E}(\theta_k^2) \simeq \frac{1}{12}$.

For the first two time slots in each block ($i = 1, 2$), the encoder transmits $X_i = \theta_i \sqrt{12P}$. Each receiver estimates its message by $\hat{\theta}_{k,i} = \theta_k + \epsilon_{k,i}$. For the first transmission ($i = 1$), only the first decoder does any processing, and it computes $\hat{\theta}_{1,1} = \frac{Y_{1,1}}{\sqrt{12P}} = \theta_1 + \frac{Z_{0,1}+Z_{1,1}}{\sqrt{12P}}$. The first decoder ignores the second transmission and sets $\hat{\theta}_{1,2} = \hat{\theta}_{1,1}$ in preparation of the rest of the algorithm. For the second transmission ($i = 2$), only the second decoder does any processing, and it computes $\hat{\theta}_{2,2} = \frac{Y_{2,2}}{\sqrt{12P}} = \theta_2 + \frac{Z_2+Z_{2,2}}{\sqrt{12P}}$. At each time i, the encoder receives via feedback the values $Y_{k,i-1}$ from the decoders, with $Y_{k,0} = 0$. Since the encoder knows θ_k, it can compute $\epsilon_{k,i}$ at each step. Defining $\alpha_{k,i} = \mathbf{E}[\epsilon_{k,i}^2]$ and ρ_i as in Equation 3, then $\alpha_{k,2} = \frac{N_{0,2}+N_{k,2}}{12\ P}$ and $\rho_2 = 0$.

For transmissions $3 \leq i \leq N$, the encoder sends

$$X_i = \left[\frac{\epsilon_{1,i-1}}{\sqrt{\alpha_{1,i-1}}} + \frac{\epsilon_{2,i-1}}{\sqrt{\alpha_{2,i-1}}}\, g\ \mathrm{sgn}(\rho_{i-1}) \right] \sqrt{\frac{P}{D_{i-1}}}\ ,$$

where g is a nonnegative number which trades off between the receivers, $\mathrm{sgn}(x)$ is defined as in Equation 2, and $D_i = 1 + g^2 + 2g|\rho_i|$ is a normalization parameter for satisfying the average power constraint. The receivers get $Y_{k,i} = X_i + Z_{0,i} + Z_{k,i}$ and form their estimates as linear combinations of the old and new information via

$$\hat{\theta}_{k,i} = \hat{\theta}_{k,i-1} - \frac{\mathbf{E}[\epsilon_{k,i-1} Y_{k,i}]}{\mathbf{E}[Y_{k,i}^2]} Y_{k,i}\ .$$

This process is continued until time $i = N$, or when a predetermined variance threshold as been met, at which point the receivers will use the closest available true value of θ_k as the final estimate and reverse computes m_{kj} from Equation 1. The rate bounds are found by examining the probability of error at the receiver required for successful decoding.

In the same paper, Ozarow and Leung also derive an outer bound on the capacity of the AWGN BC with FB. This is done by allowing one receiver (say receiver 1) to obtain both Y_1 and Y_2, and hence form a physically degraded BC with FB for the Markov chain $X \leftrightarrow (Y_1, Y_2) \leftrightarrow Y_2$. This process can be reversed by allowing the other receiver (2) to obtain both received values. Since the channel capacity is known for physically degraded BC's with FB, the outer bound on the capacity for the AWGN BC with FB can be found as the intersection of these two regions via the following theorem.

Theorem 4 (Ozarow and Leung [3]) *For the AWGN BC with FB and $k = 2$ encoders as shown in Figures 4 and 2(b), an outer bound on the capacity region for this channel is $\mathcal{C}_{AWGN,BC,Outer} \subseteq \mathcal{C}_1 \bigcap \mathcal{C}_2$, with*

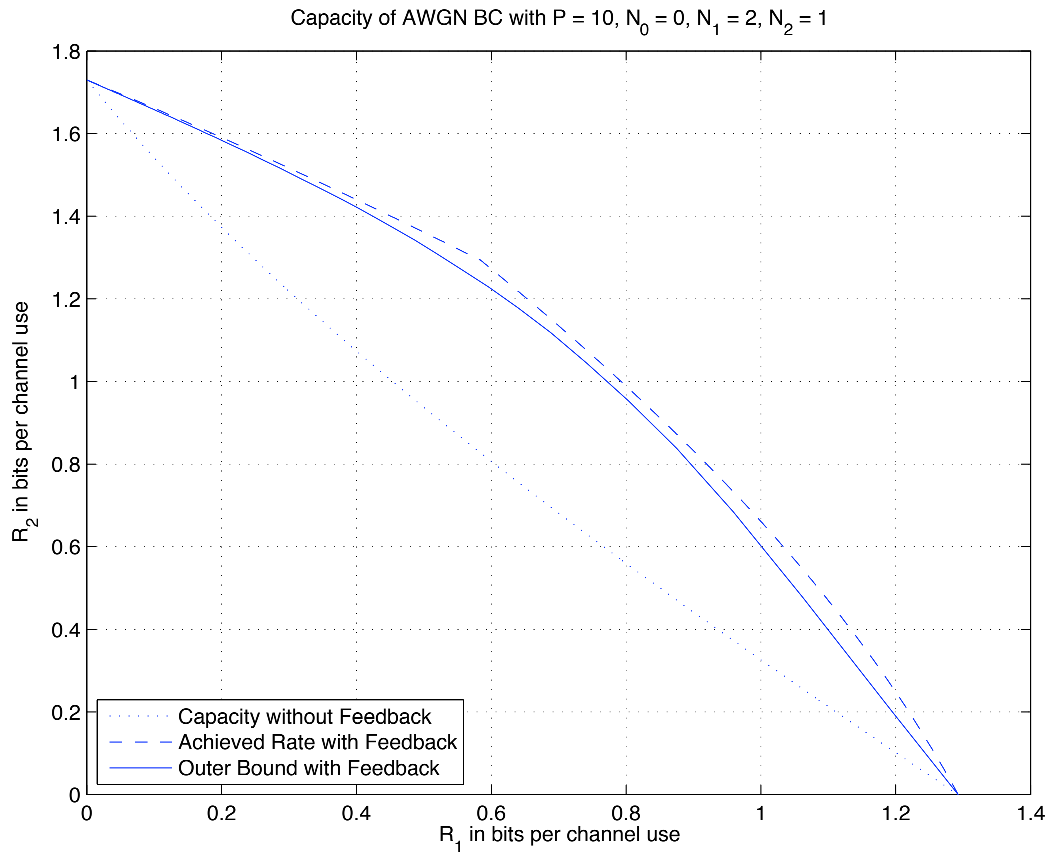


Figure 5: Rate Regions for the AWGN BC with and without FB

$$\mathcal{C}_1 = \bigcup_{0 \le \alpha \le 1} \Bigg\{ (R_1, R_2) \ : \ R_1 \le \frac{1}{2} \log \left[1 + \frac{\alpha P}{N_0 + \frac{N_1 N_2}{N_1 + N_2}} \right] , \quad R_2 \le \frac{1}{2} \log \left[1 + \frac{(1-\alpha) P}{\alpha P + N_0 + N_2} \right] \Bigg\} ,$$

$$\mathcal{C}_2 = \bigcup_{0 \le \alpha \le 1} \Bigg\{ (R_1, R_2) \ : \ R_1 \le \frac{1}{2} \log \left[1 + \frac{(1-\alpha) P}{\alpha P + N_0 + N_1} \right] , \quad R_2 \le \frac{1}{2} \log \left[1 + \frac{\alpha P}{N_0 + \frac{N_1 N_2}{N_1 + N_2}} \right] \Bigg\} .$$

Figure 5 shows rate regions for the AWGN BC with and without FB. Note that, at least for the 2 user system, FB provides rate gains relative to no FB.

3 Bibliography

1. Thomas M. Cover and Cyril S. K. Leung-Yan-Cheong, "An Achievable Rate Region for the Multiple-Access Channel with Feedback," *IEEE Trans. Inform. Theory*, vol. 27, no. 3, pp. 292–298, May 1981.

2. Lawrence H. Ozarow, "The Capacity of the White Gaussian Multiple Access Channel with Feedback," *IEEE Trans. Inform. Theory*, vol. 30, no. 4, pp. 623–629, July 1984.

3. Lawrence H. Ozarow and Cyril S. K. Leung-Yan-Cheong, "An Achievable Region and Outer Bound for the Gaussian Broadcast Channel with Feedback," *IEEE Trans. Inform. Theory*, vol. 30, no. 4, pp. 667–671, July 1984.

4. Claude E. Shannon, "The Zero Error Capacity of a Noisy Channel," *IEEE Trans. Inform. Theory*, vol. 2, no. 3, pp. 8–19, Sept. 1956.

5. P. E. Green, Jr., "Feedback communication systems," in *Lectures on Communication System Theory*, E. J. Baghdady, Ed. New York: McGraw–Hill, 1961, pp. 345–368.

6. Andrew J. Viterbi, "The Effect of Sequential Decision Feedback on Communication Over the Gaussian Channel," *Information and Control*, vol. 8, no. 1, pp. 80–92, Feb. 1965.

7. George L. Turin, "Signal Design for Sequential Detection Systems with Feedback," *IEEE Trans. Inform. Theory*, vol. 11, no. 3, pp. 401–408, July 1965.

8. Fady Alajaji, "Feedback does not increase the capacity of discrete channels with additive noise," *IEEE Trans. Inform. Theory*, vol. 41, no. 2, pp. 546–549, Mar. 1995.

9. J. Pieter M. Schalkwijk and T. Kailath, "A Coding Scheme for Additive Noise Channels with Feedback – Part I: No Bandwidth Constraint," *IEEE Trans. Inform. Theory*, vol. 12, no. 2, pp. 172–182, Apr. 1966.

10. J. Pieter M. Schalkwijk, "A Coding Scheme for Additive Noise Channels with Feedback – Part II: Band-Limited Signals," *IEEE Trans. Inform. Theory*, vol. 12, no. 2, pp. 183–189, Apr. 1966.

11. Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*. John Wiley, New York, 1990.

12. Thomas M. Cover and Sandeep Pombra, "Gaussian feedback capacity," *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 37–43, Jan. 1989.

13. P. M. Ebert, "The Capacity of the Gaussian Channel with Feedback," *Bell Systems Technical Journal*, vol. 49, pp. 1705–1712, Oct. 1972.

14. Thomas M. Cover, "Conjecture: Feedback Doesnt Help Much," in *Open Problems in Communication and Computation*, Thomas M. Cover and B. Gopinath, Ed. New York: Springer–Verlag, 1987, pp. 70–71.

15. Erik Ordentlich, "On the Factor-of-Two Bound for Gaussian Multiple-Access Channels with Feedback," *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pp. 2231–2235, Nov. 1996.

16. Han Wu Chen and Kenjiro Yanagi, “Refinements of the Half-Bit and Factor-of-Two Bounds for Capacity in Gaussian Channel with Feedback,” *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 319–325, Jan. 1999.

17. Young-Han Kim, “A Counterexample to Covers $2P$ Conjecture on Gaussian Feedback Capacity,” *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3792–3793, Aug. 2006.

18. N. Thomas Gaardner and Jack K. Wolf, “The Capacity Region of a Multiple-Access Discrete Memoryless Channel Can Increase with Feedback,” *IEEE Trans. Inform. Theory*, vol. 21, no. 1, pp. 100–102, Jan. 1975.

19. Aydano B. Carleial, “On the capacity of the Gaussian multiple access channel with feedback,” *Revista Brasileira de Tecnologia*, vol. 27, pp. 135–145, July-Sept. 1977, Published in 1978.

20. ——, “Multiple-Access Channels with Different Generalized Feedback Signals,” *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 841–850, Nov. 1982.

21. Lawrence Howard Ozarow, “Coding and Capacity for Additive White Gaussian Noise Multi-User Channels with Feedback,” Ph.D. dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, USA, 1979.

22. Lawrence H. Ozarow and Cyril S. K. Leung-Yan-Cheong, “Deterministic Feedback Coding Schemes for the Additive White Gaussian Noise Broadcast and Multiple-Access Channels,” *NASA STI/Recon Technical Report N*, vol. 79, pp. 29 404–29 434, Jan. 1979.

23. Joy A. Thomas, “Feedback Can at Most Double Gaussian Multiple Access Channel Capacity,” *IEEE Trans. Inform. Theory*, vol. 33, no. 5, pp. 711–716, Sept. 1987.

24. Chao-Ming Zeng, Federico Kuhlmann, and Andres Buzo, “Achievable Proof of Some Multiuser Channel Coding Theorems using Backward Decoding,” *IEEE Trans. Inform. Theory*, vol. 35, no. 6, pp. 1160–1165, Nov. 1989.

25. Sandeep Pombra and Thomas M. Cover, “Non white gaussian multiple access channels with feedback,” *IEEE Trans. Inform. Theory*, vol. 40, no. 3, pp. 885–892, May 1994.

26. Maria-Stella Iacobucci and Maria-Gabriella Di Benedetto, “A Feedback Code for the Multiple Access Channel (MAC): a Case Study,” in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM), Communication Theory Mini-Conference (CTMC)*, Phoenix, AZ, USA, Nov. 1997, pp. 128–132.

27. ——, “Design and Performance of a Code for the Multiple Access Channel with Feedback,” in *Proceedings of the IEEE International Symposium on Information Theory*, Cambridge, MA, USA, Aug. 1988, p. 125.

28. Gerhard Kramer, “Feedback Strategies for a Class of Two-User Multiple-Access Channels,” *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2054–2059, Sept. 1999.

29. Thomas M. Cover, “Broadcast Channels,” *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

30. Patrick P. Bergmans, “A Simple Converse for Broadcast Channels with Additive White Gaussian Noise,” *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 279–280, Mar. 1974.

31. William J. Leighton III and Harry H. Tan, “Capacity Region of Degraded Broadcast Channels with Feedback,” *Information Sciences*, vol. 13, no. 2, pp. 167–177, 1977.

32. Abbas A. El Gamal, “Broadcast Channels With and Without Feedback,” in *Conference Record of the Eleventh Annual Asilomar Conference on Circuits, Systems, and Computers*, Shu-Park Chan, Ed., New York, Nov. 1978, pp. 180–183.

33. ——, “The Feedback Capacity of Degraded Broadcast Channels,” *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 379–381, May 1978.

34. ——, “The Capacity of the Physically Degraded Gaussian Broadcast Channel with Feedback,” *IEEE Trans. Inform. Theory*, vol. 27, no. 4, pp. 508–511, July 1981.

35. Gunter Dueck, “Partial Feedback for Two-Way and Broadcast Channels,” *Information and Control*, vol. 46, no. 1, pp. 1–15, 1980.

36. Thomas M. Cover, “Comments on Broadcast Channels,” *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2524–2530, Oct. 1998.

37. Gerhard Kramer, “Feedback Strategies for White Gaussian Interference Networks,” *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1423–1438, June 2002.

38. W. Wu, S. Vishwanath, and A. Arapostathis, “Feedback Strategies of Gaussian Interference Networks: Duality, Sum Capacity and Dynamic Team Problems,” in *Proceedings of the 43rd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, Sept. 2005, pp. 926–935.

39. Gerhard Kramer and Michael Gastpar, “Dependence Balance and the Gaussian Multiaccess Channel with Feedback,” in *Proceedings of the IEEE Information Theory Workshop*, Punta del Este, Uruguay, Mar. 2006, pp. 198–202.

40. Andries P. Hekstra and Frans M. J. Willems, “Dependence Balance Bounds for Single-Output Two-Way Channels,” *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.

41. Edward C. Van der Meulen, “A Survey of Multi-Way Channels in Information Theory: 1961-1976,” *IEEE Trans. Inform. Theory*, vol. 23, no. 1, pp. 1–37, Jan. 1977.

42. Abbas A. El Gamal and Thomas M. Cover, “Multiple User Information Theory,” *Proc. IEEE*, vol. 68, no. 12, pp. 1466–1483, Dec. 1980.

43. Masoud Salehi, “Improved Computational Efficiency in Derivation of Capacity Regions in Network Information Theory,” *Computers and Electrical Engineering*, vol. 2, no. 4, pp. 247–256, July 1996.