THE FEEDBACK CAPACITY OF DEGRADED BROADCAST CHANNELS

BY

ABBAS EL GAMAL

TECHNICAL REPORT NO. 21
SEPTEMBER 1, 1976

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION
GRANT ENG 76-03684

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THE FEEDBACK CAPACITY OF DEGRADED BROADCAST CHANNELS

by

Abbas El Gamal

TECHNICAL REPORT NO. 21
SEPTEMBER 1, 1976

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION
GRANT ENG 76-03684

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
The Feedback Capacity of Degraded Broadcast Channels

Abbas El Gamal†

ABSTRACT

This correspondence establishes that the capacity region of a discrete memoryless degraded broadcast channel is not increased by feedback.

†Stanford University. This work was partially supported by NSF Grant ENG 76-03684
1. Introduction

The capacity region of the discrete memoryless degraded broadcast channel (Cover [1]) was established in [2], [3], [4], and [6]. Bergmans [2] exhibited an achievable rate region. A converse for the binary symmetric broadcast channel was established by Wyner and Ziv [3]. Gallager [4] then proved a converse for the general discrete memoryless degraded broadcast channel. An alternative converse was given by Ahlswede [6]. Using methods similar to [4], it will be shown that the capacity region is unchanged by feedback.

2. Preliminaries and Definitions

The model under investigation is shown in Figure 1.

---

Figure 1.
There are two sources, the first producing an integer  \( W_1 \in \mathcal{M}_1 = \{1, \ldots, M_1\} \) and the second an integer  \( W_2 \in \mathcal{M}_2 = \{1, \ldots, M_2\} \). At the  \( n \)-th transmission the encoder maps the pair \( \{W_1, W_2\} \) and the past outputs \( \{Z_1, \ldots, Z_{n-1}\} \) and \( \{Y_1, \ldots, Y_{n-1}\} \) into \( X_n \). Thus,

\[
X_n = f_n(W_1, W_2, Z_1, \ldots, Z_{n-1}, Y_1, \ldots, Y_{n-1}), \quad n=1,2,\ldots,N
\]  

(1)

The channel is discrete memoryless with finite alphabet  \( x \in \mathcal{X} = \{1, \ldots, l\} \),  \( y \in \mathcal{Y} = \{1, \ldots, l\} \),  \( z \in \mathcal{Z} = \{1, \ldots, k\} \). The probability transition matrices of the channel are \( \{P_1(y|x)\} \) and \( \{P_2(z|x)\} \). By the degraded assumption, there exists a third probability transition matrix \( \{P_3(z|y)\} \) such that

\[
P_2(z|x) = \sum_{y \in \mathcal{Y}} P_1(y|x)P_3(z|y), \quad \forall x \in \mathcal{X}, \quad z \in \mathcal{Z}.
\]

Define an \( \left( M_1, M_2, N, \lambda_N \right) \) code for the channel as set of functions \( \{f_n\} \) together with the associated set of codewords generated as in (1), and two decoding maps,

\[
g_1: \mathcal{Y} \rightarrow \mathcal{M}_1, \quad g_2: \mathcal{Z} \rightarrow \mathcal{M}_2
\]

such that

\[
\max \left\{ \frac{p^{N}_{e,1}, p^{N}_{e,2}}{e,1, e,2} \right\} \leq \lambda_N
\]

where,

\[
p^{N}_{e,1} = \sum_{W_1 \in \mathcal{M}_1} \frac{1}{M_1} \cdot p \left( g_1(Y) = W_1 | W_1 \text{ was sent} \right)
\]

\[
p^{N}_{e,2} = \sum_{W_2 \in \mathcal{M}_2} \frac{1}{M_2} \cdot p \left( g_2(Z) = W_2 | W_2 \text{ was sent} \right)
\]
are the average probabilities of error.

A rate pair \((R_1, R_2)\) is said to be achievable if there exists a sequence of \(\left(\frac{nR_1}{2}, \frac{nR_2}{2}, n, \lambda_n\right)\) codes with \(\lambda_n \to 0\).

The capacity region is defined to be the closure of the set of all achievable \((R_1, R_2)\) pairs.

As established by [2], [3], [4], [6], the capacity region for the channel without feedback is given by the following:

**Theorem 1:** Let \(U, X, Y, Z\) be a joint ensemble such that \(X, Y, Z\) are as before and the number of sample points in \(U\) is \(\min \{I_{x,y}, I_{y,z}\}\). Then the set of all pairs \(\{(I(X;Y|U), I(U;Z))\}\) is the capacity region.

An equivalent characterization is \(R^* = \{ (R_1, R_2) : R_2 + \lambda R_1 \leq C(\lambda), \forall \lambda > 0 \}\), where \(C(\lambda) = \max_{Q_1(u)Q_2(x|u)} \left\{ I(U;Z) + \lambda I(X;Y|U) \right\} \). It will be shown that \(C(\lambda)\) is unchanged when feedback is added.

3. Feedback Converse

**Theorem 2** (Converse): If for some \(\lambda > 0, \varepsilon > 0\) we have

\[
R_2 + \lambda R_1 \geq C(\lambda) + \varepsilon
\]  

then there exists \(\varepsilon > 0\) such that

\[
\max \left\{ p_{e,1}^N, p_{e,2}^N \right\} \geq \varepsilon \text{ for all } N.
\]

**Proof:** The probability mass function on the joint ensemble \(W_1, W_2, X, Y, Z\), is given by

\[
p\left(\mathbf{w}_1, \mathbf{w}_2, \mathbf{x}, \mathbf{y}, \mathbf{z}\right) = \frac{1}{M_1 M_2} \prod_{n=1}^{N} q_n \left( x_n, w_1, w_2, z_1, \ldots, z_{n-1}, y_1, \ldots, y_{n-1} \right)
\]

\[
\cdot \prod_{n=1}^{N} p_1 \left( y_n | x_n \right) \prod_{n=1}^{N} p_3 \left( z_n | y_n \right)
\]

\[-3-\]
It follows that

\[(i) \quad I(W_2;Z) = H(W_2) - H(W_2|Z)\]
\[= NR_2 - H(W_2|Z),\]

\[(ii) \quad I(W_1;Y|W_2) = H(W_1) - H(W_1|Y,W_2)\]
\[\geq NR_1 - H(W_1|Y),\]

where \(R_i\) is the rate of source \(i\), \(i = 1, 2\).

From (i), (ii) it follows that,

\[N(R_2 + \lambda R_1) \leq \left[ I(W_2;Z) + \lambda I(W_1;Y|W_2) \right] +\]
\[\left[ H(W_2|Z) + \lambda H(W_1|Y) \right]\]  \hspace{1cm} (4)

We shall need a lemma that relates the first bracketed quantity to \(C(\lambda)\).

**Lemma 3:** For all \(\lambda \geq 0\),

\[I(W_2;Z) + \lambda I(W_1;Y|W_2) \leq \sum_{n=1}^{N} \left[ I(U_n;Z_n) + \lambda I(X_n;Y_n|U_n) \right]\]

\[\leq NC(\lambda),\]

where \(U_n \triangleq (W_2,Y_1,\ldots,Y_{n-1},Z_1,\ldots,Z_{n-1})\).  \hspace{1cm} (5)

**Proof:** A similar lemma was proved in [4] for the channel without feedback. We follow parallel steps to [4] in order to show the necessity of introducing a different definition for \(U_n\) as well as adding an additional inequality that demonstrates the intuitive fact that even if the receiver \(Y\) were given the sequence \(Z\), the achievable rate region would remain unchanged.

-4-
Now consider

\[
I(W_2;Z) = \sum_{n=1}^{N} I(W_2;Z_{n} | Z_1, \ldots, Z_{n-1})
\]

\[
= \sum_{n=1}^{N} H(Z_{n} | Z_1, \ldots, Z_{n-1}) - H(Z_{n} | W_2, Z_1, \ldots, Z_{n-1})
\]

\[
\leq \sum_{n=1}^{N} H(Z_{n}) - H(Z_{n} | W_2, Z_1, \ldots, Z_{n-1}, Y_1, \ldots, Y_{n-1})
\]

\[
= \sum_{n=1}^{N} I(Z_{n};U_{n})
\]

The change in Equation (5) from [4] in which \( U_n = (W_2, Y_1, \ldots, Y_{n-1}) \) is necessary since, with feedback, \( Z_n \) and \( Z_1, \ldots, Z_{n-1} \) are not necessarily independent given \( Y_1, \ldots, Y_{n-1} \).

Next, we would like to show that

\[
I(W_1;Y|W_2) \leq \sum_{n=1}^{N} I(X_n;Y_n|U_n)
\]

Consider the seemingly loose inequality,

\[
I(W_1;Y|W_2) \leq I(W_1;Y,Z|W_2)
\]

\[
= \sum_{n=1}^{N} I(W_1;Y_n,Z_n|W_2,Y_1,\ldots,Y_{n-1},Z_1,\ldots,Z_{n-1})
\]

\[
= \sum_{n=1}^{N} I(W_1;Y_n,Z_n|U_n)
\]

Applying the data processing inequality we get

\[
\sum_{n=1}^{N} I(W_1;Y_n,Z_n|U_n) \leq \sum_{n=1}^{N} I(W_1,X_n;Y_n,Z_n|U_n)
\]
\[
\sum_{n=1}^{N} H(Y_n, Z_n | U_n) = H(Y_n, Z_n | U_n, X_n, W_1)
\]

But by the discrete memoryless assumption, \((Y_n, Z_n)\) and \((U_n, W_1)\) are conditionally independent given \(X_n\). Thus

\[
\sum_{n=1}^{N} I(W_1, X_n; Y_n, Z_n | U_n) = \sum_{n=1}^{N} I(X_n; Y_n, Z_n | U_n)
\]

Now, since

\[
I(X_n; Y_n, Z_n | U_n) = I(X_n; Y_n | U_n) + I(X_n; Z_n | Y_n, U_n)
\]

it remains to show that \(I(X_n; Z_n | Y_n, U_n) = 0\). But this is true because \(U_n \rightarrow X_n \rightarrow Y_n \rightarrow Z_n\) form a Markov chain in this order.

Hence

\[
H(Z_n | Y_n, U_n) = H(Z_n | Y_n, U_n, X_n) = H(Z_n | Y_n)
\]

This establishes the required result, i.e., that

\[
I(W_1; Y | W_2) \leq \sum_{n=1}^{N} I(X_n; Y_n | U_n)
\]

and Lemma (3) is proved.

Now, combining Lemma (3) and Eq. (4) we obtain

\[
N(R_2 + \lambda R_1) \leq NC(\lambda) + \left[ H(W_2 | Z) + \lambda H(W_1 | Y) \right]
\]

Applying Fano's inequality, we have

\[
N(R_2 + \lambda R_1) \leq NC(\lambda) + \left[ NR_2 P^{N}_{e_2} + h \left( \frac{P^{N}_{e_2}}{R_2} \right) + \lambda NR_1 P^{N}_{e_1} + \lambda h \left( \frac{P^{N}_{e_1}}{R_1} \right) \right].
\]
Introducing assumption (2) together with some algebraic steps we obtain the required result. This completes the proof of Theorem (2).

4. Discussion:

The result is intuitively clear since feedback merely adds degraded forms of what the transmitter already knows. It is also consistent with Shannon's result on the discrete memoryless channel with feedback [5]. We believe that a similar result can be proved for any discrete memoryless broadcast channel, the difficulty being the lack of a simple description for the capacity region of this channel. It is also important to point out that in the case of multiple access channels, feedback does increase the channel capacity.

Acknowledgement: The author would like to thank Prof. Thomas Cover for discussions during the preparation of this paper.
References


