

Variable Length Codes for Degraded Broadcast Channels

Stéphane Musy

School of Computer and Communication Sciences, EPFL

CH-1015 Lausanne, Switzerland

Email: stephane.musy@epfl.ch

Abstract—This paper investigates the achievable rates using variable length codes when transmitting independent information over a degraded broadcast channel. In this note, we define the transmission rates from the perspective of the receivers and allow the decoders to make their decision at a different instant of time. We give an outer bound to the region of achievable rates, as well as examples of code that achieve this bound in some settings.

I. INTRODUCTION

When a user wishes to send information to a single receiver, it is known that the use of a variable length code does not allow to achieve a better transmission rate than the one achieved with a fixed length code [1]. Variable length coding is rather used to improve the error exponent of the communication scheme.

In the multiple-user setting, when a user wishes to transmit information to multiple receivers, a simple argument shows that, if we require that the receivers decode at the same instant of time, the set of achievable rates is the same for variable and fixed length codes.¹

In this paper, we focus on discrete memoryless broadcast channels with two receivers and consider the transmission of independent messages to each of them. As in [2], we let the codewords be infinite sequences and define the rates from the perspective of the receivers, and capture the trade-off between the amount of information received with the “timeliness” of the information.² Let us denote by M the number of possible messages for one receiver and by $E[N]$ its average decoding time, then we define the transmission rate to that receiver by $\frac{\log M}{E[N]}$. The expectation operator is motivated by the following argument. Suppose that, to send information to a receiver, we use the transmission method a large number n of times. The corresponding rate experienced by this receiver is equal to $\frac{n \log M}{l_1 + l_2 + \dots + l_n}$, where l_i represents the transmission duration of the i -th transmission. Hence, by the law of large numbers, the average transmission rate approaches $\frac{\log M}{E[N]}$ with probability one, as n gets large. The same definition is made in [3]. Notice

The work presented in this paper was partially supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

¹To the contrary assume that such a code exists, let $E[N]$ be its expected length, then, by the law of large number the total length of n successive transmissions is very likely to be less than $n[E[N] + \epsilon]$. Thus, a fixed length code of this length will achieve almost the same rate with small probability of error.

²We start counting time from the beginning of transmission for both receivers.

that, contrary to the setting of [2], there is no schedule that controls the sequence of channel outputs seen by the receivers.

It was observed by T. Cover in [4], that if we allow the receivers to decode at a different instant of time, the channel capacity of each link can be (simultaneously) achieved. This is done by sending the messages to each receiver in a separated period of time, and letting the ratio between the sizes of messages grow appropriately as they go to infinity. In our setting, this requires that the ratio between the average transmission time of each receiver to be infinite. The objective of this paper, is to characterize the region of achievable rates for bounded values of this ratio.

In the next section, we give the definition of a variable length code, along with the associated notions of reliability and capacity region. In section III we show an outer bound to the region of achievable rates. Then, in section IV, we analyze this bound and present examples of coding scheme that can achieve the outer bound.

II. DEFINITIONS

We consider sending independent information over a memoryless degraded broadcast channel. There are two independent sources, one producing a message $W_1 \in \{1, 2, \dots, M_1\}$ and the other producing a message $W_2 \in \{1, 2, \dots, M_2\}$. The channel consists of an input alphabet \mathcal{X} , two output alphabets \mathcal{Y} and \mathcal{Z} , and a probability transition function $p(y, z|x)$. By the memorylessness of the channel we have, for any n , $p(y^n, z^n|x^n) = \prod_{i=1}^n p(y_i, z_i|x_i)$, where $x^n \in \mathcal{X}^n$, $y^n \in \mathcal{Y}^n$ and $z^n \in \mathcal{Z}^n$.

Let N_1 be a stopping time w.r.t. (with respect to) $\{Y_i\}_{i \geq 1}$, the sequence of received values at the strong receiver. And, let N_2 be a stopping time w.r.t. $\{Z_i\}_{i \geq 1}$, the sequence of received values at the weak receiver. We define a (M_1, M_2, N_1, N_2) *variable length code* of rates $\frac{\log M_1}{E[N_1]}$ and $\frac{\log M_2}{E[N_2]}$, as a sequence of mappings $\{x_i(W_1, W_2)\}_{i \geq 1}$, where each x_i is a function of W_1 and W_2 , and two decoding functions (two decoders) w.r.t. the decoding times N_1 and N_2 ,³

$$g_1 : \mathcal{Y}^{N_1} \rightarrow \{1, 2, \dots, M_1\}$$

³For deterministic stopping rules, we can represent the set of all output sequences for which a decision is made, at each decoder, as the leaves of a complete $|\mathcal{Y}|$ -ary (resp. $|\mathcal{Z}|$ -ary) tree. The leaves have a label from the set of messages. Each decoder starts climbing the tree from the root. At each time it chooses the branch that corresponds to the received symbol. When a leaf is reached, the decoder makes a decision as indicated by the label of the leaf.

and

$$g_2 : \mathcal{Z}^{N_2} \rightarrow \{1, 2, \dots, M_2\}.$$

We define the average probability of error as the probability that the decoded messages are not equal to the transmitted messages, i.e.,

$$P_e = \Pr\{g_1(Y^{N_1}) \neq W_1 \text{ or } g_2(Z^{N_2}) \neq W_2\}$$

when (W_1, W_2) are assumed to be uniformly distributed over $\{1, 2, \dots, M_1\} \times \{1, 2, \dots, M_2\}$.

A rate pair (R_1, R_2) is said to be *achievable* if for all $\epsilon > 0$, there exists a (M_1, M_2, N_1, N_2) variable length code with $\frac{\log M_1}{E[N_1]} \geq R_1$, $\frac{\log M_2}{E[N_2]} \geq R_2$ and $P_e < \epsilon$.

The *capacity region* of the broadcast channel is the closure of the set of achievable rates. Notice that, with this definition, the capacity region is simply given by the rectangle $[0, C_1] \times [0, C_2]$, where $C_1 \triangleq \max_{p(x)} I(X; Y)$ is the (usual) channel capacity of the strong link, and $C_2 \triangleq \max_{p(x)} I(X; Z)$ is the channel capacity of the weak link.

As mentioned in the introduction, the argument presented in [4], that demonstrates the achievability of $[0, C_1] \times [0, C_2]$, requires making $E[N_1]/E[N_2]$ approach 0 (or infinity). Therefore, in the next section, we show a theorem that gives an outer bound on the capacity region, with a restriction on $E[N_1]$ and $E[N_2]$.

III. OUTER REGION

In order to obtain an outer bound on the capacity region we state two lemmas. First, let us define $N = \min(N_1, N_2)$. Observe that N is a stopping time w.r.t. $\{(Y_i, Z_i)\}_{i \geq 1}$, but it is not (in general) a stopping time w.r.t. individual channel outputs.

Lemma 3.1: The following inequalities hold:

$$\begin{aligned} I(W_2; Z^N) &\leq E[N]I(U; Z) + \log(eE[N]) \\ I(W_1; Y^N | W_2) &\leq E[N]I(X : Y | U) + \log(eE[N]), \end{aligned}$$

for some joint distribution $p(u)p(x|u)p(y, z|x)$.

Proof: Let $\mu_i = 1\{N \geq i\}$,⁴ from the chain rule for mutual information, we have

$$\begin{aligned} I(W_2; Z^N) &= I(W_2; Z_1\mu_1, \mu_1, \dots, Z_n\mu_n, \mu_n, \dots) \\ &= I(W_2; \mu_1) + I(W_2; Z_1\mu_1 | \mu_1) + \dots \\ &\quad + I(W_2; \mu_n | (Z\mu)^{n-1}, \mu^{n-1}) \\ &\quad + I(W_2; Z_n\mu_n | (Z\mu)^{n-1}, \mu^n) + \dots \\ &= \sum_{i=1}^{\infty} I(W_2; \mu_i | (Z\mu)^{i-1}, \mu^{i-1}) \\ &\quad + \sum_{i=1}^{\infty} I(W_2; Z_i\mu_i | (Z\mu)^{i-1}, \mu^i). \end{aligned}$$

⁴Where $1\{N \geq i\}$ is equal to 1 if $N \geq i$ and equal to 0 otherwise. Also, we define $Z_i\mu_i$ as being equal to Z_i if $N \geq i$ and equal to 0 otherwise.

The first summation can be upper bounded as

$$\begin{aligned} \sum_{i=1}^{\infty} I(W_2; \mu_i | (Z\mu)^{i-1}, \mu^{i-1}) &\leq \sum_{i=1}^{\infty} H(\mu_i | \mu^{i-1}) \\ &= H(\mu_1, \mu_2, \dots) \\ &= H(N) \\ &\leq \log(eE[N]), \end{aligned}$$

where the last inequality is proved in [1] and [5, §3].

For the second summation, we get

$$\begin{aligned} I(W_2; Z_i\mu_i | (Z\mu)^{i-1}, \mu^i) &= H(Z_i\mu_i | (Z\mu)^{i-1}, \mu^i) - H(Z_i\mu_i | (Z\mu)^{i-1}, \mu^i, W_2) \\ &\leq H(Z_i\mu_i | \mu_i) - H(Z_i\mu_i | (Y\mu)^{i-1}, (Z\mu)^{i-1}, \mu^i, W_2) \\ &= H(Z_i\mu_i | \mu_i) - H(Z_i\mu_i | (Y\mu)^{i-1}, (Z\mu)^{i-1}, \mu_i, W_2) \\ &= \Pr(\mu_i = 1) [H(Z_i | \mu_i = 1) \\ &\quad - H(Z_i | Y^{i-1}, Z^{i-1}, \mu_i = 1, W_2)] \\ &= \Pr(N \geq i) I(U_i; Z_i | \mu_i = 1), \end{aligned}$$

where $U_i = \{Y^{i-1}, Z^{i-1}, W_2\}$.

Observe that $p(z_i | x_i, \mu_i = 1) = p(z_i | x_i)$, thus, with a slight abuse of notation, we can write $I(U_i; Z_i | \mu_i = 1) = I(U_i; Z_i)$, with $p(u_i) \triangleq p(u_i | \mu_i = 1)$ and $p(x_i | u_i) \triangleq p(x_i | u_i, \mu_i = 1)$. Hence, we obtain

$$\begin{aligned} I(W_2; Z_i\mu_i | (Z\mu)^{i-1}, \mu^i) &\leq \sum_{i=1}^{\infty} \Pr(N \geq i) I(U_i; Z_i) \\ &= E[N] \sum_{i=1}^{\infty} \frac{\Pr(N \geq i)}{E[N]} I(U_i; Z_i). \end{aligned}$$

Now let $a_i = \frac{\Pr(N \geq i)}{E[N]}$, note that $a_i \geq 0$ for all i , and $\sum_i a_i = 1$. Thus, we can define an integer random variable Q by setting $\Pr(Q = i) = a_i$, for all $i \in \{1, 2, \dots\}$. Using this, the preceding equation becomes

$$\begin{aligned} I(W_2; Z_i\mu_i | (Z\mu)^{i-1}, \mu^i) &\leq E[N] \sum_{i=1}^{\infty} \Pr(Q = i) I(U_Q; Z_Q | Q = i) \\ &= E[N] I(U_Q; Z_Q | Q) \\ &\leq E[N] I(U; Z), \end{aligned}$$

where $U \triangleq \{U_Q, Q\}$ and $Z \triangleq Z_Q$ are new random variables, the distributions of U_Q and Z_Q depend on Q in the same way as the distributions of U_i and Z_i depend on i .

Then, consider

$$\begin{aligned}
I(W_1; Y^N | W_2) &= I(W_1; Y_1 \mu_1, \mu_1, \dots, Y_n \mu_n, \mu_n, \dots | W_2) \\
&= I(W_1; \mu_1 | W_2) + I(W_1; Y_1 \mu_1 | \mu_1, W_2) + \dots \\
&\quad + I(W_1; \mu_n | (Y \mu)^{n-1}, \mu^{n-1}, W_2) \\
&\quad + I(W_1; Y_n \mu_n | (Y \mu)^{n-1}, \mu^n, W_2) + \dots \\
&= \sum_{i=1}^{\infty} I(W_1; \mu_i | (Y \mu)^{i-1}, \mu^{i-1}, W_2) \\
&\quad + \sum_{i=1}^{\infty} I(W_1; Y_i \mu_i | (Y \mu)^{i-1}, \mu^i, W_2).
\end{aligned}$$

As previously, we may upper bound the first summation as

$$\begin{aligned}
\sum_{i=1}^{\infty} I(W_1; \mu_i | (Y \mu)^{i-1}, \mu^{i-1}, W_2) &\leq \sum_{i=1}^{\infty} H(\mu_i | \mu^{i-1}) \\
&= H(\mu_1, \mu_2, \dots) \\
&= H(N) \\
&\leq \log(eE[N]).
\end{aligned}$$

And, the i th term in the second summation is

$$\begin{aligned}
&I(W_1; Y_i \mu_i | (Y \mu)^{i-1}, \mu^i, W_2) \\
&= H(Y_i \mu_i | (Y \mu)^{i-1}, \mu^i, W_2) \\
&\quad - H(Y_i \mu_i | (Y \mu)^{i-1}, \mu^i, W_1, W_2) \\
&\stackrel{(a)}{=} H(Y_i \mu_i | (Y \mu)^{i-1}, (Z \mu)^{i-1}, \mu^i, W_2) \\
&\quad - H(Y_i \mu_i | (Y \mu)^{i-1}, (Z \mu)^{i-1}, \mu^i, W_1, W_2) \\
&\stackrel{(b)}{=} H(Y_i \mu_i | (Y \mu)^{i-1}, (Z \mu)^{i-1}, \mu_i, W_2) \\
&\quad - H(Y_i \mu_i | (Y \mu)^{i-1}, (Z \mu)^{i-1}, X_i \mu_i, \mu_i, W_2) \\
&= \Pr(\mu_i = 1) [H(Y_i | Y^{i-1}, Z^{i-1}, \mu_i = 1, W_2) \\
&\quad - H(Y_i | Y^{i-1}, Z^{i-1}, X_i, \mu_i = 1, W_2)] \\
&= \Pr(N \geq i) I(X_i; Y_i | U_i, \mu_i = 1),
\end{aligned}$$

where in (a) we use the fact that the channel is degraded, and (b) follows since X_i is a function of $(W_1, W_2, Y^{i-1}, Z^{i-1})$, and then given X_i , Y_i is independent of W_1 .

Here, following the steps done for the preceding inequality, we let $p(u_i) \triangleq p(u_i | \mu_i = 1)$ and $p(x_i | u_i) \triangleq p(x_i | u_i, \mu_i = 1)$, then since $p(y_i | x_i, u_i, \mu_i = 1) = p(y_i | x_i)$, we have $I(X_i; Y_i | U_i, \mu_i = 1) = I(X_i; Y_i | U_i)$. Hence, we get

$$\begin{aligned}
&I(W_1; Y_i \mu_i | (Y \mu)^{i-1}, \mu^i, W_2) \\
&\leq \sum_{i=1}^{\infty} \Pr(N \geq i) I(X_i; Y_i | U_i) \\
&= E[N] \sum_{i=1}^{\infty} \frac{\Pr(N \geq i)}{E[N]} I(X_i; Y_i | U_i).
\end{aligned}$$

Now, introducing the random variable Q defined earlier, we may write

$$\begin{aligned}
&I(W_1; Y_i \mu_i | (Y \mu)^{i-1}, \mu^i, W_2) \\
&\leq E[N] \sum_{i=1}^{\infty} \Pr(Q = i) I(X_Q; Y_Q | U_Q, Q = i) \\
&= E[N] I(X; Y | U),
\end{aligned}$$

where $U \triangleq \{U_Q, Q\}$, and $X \triangleq X_Q$ and $Y \triangleq Y_Q$ are new random variables, whose distributions depend on Q in the same way as the distributions of X_i and Y_i depend on i . Notice that $U \rightarrow X \rightarrow (Y, Z)$ forms a Markov chain. Therefore, we obtain

$$\begin{aligned}
I(W_2; Z^N) &\leq E[N] I(U; Z) + \log(eE[N]) \\
I(W_1; Y^N | W_2) &\leq E[N] I(X; Y | U) + \log(eE[N]),
\end{aligned}$$

for some joint distribution $p(u)p(x|u)p(y, z|x)$. ■

We state the next lemma without proof, the main ideas being presented in the previous lemma.

Lemma 3.2: We have the following inequalities:

$$\begin{aligned}
I(W_1; Y_{N+1}^{N_1} | Y^N, W_2) &\leq E[N_1 - N] C_1 \\
&\quad + \log(eE[N_1 - N]) \\
I(W_2; Z_{N+1}^{N_2} | Z^N) &\leq E[N_2 - N] C_2 \\
&\quad + \log(eE[N_2 - N]).
\end{aligned}$$

The following theorem shows an outer bound on the region of achievable rates.

Theorem 3.3: (Outer bound) Let us denote by \mathcal{C}_{r_1, r_2} , the set of rates achievable by using variable length codes for which $\frac{E[N]}{E[N_1]} \geq r_1$ and $\frac{E[N]}{E[N_2]} \geq r_2$. Then, any rate pairs $(R_1, R_2) \in \mathcal{C}_{r_1, r_2}$ must satisfy

$$\begin{aligned}
R_1 &\leq r_1 I(X; Y | U) + (1 - r_1) C_1 \\
R_2 &\leq r_2 I(U; Z) + (1 - r_2) C_2,
\end{aligned}$$

for some joint distribution $p(u)p(x|u)p(y, z|x)$, with the cardinality of the auxiliary random variable bounded by $|\mathcal{U}| \leq \min(|\mathcal{X}|, |\mathcal{Y}|, |\mathcal{Z}|)$.

Proof: Let W_i be uniformly distributed over $\{1, 2, \dots, M_i\}$, $i = 1, 2$. Then,

$$\begin{aligned}
I(W_2; Z^{N_2}) &= H(W_2) - H(W_2 | Z^{N_2}) \\
&= E[N_2] R_2 - H(W_2 | Z^{N_2}),
\end{aligned}$$

and

$$\begin{aligned}
I(W_1; Y^{N_1} | W_2) &= H(W_1 | W_2) - H(W_1 | Y^{N_1}, W_2) \\
&\geq E[N_1] R_1 - H(W_1 | Y^{N_1}).
\end{aligned}$$

Thus, using Fano's inequality, we have

$$\begin{aligned}
E[N_2](R_2 - \epsilon) &\leq I(W_2; Z^{N_2}) \\
E[N_1](R_1 - \epsilon) &\leq I(W_1; Y^{N_1} | W_2),
\end{aligned}$$

where $\epsilon \rightarrow 0$ as $P_e \rightarrow 0$.

From the chain rule for mutual information, we can write

$$I(W_2; Z^{N_2}) = I(W_2; Z^N) + I(W_2; Z_{N+1}^{N_2} | Z^N),$$

and

$$I(W_1; Y^{N_1} | W_2) = I(W_1; Y^N | W_2) + I(W_1; Y_{N+1}^{N_1} | Y^N, W_2).$$

Then, applying Lemma 3.1 and Lemma 3.2, we get

$$\begin{aligned} I(W_2; Z^{N_2}) &\leq E[N]I(U; Z) + E[N_2 - N]C_2 \\ &\quad + \log(eE[N]) + \log(eE[N_2 - N]) \\ I(W_1; Y^{N_1} | W_2) &\leq E[N]I(X; Y | U) + E[N_1 - N]C_1 \\ &\quad + \log(eE[N]) + \log(eE[N_1 - N]), \end{aligned}$$

for some joint distribution $p(u)p(x|u)p(y, z|x)$.

Hence,

$$\begin{aligned} E[N_1](R_1 - \epsilon) &\leq E[N]I(X; Y | U) + E[N_1 - N]C_1 \\ &\quad + \log(eE[N]) + \log(eE[N_1 - N]) \\ E[N_2](R_2 - \epsilon) &\leq E[N]I(U; Z) + E[N_2 - N]C_2 \\ &\quad + \log(eE[N]) + \log(eE[N_2 - N]), \end{aligned}$$

for some joint distribution $p(u)p(x|u)p(y, z|x)$.

Dividing by $E[N_1]$ and letting $E[N_1] \rightarrow \infty$ in the first inequality, and dividing by $E[N_2]$ and letting $E[N_2] \rightarrow \infty$ in the second one, we obtain ⁵

$$\begin{aligned} R_1 &\leq r_1 I(X; Y | U) + (1 - r_1)C_1 \\ R_2 &\leq r_2 I(U; Z) + (1 - r_2)C_2, \end{aligned}$$

for some joint distribution $p(u)p(x|u)p(y, z|x)$. The cardinality bounds for the auxiliary random variable U can be derived using standard methods from convex set theory. ■

Notice that the region \mathcal{C}_{r_1, r_2} is defined for variable length coding schemes verifying the restrictions on $E[N_1]$, $E[N_2]$ and $E[N]$, given by r_1 and r_2 . The next section presents some important conclusion that can be derived from the outer bound on \mathcal{C}_{r_1, r_2} .

IV. ANALYSIS AND CODING SCHEMES

To get a better insight into the meaning of the outer region found in the previous section, we focus on coding schemes that have $E[N] = E[N_1]$. ⁶ In this case the outer region is given by the set of all rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq I(X; Y | U) \\ R_2 &\leq \frac{E[N_1]}{E[N_2]} I(U; Z) + (1 - \frac{E[N_1]}{E[N_2]})C_2, \end{aligned}$$

⁵Codes achieving low probability of error may exist without requiring $E[N_1] \rightarrow \infty$ or $E[N_2] \rightarrow \infty$, but this is needed for codes achieving arbitrary low probability of error.

⁶This means that the strong receiver always makes a decision before the weak one.

for some joint distribution $p(u)p(x|u)p(y, z|x)$. ⁷

Note that this region is composed of the usual (block code) capacity region for degraded broadcast channels \mathcal{R}_{DBC} , with a scaling factor on R_2 , plus a fraction of C_2 . It turns out that any rate pairs in this region can be achieved, by choosing $E[N_1]$ and $E[N_2]$ large enough, and using a block code of length $E[N_1]$ achieving the corresponding rate pair in \mathcal{R}_{DBC} , ⁸ followed by a capacity achieving code of length $E[N_2] - E[N_1]$, for the link to the weakest receiver. Note that, the number of messages for the second receiver behaves like $\log M_2 = E[N_1]R_2' + (E[N_2] - E[N_1])C_2$, where $R_2' \in \mathcal{R}_{DBC}$ is the rate at which the codewords for the weak receiver are generated, during the transmission of the first block code.

This shows that among all coding schemes with $E[N] = E[N_1]$, the best one is composed of two successive block codes. Thus, the possibility to employ variable length coding gives no real improvement over block coding. The gain in the achievable rates comes from the fact that the receivers decode their message at a different instant of time. The same conclusion can be derived if $E[N] = E[N_2]$.

Concerning coding schemes with an arbitrary $E[N]$, the best outer bound is obtained by minimizing $E[N]$. We know that with high probability $N_1 \geq \frac{\log M_1}{C_1}$ and $N_2 \geq \frac{\log M_2}{C_2}$, thus we may write $E[N] \geq \min(\frac{\log M_1}{C_1}, \frac{\log M_2}{C_2})$. In the special case when $\frac{\log M_1}{C_1} = \frac{\log M_2}{C_2}$, we can rewrite the outer bound as

$$\begin{aligned} R_1 &\leq \frac{C_1}{2 - \frac{I(X; Y | U)}{C_1}} \\ R_2 &\leq \frac{C_2}{2 - \frac{I(U; Z)}{C_2}}, \end{aligned}$$

for some joint distribution $p(u)p(x|u)p(y, z|x)$. ⁹ Observe that, this outer bound is valid for codes having $E[N_1]$ and $E[N_2]$ verifying the condition on their ratio, implicitly given by $\frac{\log M_1}{\log M_2} = \frac{C_1}{C_2}$.

Now, consider the following random coding scheme. Generate two capacity achieving (block) codes, one for the strong link C_1 of length $\frac{\log M_1}{C_1}$, and one for the weak link C_2 of length $\frac{\log M_2}{C_2}$. Then, to transmit a message pair $(w_1, w_2) \in (W_1, W_2)$, with probability p send the codeword in C_1 corresponding to w_1 , followed by the codeword in C_2 corresponding to w_2 . And, with probability $1 - p \triangleq \bar{p}$ send it in the reverse order (C_2 followed by C_1). For M_1 and M_2 large enough, this coding scheme achieve the following rates

$$\begin{aligned} R_1' &= \frac{\log M_1}{\frac{\log M_1}{C_1} + \bar{p}\frac{\log M_2}{C_2}} = \frac{C_1}{1 + \bar{p}\frac{C_1 \log M_2}{C_2 \log M_1}} \\ R_2' &= \frac{\log M_2}{\frac{\log M_2}{C_2} + p\frac{\log M_1}{C_1}} = \frac{C_2}{1 + p\frac{C_2 \log M_1}{C_1 \log M_2}}, \end{aligned}$$

with $p \in [0, 1]$.

⁷From now on, we omit to mention the cardinality bound on the auxiliary random variable U .

⁸For a careful definition and analysis of block codes and broadcast channels, the reader is referred to [6] and the references therein.

⁹This region is not necessarily convex.

Assuming that $\frac{\log M_1}{C_1} = \frac{\log M_2}{C_2}$, we have that the region of all rate pairs (R_1, R_2) , satisfying

$$R_1 = \frac{C_1}{1+p}$$

$$R_2 = \frac{C_2}{1+p},$$

for some $p \in [0, 1]$, is achievable. This region can be related to the outer region in the special case given previously. We see that the corner points of this outer region are achieved. Furthermore, if $C_1 = C_2$ (the statistics over each link are the same), the two regions coincide.

Again, in this example, the improvement of using variable length codes is only apparent through the possibility of sending the messages at different periods of time.

V. REMARKS AND CONCLUSION

For variable length coding over a degraded broadcast channel, we introduced a new notion of capacity region with a receiver centric definition of the transmission rates. We derived an outer bound on this region, that capture the variability in the receiver decoding times.

Through examples of coding scheme, we motivated that the gain in using variable length codes essentially comes from the possibility for the receivers to decode at a different instant of time.

The setup of this paper can be extended to allow an immediate and noiseless feedback from the receivers to the transmitter. In the case when the degradation is physical, the outer bound remains valid.¹⁰ This gives an equivalent to the result in [7], when the encoder is able to use variable length codes.

ACKNOWLEDGMENT

The author wishes to thank Emre Telatar for insightful discussions and helpful comments.

REFERENCES

- [1] I. Csiszár, "On the capacity of noisy channels with arbitrary signal costs," *Problems of Control and Information Theory*, Vol. 2 (2-4), pp. 283-304, 1973.
- [2] S. Shamai, E. Telatar and S. Verdú, "Fountain Capacity," in *Proc. IEEE Int. Symp. Inf. Theory*, pp. 1881-1884, Seattle, 2006.
- [3] A. Tchamkerten, "Feedback Communication over Unknown Channels," Ph.D. Thesis, Information Theory Laboratory, EPFL, March 2005.
- [4] T. Cover, "Comments on Broadcast Channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2524-2530, October 1998.
- [5] I. Csiszár and J. Körner, "Information Theory: Coding Theorems for Discrete Memoryless Systems," Academic Press, New York, 1981.
- [6] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 2006.
- [7] A. El Gamal, "The Feedback Capacity of Degraded Broadcast Channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 379-381, May 1978.

¹⁰Each x_i becomes dependent of Y^{i-1} and Z^{i-1} , and the proof of theorem 3.3 follows with minor changes.