Outer Bounds on the Admissible Source Region for Broadcast Channels with Dependent Sources

Gerhard Kramer  
Dept. of Electrical Engineering  
University of Southern California  
Los Angeles, CA 90089  
gkramer@usc.edu

Yingbin Liang  
Dept. of Electrical Engineering  
University of Hawaii  
Honolulu, HI 96822  
yingbinl@hawaii.edu

Shlomo Shamai (Shitz)  
Dept. of Electrical Engineering  
Technion - Israel Institute of Technology  
Technion City, Haifa 32000, Israel  
sshlomo@ee.technion.ac.il

Abstract— Outer bounds on the admissible source region for broadcast channels with dependent sources are developed and used to prove capacity results for several classes of sources and channels.

I. INTRODUCTION

Consider a two-receiver broadcast channel (BC), say \( P_{YZ|X}(\cdot) \), and a discrete memoryless source \((S, T)\) with finite alphabet \(S \times T\). Source \((S, T)\) is said to be admissible for this BC if for any \(\lambda, 0 < \lambda < 1\), and for large enough \(n\) there is a code with length-\(n\) codewords such that \(P_{e1} \leq \lambda\) and \(P_{e2} \leq \lambda\), where \(P_{e1}\) and \(P_{e2}\) are the respective error probabilities for receivers 1 and 2. The set of all admissible sources is called the admissible source region.

Han and Costa [1] developed a coding strategy that admits certain sources. Let \(K = f(S) = g(T)\) be the common variable in the sense of Gacs and Körner (and also Witsenhausen), and consider auxiliary random variables \(W, U, V\) that satisfy the Markov chain property

\[
ST - WUV - X - YZ. \tag{1}
\]

Then the source \((S, T)\) is admissible if (see [1], [2])

\[
H(S) < I(SWU[Y]; Y) - I(T; WU|S) \tag{2}
\]

\[
H(T) < I(TWV; Z) - I(S; WV|T) \tag{3}
\]

\[
H(ST) < \min\{I(KW; Y), I(KW; Z)\} + I(SU; YKW) + I(TV; Z|KW) - I(SU; TV|KW) \tag{4}
\]

\[
H(ST) < I(SWUV; Y) + I(TWV; Z) - I(SU; TV|KW) - I(ST; KW). \tag{5}
\]

Our paper is concerned with "outer bounds" on the set of admissible sources, i.e., we wish to determine a superset of the admissible source region. To do so, we borrow tools from [3] and modify them to include dependent sources. We remark that a more general class of outer bounds for BCs is presented in [4]. Nevertheless, we use the approach of [3] because the auxiliary random variables are simpler and they can be used to prove capacity theorems. For example, it is not obvious how to choose the auxiliary random variables of [4] to even determine the capacity region of degraded broadcast channels.

II. OUTER BOUNDS ON THE ADMISSIBLE SOURCE REGION

Consider the source sequences \(S^m, T^m\) where \(m\) may be different than the code length \(n\). Let \(R = n/m\) be the source-channel "bandwidth expansion." Further let the auxiliary random variables \((K, S, T)\) have the same distribution as \((K^m, S^m, T^m)\) and let them form the Markov chain

\[
KSTUV - X - YZ. \tag{6}
\]

We may also let \(X\) be a deterministic function of \((S, T)\). We have the following outer bound.

**Theorem 1:** An admissible source \((S, T)\) satisfies the following bounds for some Markov chain [4]:

\[
H(K)/R \leq \min\{I(K; Y|U), I(K; Z|V)\} \tag{7}
\]

\[
H(S)/R \leq I(\tilde{S}; Y|U) \tag{8}
\]

\[
H(T)/R \leq I(\tilde{T}; Z|V) \tag{9}
\]

\[
H(ST)/R \leq I(\tilde{S}; Y|\tilde{T}UV) + I(\tilde{T}U; Z|V) \tag{10}
\]

\[
H(ST)/R \leq I(\tilde{T}; Z|\tilde{S}UV) + I(\tilde{S}V; YU) \tag{11}
\]

\[
H(ST)/R \leq I(\tilde{K}U; Y) \tag{12}
\]

\[
+ I(\tilde{S}; Y|\tilde{T}UV) + I(\tilde{T}; Z|\tilde{K}U). \tag{13}
\]

**Remark 1:** Consider the case \(R = 1, K, S, T\) statistically independent, \(KST\) statistically independent of \(UV\), and \((H(K), H(S), H(T)) = (R_0, R_1, R_2)\). The bounds \(7-13\) are then a subset of those given in [3, Thm. 1].

**Theorem 2:** A weaker set of conditions than \(7-13\) is as follows, where \(Q = (K, U, V)\):

\[
H(K)/R \leq \min\{I(Q; Y), I(Q; Z)\} \tag{14}
\]

\[
H(S)/R \leq I(Q\tilde{S}; Y) \tag{15}
\]

\[
H(T)/R \leq I(Q\tilde{T}; Z) \tag{16}
\]

\[
H(ST)/R \leq I(\tilde{S}; Y|Q\tilde{T}) + I(Q\tilde{T}; Z) \tag{17}
\]

\[
H(ST)/R \leq I(\tilde{T}; Z|Q\tilde{S}) + I(Q\tilde{S}; Y) \tag{18}
\]

\[
H(ST)/R \leq I(Q; Y) + I(\tilde{S}; Y|Q\tilde{T}) + I(\tilde{T}; Z|Q) \tag{19}
\]

\[
H(ST)/R \leq I(Q; Z) + I(\tilde{T}; Z|Q\tilde{S}) + I(\tilde{S}; Y|Q). \tag{20}
\]

**Remark 2:** Consider the case described in Remark 1. The bounds \(14-20\) are then identical to those in [3, Thm. 3].
III. EXAMPLES

We continue by considering only $R = 1$ for simplicity.

A. Separating Source and Channel Coding

A natural approach is to decouple $ST$ from $WUV$, i.e., choose $WUV$ independent of $ST$. Effectively, we convert the source strings to (compressed) bit strings and apply channel coding. The bounds \((23) - (25)\) are then

\[
\begin{align*}
H(S) &< I(WU;Y) \\
H(T) &< I(WV;Z) \\
H(ST) &< \min\{I(W;Y), I(W;Z)\} + I(U;Y|W) \\
&+ I(V;Z|W) - I(U;V|W) - I(S;T|K) \\
H(ST) &< I(WU;Y) + I(WV;Z) - I(U;V|W) \\
&- I(S;T|K) - I(ST;K).
\end{align*}
\]

(B) Markov Sources

Consider the case where $S - K - T$ forms a Markov chain, that is $I(S;T|K) = 0$. This situation seems to be effectively the classic case where there are three independent messages $W_0, W_1, W_2$ with $nR_0, nR_1, nR_2$ bits, respectively. In fact, for any BC for which we know the capacity region, we can match the outer bound of Theorem \ref{thm:source_region} to the admissible bounds \((21) - (24)\). For example, consider a semi-deterministic BC where $Y$ is a deterministic function of $X$. We have the following result.

**Theorem 3:** A source for which $S - K - T$ forms a Markov chain is admissible for a semi-deterministic BC where $Y$ is a deterministic function of $X$ if $ST$ satisfies

\[
\begin{align*}
H(K) &< \min\{I(W;Y), I(W;Z)\} \\
H(S) &< H(Y) \\
H(T) &< I(WV;Z) \\
H(ST) &< H(Y|WV) + I(WV;Z) \\
H(ST) &< I(W;Y) + H(Y|WV) + I(V;Z|W)
\end{align*}
\]

for some Markov chain $ST - WV - X - YZ$. Conversely, such a source is not admissible for such a BC if for every Markov chain $ST - WV - X - YZ$ the direction of one of the (strict) inequalities is reversed.

**Proof:** For admissibility, choose $U = Y$ and observe that the bounds \((25) - (28)\) imply the bounds \((21) - (24)\). For the converse, consider \((14) - (17)\) and \((19)\), identify $Q, K, T$ with $W, K, V$, respectively, and use

\[
\begin{align*}
I(Q\tilde{S};Y) &\leq H(Y) \\
I(S;Y|Q\tilde{T}) &\leq H(Y|Q\tilde{T}).
\end{align*}
\]

C. More Capable BCs

Theorem \ref{thm:source_region} does not give the admissible source region since it requires $S - K - T$ to form a Markov chain. In fact, Han and Costa showed in [1, Example 2] that separating source and channel coding (implied by decoupling $ST$ and $WUV$) is suboptimal in general. The decoupling also doesn’t seem to work for the expressions \((2) - (5)\) for an important class of channels where one might guess that it should.

Consider the class of more-capable BCs defined by the constraint that

\[
I(X;Y) \geq I(X;Z) \text{ for all } P_x.
\]

This class includes both physically and stochastically degraded BCs. If we choose $WUV$ independent of $ST$ then \((21) - (24)\) still exhibit the rate loss $I(S;T|K)$ in \((23) - (24)\). However, choosing $W$ differently we have the following result.

**Theorem 4:** The admissible source region for more-capable BCs is defined by the bounds

\[
\begin{align*}
H(ST) &\leq I(X;Y) \\
H(T) &\leq I(\tilde{W};Z) \\
H(ST) &\leq I(X;Y|\tilde{W}) + I(\tilde{W};Z)
\end{align*}
\]

for some Markov chain $ST - \tilde{W} - X - YZ$.

**Proof:** We choose

\[
\begin{align*}
W &= T\tilde{W} \text{ where } ST \text{ is independent of } \tilde{W}UV \\
U &= X = \text{ a noisy function of } \tilde{W} \\
V &= \text{ a constant.}
\end{align*}
\]

The bounds \((2) - (5)\) are then simply \((33) - (35)\). Furthermore, for any Markov chain $ST - W - X - YZ$ we can achieve the right-hand sides of \((33) - (35)\) since these depend on $ST$ only through $W$. For the converse, consider \((16) - (18)\) and identify $Q\tilde{T}, \tilde{S}$ with $\tilde{W}, U$, respectively. The bounds \((33) - (35)\) follow from \((16) - (18)\) and the bounds

\[
\begin{align*}
I(\tilde{T};Z|Q\tilde{S}) &\leq I(X;Z|Q\tilde{S}) \\
&= \sum P_{Q\tilde{S}}(ab)I(X;Z|Q\tilde{S} = ab) \\
&\leq \sum P_{Q\tilde{S}}(ab)I(X;Y|Q\tilde{S} = ab) \\
&= I(X;Y|Q\tilde{S})
\end{align*}
\]

where the second inequality follows by applying \((32)\). 

We remark that the choice $W = TW$ does, in fact, permit separating source and channel coding because $\tilde{W}$ is independent of $T$. The coding approach is to simply compress $T$ to its entropy-rate $H(T)$ and consider the resulting bits as a common message. Next, compress $S$ to the conditional entropy-rate $H(S|T)$ and consider the resulting bits as a private message. Decoder $Y$ first decodes and decompresses $T$ and then decodes and decompresses $S$. Decoder $Z$ decodes and decompresses $T$ only. This natural approach is included in the Han-Costa coding method, but only indirectly.

IV. PROOF OF THEOREM \ref{thm:source_region}

Let $|S|$ be the cardinality of the set $S$ and let $(\hat{K}, \hat{S}, \hat{T}) = (K^m, S^m, T^m)$. Fano’s inequality [5, Sec. 2.11] gives

\[
\begin{align*}
H(K^m|Y^n) &\leq H(S^m|Y^n) \leq P_{e_1} \cdot m \log_2 |S| + 1 \\
H(K^m|Z^n) &\leq H(T^m|Z^n) \leq P_{e_2} \cdot m \log_2 |T| + 1
\end{align*}
\]
Let $\delta_1 = P_{e1} \log_2 |S| + 1/m$ and $\delta_2 = P_{e2} \log_2 |S| + 1/m$, and observe that reliable communication ($P_{e1} \to 0$ and $P_{e2} \to 0$) means that $\delta_1 \to 0$ and $\delta_2 \to 0$ as $m \to \infty$. We define the following auxiliary random variables

$$U_i = Y^{i-1}, \quad V_i = Z^n_{i+1}$$

$$P_i(i) = \frac{1}{n}, \quad i = 1, 2, \ldots, n$$

$$U = (U_i, I), \quad V = (V_i, I)$$

$$X = X_i, \quad Y = Y_i, \quad Z = Z$$

and observe that we have the Markov chain

$$KSTUV = X - YZ.$$  \hspace{1cm} (46)

Furthermore, the definition (45) of $XYZ$ is consistent with our BC in the sense that $P_{Y|X}(\cdot)$ is the same in Sec. [I]. Fano’s inequality (30) implies

$$m[H(K) - \delta_1] \leq H(K^m) - H(K^m|Y^n)$$

$$= I(K^m; Y^n)$$

$$= \sum_{i=1}^n I(K^m; Y_i|Y^{i-1})$$

$$= \sum_{i=1}^n I(\tilde{K}; Y_i|U_i)$$

$$= n I(\tilde{K}; Y|U)$$

where the last step follows from Lemma 1. Continuing, we have

$$m[H(ST) - \delta_1 - \delta_2] \leq \sum_{i=1}^n \left[ I(S^n; Y_i|T^n Y^{i-1}) + I(T^n Y^{i-1}; Z_i|Z^n_{i+1}) - I(Y_i; Z^n_{i+1}|T^n Y^{i-1}) \right]$$

$$= \sum_{i=1}^n \left[ -H(Y_i|T^n S^n Y^{i-1}) + I(T^n Y^{i-1}; Z_i|Z^n_{i+1}) + H(Y_i|T^n Y^{i-1} Z^n_{i+1}) \right]$$

$$\leq \sum_{i=1}^n \left[ -H(Y_i|T^n S^n Y^{i-1} Z^n_{i+1}) + I(T^n Y^{i-1}; Z_i|Z^n_{i+1}) + H(Y_i|T^n Y^{i-1} Z^n_{i+1}) \right]$$

$$= \sum_{i=1}^n \left[ I(\tilde{S}; Y_i|\tilde{T}U_i) + I(\tilde{T}U_i; Z_i|V_i) \right]$$

$$= n \left[ I(\tilde{S}; Y|\tilde{T}UV) + I(\tilde{T}U; Z|V) \right].$$

Similar arguments give

$$m[H(ST) - \delta_1 - \delta_2] \leq n \left[ I(\tilde{T}; Z|\tilde{S}UV) + I(\tilde{S}V; Y|U) \right].$$  \hspace{1cm} (68)

Next, consider the following steps:

$$m[H(S) + H(T) - \delta_1 - \delta_2]$$

$$\leq I(S^n; Y^n) + I(T^n; Z^n)$$

$$\leq I(K^m; Y^n) + I(S^n; Y^n|K^m) + I(K^m; T^n) + I(T^n; Z^n|K^m)$$

$$\leq I(K^m; Y^n) + I(S^n; T^n|K^m) + I(S^n; Y^n|K^m T^n) + I(K^m; T^n) + I(T^n; Z^n|K^m)$$

$$\leq I(K^m S^n T^n; Y^n) - I(T^n; Y^n|K^m) + I(T^n; Z^n|K^m)$$

$$+ m(I(S; T|K) + H(K)).$$  \hspace{1cm} (73)

The reader can check that

$$I(S; T|K) + H(K) = I(S; T).$$  \hspace{1cm} (74)

The first expression in (73) is bounded as

$$I(K^m S^n T^n; Y^n) = \sum_{i=1}^n I(K^m S^n T^n; Y_i|Y^{i-1})$$

$$= \sum_{i=1}^n I(\tilde{K}S\tilde{T}; Y_i|U_i)$$

$$\leq \sum_{i=1}^n I(\tilde{K}S\tilde{T}U_i; Y_i).$$  \hspace{1cm} (77)
The second and third expressions in (73) can be manipulated as follows:

\[-I(T_m^n; Y^n | K^m) + I(T_m^n; Z^n | K^m)\] (78)

\[= \sum_{i=1}^{n} \left[ -I(T_m^n; Y_i | K^m Y^{i-1}) + I(T_m^n; Z_i | K^m Z^{i+1}_i) \right] \] (79)

\[= \sum_{i=1}^{n} \left[ -I(T_m^n Z^n_i; Y_i | K^m Y^{i-1}) + I(Z^n_i; Y_i | K^m T^n M Y^{i-1}) + I(T_m^n Y^{i-1}; Z_i | K^m Z^{i+1}_i) - I(Y^{i-1}; Z_i | K^m T^n M Z^{i+1}_i) \right] \] (80)

\[= \sum_{i=1}^{n} \left[ -I(T_m^n Z^n_i; Y_i | K^m Y^{i-1}) + I(T_m^n Y^{i-1}; Z_i | K^m Z^{i+1}_i) \right] \] (81)

where the last step follows from Lemma 2. Continuing, we have

\[-I(T_m^n; Y^n | K^m) + I(T_m^n; Z^n | K^m)\] (82)

\[= \sum_{i=1}^{n} \left[ -I(Z^n_i; Y_i | K^m Y^{i-1}) - I(T_m^n; Y_i | K^m Y^{i-1}, Z^n_{i+1}) + I(Z^n_i; Y_i | K^m Z^{i+1}_i) + I(T_m^n; Z_i | K^m Z^{i+1}_i Y^{i-1}) \right] \] (83)

\[= \sum_{i=1}^{n} \left[ -I(T_m^n; Y_i | K^m Y^{i-1} Z^n_{i+1}) + I(T_m^n; Z_i | K^m Z^{i+1}_i Y^{i-1}) \right] \] (84)

\[= \sum_{i=1}^{n} \left[ -I(T_m^n; Y_i | K^m U_i V_i) + I(T_m^n; Z_i | K^m U_i V_i) \right] \] (85)

where the second last step follows from Lemma 2.

We substitute (74), (77), and (85) into (73) and obtain

\[m[H(ST) - \delta_1 - \delta_2] \leq n \left[ I(\tilde{K} U V; Y) + I(\tilde{S}; Y | \tilde{T} U V) + I(\tilde{T}; Z | \tilde{K} U V) \right] \] (91)

REFERENCES

[1] T. S. Han and M. H. M. Costa, “Broadcast channels with arbitrarily correlated sources,” IEEE Trans. Inf. Theory, vol. 33, no. 5, pp. 641–650, Sep. 1987.

[2] G. Kramer and C. Nair, “Comments on "Broadcast channels with arbitrarily correlated sources",” in Proc. 2009 IEEE Int. Symp. Inf. Theory, Seoul, South Korea, June 28 - July 3 2009.

[3] Y. Liang, G. Kramer, and S. S. (Shitz), “Capacity outer bounds for broadcast channels,” in Proc. 2008 IEEE Inf. Theory Workshop, Porto, Portugal, May 5-9 2008, pp. 2–4.

[4] C. Nair, “An outer bound for 2-receiver discrete memoryless broadcast channels,” arXiv.org e-print service, vol. abs/0807.3593, 2008.

[5] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York: Wiley, 1991.

[6] G. Kramer, “Topics in multi-user information theory,” Foundations and Trends in Communications and Information Theory, vol. 4, no. 4-5, pp. 265–444, 2007.

[7] I. Csiszár and J. Körner, “Broadcast channels with confidential messages,” IEEE Trans. Inf. Theory, vol. 24, no. 3, pp. 339–348, May 1978.