Comments on “Broadcast Channels with Arbitrarily Correlated Sources”

Gerhard Kramer
Department of Electrical Engineering
University of Southern California
Los Angeles, CA, USA
Email: gkramer@usc.edu

Chandra Nair
Department of Information Engineering
Chinese University of Hong Kong
Sha Tin, N.T., Hong Kong
Email: chandra@ie.cuhk.edu.hk

***Abstract*—The Marton-Gelfand-Pinsker inner bound on the capacity region of broadcast channels was extended by Han-Costa to include arbitrarily correlated sources where the capacity region is replaced by an admissible source region. The main arguments of Han-Costa are correct but unfortunately the authors overlooked an inequality in their derivation. The corrected region is presented and the absence of the omitted inequality is shown to sometimes admit sources that are not admissible.**

***Index Terms*—broadcast channels, capacity, correlated sources**

I. Introduction

We borrow terminology from [1] with minor modifications. Consider a two-receiver broadcast channel (BC), say ω, with correlated, or more precisely dependent, sources (S, T). Source (S, T) is said to be *admissible* for this BC if for any λ, $0 < \lambda < 1$, and for large enough n there is a code with length-n codewords such that $Pe_1 \leq \lambda$ and $Pe_2 \leq \lambda$, where Pe_1 and Pe_2 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 developed a (purported) subset of the admissible source region for BCs with arbitrarily correlated sources in [1, Theorem 1 and Example 1]. We observe that the main arguments in [1] are valid but the authors unfortunately overlooked an inequality in one of the final steps of their proof. The corrected versions of [1, Theorem 1 and Example 1] are presented in Sec. II. On the other hand, since the admissible source region is not known in general, it is not a priori clear whether or not the Han-Costa source set is in fact a subset of the admissible source region after all. We rule out this possibility by giving two examples where the Han-Costa source set includes sources that are not admissible.

II. Revised Theorem 1 and Example 1 in [1]

The wording of the theorem and the example below are taken with minor modifications from [1].

Theorem 1 (revised from [1]): Suppose that a broadcast channel ω and a source (S, T) are given, and let $K = f(S) = g(T)$ be the common variable in the sense of Gacs and Körner (and also Witsenhausen). If there exist auxiliary random variables W, U, V (with values in finite sets) that satisfy the Markov chain property

$$ST - WUV - X - Y_1Y_2 \tag{1}$$

and the inequalities

$$H(S) \leq I(SWU; Y_1) - I(T; WU|S) \tag{2}$$

$$H(T) \leq I(TWV; Y_2) - I(S; WV|T) \tag{3}$$

$$\begin{aligned} H(ST) \leq \min\{I(KW; Y_1), I(KW; Y_2)\} + I(SU; Y_1|KW) \\ + I(TV; Y_2|KW) - I(SU; TV|KW) \end{aligned} \tag{4}$$

$$\begin{aligned} H(ST) \leq I(SWU; Y_1) + I(TWV; Y_2) - I(SU; TV|KW) \\ - I(ST; KW). \end{aligned} \tag{5}$$

then the source (S, T) is admissible for the channel ω. Here, X is an input variable with values in the input alphabet $\mathcal{X}$, and Y_1, Y_2 are the output variables with values in the output alphabets $\mathcal{Y}_1, \mathcal{Y}_2$, respectively, induced by X via ω.

Example 1 (revised from [1]): Consider sources with $S = (S_0, K), T = (T_0, K)$, where S_0, T_0, K are statistically independent, and where $H(K) = R_0$, $H(S_0) = R_1$, $H(T_0) = R_2$. If we choose WUV to be independent of ST, then the conditions of Theorem 1 reduce to the Markov chain property

$$WUV - X - Y_1Y_2 \tag{6}$$

and the inequalities

$$R_0 + R_1 < I(WU; Y_1) \tag{7}$$

$$R_0 + R_2 < I(WV; Y_2) \tag{8}$$

$$\begin{aligned} R_0 + R_1 + R_2 < \min\{I(W; Y_1), I(W; Y_2)\} + I(U; Y_1|W) \\ + I(V; Y_2|W) - I(U; V|W) \end{aligned} \tag{9}$$

$$2R_0 + R_1 + R_2 < I(WU; Y_1) + I(WV; Y_2) - I(U; V|W). \tag{10}$$

Remark 1: Inequalities (5) and (10) are missing in [1]. Note that the revised Example 1 is a special case of a more general result that appeared in the Ph.D. thesis of Y. Liang in 2005 (see [8, p. 89, Remark 10] and [2, Theorem 5]). Note also that the rate region (6)-(10) was shown to be equivalent to the Marton-Gelfand-Pinsker region in [3] (what we call the “Marton-Gelfand-Pinsker region” is given in [4, Theorem 1] and [5, p. 391, Problem 10(c)]).

A. Solving the Case of the Missing Inequality

In [1, p. 647], the authors derive the following valid inequalities, see Equations (3.34)-(3.37):

$$\begin{aligned}
&H(S|KW)+H(T|KW)+H(K)\\
&\qquad < I(TVW;Y_2)+I(SU;Y_1|KW)-\rho_0-\rho_1-\rho_2\\
&H(T|KW)+H(K) < I(TVW;Y_2)-\rho_0-\rho_2\\
&H(S|KW)+H(T|KW)+H(K)\\
&\qquad < I(SUW;Y_1)+I(TV;Y_2|KW)-\rho_0-\rho_1-\rho_2\\
&H(S|KW)+H(K) < I(SUW;Y_1)-\rho_0-\rho_1.
\end{aligned}$$

They next eliminate the variables ρ_0, ρ_1, ρ_2 using the following inequalities in [1, p. 645], see Equations (3.5)-(3.8):

$$\begin{aligned}
\rho_0 &> I(ST;W|K)\\
\rho_1 &> I(T;U|SW)\\
\rho_2 &> I(S;V|TW)\\
\rho_1+\rho_2 &> I(SU;TV|W)-I(S;T|W).
\end{aligned}$$

The oversight occurs in this elimination. Eliminating ρ_0, ρ_1, ρ_2 and removing redundant inequalities we obtain the bounds in [1, (3.38)-(3.41)]:

$$\begin{aligned}
&H(S|KW)+H(T|KW)+H(K)+I(ST;W|K)\\
&\qquad < I(TVW;Y_2)+I(SU;Y_1|KW)-A && (11)\\
&H(T|KW)+H(K)\\
&\qquad < I(TVW;Y_2)-I(ST;W|K)-I(S;V|TW) && (12)\\
&H(T|KW)+H(S|KW)+H(K)+I(ST;W|K)\\
&\qquad < I(SUW;Y_1)+I(TV;Y_2|KW)-A && (13)\\
&H(S|KW)+H(K)\\
&\qquad < I(SUW;Y_1)-I(ST;W|K)-I(T;U|SW) && (14)
\end{aligned}$$

where $A = I(SU;TV|W) - I(S;T|W)$, as well as the bound

$$\begin{aligned}
&H(S|KW)+H(T|KW)+2H(K)+2I(ST;W|K)\\
&\qquad < I(SUW;Y_1)+I(TVW;Y_2)-A. && (15)
\end{aligned}$$

It is the inequality (15) that was omitted in [1].

Continuing as in [1, p. 647], we obtain the revised Theorem 1 by using the equalities

$$\begin{aligned}
&H(S|KW)+H(K) = H(S)-I(S;W|K)\\
&H(T|KW)+H(K) = H(T)-I(T;W|K)\\
&H(S|KW)+H(T|KW)+H(K)\\
&\quad = H(ST)+I(S;T|K)-I(S;W|K)-I(T;W|K).
\end{aligned}$$

B. Counterexample II-B

Since the set-up of Example 1 is a well-studied and important case, we explore the following question: If we remove the inequality (10) then is the resulting rate region (the Han-Costa region of [1, Example 1]) always achievable? We develop two counterexamples to show that this is not the case. The reader will notice that the counterexamples are closely related. We present them both for reasons that will become clear in Sec. III.

As a first counterexample, consider the deterministic BC

$$(Y_1, Y_2) = \begin{cases} (0,0), & X=0\\ (1,0), & X=1\\ (1,1), & X=2\\ (2,1), & X=3. \end{cases} \tag{16}$$

The capacity region of a deterministic BC is known to be the union over Markov chains $W - X - Y_1Y_2$ of the non-negative rate triples (R_0, R_1, R_2) satisfying (see [5, p. 391])

$$R_0 \le \min\left(I(W;Y_1), I(W;Y_2)\right) \tag{17}$$
$$R_0 + R_1 \le H(Y_1) \tag{18}$$
$$R_0 + R_2 \le H(Y_2) \tag{19}$$
$$R_0 + R_1 + R_2 \le \min\left(I(W;Y_1), I(W;Y_2)\right) + H(Y_1Y_2|W). \tag{20}$$

We mimic the development of [6, Sec. IV]. Suppose we would like to achieve

$$R_0 + R_1 + R_2 = H(Y_1Y_2) \tag{21}$$

for the BC (16). For example, we can achieve $(R_0, R_1, R_2) = (0, 1, 1)$ and $H(Y_1Y_2) = 2$ by choosing W to be a constant and X uniform. It is easy to check that for (21) to be satisfied, one must have the double Markov relations

$$W - Y_1 - Y_2 \tag{22}$$
$$W - Y_2 - Y_1 \tag{23}$$

in the expression (20).

Suppose next that we would like to achieve

$$R_0 + R_1 + R_2 = 2 \tag{24}$$

for the BC (16), as in the example we just considered. Obviously, the input X must be uniform, and for this choice of X one can check that the joint distribution of (Y_1, Y_2) is indecomposable in the sense of [5, p. 350]. This further implies, by [5, p. 402], that

$$W \text{ is independent of } Y_1Y_2 \tag{25}$$

and therefore, by (17), that $R_0 = 0$. One can further check that with uniform X, but without the bound (17), the following rate-triple is permitted

$$(R_0, R_1, R_2) = (1/2, 1, 1/2) \tag{26}$$

Thus, the bound (17) is needed because the rate-triple (26) is not achievable.

Finally, note that we are further suggesting that one replace (17) with the bound (10) where $U = Y_1$ and $V = Y_2$, i.e, with

$$2R_0 + R_1 + R_2 \le I(W;Y_1) + I(W;Y_2) + H(Y_1Y_2|W). \tag{27}$$

For example, if $R_0 + R_1 + R_2 = 2$ then from (25) and (27) we see that we must have

$$2R_0 + R_1 + R_2 \le H(Y_1Y_2) = 2 \tag{28}$$

so that $R_0 = 0$. Summarizing, we need to add the bound (17) *or* the bound (10) to the bounds (2.14)-(2.16) in [1]. The equivalence of adding either bound was proved for general broadcast channels in [3].

C. Counterexample II-C

Consider the Blackwell BC shown in Fig. 1. This channel is deterministic so the capacity region is given by (17)-(20) where $W - X - Y_1Y_2$ forms a Markov chain. We have the following lemma that is closely related to Counterexample II-B.

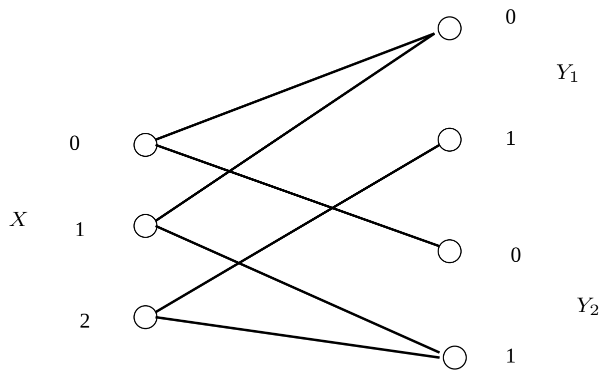

Fig. 1. Blackwell Channel

Lemma 1: If $W - X - Y_1Y_2$ forms a Markov chain and

$$\min\left(I(W;Y_1), I(W;Y_2)\right) + H(Y_1Y_2|W) = H(Y_1Y_2)$$

for the Blackwell channel in Fig. 1 then the random variables W and X are independent.

Before we prove Lemma 1, we claim that the lemma provides a counterexample to our question posed above. In particular, the lemma implies that if $R_0+R_1+R_2 \to H(Y_1Y_2)$, then we must have $R_0\big(\le \min\left(I(W;Y_1), I(W;Y_2)\right)\big) \to 0$. Thus the triple $R_0 = I(Y_1;Y_2) - \epsilon$, $R_1 = H(Y_1|Y_2), R_2 = H(Y_2|Y_1)$, is not achievable for ϵ small enough. However, this rate-triple is permitted by [1, Example 1].

Proof: (Lemma 1) The equality

$$\min\left(I(W;Y_1), I(W;Y_2)\right) + H(Y_1Y_2|W) = H(Y_1Y_2)$$

implies that $I(W;Y_1) = I(W;Y_2) = I(W;Y_1Y_2)$. Using the observation that for a Blackwell channel X is a deterministic function of Y_1Y_2, we have the following equalities

$$I(W;Y_1) = I(W;X) \tag{29}$$
$$I(W;Y_2) = I(W;X). \tag{30}$$

From (29) and the Markov relationship $W - X - Y_1$ we see that $I(W;X|Y_1) = 0$ and therefore

$$\mathrm{P}(X = 0|Y_1 = 0, W = w) = \mathrm{P}(X = 0|Y_1 = 0). \tag{31}$$

For the Blackwell Channel, (31) is equivalent to

$$\frac{\mathrm{P}(X=0)}{\mathrm{P}(X=0)+\mathrm{P}(X=1)} = \frac{\mathrm{P}(X=0|W=w)}{\mathrm{P}(X=0|W=w)+\mathrm{P}(X=1|W=w)}.$$

Thus we obtain

$$\frac{\mathrm{P}(X=0)}{\mathrm{P}(X=1)} = \frac{\mathrm{P}(X=0|W=w)}{\mathrm{P}(X=1|W=w)}. \tag{32}$$

Similarly starting from (30) and the Markov relationship $W - X - Y_2$, we compute

$$\frac{\mathrm{P}(X=2)}{\mathrm{P}(X=1)} = \frac{\mathrm{P}(X=2|W=w)}{\mathrm{P}(X=1|W=w)}. \tag{33}$$

From (32) and (33) we deduce

$$\mathrm{P}(X = i|W = w) = \mathrm{P}(X = i), \text{ for } i = 0, 1, 2.$$

This concludes the proof of Lemma 1. ∎

III. Historical Remarks

The revised Theorem 1 and Example II-B were developed by G. Kramer in the summer of 2005. His motivation was that S. Shamai pointed out to him that the potential improvement (6)-(10) of the Marton-Gelfand-Pinsker region that appeared in the Ph.D. thesis of Y. Liang (see [8, p. 89, Remark 10] and [2, Theorem 5]) was superseded by the earlier results of Han-Costa [1]. Kramer communicated the revised Theorem 1 and Example II-B to Shamai and Han in August 2005 via email but did not otherwise document the results.

In 2008, Y.-H. Kim queried C. Nair about the validity of the results in Han-Costa [1]. Nair independently discovered and corrected the error of [1] in 2008 and developed Example II-C. He forwarded a write-up of his results to A. El Gamal and M. Costa. Costa forwarded the write-up to Han, who then replied back with the earlier communication by Kramer. This eventually led to the current joint paper.

Acknowledgement

The authors wish to thank S. Shamai, T. S. Han, M. Costa, A. El Gamal, and Y.-H. Kim for their roles in bringing these results together.

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] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.

[3] Y. Liang, G. Kramer, and H. V. Poor, "Equivalence of two inner bounds on the capacity region of the broadcast channel," in *Proc. 46th Annual Allerton Conf. on Commun., Control, and Computing*, Monticello, IL, Sept. 23–26 2008.

[4] S. I. Gel'fand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Inform. Transm.*, vol. 16, no. 1, pp. 17–25, Jan.-Mar. 1980.

[5] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Channels*. Budapest: Akadémiai Kiadó, 1981.

[6] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2425–2432, June 2006.

[7] T. S. Han, "The capacity region for the deterministic broadcast channel with a common message," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 122–125, Jan. 1981.

[8] Y. Liang, "Multiuser communications with relaying and user cooperation," Ph.D. dissertation, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Illinois, 2005.