A New Outer Bound to the Capacity Region of Deterministic-Code Discrete Memoryless Arbitrary Varying General Broadcast Channel

Amin Aminzadeh Gohari
Venkat Anantharam

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2008-81
http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-81.html

June 9, 2008

A New Outer Bound to the Capacity Region of Deterministic-Code Discrete Memoryless Arbitrary Varying General Broadcast Channel

Amin Aminzadeh Gohari[1] and Venkat Anantharam [1]

[1] Department of Electrical Engineering and Computer Science

University of California, Berkeley

{aminzade,ananth}@eecs.berkeley.edu

Abstract

In this paper we apply the "potential function method" introduced by the authors in [1] and [2] to prove a new outer bound on the capacity region of the deterministic-code arbitrarily varying general broadcast channel. Specializing by removing the variability of the channel, our outer bound gives another proof of the recent result of Liang, Kramer and Shamai, which is the currently best upper bound on the capacity region of the traditional broadcast channel [3].

I. Introduction

Broadcast channels form basic building blocks of many wireless system models. A broadcast channel is a single-input, multi-output system whose goal is to model reliable communication of sets of messages from a transmitter to different sets of receivers. In some practical scenarios the channel parameters may be unknown, imprecise, or subject to variations from one symbol transmission to the next one. An arbitrarily varying channel (AVC) models such a discrete memoryless channel. It is assumed that the channel parameters admit no statistical description and any code over this channel must have guaranteed performance under the worst possible choice of the channel parameters.

We consider only two-receiver arbitrarily varying general broadcast channels in this paper. A two-receiver broadcast channel is characterized by the conditional distribution $q(y,z|x)$ where X is the input to the channel and Y and Z are the outputs of the channel at the two receivers. In a general broadcast channel, the transmitter has a common message and two private messages for the two receivers. Roughly speaking, the capacity region of the general broadcast channel is the set of all triples (R_0, R_1, R_2) for which there exists a strategy for the transmitter to send R_0 common bits per channel use to both the receivers, R_1 private bits per channel use to receiver Y and R_2 private bits per channel use to receiver Z.

An arbitrarily varying general broadcast channel is characterized by the conditional distribution $q(y, z|x, s)$ where X is the input of the transmitter to the channel, S is the state parameter of the channel (that can vary in an arbitrary way throughout the communication) and Y and Z are the outputs of the channel at the two receivers. Roughly speaking, the capacity region of an arbitrarily varying general broadcast channel is the set of all triples (R_0, R_1, R_2) for which there exists a strategy for the transmitter to send R_0 common bits per channel use to both the receivers, R_1 private bits per channel use to receiver Y and R_2 private bits per channel use to receiver Z no matter how the state of the channel varies over time. The transmitted messages should be recoverable by the receivers with high probability. Depending on the model, either an average probability of error, or a maximal probability of error constraint at the receivers is imposed. Furthermore, sometimes it is assumed that there are common private random bits shared between the transmitter and the receivers. Depending on the choice of model, different notions of capacity can be defined. In this paper we assume that no shared common randomness is provided to the transmitter and the receivers (deterministic-code arbitrarily varying general broadcast channels), and that the receivers are required to find the intended messages under an average probability of error constraint (see section 2 for a formal definition).

The capacity region of a broadcast channel is not known when the channel parameters are fixed except in certain special cases; less is known when the channel parameters vary arbitrarily. The best known inner bound for the two receiver general broadcast channel is due to Marton [4]. The best outer bound is due to Liang, Kramer and Shamai [3]; it is not however known whether this bound strictly improves on the earlier outer bound of Nair and El Gamal [5]. For arbitrarily varying general broadcast channels (AVGBC), the best known inner bound, as far as we are aware, belongs to Jahn [6]. For the family of degraded message sets[1], Hof and Bross found a new inner bound on the capacity region of the AVGBC under state and input constraints. We are not aware of any previous work discussing any interesting outer bounds on the capacity region of an AVGBC [7].

In this paper, we consider the capacity region of the AVC general broadcast channels when no shared common randomness is provided to the transmitter and the receivers. The capacity region is defined as the average probability of error over messages; rate-tuples in the region need to be achievable uniformly over

[1] We do not consider the degraded message set restriction here; for a definition see [7]

the channel parameters (which can vary symbol by symbol). We apply the "potential function method" introduced by the authors in [1] and [2] to prove a new outer bound on the capacity region of AVC general broadcast channels.

A sketch of the "potential function method" is as follows: we consider the set of all joint distributions on products of four finite sets which represent, roughly speaking, the knowledge of the two receivers and the transmitter, and the history of broadcast channel parameter choices at some stage of the communication. We then identify properties of a function on such distributions which would need to be satisfied in one step of the communication for it to give rise to an outer bound. For details, see the statement of theorem 1 or see [1] and [2].

The outline of this paper is as follows. In section II, we introduce the basic notations and definitions used in this paper. Section III contains the main results of this paper followed by section IV which gives formal proofs for the results. The appendix completes the proof of theorem 2 of section IV.

II. Definitions and Notation

Throughout this paper we assume that each random variable takes values in a finite set. $\mathbb{R}_+$ denotes the set $\{x \in \mathbb{R} : x \geq 0\}$.

We represent an AVC broadcast channel by the conditional distribution $q(y, z|x, s)$ meaning that X is talking, S is the state of the channel, and Y and Z are listening. We assume that X, S, Y and Z take values from discrete sets ψ_X, ψ_S, ψ_Y and ψ_Z respectively. For any natural number n, $(\psi_X)^n$, $(\psi_S)^n$, $(\psi_Y)^n$ and $(\psi_Z)^n$ denote the n-th product sets of ψ_X, ψ_S, ψ_Y and ψ_Z.

Definition 1. Given the conditional distribution $q(y, z|x, s)$, positive real ϵ and natural numbers n, M_0, M_1, M_2, a $(n, M_0, M_1, M_2, \epsilon)$ code is the set of the following three mappings:

$$f : \{1, 2, 3, ..., M_0\} \times \{1, 2, 3, ..., M_1\} \times \{1, 2, 3, ..., M_2\} \longrightarrow (\psi_X)^n$$

$$\vartheta : (\psi_Y)^n \longrightarrow \{1, 2, 3, ..., M_0\} \times \{1, 2, 3, ..., M_1\}$$

$$\lambda : (\psi_Z)^n \longrightarrow \{1, 2, 3, ..., M_0\} \times \{1, 2, 3, ..., M_2\}$$

such that for any $s^n \in (\psi_S)^n$, the following "average probabilities of error" condition is satisfied:

Assume that L_0, L_1 and L_2 are random variables uniformly taking values from the sets $\{1, 2, 3, ..., M_0\}$, $\{1, 2, 3, ..., M_1\}$ and $\{1, 2, 3, ..., M_2\}$. Assume that $X^n = f(L_0, L_1, L_2)$. Random variables S^n, Y^n and Z^n are defined according to the following constraint:

$$p(y^n, z^n, x^n, s^n, l_0, l_1, l_2) = p(l_0, l_1, l_2, x^n).p(s^n) \prod_{i=1}^{n} q(y_i, z_i|x_i, s_i).$$

We then have the following constraints:

$$e_{\vartheta}(s^n) = \frac{1}{M_0M_1M_2}\sum_{i=1}^{M_0}\sum_{j=1}^{M_1}\sum_{k=1}^{M_2}\sum_{y^n:\vartheta(y^n)\neq(i,j)} q(y^n|f(i,j,k),s^n) \leq \epsilon$$

$$e_{\lambda}(s^n) = \frac{1}{M_0M_1M_2}\sum_{i=1}^{M_0}\sum_{j=1}^{M_1}\sum_{k=1}^{M_2}\sum_{z^n:\lambda(z^n)\neq(i,k)} q(z^n|f(i,j,k),s^n) \leq \epsilon$$

Definition 2. Given the conditional distribution $q(y,z|x,s)$, the capacity region of the deterministic-code AVC general broadcast channel, $C_{BC}(q(y,z|x,s))$, is a subset of triples of non-negative real numbers defined as follows: A triple (R_0,R_1,R_2) belongs to the capacity region of the AVC general broadcast channel if for every positive ϵ and δ and sufficiently large n, a (n,M_0,M_1,M_2,ϵ) code exists for which $\frac{1}{n}\log M_0 \geq R_0-\delta$, $\frac{1}{n}\log M_1 \geq R_1-\delta$ and $\frac{1}{n}\log M_2 \geq R_2-\delta$.

Definition 3. For any natural number c and any two sets of points K and L in $\mathbb{R}_+^c$, let $K\oplus L$ refer to their convolution: $K\oplus L = \{v_1+v_2 : v_1\in K, v_2\in L\}$. For any natural number n, let $n\otimes K$ be the addition of n K's: $K\oplus K\oplus ...\oplus K$ (n times). We also define $\frac{K}{n}$ as the set formed by shrinking K through scaling each point of it by a factor $\frac{1}{n}$: $\frac{K}{n} = \{\frac{1}{n}v : v\in K\}$

Remark. $\frac{n\otimes K}{n}$ falls inside the convex hull of K.

Definition 4. For any two points $\overrightarrow{v}_1$ and $\overrightarrow{v}_2$ in $\mathbb{R}_+^c$, we say $\overrightarrow{v}_1 \geq \overrightarrow{v}_2$ if and only if each coordinate of $\overrightarrow{v}_1$ is greater than or equal to the corresponding coordinate of $\overrightarrow{v}_2$. For a set $A\in\mathbb{R}_+^c$, the down-set $\Delta(A)$ is defined as: $\Delta(A) = \{\overrightarrow{v}\in\mathbb{R}_+^c : \overrightarrow{v}\leq\overrightarrow{w}$ for some $\overrightarrow{w}\in A\}$.

Definition 5.

For every given $p(x,s,y,z)$, we define: $\Upsilon_{p(x,s,y,z)} =$ the set of $p(w_0,w_1,w_2,u,v,x,s,y,z)$ satisfying :

$$\begin{cases} p(w_0,w_1,w_2,u,v,x) \text{ satisfies:} & p(w_0,w_1,w_2,u,v,x) = \\ & \;p(w_0)p(w_1)p(w_2)p(u,v|w_0w_1w_2)p(x|w_0,w_1,w_2,u,v); \\ & X \text{ is a deterministic function of } (W_0,W_1,W_2,U,V); \\ & X \text{ has the marginal distribution corresponding to } p(x,s,y,z); \\ \text{The following Markov chain holds:} & UVW_0W_1W_2X - X - XSYZ; \\ & XSYZ \text{ here have joint distribution } p(x,s,y,z) \end{cases}$$

Definition 6. Given $p(y,z|x,s)$ and $p(x)$, we use the notation $C_R(p(x)p(y,z|x,s))$ to denote a set whose form is motivated by Jahn's inner bound on the AVC general broadcast channel $p(y,z|x,s)$ [6]

defined below (see Theorem 2 of [6]).

$$C_R(p(x)p(y,z|x,s)) = \text{All non-negative triples of } (R_0,R_1,R_2) \text{ for which there exists}$$

$$p(w_0,w_1,w_2,x) \in \Omega \text{ with marginal the given } p(x) \text{ and such that :}$$

$$\begin{cases} R_0 \leq \inf_{p(s)} I(W_0;Y) \\ R_1 \leq \inf_{p(s)} I(W_1;Y|W_0) \\ R_0 \leq \inf_{p(s)} I(W_0;Z) \\ R_2 \leq \inf_{p(s)} I(W_2;Z|W_0) \\ R_1 \leq \inf_{p(s)} I(W_1;Y) \\ R_2 \leq \inf_{p(s)} I(W_2;Z) \\ R_0+R_1 \leq \inf_{p(s)} I(W_0W_1;Y) \\ R_0+R_2 \leq \inf_{p(s)} I(W_0W_2;Z) \\ R_0+R_1+R_2 \leq \inf_{p(s)} I(W_1;Y|W_0W_2)+I(W_0W_2;Z) \\ R_0+R_1+R_2 \leq \inf_{p(s)} I(W_2;Z|W_0W_1)+I(W_0W_1;Y) \\ R_0+R_1+R_2 \leq \inf_{p(s)} I(W_0;Y)+I(W_1;Y|W_0W_2)+I(W_2;Z|W_0) \\ R_0+R_1+R_2 \leq \inf_{p(s)} I(W_0;Z)+I(W_2;Z|W_0W_1)+I(W_1;Y|W_0) \end{cases} \tag{1}$$

where $W_0, W_1, W_2, X, S, Y, Z$ have joint distribution $p(w_0,w_1,w_2,x).p(y,z|x,s)p(s)$ for arbitrary $p(s)$; and Ω is defined as follows:

$\Omega =$ the set of $p(w_0,w_1,w_2,x)$ satisfying :

$$\begin{cases} p(w_0,w_1,w_2,x) \text{ satisfies:} \quad & p(w_0,w_1,w_2,x) = \\ & p(w_0)p(w_1)p(w_2)p(x|w_0,w_1,w_2); \\ & X \text{ is a deterministic function of } (W_0,W_1,W_2); \end{cases}$$

To compare this with the inner bound of Jahn (Theorem 2 of [6]), replace U^c with W_0, U^y with W_1 and U^z with W_2 (where U^c, U^y and U^z are defined in [6]). Other differences are the following:

- We require U^y, U^z and U^c to be independent of each other
- The constraints on R_0 were strengthened by replacing $I(YU^y;U^c)$ with $I(Y;U^c)$, and $I(ZU^z;U^c)$ with $I(Z;U^c)$.
- Some extra inequalities were added (inequalities 5-12).

Finally and most importantly $C_R(p(x)p(y,z|x,s))$ is allowed to be nonempty even in cases where Jahn's inner bound is not applicable (cf. Remark IIB2 of [6]), i.e. when no common rate can be sent. This suggests that our bound can be improved further.

Definition 7. The map $\Pi : 2^{\mathbb{R}_+^{12}} \mapsto 2^{\mathbb{R}_+^3}$, from subsets of $\mathbb{R}_+^{12}$ to subsets of $\mathbb{R}_+^3$ is defined as follows: For any $A \subseteq \mathbb{R}_+^{12}$,

$$\Pi(A) = \bigcup_{(t_0,t_1,t_2,\ldots,t_{10},t_{11})\in A} \text{all non-negative triples of } (R_0,R_1,R_2) \text{ such that :}$$

$$\begin{cases} R_0 \leq \min(t_0,t_1) \\ R_1 \leq t_2 \\ R_2 \leq t_3 \\ R_0 + R_1 \leq \min(t_4,t_5) \\ R_0 + R_2 \leq \min(t_6,t_7) \\ R_0 + R_1 + R_2 \leq \min(t_8,t_9,t_{10},t_{11}) \end{cases}$$

Π will be called the projection map.

III. Statement of the results

In this section, the main claims of the paper are formally presented as Theorems 1 through 2.

Theorem 1. Let $\varphi_j(p(y,z,x,s))$ $(j = 0,1,2,\ldots)$ be a function from the set of all probability distributions defined on a product of four finite sets to subsets of $\mathbb{R}_+^{12}$. For any conditional distribution $q(y,z|x,s)$, let $\phi(q(y,z|x,s)) = \bigcup_{q(x)}\bigcap_{q(s)}\Delta\big(\varphi_1(q(s).q(x).q(y,z|x,s))\big)$ (see definition 4). The region

$$\Pi(\text{convex hull of } \phi(q(y,z|x,s)))$$

is an outer bound on $C_{BC}(q(y,z|x,s))$, the AVC general broadcast channel capacity region, if φ_j $(j = 0,1,2,\ldots)$ satisfy the following properties:

Take some arbitrary j, $p(y,z|x,s)$ and $p(x)$. Then: (please see definition 3 and 4 for definition of notations used)

1) Whenever $p(YZY'Z'|XX'SS') = p(YZ|XS).p(Y'Z'|X'S')$, $H(X'|X) = 0$ and $p(y',z'|x',s') = q(y',z'|x',s')$:

$$\bigcap_{p(ss')}\Delta\big(\varphi_{j+1}(p(yzy'z'|xx'ss')p(xx')p(ss'))\big) \subseteq \left(\bigcap_{p(s)}\Delta\big(\varphi_j(p(yz|sx)p(s)p(x))\big)\right) \oplus \phi(q(y,z|x,s));$$

2) Whenever $H(Y'|Y) = 0$ and $H(Z'|Z) = 0$:

$$\bigcap_{p(s)}\Delta\big(\varphi_j(p(y'z'|sx)p(s)p(x))\big) \subseteq \bigcap_{p(s)}\Delta\big(\varphi_j(p(yz|sx)p(s)p(x))\big);$$

3) Whenever $p(y, z, x, s) = \mathbf{1}[y = z = 0]p(s)p(x)$:

$$\varphi_j(p(y, z, x, s)) = \{(0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T\}$$

4) For any $(R_0, R_1, R_2) \in C_R(p(x)p(y, z|x, s))$:
$(R_0, R_0, R_1, R_2, R_0 + R_1, R_0 + R_1, R_0 + R_2, R_0 + R_2, R_0 + R_1 + R_2,$
$R_0 + R_1 + R_2, R_0 + R_1 + R_2, R_0 + R_1 + R_2) \in \bigcap_{p(s)} \Delta\big(\varphi_j(p(yz|sx)p(s)p(x))\big).$

Discussion: The domain of φ_j in Theorem 1 is the set of *all* probability distributions on *all* products of four finite sets. Given $p(y, z|x, s)$ and $p(x)$, for each $j \geq 1$, the quantity $\bigcap_{p(s)} \Delta\big(\varphi_j(p(y, z|x, s)p(x)p(s))\big)$ can be intuitively understood as representing the set of 12-tuples $(R_0, R_0, R_1, R_2, R_0+R_1, R_0+R_1, R_0+R_2, R_0+R_2, R_0+R_1+R_2, R_0+R_1+R_2, R_0+R_1+R_2, R_0+R_1+R_2)$ where (R_0, R_1, R_2) belongs to an outer bound on the capacity region of a broadcast channel with description $p(y, z|x, s)$ and specified input distribution $p(x)$; the new AVC channel $p(y, z|x, s)$ should be effectively created on the first j uses of the original broadcast channel $q(y, z|x, s)$. With this rough picture in mind, condition 1 can be understood as saying that having already insisted on working with a j-channel use $p(y, z|x, s)$, one more use of the channel can at most buy us the broadcast capacity on a per use basis. Condition 2 says that further insistence on working with a distribution that results from information reduction by the receivers cannot increase the per channel use broadcast region. Condition 3 specifies the region when Y and Z are constant. The right hand side of condition 4 is just a convenient expression that is easily seen to be an inner bound on the corresponding constrained broadcast rate; other such expressions would have worked as well. ●

Theorem 2. Given any AVC-Broadcast channel $q(y, z|x, s)$, the following region forms an outer bound on the capacity region of the broadcast channel:

$$\zeta(q(y, z|x, s)) =$$

$$\bigcup_{p(x)} \bigcap_{p(s)} \bigcup_{p(w_0,w_1,w_2,u,v,x,s,y,z)\in\Upsilon_{p(s)p(x)q(y,z|x,s)}}$$

$$\begin{cases} R_0 \geq 0, R_1 \geq 0, R_2 \geq 0; \\ R_0 \leq \min\{I(W_0;Y|U), I(W_0,Z|V)\}; \\ R_1 \leq I(W_1;Y|U)); \\ R_2 \leq I(W_2;Z|V)); \\ R_0 + R_1 \leq \min(I(W_0W_1;Y|U), I(W_1;Y|W_0UV) + I(W_0U;Z|V)); \\ R_0 + R_2 \leq \min(I(W_0W_2;Z|V), I(W_2;Z|W_0UV) + I(W_0V;Y|U)); \\ R_0 + R_1 + R_2 \leq I(W_1;Y|W_0W_2UV) + I(W_0W_2U;Z|V); \\ R_0 + R_1 + R_2 \leq I(W_2;Z|W_0W_1UV) + I(W_0W_1V;Y|U); \\ R_0 + R_1 + R_2 \leq I(W_0UV;Y) + I(W_1;Y|W_0W_2UV) + I(W_2;Z|W_0UV); \\ R_0 + R_1 + R_2 \leq I(W_0UV;Z) + I(W_2;Z|W_0W_1UV) + I(W_1;Y|W_0UV). \end{cases}$$

Remark: If $q(y,z|x,s) = q(y,z|x)$, the above outer bound reduces to that of Liang, Kramer and Shamai [3]. Please note that we have removed the constraint in [3] on W_0, W_1 and W_2 being uniform. The Liang, Kramer and Shamai region with or without this constraint is the same. This is because given any $(W_0, W_1, W_2, U, V, X, S, Y, Z)$ with joint distribution

$$p(w_0)p(w_1)p(w_2)p(u,v|w_0,w_1,w_2)p(x|u,v,w_0,w_1,w_2)p(s)q(y,z|x,s)$$

where $H(X|U,V,W_0,W_1,W_2) = 0$, one can find $(\widetilde{W_0}, \widetilde{W_1}, \widetilde{W_2}, \widetilde{U}, \widetilde{V}, \widetilde{X}, \widetilde{S}, \widetilde{Y}, \widetilde{Z})$ with joint distribution $p(\widetilde{w_0})p(\widetilde{w_1})p(\widetilde{w_2})p(\widetilde{u},\widetilde{v}|\widetilde{w_0},\widetilde{w_1},\widetilde{w_2})p(\widetilde{x}|\widetilde{u},\widetilde{v},\widetilde{w_0},\widetilde{w_1},\widetilde{w_2})p(\widetilde{s})q(\widetilde{y},\widetilde{z}|\widetilde{x},\widetilde{s})$ where $\widetilde{W_0}, \widetilde{W_1}$ and $\widetilde{W_2}$ are uniform and $H(\widetilde{X}|\widetilde{U},\widetilde{V},\widetilde{W_0},\widetilde{W_1},\widetilde{W_2}) = 0$ such that there are W_0', W_1' and W_2' with $H(W_0'|\widetilde{W_0}) = H(W_1'|\widetilde{W_1}) = H(W_2'|\widetilde{W_2}) = 0$ and $(W_0', W_1', W_2', \widetilde{U}, \widetilde{V}, \widetilde{X}, \widetilde{S}, \widetilde{Y}, \widetilde{Z})$ being arbitrarily close to $(W_0, W_1, W_2, U, V, X, S, Y, Z)$ in total variation. To do this, take $\widetilde{W_0}, \widetilde{W_1}$ and $\widetilde{W_2}$ independent and uniform on large finite sets. Create W_0', W_1' and W_2' respectively with a joint distribution close to $p(w_0,w_1,w_2)$ in total variation. Then use the same channels $p(u,v|w_0,w_1,w_2)$, $p(x|u,v,w_0,w_1,w_2)$ as in the original joint distribution to create $\widetilde{U}, \widetilde{V}, \widetilde{X}, \widetilde{S}, \widetilde{Y}$ and $\widetilde{Z}$.

IV. Proofs of Theorems 1-2

Proof of Theorems 1: Take a triple (t_0, t_1, t_2) inside the broadcast channel rate region for the channel $q(y,z|x,s)$. Take some positive ϵ and δ, and consider a $(n, M_0, M_1, M_2, \epsilon)$ code such that $\frac{1}{n}\log M_0 \geq t_0 - \delta$, $\frac{1}{n}\log M_1 \geq t_1 - \delta$ and $\frac{1}{n}\log M_2 \geq t_2 - \delta$.

Let L_0, L_1, and L_2 denote the three messages that the encoder is transmitting to the two receivers; this would imply that $H(L_i) = \frac{1}{n}\log M_i$ for $i = 0, 1, 2$. Define the random variable X as $X = (L_0, L_1, L_2)$. Also, let X_1', X_2', ..., X_n' represent the inputs by the encoder at the broadcast channel, i.e. $(X_1', X_2', ..., X_n') = f(X)$. Clearly $H(X_i'|X) = 0$. We also let the random variables S_1, S_2, ..., S_n denote the adversary's input to the broadcast channel. When X_i''s and S_i's are inserted at the input of the broadcast channel the Y-party receives Y_1', Y_2', ..., Y_n' and the Z-party receives Z_1', Z_2', ..., Z_n'. Let $Y = (Y_1', Y_2', ..., Y_n')$, $Z = (Z_1', Z_2', ..., Z_n')$ and $S = (S_1, S_2, ..., S_n)$. The decoding rule ensures that for any realization of S_1, S_2, ..., S_n, the Y-party and the Z-party are able to compute

$$(L_0', L_1') = \vartheta(Y_1', Y_2', ..., Y_n')$$

$$(\widehat{L}_0, \widehat{L}_2) = \lambda(Z_1', Z_2', ..., Z_n')$$

such that $p((L_0, L_1) = (L_0', L_1')) \geq 1 - \epsilon$ and $p((L_0, L_2) = (\widehat{L}_0, \widehat{L}_2)) \geq 1 - \epsilon$.

Lastly, let random variable S_0 be independent of all random variables mentioned above. We define random variables Y_0' and Z_0' as $p(y_0', z_0'|x, s_0) = \mathbf{1}[y_0' = z_0' = 0]$.

Using the properties of $\varphi_j(.)$, we have:

$$n \otimes \phi(q(y, z|x, s))$$

$$=^i n \otimes \phi(q(y, z|x, s)) \oplus \bigcap_{p(s_0)} \Delta\big(\varphi_0\big(p(y_0', z_0'|x, s_0)p(x)p(s_0)\big)\big)$$

$$\supseteq^{ii} [(n-1) \otimes \phi(q(y, z|x, s))] \oplus \bigcap_{p(s_0, s_1)} \Delta\big(\varphi_1\big(p(y_0'y_1', z_0'z_1'|x, s_0s_1)p(x)p(s_0s_1)\big)\big)$$

$$\supseteq^{iii} [(n-2) \otimes \phi(q(y, z|x, s))] \oplus \bigcap_{p(s_0, s_1, s_2)} \Delta\big(\varphi_2\big(p(y_0'y_1'y_2', z_0'z_1'z_2'|x, s_0s_1s_2)p(x)p(s_0s_1s_2)\big)\big)$$

$$...$$

$$\supseteq \bigcap_{p(s_0, s_1, s_2, ..., s_n)} \Delta\big(\varphi_n\big(p(y_0'y_1'y_2'...y_n', z_0'z_1'z_2'...z_n'|x, s_0s_1s_2...s_n)p(x)p(s_0s_1s_2...s_n)\big)\big)$$

$$\supseteq^{iv} \bigcap_{p(s_0, s_1, s_2, ..., s_n)} \Delta\big(\varphi_n\big(p(L_0'L_1', \widehat{L}_0\widehat{L}_2|x, s_0s_1s_2...s_n)p(x)p(s_0s_1s_2...s_n)\big)\big)$$

where in i we have used property 3;

in ii we have used property 1 because

$$p(y_0'y_1'z_0'z_1'|xs_0s_1) = p(y_0'y_1'z_0'z_1'|xx_1's_0s_1) = p(y_0'z_0'|xs_0).p(y_1'z_1'|x_1's_1)$$

and furthermore $p(y_1'z_1'|x_1's_1) = q(y_1'z_1'|x_1's_1)$;

in iii we have used property 1 because

$$p(y_0'y_1'y_2'z_0'z_1'z_2'|xs_0s_1s_2) = p(y_0'y_1'y_2'z_0'z_1'z_2'|xx_1'x_2's_0s_1s_2) = p(y_0'y_1'z_0'z_1'|xx_1's_0s_1).p(y_2'z_2'|x_2's_2)$$

and furthermore $p(y'_2z'_2|x'_2s_2) = q(y'_2z'_2|x'_2s_2)$;
in iv, we have used property number 2 because $H(L'_0L'_1|Y'_0Y'_1Y'_2...Y'_n) = 0$ and $H(\widehat{L}_0\widehat{L}_2|Z'_0Z'_1Z'_2...Z'_n) = 0$.

We therefore have:

$$\bigcap_{p(s_0,s_1,s_2,...,s_n)} \Delta\big(\varphi_n\big(p(L'_0L'_1,\widehat{L}_0\widehat{L}_2|L_0L_1L_2,s_0s_1s_2...s_n)p(L_0L_1L_2)p(s_0s_1s_2...s_n)\big)\big) \subseteq$$

$$n \otimes \phi(q(y,z|x,s))$$

Now, we would like to use property 4 and definition 6 on the conditional distribution with parameters $W_1 = L_1$, $W_2 = L_2$, $W_0 = L_0$. Both equations $(L_0, L_1) = (L'_0, L'_1)$ and $(L_0, L_2) = (\widehat{L}_0, \widehat{L}_2)$ are valid with probability at least $1 - \epsilon$ for every choice of $s_0s_1s_2...s_n$. The Fano inequality implies that

$$\big(H(L_0)-O(n\epsilon), H(L_1)-O(n\epsilon), H(L_2)-O(n\epsilon)\big) \in C_R(p(L_0L_1L_2)p(L'_0L'_1,\widehat{L}_0\widehat{L}_2|L_0L_1L_2,s_0s_1...s_n)).$$

Therefore

$$\Big(H(L_0) - O(n\epsilon),$$

$$H(L_0) - O(n\epsilon),$$

$$H(L_1) - O(n\epsilon),$$

$$H(L_2) - O(n\epsilon),$$

$$H(L_0) + H(L_1) - O(n\epsilon),$$

$$H(L_0) + H(L_1) - O(n\epsilon),$$

$$H(L_0) + H(L_2) - O(n\epsilon),$$

$$H(L_0) + H(L_2) - O(n\epsilon),$$

$$H(L_0) + H(L_1) + H(L_2) - O(n\epsilon),$$

$$H(L_0) + H(L_1) + H(L_2) - O(n\epsilon),$$

$$H(L_0) + H(L_1) + H(L_2) - O(n\epsilon),$$

$$H(L_0) + H(L_1) + H(L_2) - O(n\epsilon)\Big) \in$$

$$\bigcap_{p(s_0,s_1,s_2,...,s_n)} \Delta\big(\varphi_n(p(L'_0L'_1,\widehat{L}_0\widehat{L}_2|x,s_0s_1s_2...s_n)p(x)p(s_0s_1s_2...s_n))\big) \subseteq$$

$$n \otimes \phi(q(y,z|x,s))$$

We will be done by letting $\epsilon \rightarrow 0$ and noting that $\frac{n\otimes\phi(q(y,z|x,s))}{n}$ falls inside the convex hull of $\phi(q(y,z|x,s))$.

Proof of Theorem 2. It can be observed that $\zeta(q(y,z|x,s))$ can be written as $\Pi(\phi(q(y,z|x,s)))$ where

$$\phi(q(y,z|x,s)) =$$

$$\bigcup_{p(x)}\bigcap_{p(s)}\bigcup_{p(w_0,w_1,w_2,u,v,x,s,y,z)\in\Upsilon_{p(s)p(x)q(y,z|x,s)}}$$

$$\Delta\Bigg(\big\{\big(I(W_0;Y|U),$$

$$I(W_0,Z|V),$$

$$I(W_1;Y|U)),$$

$$I(W_2;Z|V)),$$

$$I(W_0W_1;Y|U),$$

$$I(W_1;Y|W_0UV)+I(W_0U;Z|V)),$$

$$I(W_0W_2;Z|V),$$

$$I(W_2;Z|W_0UV)+I(W_0V;Y|U)),$$

$$I(W_1;Y|W_0W_2UV)+I(W_0W_2U;Z|V),$$

$$I(W_2;Z|W_0W_1UV)+I(W_0W_1V;Y|U),$$

$$I(W_0UV;Y)+I(W_1;Y|W_0W_2UV)+I(W_2;Z|W_0UV),$$

$$I(W_0UV;Z)+I(W_2;Z|W_0W_1UV)+I(W_1;Y|W_0UV)\big)\big\}\Bigg)$$

In the appendix, we have shown that $\phi(q(y,z|x,s))$ takes values in convex sets. In order to use Theorem 1, we still need to define $\varphi_j(p(y,z,x,s))$ $(j = 0,1,2,...)$ consistently with the above definition of $\phi(q(y,z|x,s))$. This would be straightforward by taking, for any joint distribution $p(y,z,x,s)$,

$$\varphi_j(p(y,z,x,s)) = \bigcup_{p(w_0,w_1,w_2,u,v,x,s,y,z)\in\Upsilon_{p(y,z,x,s)}}$$

$$\begin{gathered}
\Delta\Big(\big\{\big(I(W_0;Y|U), \\
I(W_0,Z|V), \\
I(W_1;Y|U)), \\
I(W_2;Z|V)), \\
I(W_0W_1;Y|U), \\
I(W_1;Y|W_0UV)+I(W_0U;Z|V)), \\
I(W_0W_2;Z|V), \\
I(W_2;Z|W_0UV)+I(W_0V;Y|U)), \\
I(W_1;Y|W_0W_2UV)+I(W_0W_2U;Z|V), \\
I(W_2;Z|W_0W_1UV)+I(W_0W_1V;Y|U), \\
I(W_0UV;Y)+I(W_1;Y|W_0W_2UV)+I(W_2;Z|W_0UV), \\
I(W_0UV;Z)+I(W_2;Z|W_0W_1UV)+I(W_1;Y|W_0UV)\big)\big\}\Big)
\end{gathered}$$

Please note that $\varphi_j(.)$ as defined above is a down-set, i.e. $\varphi_j(.) = \Delta(\varphi_j(.))$.

Now, we will prove that $\varphi_j(.)$ ($j = 0, 1, 2, ...$) satisfies the properties of Theorem 1.

Property number 1. Given $p(y, z|x, s)$ and $p(x)$, we need to show that if $p(YZY'Z'|XX'SS') = p(YZ|XS).p(Y'Z'|X'S')$, $H(X'|X) = 0$ and $p(y', z'|x', s') = q(y', z'|x', s')$, then

$$\bigcap_{p(ss')} \Delta\big(\varphi_{j+1}(p(yzy'z'|xx'ss')p(xx')p(ss'))\big) \subseteq \bigcap_{p(s)} \Delta\big(\varphi_j(p(y, z|x, s)p(s)p(x))\big)\big) \oplus \phi(q(y, z|x, s)).$$

Since the distribution of X is given and $H(X'|X) = 0$, $p(x')$ would be fixed. Since $\phi(q(y, z|x, s)) = \bigcup_{q(x)} \bigcap_{p(s')} \Delta\big(\varphi_1(p(s').q(x').q(y', z'|x', s'))\big) \supseteq \bigcap_{p(s')} \Delta\big(\varphi_1(p(s').p(x').q(y', z'|x', s'))\big)$, it would be enough to prove that

$$\begin{gathered}
\bigcap_{p(ss')} \Delta\big(\varphi_{j+1}(p(yzy'z'|xx'ss')p(xx')p(ss'))\big) \subseteq \\
\big(\bigcap_{p(s)} \Delta\big(\varphi_j(p(y, z|x, s)p(s)p(x))\big)\big) \oplus \big(\bigcap_{p(s')} \Delta\big(\varphi_1(p(s').p(x').q(y', z'|x', s'))\big)\big).
\end{gathered}$$

In all of the three terms the expression begins by taking intersection over choices of the distribution of the state variable. Take some arbitrary $p(s)$ and $p(s')$ on the right hand side. In the above inequality we are taking infimum over all possible joint distributions of SS'; if we restrict to $p(ss') = p(s).p(s')$ the expression on the left hand side would increase. Therefore the above inequality would be valid if one can show the following:

$$\begin{gathered}\Delta\big(\varphi_{j+1}\big(p(yzy'z'|xx'ss')p(xx')p(s)p(s')\big)\big) \subseteq \\ \Delta\big(\varphi_j\big(p(y,z|x,s)p(s)p(x)\big)\big) \oplus \Delta\big(\varphi_1\big(p(s').p(x').q(y',z'|x',s')\big)\big).\end{gathered}$$

Now, take an arbitrary point $\overrightarrow{v}$ inside $\varphi_{j+1}\big(p(yzy'z'|xx'ss')p(xx')p(s)p(s')\big)$. We would like to prove that there exists $\overrightarrow{v}_1 \in \Delta\big(\varphi_1\big(p(s').p(x').q(y',z'|x',s')\big)\big)$ and $\overrightarrow{v}_2 \in \Delta\big(\varphi_j\big(p(s)p(x)p(yz|xs)\big)\big)$ such that $\overrightarrow{v}_1 + \overrightarrow{v}_2 \geq \overrightarrow{v}$

There exists some U, V, W_0, W_1, W_2 created from XX' satisfying $UVW_0W_1W_2 - XX' - SS'XX' - YY'ZZ'$, and with W_0, W_1, W_2 being independent of each other and X being a deterministic function of (W_0, W_1, W_2, U, V) such that

$$\begin{gathered}\overrightarrow{v} = \Big(I(W_0;YY'|U), \\ I(W_0,ZZ'|V), \\ I(W_1;YY'|U)), \\ I(W_2;ZZ'|V)), \\ I(W_0W_1;YY'|U), \\ I(W_1;YY'|W_0UV) + I(W_0U;ZZ'|V)), \\ I(W_0W_2;ZZ'|V), \\ I(W_2;ZZ'|W_0UV) + I(W_0V;YY'|U)), \\ I(W_1;YY'|W_0W_2UV) + I(W_0W_2U;ZZ'|V), \\ I(W_2;ZZ'|W_0W_1UV) + I(W_0W_1V;YY'|U), \\ I(W_0UV;YY') + I(W_1;YY'|W_0W_2UV) + I(W_2;ZZ'|W_0UV), \\ I(W_0UV;ZZ') + I(W_2;ZZ'|W_0W_1UV) + I(W_1;YY'|W_0UV)\Big)\end{gathered}$$

Let $W_0' = \widetilde{W_0} = W_0$, $W_1' = \widetilde{W_1} = W_1$, $W_2' = \widetilde{W_2} = W_2$, $V' = VY$, $U' = U$, $\widetilde{V} = V$ $\widetilde{U} = UZ'$. The following properties hold:

- The Markov chain $U'V'W_0'W_1'W_2' - X' - X'S' - Y'Z'$ holds.
- W_0', W_1', W_2' are independent of each other; X' is a deterministic function of $(W_0', W_1', W_2', U', V')$
- The Markov chain $\widetilde{U}\widetilde{V}\widetilde{W_0}\widetilde{W_1}\widetilde{W_2} - X - XS - YZ$ holds.
- $\widetilde{W_0}, \widetilde{W_1}, \widetilde{W_2}$ are independent of each other; X is a deterministic function of $(\widetilde{W_0}, \widetilde{W_1}, \widetilde{W_2}, \widetilde{U}, \widetilde{V})$

We can define two points $\overrightarrow{v}_1 \in \Delta\big(\varphi_1(p(s').p(x').q(y', z'|x', s'))\big)$ and $\overrightarrow{v}_2 \in \Delta\big(\varphi_j(p(yz|xs)p(s)p(x))\big)$ using the above auxiliary random variables. It can be easily seen that $\overrightarrow{v}_1 + \overrightarrow{v}_2$ is coordinate by coordinate greater than or equal to $\overrightarrow{v}$. ●

Property number 2. This is clear from the definition of φ_j.

Property number 3. This is clear from the definition of φ_j.

Property number 4. We need to show that for any $(R_0, R_1, R_2) \in C_R(p(x)p(y, z|x, s))$:

$$(R_0, R_0, R_1, R_2, R_0 + R_1, R_0 + R_1, R_0 + R_2, R_0 + R_2,$$

$$R_0 + R_1 + R_2, R_0 + R_1 + R_2, R_0 + R_1 + R_2, R_0 + R_1 + R_2) \in \bigcap_{p(s)} \Delta\big(\varphi_j(p(yz|xs)p(s)p(x))\big).$$

Take some arbitrary $(R_0, R_1, R_2) \in C_R(p(x)p(y, z|x, s))$. We know that there is $p(w_0, w_1, w_2, x) \in \Omega$ with marginal $p(x)$ such that for every $p(s)$ when we have $(W_0, W_1, W_2, X, S, Y, Z) \sim p(w_0, w_1, w_2, x).p(s).p(y, z|x, s)$ we have inequalities (1) with right hand side $p(s)$. It is necessary and sufficient to prove that for any $p(s)$,

$$(R_0, R_0, R_1, R_2, R_0 + R_1, R_0 + R_1, R_0 + R_2, R_0 + R_2,$$

$$R_0 + R_1 + R_2, R_0 + R_1 + R_2, R_0 + R_1 + R_2, R_0 + R_1 + R_2) \in \Delta\big(\varphi_j(p(yz|xs)p(s)p(x))\big).$$

Let $W_0' = W_0$, $W_1' = W_1$, $W_2' = W_2$, $U' = \emptyset$ and $V' = \emptyset$. It can be seen that $p(w_0', w_1', w_2', u', v', x, s, y, z) \in \Upsilon_{p(yz|xs)p(s)p(x)}$. This joint distribution in $\Upsilon_{p(yz|xs)p(s)p(x)}$ specifies a point in $\varphi_j(p(yz|xs)p(s)p(x))$ which pointwise dominates the point $\{(R_0, R_0, R_1, R_2, R_0 + R_1, R_0 + R_1, R_0 + R_2, R_0 + R_2, , R_0 + R_1 + R_2, R_0 + R_1 + R_2, R_0 + R_1 + R_2, R_0 + R_1 + R_2)\}$ ●

V. Acknowledgement

The authors would like to thank TRUST (The Team for Research in Ubiquitous Secure Technology), which receives support from the National Science Foundation (NSF award number CCF-0424422) and the following organizations: Cisco, ESCHER, HP, IBM, Intel, Microsoft, ORNL, Pirelli, Qualcomm, Sun, Symantec, Telecom Italia and United Technologies, for their support of this work. The research was also partially supported by NSF grants CCF-0500023, CCF-0635372, and CNS-0627161.

VI. APPENDIX

In this appendix we show that $\phi(q(y,z|x,s))$ is convex for each $q(y,z|x,s)$.

Take two vectors $\overrightarrow{a}=(a_0,a_1,a_2,...,a_{11})$ and $\overrightarrow{b}=(b_0,b_1,b_2,...,b_{11})$ in this region. Since $\overrightarrow{a}$ is inside $\phi(q(y,z|x,s))$, distribution $p(x_a)$ exists such that for all $p(s_a)$ a distribution

$$p(u_a v_a w_{0a} w_{1a} w_{2a} x_a)p(s_a)q(y_a,z_a|x_a,s_a)\in\Upsilon_{p(x_a)p(s_a)q(y_a,z_a|x_a,s_a)}$$

exists for which the following inequalities are satisfied:

- $a_0\le I(W_{0a};Y_a|U_a)$
- $a_1\le I(W_{0a},Z_a|V_a)$
- $a_2\le I(W_{1a};Y_a|U_a))$
- $a_3\le I(W_{2a};Z_a|V_a))$
- $a_4\le I(W_{0a}W_{1a};Y_a|U_a)$
- $a_5\le I(W_{1a};Y_a|W_{0a}U_aV_a)+I(W_{0a}U_a;Z_a|V_a))$
- $a_6\le I(W_{0a}W_{2a};Z_a|V_a)$
- $a_7\le I(W_{2a};Z_a|W_{0a}U_aV_a)+I(W_{0a}V_a;Y_a|U_a))$
- $a_8\le I(W_{1a};Y_a|W_{0a}W_{2a}U_aV_a)+I(W_{0a}W_{2a}U_a;Z_a|V_a)$
- $a_9\le I(W_{2a};Z_a|W_{0a}W_{1a}U_aV_a)+I(W_{0a}W_{1a}V_a;Y_a|U_a)$
- $a_{10}\le I(W_{0a}U_aV_a;Y_a)+I(W_{1a};Y_a|W_{0a}W_{2a}U_aV_a)+I(W_{2a};Z_a|W_{0a}U_aV_a)$
- $a_{11}\le I(W_{0a}U_aV_a;Z_a)+I(W_{2a};Z_a|W_{0a}W_{1a}U_aV_a)+I(W_{1a};Y_a|W_{0a}U_aV_a)$

A similar statement holds for $(b_0,b_1,...,b_{11})$ involving random variables $(U_b,V_b,W_{0b},W_{1b},W_{2b},X_b,S_b,Y_b,Z_b)$.

Let $p(\widetilde{x})=0.5*p(x_a)+0.5*p(x_b)$. Take an arbitrary $p(s)$. We consider the case $S_a\sim p(s)$ and $S_b\sim p(s)$.

Without loss of generality, one can assume that $(U_a,V_a,W_{0a},W_{1a},W_{2a},X_a,S_a,Y_a,Z_a)$ is independent of $(U_b,V_b,W_{0b},W_{1b},W_{2b},X_b,S_b,Y_b,Z_b)$.

Take a binary and uniform random variable T on $\{0,1\}$ that is independent of all the above mentioned random variables and let $(U,V,W_0,W_1,W_2,X,S,Y,Z)$ be equal to

$$(TU_a,TV_a,TW_{0a}W_{0b},W_{1a}W_{1b},W_{2a}W_{2b},X_a,S_a,Y_a,Z_a)$$

if $T=0$,
and be equal to

$$(TU_b,TV_b,TW_{0a}W_{0b},W_{1a}W_{1b},W_{2a}W_{2b},X_b,S_b,Y_b,Z_b)$$

if $T=1$.

X has the distribution $p(x)$ we started with and S has the distribution $p(s)$ we started with. Furthermore, one can verify that $p(yz|xs) = q(yz|xs)$ and that random variables $(U, V, W_0, W_1, W_2, X, S, Y, Z)$ have joint distribution $p(u, v, w_0, w_1, w_2, x)p(s)q(yz|xs)$ belonging to $\Upsilon_{p(s)p(x)q(yz|xs)}$.

It can be verified that this choice of variables gives us a point in $\phi(q(y, z|x, s))$ that pointwise dominates $\frac{1}{2}\overrightarrow{a} + \frac{1}{2}\overrightarrow{b}$. The proof is complete when we note that our choice of $p(s)$ was arbitrary.

References

[1] Amin A. Gohari and V. Anantharam, "Information-Theoretic Key Agreement of Multiple Terminals – Part I: Source Model," *Preprint*, Dec. 2007. Available at http://www.eecs.berkeley.edu/∼aminzade/SourceModel.pdf

[2] Amin A. Gohari and V. Anantharam, "Information-Theoretic Key Agreement of Multiple Terminals – Part II: Channel Model," *Preprint*, Dec. 2007. Available at http://www.eecs.berkeley.edu/∼aminzade/ChannelModel.pdf

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

[4] K Marton, "A coding theorem for the discrete memoryless broadcast channel," IEEE Trans. Inform. Theory, 25(3): 306-311 (1979).

[5] C. Nair and A. A. El Gamal, "An outer bound to the capacity region of the broadcast channel," IEEE Trans. Inform. Theory, 53(1): 350-355 (2007).

[6] J. Jahn, "Coding of arbitrarily varying multiuser channels," IEEE Trans. Inform. Theory, 27(2): 212-226 (1981).

[7] E. Hof, Shraga I. Bross: "On the Deterministic-Code Capacity of the Two-User Discrete Memoryless Arbitrarily Varying General Broadcast Channel With Degraded Message Sets," IEEE Trans. Inform. Theory 52(11): 5023-5044 (2006)

[8] Y. Liang, G. Kramer, "Rate Regions for Relay Broadcast Channels," IEEE Transactions on Information Theory 53(10): 3517-3535 (2007)

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.