ACTA MATHEMATICA VIETNAMICA
Volume 17, Number 1, 1992

SEMI-CONTINUOUS ARBITRARILY VARYING CHANNELS WITH GENERAL STATE CONSTRAINTS

BUI VAN THANH

Introduction

This paper is a continuation of our earlier paper [3], where the concept of capacity of arbitrarily varying channels under general state constraints was introduced. The state constraints were expressed in terms of types of state sequences. The problem rose from earlier papers of Csiszár and Narayan [1], [2], who considered the state constraints given in terms of a function. The motivation was the need of study of some communication situations which involve several state constraints of such type. We also gave the exact formula for the capacity of discrete AVC's with finite state set.

The AVC's having continuous alphabets and set of states are the most important, but they are relatively less understood. In this paper we shall extend the results of [3] to memoryless semi-continuous AVC's, that is, the AVC's with finite alphabet $\mathcal{X}$ and general output alphabet $\mathcal{Y}$ and state set $\mathcal{S}$.

Dropping the assumption on finiteness of $\mathcal{S}$ and $\mathcal{Y}$ presents no difficulties. Dealing with infinite $\mathcal{X}$ is more difficult and to get satisfactory results for that case we shall need strong regularity assumptions. This is not suprising because for infinite input alphabets, no general solution is known, even to the simpler compound channel capacity problem.

Given the (memoryless) semi-continuous $AVC\{\mathcal{W}\}$ with finite input alphabet $\mathcal{X}$, general output alphabet $\mathcal{Y}$ and the set of states $\mathcal{S}$ we shall adopt the following hypotheses.

Received March 14, 1991

(H.1) $\mathcal{X}$ is a finite set, while $(\mathcal{Y}, \mathcal{B}), (\mathcal{S}, \mathcal{C})$ are Polish spaces, i.e. separable, complete metric spaces, where $\mathcal{B}$ and $\mathcal{C}$ are the σ- algebras of all Borel subsets of $\mathcal{Y}$ and $\mathcal{S}$, respectively. The distributions on these spaces are defined on their σ-algebras of Borel subsets. The topology for distributions will be of weak convergence. Thus a sequence of distributions Q_n on $\mathcal{S}$ is said to be convergent to the distribution Q if and only if

$$\int f(s)Q_n(ds) \to \int f(s)Q(ds) \tag{1}$$

for all continuous bounded real-valued funtions f defined on S.

(H.2) $W(\cdot|x,s)$, as a distribution-valued function of s, is continuous for every fixed $x \in \mathcal{X}$, i.e., for every $x \in \mathcal{X}, s_n \to s$ implies the convergence of $W(\cdot|x, s_n)$ to $W(\cdot|x,s)$ in the above sense.

We recall that the space $\mathcal{M}(\mathcal{S})$ of all probability measures on a Polish space $(\mathcal{S}, \mathcal{C})$ endowed with weak topology is separable and metrizable by the *Prohorov* metric, i.e. the metric defined by

$$d(Q,Q') = \inf\{\epsilon : Q(B) \leq Q'(B^\epsilon) + \epsilon, Q'(B) \leq Q(B^\epsilon) + \epsilon\}, \tag{2}$$

where $B^\epsilon = \{s : \min_{s' \in B} \rho(s,s') = \rho(s,B) < \epsilon\}$ and ρ is the metric on $\mathcal{S}$ (Billingsley [5]). Clearly, a sequence of distributions Q_n converges to Q in this topology exactly when (1) is satisfied.

Note that Hypothesis (H.2) is weaker than the continuity of $W(A|x,s)$ as a function of s for every fixed x and Borel set A. For example, the *semi-continuous noiseless adder* AVC defined by $\mathcal{Y} = \mathcal{S} = R, \mathcal{X} \subset R$ being some *finite* set of R and $W(\cdot|x,s)$ - the point mass at $x+s$, obviously satisfies (H.2), while $W(A|x,s)$ as a function of s has discontinuity for every non-trivial set $A \subset R$.

(H.3) State constraints are defined in terms of a *convex, compact* (in weak topology) subset Π of $\mathcal{M}(\mathcal{S})$ which has non-empty interior in $\mathcal{M}(\mathcal{S})$. A state sequence $s = (s_1, \ldots, s_n)$ will be called an *admissible* state sequence if its empirical distribution defined by

$$P_s(B) = \frac{1}{n}\sum_{i=1}^{n} \chi_{s_i}(B), B \in C$$

belongs to Π, where $\chi_{s_i}(\cdot)$ denotes the point mass at s_i.

The concept of codes, achievable rates and capacity under state constraint Π can be now extended to semi-continuous AVC's in a straightfoward manner. Namely, if a subset $\Pi \subset \mathcal{M}(\mathcal{S})$ is given, the average error probability of the code (f, ϕ) under state constraint Π is defined by

$$\bar{e}(\Pi) = \bar{e}(\Pi; f, \phi) = \max_{s: P_s \in \Pi} \sum_{m=1}^{|\mathcal{M}|} W^n(\{\phi(Y) \neq m\}|f(m), s). \tag{3}$$

A number $R > 0$ is called an *achievable* rate for the given AVC under state constraint Π if for every $\epsilon > 0, \delta > 0$ and all sufficiently large n there exists an n-length block code with N codewords and

$$\frac{1}{n}\log N > R - \delta, \qquad \bar{e}(\Pi) \leq \epsilon. \tag{4}$$

The supremum of such achievable rates is called the *capacity* of the AVC under state constraint Π and will be denoted by $C(\Pi)$.

For any distribution Q on S let $W_Q(\cdot|x)$ denote the distribution on $\mathcal{Y}$ defined by

$$W_Q(\cdot|x) = \int W(\cdot|x, s)Q(ds). \tag{5}$$

Then, for any input distribution P on $\mathcal{X}$ we define $I(P, W_Q)$ as the mutual information of random variables X and Y such that the distribution of X is P and the conditional distribution of Y given X is W_Q. For finite input alphabet $\mathcal{X}$, $I(P, W_Q)$ is concave continuous in P and convex lower-semicontinuous in Q. For basic properties of the mutual information of random variables see e.g. Pinsker [7].

As in the discrete case, we now define

$$I(P, \Pi) = \min_{Q \in \Pi} I(P, W_Q). \tag{6}$$

Here a lower-semicontinuous function is to be minimized over a compact set, thus the minimum is always attained. Since $I(P,\Pi)$ is the minimum of concave functions $I(P, W_Q)$, it is a concave upper semi-continuous function of P. For a given semi-continuous AVC we define the *symmetrization set* $\mathcal{U}$ as the set of all $|\mathcal{X}|$-tuples of distributions $U(\cdot|x)$ on $\mathcal{S}$ which satisfy the condition

$$\int W(\cdot|x,s)U(ds|x') = \int W(\cdot|x',s)U(ds|x) \tag{7}$$

for every $x, x' \in \mathcal{X}$. This equality of distributions is interpreted either as the equality of both sides when replacing the dot by any Borel set in $\mathcal{B}$ or, equivalently, as the equality of integrals

$$\int\int g(y)U(ds|x')W(dy|x,s) = \int\int g(y)U(ds|x)W(dy|x',s) \tag{8}$$

for every *bounded, uniformly continuous function* g defined on $\mathcal{Y}$ (Parthasarsthy [6, Ch. II, Theorem 5.9]). Also, for any $\alpha > 0$ we define $\Pi(\alpha)$ as the open α-neighborhood of Π :

$$\Pi(\alpha) = \{Q : \min_{Q' \in \Pi} d(Q,Q') = d(Q,\Pi) < \alpha\}. \tag{9}$$

For every input distribution P on $\mathcal{X}$ and any $U \in \mathcal{U}, PU$ denotes the marginal distribution on $\mathcal{S}$:

$$PU(\cdot) = \sum_{x \in \mathcal{X}} P(x)U(\cdot|x).$$

2. Semi-continuous AVC's with general state constraints

First, we consider the case when $\mathcal{Y}$ is finite and only $\mathcal{S}$ is allowed to be a Polish space. In this case Hypothesis (H.2) is reduced to a simpler form. Namely, it means that for every fixed $x \in \mathcal{X}, y \in \mathcal{Y}, W(y|x,s)$ is a continuous function of s. Further, equation (8) means that

$$\int W(y|x,s)U(ds|x') = \int W(y|x',s)U(ds|x) \tag{10}$$

for every $x, x' \in \mathcal{X}$ and $y \in \mathcal{Y}$.

We shall use the following key theorem proved in [3]. For any discrete AVC $\{\mathcal{W}\}$, let $\mathcal{U}$ denote its symmetrization set, i.e. the set of the channels satisfying (10) with sums instead of integrals. Further, $\Pi(\alpha)$ denotes the open α-neighborhood of Π with the usual Euclidean distance.

THEOREM KT. *For any $\beta > 0, \alpha > 0$, let P be a type of input sequences such that $\min_{x\in\mathcal{X}} P(x) \geq \beta$ and $PU \notin \Pi(\alpha)$ for every $U \in \mathcal{U}$. Then for any $\delta > 0$, there exist numbers $\gamma = \gamma(\alpha, \beta, \delta) > 0$ and $n_0 = n_0(\alpha, \beta, \delta)$ such that for all $n \geq n_0$, there exists a code with N codewords $x_1, x_2, \ldots, x_N$, each of type P, satisfying*

$$\frac{1}{n}\log N > I(P, \Pi) - \delta, \qquad \bar{e}(\Pi) \leq \exp(-n\gamma).$$

The existence of good code in Theorem KT is guaranteed by the following fact ([2, Lemma 3]): For any $\epsilon > 0, n \geq n_0(\epsilon), N \geq \exp(n\epsilon)$ and type P, there exist N codewords $x_1, x_2, \ldots, x_N$ in $\mathcal{X}^n$, each of type P, such that for every $x \in \mathcal{X}^n, s \in \mathcal{S}^n$ and every joint type $P_{XX'S}$ with $P_X = P_{X'} = P$ we have

a) $|\{j : (x, x_j, s) \in \tau_{XY}\}| \leq \exp\{n(|R - I(X' \wedge XS)|^+ + \epsilon)\}$,

b) $\frac{1}{N}|\{j : (x_j, s) \in \tau_{XS}\}| \leq \exp(-n\epsilon/2)$ if $I(X \wedge S) > \epsilon$,

c) $\frac{1}{n}|j : (x_j, x_i, s) \in \tau_{XX'S}$ for some $i \neq j\}| \leq \exp(-n\epsilon/2)$

if $I(X \wedge X'S) > |R - I(X' \wedge S)|^+ + \epsilon$, where $R = (1/n)\log N$ and $a^+ = \max(a, 0)$.

It is worthy to observe that Lemma 3 of [2] was proved as follows. The codewords $x_1, x_2, \ldots, x_N$ were chosen randomly and independently according to the uniform distribution on the set τ_X, the set of all sequences x having type P. The sequences $x \in \mathcal{X}^n$ and $s \in \mathcal{S}^n$ were first assumed to be fixed. Then, it was shown that the probability that a), b), c) do not simultaneously hold tends to 0 doubly exponentially. As the number of all possible combinations of sequences $x \in \mathcal{X}^n, s \in \mathcal{S}^n$ and joint types $P_{XX'S}$ only grows exponentially with n, it was ensured that the probability of the simultaneous validity of a), b) and c) tends to 1 as $n \to \infty$. Therefore any realization of the random choosing satisfying all these inequalities is a proper choise for $\{x_1, x_2, \ldots, x_N\}$.

Now if we drop the assumption that $\mathcal{S}$ is fixed and allow that $\mathcal{S} = \mathcal{S}_{(n)}$ depends on n, but its cardinality $|\mathcal{S}_{(n)}|$ only polynomially grows with n, the number of all possible combinations of sequences and joint types still grows slower than doubly exponentially. Thus the previous assertion remains valid.

After these preparations we can state the following theorem.

THEOREM 1. *Let a discrete* $AVCW = \{W(y|x,s)\}$ *be given with assumptions that the set of states* $\mathcal{S}$ *is a Polish space and the set of state constraints* Π *is convex and compact in the weak topology. Further let* P *be any strictly positive type on* $\mathcal{X}$ *for which* $\alpha > 0$ *can be found such that* $PU \notin \Pi(\alpha)$ *for each* $U \in \mathcal{U}$, *where* $\mathcal{U}$ *and* $\Pi(\alpha)$ *are defined as in (10) and (9). Then for any* $\delta > 0$ *there exists a number* $\gamma > 0$ *such that for all sufficiently large n there exists a length n bloc code* (f, ϕ) *with* N *codewords* $x_1, x_2, \ldots, x_N$, *each of type P, such that*

$$\frac{1}{n}\log N > I(P,\Pi) - \delta, \quad \max_{s:P_s \in \Pi} \bar{e}(s) \leq \exp(-n\gamma).$$

PROOF: The idea is to reduce the problem to the case of finite set of states by an approximation argument. First, let $\mathcal{D}$ be any finite subset of $\mathcal{S}$. Define the set $\Pi_{\mathcal{D}}$ as the subset of Π whose elements have support contained in $\mathcal{D}$. It is easy to show that $\Pi_{\mathcal{D}}$ is compact, and hence closed in the topology generated by the usual Euclidean metric.

Next, let $\mathcal{U}_{\mathcal{D}}$ be the set of channels U : $\mathcal{X} \to \mathcal{D}$ which satisfy

$$\sum_{s\in\mathcal{D}} W(y|x,s)U(s|x') = \sum_{s\in\mathcal{D}} W(y|x',s)U(s|x) \tag{11}$$

for every $x, x' \in \mathcal{X}$ and $y \in \mathcal{Y}$. From (7) and (11) we see that $\Pi_{\mathcal{D}}(\alpha) \subset \Pi(\alpha)$. Thus $PU \notin \Pi_{\mathcal{D}}(\alpha)$ for every $U_p \in \mathcal{U}$, therefore $\Pi_{\mathcal{D}}(\alpha)$ also means the open α-neighborhood of $\Pi_{\mathcal{D}}$ with the Euclidean metric.

We now follow the approximation argument in Csiszár and Korner [4]. For fixed n we subdivide the $|\mathcal{X}||\mathcal{Y}|$-dimensional unit cube into sub-cubes of edge of length n^{-4} and pick a matrix W in such a way that each sub-cube contains at least one of such matrices. Let $\mathcal{S}_{(n)}$ be the set of indices of the matrices picked in this way. The set $S_{(n)}$ has size $|\mathcal{S}_{(n)}| \leq n^{4|\mathcal{X}||\mathcal{Y}|}$. Thus for $\delta > 0$ and for n large enough, we apply Theorem KT with the chosen set $\mathcal{S}_{(n)}$

playing the role of the set $\mathcal{D}$ above to get a code (f_n, ϕ_n) with N codewords, each of type P, such that

$$\frac{1}{n}\log N > I(P, \Pi_{\mathcal{S}_{(n)}}) - \delta, \qquad \max_{s:P_s \in \Pi_{\mathcal{S}_{(n)}}} \overline{e}(s) \le \exp(-n\gamma), \tag{12}$$

where γ is a positive number. Since $\Pi_{\mathcal{S}_{(n)}} \subset \Pi$, we have $(1/n)\log N > I(P, \Pi) - \delta$.

For the code (f_n, ϕ_n) and for every $s \in \mathcal{S}^n$, there exists an $s' \in \mathcal{S}^n_{(n)}$ satisfying

$$e_m(s; f_n, \phi_n) = \sum_{y:\phi(y)\neq m} W^n(y|f_n(m), s) \le \sum_{y:\phi(y)\neq m} W^n(y|f_n(m), s') + \frac{|\mathcal{Y}|}{n}.$$

Therefore

$$\max_{s:P_s \in \Pi} \overline{e}(s) \le 2 \max_{s':P_{s'} \in \Pi_{\mathcal{S}^n_{(n)}}} \overline{e}(s') + \frac{|\mathcal{Y}|}{n}. \tag{13}$$

From (12) and (13) we obtain that for $n \ge n_1(\epsilon), \max_{s:P_s \in \Pi} \overline{e}(s) \le \epsilon$. This proves our Theorem.

We now turn to the general case when both $\mathcal{Y}$ and $\mathcal{S}$ are Polish spaces. Consider a partition $\mathcal{A} = (A_1, \ldots, A_k)$ of $\mathcal{Y}$ into disjoint Borel sets. By $W^{\mathcal{A}}$ we denote the quantized AVC defined by replacing $\mathcal{Y}$ by the set $\mathcal{Y}' = \{1, 2, \ldots, k\}$ and setting

$$W^{\mathcal{A}}(j|x, s) = W(A_j/x, s), \quad j = 1, \ldots, k. \tag{14}$$

Further, for any distribution Q on $\mathcal{S}$, let $W_Q^{\mathcal{A}} : \mathcal{X} \to \mathcal{Y}'$ be the discrete channel defined by

$$W_Q^{\mathcal{A}}(j|x) = W_q(A_j|x) = \int W(A_j|x, s)Q(ds) \quad j = 1, \ldots, k \tag{15}$$

and set

$$I^{\mathcal{A}}(P, \Pi) = \min_{Q \in \Pi} I(P, W_Q^{\mathcal{A}}) \tag{16}$$

LEMMA 1. $I(P, \Pi) = \sup I^{\mathcal{A}}(P, \Pi)$, *where the supremum is taken over all partitions* $\mathcal{A}$ *of* $\mathcal{Y}$.

PROOF: First we observe that by definition, $I(P, W_Q)$ is the supremum of discrete approximations $I(P, W_Q^{\mathcal{A}})$ for all finite partitions $\mathcal{A}$ of $\mathcal{Y}$ into disjoint Borel sets. It may be not possible, however, to find partitions $\mathcal{A}$ that make $I(P, W_Q^{\mathcal{A}})$ uniformly close to $I(P, W_Q)$ for all Q in Π. To circumvent this difficulty we notice that by the Dobrushin theorem ([7. Theorem 2.1.1]), for a fixed $Q, I(P, W_Q) = \sup_{\mathcal{A}} I(P, W_Q^{\mathcal{A}})$ holds even when the supremum is restricted to those partitions whose atoms A_j are $W_Q(.|x)$- continuity sets for every j and $x \in \mathcal{X}$. For fixed Q and any η there exists a partition of the latter kind such that

$$I(P, W_Q^{\mathcal{A}}) > I(P, W_Q) - \eta/2 \geq I(P, \Pi) - \eta/2. \tag{17}$$

As $W_Q(.|x)$ depends continuously on Q, every $Q \in \Pi$ has a neighborhood such that if Q' is in it, we have

$$I(P, W_{Q'}^{\mathcal{A}}) > I(P, \Pi) - \eta. \tag{18}$$

Since Π is compact, it follows that a finite number of these neighborhoods with "centers" $Q_1, Q_2, \ldots, Q_M$, say, exist, whose union still covers Π. Then, by letting $\mathcal{A}$ to be a common refinement of the partitions corresponding to $Q_1, \ldots, Q_M$,(18) will hold for every $Q' \in \Pi$. This proves that

$$I(P, \Pi) \leq \sup_{\mathcal{A}} I^{\mathcal{A}}(P, \Pi).$$

Since the opposite inequality is obvious, our lemma is proved.

Now for any partition $\mathcal{A} = (A_1, \ldots, A_k)$ of $\mathcal{Y}$, let $\mathcal{U}^{\mathcal{A}}$ denote the set of all $|\mathcal{X}|$-tuples of distributions $U(\cdot|x)$ on S such that for all atoms A_j of $\mathcal{A}$ and every x, x' in $\mathcal{X}$,

$$\int W(A_j|x, s)U(ds|x') = \int W(A_j|x', s)U(ds|x). \tag{19}$$

Clearly, $\mathcal{U} \subset \mathcal{U}^{\mathcal{A}}$ for every partition $\mathcal{A}$ of $\mathcal{Y}$. We can now state our main theorem for semicontinuous AVC's.

THEOREM 2. *Let a semi-continuous $AVC\{W(\cdot|x, s)\}$ be given with the state constraint set Π which is convex and compact. Suppose that $\mathcal{Y} = R^m$, the* *m-dimensional Euclidian space. Further, let P be a strictly positive type on $\mathcal{X}$ for which $PU \notin \Pi(\alpha)$ for every $U \in \mathcal{U}$ and for some $\alpha > 0$. Then for any $\delta > 0$ there is a number $\gamma > 0$ such that for all sufficiently large n there exists a length n block code with N codewords, each of type P, such that*

$$\frac{1}{n}\log N > I(P,\Pi) - \delta, \quad \max_{s:P_s \in \Pi} \bar{e}(s) \leq (-n\gamma).$$

PROOF: We will reduce the problem to the finite output alphabet case. For any $\delta > 0$, let $\mathcal{A}_0$ be a partition for which

$$I^{\mathcal{A}_0}(P,\Pi) > I(P,\Pi) - \delta/2.$$

We claim that there exists $\alpha' > 0$ and a partition $\mathcal{A}_1$ of $\mathcal{Y}$ which is a refinement of $\mathcal{A}_0$ such that $PU \notin \Pi(\alpha')$ for every $U \in \mathcal{U}^{\mathcal{A}_1}$. Proceeding indirectly, suppose that for every n and every partition $\mathcal{A}$ of $\mathcal{Y}$ which is a refinement of $\mathcal{A}_0$, there is $U \in \mathcal{U}^{\mathcal{A}}$ satisfying $PU \in \Pi(1/n)$.

First, let $M_n \subset R^m, n = 1,2,...,$ be the closed spheres of center 0 and radius n. By $\mathcal{A}_n$ we denote the partition $(A_{n0}, A_{n1}, ..., A_{nk})$ of R^m , where $A_{n0} = R^M - M_n$, $\cup_{j=1}^k A_{nj} = M_n$ and each of A_{nj} , $j = 1,2,...,k$, has diameter less than $1/n$. Making the common refinement from $\mathcal{A}_0$ and $\mathcal{A}_n$ if necessary, we can suppose that $\mathcal{A}_n$ is a refinement of $\mathcal{A}_0$. By indirect assumption we can choose a subsequence of $U_n \in \mathcal{U}^{\mathcal{A}_n}$ such that $Q_n = PU_n \in \Pi(1/n)$. Since Π is compact, we can choose a subsequence of Q_n which converges to some $Q \in \Pi$. Without loss of generality we can assume that Q_n converges to Q. Then the sequence PU_n converges to Q, too. Now, since Q and PU_n are individually tight and PU_n is convergent, the sequence PU_n is uniformly tight, (Billingsley, [5]). Thus, for any $\epsilon > 0$, a compact set $K \in \mathcal{C}$ can be found such that $PU_n(K) > 1 - \epsilon$ for every n. Denoting $\beta = \min\{P(x), x \in \mathcal{X}\} > 0$, we have

$$U_n(\mathcal{S} - K|x) \leq \frac{1}{\beta}\sum P(x)U_n(\mathcal{S} - K|x) < \epsilon/\beta.$$

This means that the family $U_n(.|x), x \in \mathcal{X}$, is also tight. Then we can choose a subsequence U_{n_j} of U_n converging to some channel U_0 and $PU_0 = Q \in \Pi$.

Again, without loss of generality we can assume that $U_n \to U_0 = Q$. Therefore $PU_n \to PU_0 = Q$. We shall show $U \in \mathcal{U}$ by verifying (8).

Let g be any bounded, uniformly continuous function on $\mathcal{Y}$ and let $\epsilon > 0$ be arbitrary. Since $U_n \to U_0$, it follows that the family $\{U_0, U_1, ...\}$ is tight. Thus, for $\epsilon > 0$ we can find a compact set $K_\epsilon \subset \mathcal{S}$ such that

$$U_n(\mathcal{S} - K_\epsilon | x) < \epsilon \tag{20}$$

for each n and $x \in \mathcal{X}$. Then, as K_ϵ is compact, the family $\{W(\cdot|x,s), x \in \mathcal{X}, s \in K_\epsilon\}$ is tight. Thus, we can choose a compact set $M_\epsilon \subset R^n$ such that

$$W(R^m - M_\epsilon | x, s) < \epsilon \tag{21}$$

for every $x \in \mathcal{X}, s \in K_\epsilon$.

Choose n so large that $|g(y) - g(y')| < \epsilon$ for all $y, y' \in R^m$ satisfying $|y - y'| < 1/n$. Then for $n \geq n_0$, where $n_0 = n_0(\epsilon)$ is sufficiently large, we have

$$\left| \int_{M_n} \int_{\mathcal{S}} W(dy|x,s) U_n(ds|x') - \sum_{j=1}^{k} g_j \int W(A_{nj}|x',s) U_n(ds|x) \right| < \frac{\epsilon}{2},$$

$$\left| \int_{M_n} \int_{\mathcal{S}} g(y) W(dy|x',s) U_n(ds|x) - \sum_{j=1}^{k} g_j \int W(A_{nj}|x,s) U_n(ds|x') \right| < \frac{\epsilon}{2}$$

for all $x, x' \in \mathcal{X}$, where $g_j = g(y_j)$ for some $y_j \in A_{nj}, j = 1, ..., k$. Since $U_n \in \mathcal{U}^{A_n}$, we obtain

$$\left| \int_{M_n} \int_{\mathcal{S}} g(y) W(dy|x,s) U_n(ds|x') - \int_{M_n} \int_{\mathcal{S}} g(y) W(dy|x',s) U_n(ds|x) \right| < \epsilon \tag{22}$$

if $n \geq n_0$. From (20) and (21) we have

$$\left| \int_{R^m - M_\epsilon} \int_{\mathcal{S}} U_n(ds|x') W(dy|x,s) + \int_{R^m - M_\epsilon} \int_{\mathcal{S}} U_n(ds|x) W(dy|x',s) \right| \leq$$

$$\leq \left| \int_{K_\epsilon} \int_{R^m - M_\epsilon} W(dy|x,s) U_n(ds|x') + \int_{K_\epsilon} \int_{R^m - M_\epsilon} W(dy|x',s) U_n(ds|x) \right| +$$

$$+ \left| \int_{\mathcal{S} - K_\epsilon} \int_{R^m - M_\epsilon} W(dy|x,s) U_n(ds|x') + \int_{\mathcal{S} - K_\epsilon} \int_{R^m - M_\epsilon} W(dy|x',s) U_n(ds|x) \right|$$

$$\leq \left| \int_{S-K_\epsilon} U_n(ds|x') + \int_{S-K_\epsilon} U_n(ds|x) \right| + 2\epsilon \leq 4\epsilon. \quad (23\ )$$

Further, we observe that every compact set in R^m is bounded, hence for sufficiently large n, $M_\epsilon \subset M_n$. From (23), letting $|g| \leq L$, we obtain

$$\left| \int_{R^m - M_n} \int_S g(y) U_n(ds|x') W(dy|x,s) - \int_{R^m - M_n} \int_S g(y) U_n(ds|x) W(dy|x',s) \right| \leq$$

$$\leq L \left| \int_{R^m - M_\epsilon} \int_S U_n(ds|x') W(dy|x,s) + \int_{R^m - M_\epsilon} \int_S U_n(ds|x) W(dy|x',s) \right| \leq 4L\epsilon. \quad (24)$$

Combining (22) and (24), it follows that

$$\left| \int_{R^m} \int S g(y) U_n(ds|x') W(dy|x,s) - \int_{R^m} \int_S g(y) U_n(ds|x) W(dy|x',s) \right| \leq$$
$$\leq (4L+1)\epsilon, \quad (25)$$

for all n sufficiently large.

Now, since $U_n \to U_0$ by Hypothesis (H.2), we get

$$\left| \int \int g(y) U_0(ds|x) W(dy|x,s) - \int \int g(y) U_n(ds|x') W(dy|x,s) \right| < \epsilon,$$

$$\left| \int \int g(y) U_0(ds|x') W(dy|x',s) - \int \int g(y) U_n(ds|x) W(dy|x',s) \right| < \epsilon \quad (26)$$

for all n larger than, say, $n_1(\epsilon) \geq n_0(\epsilon)$. Finally, combining (25) and (26) we obtain

$$\left| \int_{\mathcal{Y}} \int_S g(y) U_0(ds|x') W(dy|x,s) - \int_{R^m} \int_S g(y) U_0(ds|x) W(dy|x',s) \right| \leq (4L+3)\epsilon.$$

Since ϵ is arbitrarily small, we conclude that

$$\int g(y) W_{U_0}(dy|x,x') = \int g(y) W_{U_0}(dy|x',x)$$

for every bounded, uniformly continuous function $g \in C_b(\mathcal{Y})$ and $x, x' \in \mathcal{X}$, or $U \in \mathcal{U}$, which proves our claim.

Take into account the fact that if $\mathcal{A}_1$is finer than $\mathcal{A}_0$, then $I(P, W_Q^{\mathcal{A}_0}) \geq I(P, W_Q^{\mathcal{A}_0})$, we have got a partition $\mathcal{A}_1 = (A_1, ..., A_k)$ satisfying

(i) $I^{\mathcal{A}_1}(P, \Pi) > I(P, \Pi) - \delta/2$,

(ii) $PU \notin \Pi(\alpha')$ for all $U \in \mathcal{U}$ and for some $\alpha' > 0$.

Consider now the discrete $AVCW^{A_1} : \mathcal{X} \to \mathcal{Y}' = \{1, ..., k\}$ defined by (14). Applying Theorem 1, we see that for $\delta/2$ there exists $\gamma > 0$ such that for all sufficiently large n, there is an n-length block code (f'_n, ϕ'_n) with N codewords, each of type P, such that

$$\frac{1}{n}\log N > I^{A_1}(P, \Pi) - \delta/2 > I(P, \Pi) - \delta, \quad \bar{e}'(\Pi) \leq \exp(-n\gamma), \tag{27}$$

where $\bar{e}'(\Pi)$ denotes the average probability of error of (f', ϕ').

Next, we construct a new code (f_n, ϕ_n) for the given semi-continuous AVC from (f'_n, ϕ'_n). Let $f_n = f'_n$. Further, for each $y \in \mathcal{Y}^n, y \in A_{j_1} \times \cdots \times A_{j_k}$ for some $j = (j_1, \ldots, j_n) \in \mathcal{Y}'^n$. In this case we set $\phi_n(y) = \phi'_n(j)$. For this code,

$$e_m(s) = \sum_{\phi_n(y) \neq m} W^n(y|f_n(m), s) = \sum_{y \in A_{j_1} \times \cdots \times A_{j_n}} W^n(y|f'_n(m), s) =$$

$$\sum_{\phi'_n(j) \neq m} W^n(A_{j_1} \times \cdots \times A_{j_k}|f'_n, s) = \sum_{\phi'_n(j) \neq m} (W^{A_1})^n(j|f'_n(m), s) = e'_m(s).$$

Therefore, by (27), $\bar{e}(\Pi) < \exp -n\gamma)$. Thus Theorem 2 is proved.

The above Theorem 2 is rather general with only the assumption that the output alphabet is a finite dimensional Euclidean space. However, this is satisfactory to model several communication situations of most practical importance. In order to prove the same result for general output alphabet we must impose a rather strong condition on the set of states or on the constraint set. Namely, we shall suppose that $\mathcal{S}$ is a compact space. It is suitable, however, to model such real situations when the state set is a bounded, compact set in Euclidean space. Using the idea of Theorem 2, we can prove the following

THEOREM 3. *Suppose that the set of states $\mathcal{S}$ is compact, while $\mathcal{Y}$ is an arbitrary Polish space. Then the statement of Theorem 2 remains true.*

REMARK. If one of the following conditions is satisfied, then the statement of Theorem 2 remains valid for general $\mathcal{Y}$ and $\mathcal{S}$:

1. Compactness Condition: There exists a small number $\xi > 0$ such that the closed ξ-neighborhood $cl(\Pi(\xi))$ of Π is compact, (cl(A) denotes the closure of the set A).

2. $\mathcal{Y}$ is σ-compact, i.e. there are compact sets K_n such that $\mathcal{Y} = \cup_n^\infty K_n$, and $\mathcal{Y}$ possesses the following property:

(P) For every compact subset K of $\mathcal{Y}$ there is an n such that $K_n \supset K$.

For the semi-continuous case, we set again

$$\mathcal{L}_\alpha = \{P : PU \notin \Pi(\alpha) \text{ for every } U \in \mathcal{U}\},$$

$$\mathcal{L}^+ = \{P : PU \notin int\Pi \text{ for every } U \in \mathcal{U}\},$$

$$\mathcal{L}^- = \{P : PU \notin \Pi \text{ for every } U \in \mathcal{U}\},$$

where $Int\Pi$ denotes the interior of Π, and let

$$C_\alpha = \max_{P \in \mathcal{L}_\alpha} I(P, \Pi), \quad C_0 = \lim_{\alpha \to 0} C_\alpha, \quad C^- = \sup_{P \in \mathcal{L}^-} I(P, \Pi).$$

We have the following theorem, whose proof will be omitted.

THEOREM 4. *i) If $\mathcal{L}^+ = \emptyset$, then $C(\Pi) = 0$,*

ii) Suppose either $\mathcal{S}$ is compact or $\mathcal{Y} = R^m$. Then $C(\Pi) \geq C^-$ if one of the following conditions is satisfied :

a) $\mathcal{L}_\alpha \neq \emptyset$ for some $\alpha > 0$,

b)$\mathcal{L}^- \neq \emptyset$ and $\mathcal{L}^-$ is an open set in $\mathcal{M}(\mathcal{X})$,

c) $\mathcal{L}^- \neq \emptyset$ and the sets $\{PU : U \in \mathcal{U}\}$ are closed for each P.

The above theorem only gives a lower bound on the capacity. For proving the converse part we must impose further conditions. Under rather strong conditions we cannot get the exact formula for the capacity. Only a lower and upper bound can be obtained for it (see Section 4). In the following section we will consider the most important case when the state constraints are expressed in term of a continuous function.

3. Semi-continuous AVC's with constraints set given by a continuous function

Instead of (H.3) we now adopt the following hypothesis :

(H.3') Let ℓ be a real-valued continuous function defined on $\mathcal{S}$ whose level sets $\{s : \ell(s) \leq K\}$ are compact for every K . A state sequence $s = (s_1, \ldots, s_n) \in \mathcal{S}^n$ will be called an admissible state sequence if $(1/n)\sum_{n=1}^{n} \ell(s_i) \leq \Lambda$, where Λ is a fixed number. Here, the state sequence s is admissible if $P_s \in \Pi_\Lambda$, where

$$\Pi_\Lambda = \{Q : \int \ell(s)Q(ds) \leq \Lambda\}. \tag{28}$$

Note that unless $\mathcal{S}$ is compact, (H.3) requires that ℓ is unbounded. The following lemma will show that Π_Λ is compact, hence the previous results can be applied.

LEMMA 2. *For every $\Lambda \geq 0$, Π_Λ is compact in $\mathcal{M}(\mathcal{S})$.*

PROOF: The proof is standard, therefore it is omitted.

For every P we define

$$I(P, \Lambda) = I(P, \Pi_\Lambda), \tag{29}$$

$$\Lambda_0(P) = \inf_{U \in \mathcal{U}} \sum_{x \in \mathcal{X}} P(x) \int \ell(s)U(ds|x). \tag{30}$$

We also denote the capacity under state constraint Π_Λ by $C(\Lambda)$. The function $\Lambda_0(P)$ is concave, hence it is lower semi-continuous in P and continuous in the set of all strictly positive P. We also define

$$\Lambda_0 = \sup_{P \in \mathcal{M}(\mathcal{X})} \Lambda_0(P). \tag{31}$$

Now suppose that P is a strictly positive distribution on the input alphabet $\mathcal{X}$ and $\Lambda_0(P) > \Lambda$. If for every n there is an $U_n \in \mathcal{U}$ satisfying $PU_n \in \Pi_\Lambda(1/n)$, then by the same reasoning as in the proof of Theorem 2, there is a subsequence U_{nj} of U_n such that $U_{nj} \to U_0$ for some U_0 and $PU_{nj} \to PU_0 = Q \in \Pi_\Lambda$. Therefore, $\sum_{x \in \mathcal{X}} P(x) \int \ell(s) U_0(ds|x) \leq \Lambda$. On the other hand, since $\mathcal{U}$ is closed, $U_0 \in \mathcal{U}$. This contradicts our assumption

$\Lambda_0(P) > \Lambda$. This means that if $\Lambda_0(P) > \Lambda$, then there is a positive number $\alpha > 0$ such that $PU \notin \Pi_\Lambda(\alpha)$ for every $U \in \mathcal{U}$.

From the assumption $\Lambda_0 > \Lambda$, the concavity and the continuity of $\Lambda_0(P)$ on the set of strictly positive P , we see that there exists a strictly positive P' satisfying $\Lambda_0(P') > \Lambda$. Therefore, by the above observation , $P' \in \mathcal{L}(\alpha)$ for some $\alpha > 0$. Applying Theorem 4 we get

COROLLARY. *Suppose that either ℓ is bounded or $\mathcal{Y} = R^m$.*

If $\Lambda_0 > \Lambda$, then $C(\Lambda) \geq C_0$.

In order to prove the converse part of Coding Theorem for this case, we need the following two lemmas. For discrete AVC's, these are Lemma 2 and Lemma 3 of Csiszár and Narayan in [2], where they are proved by using Chebyshev's inequality. Here, we will use the weak law of large numbers rather than Chebyshev's inequality. From now on let τ_{P_η} denote the set of all P-typical sequences x with constant η.

LEMMA 3. *For any $\Lambda > 0, \delta > 0$ and $\epsilon < 1$ there exist numbers $\eta > 0$ and n_0 such that, for any $P \in \mathcal{M}(\mathcal{X})$, any code of blocklength $n \geq n_0$ with codewords $x_i \in \tau_{|P|_\eta}, i = 1, 2, \ldots, N$, if $\frac{1}{n} \log N \geq I(P, \Lambda) + \delta$, then $\max_{\ell(s) \leq \Lambda} \overline{e}(s) \geq \epsilon$.*

PROOF: First, we observe that $G(\Lambda) = \min_{\ell(Q) \leq \Lambda} I(P, W_Q)$ as a function of Λ, is convex and hence it is continuous. It is due to the fact that $I(P, W_Q)$ is convex in Q while $\ell(Q) = \int \ell(s) Q(ds)$ is linear. Thus, we can choose $\gamma > 0$ such that $I(P, \Lambda - \gamma) \leq I(P, \Lambda) + \frac{\delta}{2}$. Let Q be the probability measure satisfying

$$I(P, W_Q) = \min_{\ell(Q) \leq \Lambda - \gamma} I(P, W_Q) = I(P, \Lambda - \gamma)$$

with $\ell(Q) = \int \ell(s) Q(ds) \leq \Lambda - \gamma$. Consider the semi-continuous memoryless channel W_Q defined by (5). Further, consider any code with words $x_1, \ldots, x_N$ with decoder ϕ, and let $S = (S_1, \ldots, S_n)$ be the random vector with statistically independent components, each has distribution Q . For any set $B \in \mathcal{B}^n, x \in \mathcal{X}^n$, we have $EW^n(B|x, S) = W_Q^n(B|x)$. Thus

$$E\overline{e}(S) = \frac{1}{N} \sum_{m=1}^{N} e_m(S) = \frac{1}{N} \sum_{m=1}^{N} EW^n(\{y : \phi(y) \neq i\}|x_i, S) =$$

$$= \frac{1}{N}\sum_{i=1}^{N} W_Q^n(\{y : \phi(y) \neq i\}|x_i) = \bar{e}w_Q,$$

where $\bar{e}w_Q$ denotes the average probability of error of the above code used on the channel W_Q. Hence

$$\max_{\ell(s)\leq\Lambda} \bar{e}(s) \geq E\bar{e}(S) - Pr\{\ell(S_i) > \Lambda\} = \bar{e}W_Q - Pr\{\frac{1}{n}\sum_{i=1}^{n}\ell(S_i) > \Lambda\}. \quad (32)$$

Since $\ell(S_i)$ are independent and identically distributed with expectation $E\ell(S_i) = \ell(Q) \leq \Lambda - \gamma$, it follows from the law of large numbers that $Pr\{\ell(S) \geq \Lambda\} \to 0$ whenever $n \to \infty$. On the other hand, using the semi-continuous analogue of Lemma SC in [3] (which is clearly valid), we see that there is $\eta > 0$ such that if all codewords are in $\tau_{|P|_\eta}$, then $\bar{e}w_Q$ is arbitrary close to 1 if $(1/n)\log N \geq I(P,\Lambda) + \delta \geq I(P, W_Q) + \frac{\delta}{2}$ and n is large enough. This and (32) complete the proof.

LEMMA 4. *For any $\epsilon > 0$ there is $\eta > 0$ such that for any strictly positive distribution P with $\Lambda_0(P) < \Lambda$, any block code of length n block with codewords $x_i \in \tau_{[P]_\eta}, i = 1,\dots,N, N \geq 2$, has $\max_{\ell(s)\leq\Lambda}\bar{e}(s) \geq (1/4) - \epsilon$.*

PROOF: Consider any code with decoder ϕ and codewords $x_1,\dots,x_N$ such that $x_j \in \tau_{[P]_\eta}$, where η is a positive number to be chosen later independently on P. Since $\Lambda_0(P) < \Lambda$, there is a $\tau > 0$ such that $\Lambda_0(P) < \Lambda - \tau$. Let $U \in \mathcal{U}$ be the channel which satisfies

$$\sum_{x\in\mathcal{X}} P(x)\int \ell(s)U(ds|x) < \Lambda_0(P) + \frac{\tau}{4} < \Lambda - \frac{3}{4}\tau. \quad (33)$$

Consider $\mathcal{S}^n$-valued random variables $S_j = (S_{j1}, S_{j2},\dots,S_{jn}), j = 1,2,\dots,N$, with statistically independent components, where $pr\{S_{jk} \in C\} = U(C|x_{jk})$ for every $C \in \mathcal{C}$. Since $\mathcal{S}$ is separable, we see that for each pair (i,j) and every $C \in \mathcal{C}$,

$$EW^n(C|x_i, S_j) = EW^n(C|x_j, S_i).$$

Thus, for $i \neq j$, we have

$$Ee_i(S_j) + Ee_j(S_i) = \sum_{s\in\mathcal{S}^n} W^n(\{\phi(y) \neq i\}|x_i, S_j) + \sum_{s\in\mathcal{S}^n} W^n(\{\phi(y) \neq j\}|x_j, S_i)$$
$$\geq EW^n(\mathcal{Y}|x_i, S_j) \geq 1.$$

As in the proof of Lemma 4 in [2], we see that for some $j \in \{1, \ldots, N\}$

$$E\overline{e}(S_j) \geq \frac{1}{4}. \tag{34}$$

On the other hand, since P is strictly positive and $\sum_{x\in\mathcal{X}} \int \ell(s)U(ds|x) < \Lambda$, letting $\beta = \min\{P(x), x \in \mathcal{X}\}$, we have

$$E\ell(S_{jk}) = \int \ell(s)U(ds|x_{jk}) < \frac{\Lambda}{\beta}, \tag{35}$$

i.e. the expectation of each S_{jk} is finite. Thus for any $\epsilon > 0$, by the weak law of large numbers we get

$$Pr\left\{\frac{1}{|\{x_{jk} = x\}|} \sum_{\{x_{jk}=x\}} \ell(S_{jk}) > \int \ell(s)U(ds|x_{jk}) + \frac{\tau}{4}\right\} < \frac{\epsilon}{|\mathcal{X}|}, \tag{36}$$

whenever $|\{x_{jk} = x\}| \geq n_0, n_0$ is large enough. In any case, from (36) we get

$$Pr\left\{\sum_{x\in\mathcal{X}} P_{x_j}(x) \frac{1}{|\{x_{jk} = x\}|} \sum_{\{x_{jk}=x\}} \ell(S_{jk}) > \sum_{xin\mathcal{X}} \int \ell(s)U(ds|x) + \frac{\tau}{4}\right\} < \epsilon,$$

provided $n \geq n_0$.. Since $x_j \in \tau_{[P]_\eta}$, it follows that

$$Pr\left\{\sum_{x\in\mathcal{X}} P_{x_j}(x) \frac{1}{|\{x_{jk} = x\}|} \sum_{\{x_{jk}=x\}} \ell(S_{jk}) > \sum_{x\in\mathcal{X}} P(x) \int \ell(s)U(ds|x) + \right.$$
$$\left. +\eta \sum_{x\in\mathcal{X}} \int \ell(s)U(ds|x) + \frac{\tau}{4}\right\} < \epsilon. \tag{37}$$

Now, if we choose $\eta < \frac{\beta\tau}{4\Lambda|\mathcal{X}|}$, from (33) and (35) we get

$$\sum_{x\in\mathcal{X}} P(x) \int \ell(s)U(ds|x) + \eta \sum_{x\in\mathcal{X}} \int \ell(s)U(ds|x) + \frac{\tau}{4} < \Lambda - \frac{\tau}{4}. \tag{38}$$

Then (37) and (38) imply

$$Pr\{\ell(S_j) > \Lambda\} = Pr\left\{\frac{1}{n}\sum_{k=1}^{n}\ell(S_{jk}) > \Lambda\right\} \le$$

$$\le Pr\{\sum_{x\in\mathcal{X}} P_{x_j}(x)\frac{1}{|\{x_{jk}=x\}|}\sum_{\{x_{jk}=x\}}\ell(S_{jk}) > \Lambda - \frac{\tau}{4}\} < \epsilon,$$

for n sufficiently large. From this fact and (35) we get

$$\max_{s:\ell(s)>\Lambda}\overline{e}(s) \ge E\overline{e}(S_j) - Pr\{\ell(S_j) > \Lambda\} \ge \frac{1}{4} - \epsilon,$$

which proves the assertion of the lemma.

THEOREM 5. *Suppose that either ℓ is bounded or $\mathcal{Y} = R^m$. Then the capacity $C(\Lambda)$ of the AVC under state constraint Λ is*

(*i*) $C(\Lambda) = 0$ *if* $\Lambda_0 < \Lambda$,

(*ii*) $C(\Lambda) = \sup_{\Lambda_0(P)} I(P,\Lambda)$ *if* $\Lambda_0 > \Lambda$.

PROOF: To prove this theorem we can repeat the arguments used in the proof of Theorems 3 and 4 in [3]. For example, to prove (i) it suffices to observe that if $\Lambda_0 < \Lambda$, then $\Lambda_0(P) < \Lambda$ for all $P \in \mathcal{M}(\mathcal{X})$. Thus, using Lemma 4, the argument used in the proof of Theorem 3 of [3] shows that $C(\Lambda) = 0$. Using the same partition technique and Lemma 3 we get $C(\Lambda) \le \sup_{\Lambda_0 \ge \Lambda} I(P,\Lambda)$ if $\Lambda_0 > \Lambda$. On the other hand, if $\Lambda_0 > \Lambda$, then

$$C(\Lambda) \ge \sup_{\Lambda_0(P)} I(P,\Lambda).$$

This fact can be proved exactly as in the proof of Theorem 3 in [2].

REMARK: As in the discrete case, Theorem 5 still left an undecided case $\Lambda_0 = \Lambda$.

4. Concluding remarks

Theorem 5 gives the exact formula for the capacity of semi-continuous AVC under state constraints defined by terms of a continuous function whose level sets are compact. When the state constraints are expressed by empirical distributions of state sequences, we only get a lower bound on the capacity. We also have a condition under which the capacity is zero. The proof for this is based on the fact that if Q is in $Int\Pi$ (in the weak topology), then $Q^n(P_s \notin \Pi) \to 0$ as $n \to \infty$. In [8] Csiszár introduced the τ- topology (which is finer than the weak topology) and proved that even if Q is in the τ-interior of Π , then the above probability tends to 0 exponentially. From this we obtain a stronger result than Part (i) of Theorem 4, namely, $C(\Pi) = 0$ if $C^+ = 0$. Here, in the definition of the set $\mathcal{L}^+, Int_\tau\Pi$ stands instead of $Int\Pi$. On the other hand, concerning the upper bound on the capacity, we can introduce a rather strong condition : the strong convexity of the set Π with respect to the weak topology. Note that the set Π is said to be strongly convex if the segment joining any point with an interior point is in the interior of Π, except perhaps for the endpoint. A similar condition with respect to the τ_0-topology was introduced by P.Groeneboom., J.Oosterhoff and F.H.Ruymgaart in [9]. We also note that the strong convexity condition is clearly satisfied in the discrete case. If the constraint set Π is strongly convex and has nonempty interior, then $C(\Pi) \le C^+ = \max_{P\in\mathcal{L}^+} I(P,\Pi)$.

AKNOWLEGEMENT. The author expresses his gratitute to Professor Imre Csiszár for suggesting the problem and his valuable advice during the preparation of this work.

REFERENCES

[1]. I.Csiszár, P.Narayan, *Arbitrarily varying channels with constrained inputs and states*, IEEE Trans. Inform. Theory **17** (1988), 27-34.

[2]. I.Csiszár and P.Narayan, *The Capacity of the Arbitrarily Varying Channel Revisited : Positivity, Constraints*, IEEE Trans. Inform. Theory **17** (1988), 181-193.

[3]. Bui Van Thanh, *The Capacity of Arbitrarily Varying Channels under general State Constraints*, Problem of Control and Information Theory **19** (1990), 151-165.

[4]. I.Csiszár and J.Korner, *Information Theory : Coding Theorems for Discrete Memoryless Systems*, Academic Press, 1981.

[5]. P.Billingsley, *Convergence of Probability Measures*, John Willey and Sons New York, 1968.

[6]. K.R.Parathasarathy, *Probability Measures on Metric Spaces*, Academic Press, 1967.

[7]. M.S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden Day, San Fransisco, 1984.

[8]. I.Csiszár, *Sanov Property, Generalized I-Projection and a Conditional Limit Theorem*, Ann. Probab. **3** (1984), 146-158.

[9]. P.Groeneboom , J.Oosterhoff and F.H.Ruymgaart, *Large Deviation Theorems for Empirical Probability Measures*, Ann. Probab. **7** (1979), 533-586.

INSTITUTE OF INFORMATICS

NATIONAL CENTER FOR SCIENTIFIC RESEACHES

HANOI - VIETNAM.