объединенный
институт
ядерных
исследований
дубна

E5-84-713

Š.Šujan

JOINT SOURCE AND CHANNEL BLOCK CODING FOR SERIES OF DISCRETE STATIONARY CHANNELS

Submitted to "Problems of Control and Information Theory"

1984

1. INTRODUCTION

Series of channels were previously considered in connection with the information processing theorem (cf., e.g.,[2], Sections 2.3 and 4.3). The basic result is that if the (discrete) random variables X , Y , Z form a Markov chain, then the conditional entropies satisfy $H(X|Z) \leq H(X|Y)$. A natural interpretation is that the uncertainty concerning the input X never decreases during the motion of the input along subsequently joined channels. The nature of this result is rather qualitative. In particular, it does not give any information about the output rate of a channel, the latter being of primary importance when communicating the output across some other channel.

Let $[\mathfrak{U},\mu]$ be a stationary source and $([\mathfrak{X}(k),\nu_{(k)},\mathfrak{Y}(k)];\ 1\leq k\leq K)$ an ordered K-tuple of stationary channels, where all the alphabets are supposed finite. Let $I=\{\ldots,-1,0,1,\ldots\}$, and let $U=(U_i\ ;i\in I)$, $X(k)=(X(k)_i\ ;i\in I)$, and $Y(k)=(Y(k)_i\ ;i\in I)$ be generic notations for the corresponding doubly-infinite sequences of one-dimensional projections (called processes).

In this paper we address the problem of block transmission of $[\mathfrak{U},\mu]$ over the K-tuple of channels $(\nu_{(1)},\ldots,\nu_{(K)})$ in the following sense.

<u>Definition</u>. A source $[\mathfrak{U},\mu]$ is said to be block transmissible over the ordered K-tuple of channels $([\mathfrak{X}(k),\nu_{(k)},\mathfrak{Y}(k)];\ 1\leq k\leq K)$ if for any $\epsilon>0$ there is an N_0 and for any $N>N_0$ there exist block coding functions F_N :

$$\mathfrak{U}^N\to\mathfrak{X}(1)^N,\ K_N^{(k)}:\mathfrak{Y}(k)^N\to\mathfrak{X}(k+1)^N\ (1\leq k\leq K-1),$$

and $G_N:\mathfrak{Y}(K)^N\to\mathfrak{U}^N$ such that $\mathrm{Prob}[U^N\neq G_N(Y(K)^N)]\leq\epsilon$, where $Y(K)$ is the final output process of the K-tuple $(\nu_{(1)},\ldots,\nu_{(k)})$ joined by means of the mappings $K_N^{(k)}$.

2. NOTATION AND TERMINOLOGY

Let $[\mathfrak{S},\mu]$ be a source (not necessarily a stationary one) with a countable discrete alphabet $\mathfrak{S}$; that is, λ is a probability measure on Borel subsets of $\mathfrak{S}^I$ equipped with the natural product topology (cf., e.g.,[8]; by writing $E\subset\mathfrak{S}^I$ we shall always understand that E is a Borel subset of $\mathfrak{S}^I$). Given $x\in\mathfrak{S}^I$

and $m\le n$, $m,n\in I$, let $x_m^n=(x_m,\dots,x_n)$. If $m=0$, we abbreviate x_0^{n-1} as x^n and put

$$\lambda^n(E)=\lambda\{x\in\mathcal{E}^I : x^n\in E\},\ E\subset\mathcal{E}^n. \qquad (1)$$

All logarithms throughout will be to the case two. Given $\epsilon\in(0,1)$ and a positive integer N, the ϵ-N-rate of $[\mathcal{E},\lambda]$ is defined to be

$$R(\lambda;\epsilon,N)=N^{-1}\log L_N(\epsilon,\lambda), \qquad (2)$$

where

$$L_N(\epsilon,\lambda)=\min\{\|E\| : E\subset\mathcal{E}^N,\ \lambda^N(E)>1-\epsilon\} \qquad (3)$$

and $\|E\|=\mathrm{card}(E)$. With a little abuse of notation, T will denote the two-sided shift in any doubly-infinite sequence space. If $[\mathcal{E},\lambda]$ is stationary (i.e., $\lambda=\lambda T$) or, more generally, n-stationary (i.e., $\lambda=\lambda T^n$) then the limits $H_\epsilon(\lambda)=\lim_{N\to\infty} R(\lambda;\epsilon,N)$ exist except for an at most countable set of values $\epsilon\in(0,1)$, and the limit

$$H^*(\lambda)=\lim_{\epsilon\to 0+} H_\epsilon(\lambda)=\sup_{0<\epsilon<1} H_\epsilon(\lambda) \qquad (4)$$

is said to be the asymptotic rate of $[\mathcal{E},\lambda]$ (see[7], where this assertion is proved for more general sources). We let $h(\lambda)$ denote the entropy of $[\mathcal{E},\lambda]$ (that is, $h(\lambda)=h_\lambda(T)$, where $h_\lambda(T)$ is the entropy of the dynamical system (T,λ); see[1]). If $[\mathcal{E},\lambda]$ is ergodic, then the limits $H_\epsilon(\lambda)$ exist for all $\epsilon\in(0,1)$, and

$$H^*(\lambda)=H_\epsilon(\lambda)=h(\lambda);\quad 0<\epsilon<1. \qquad (5)$$

A review of these and other properties of H^* can be found in[8].

Let $\mathcal{E},\mathcal{F}$ be finite sets. By definition, a channel $[\mathcal{E},\omega,\mathcal{F}]$ is a measurable family $\omega=(\omega_x; x\in\mathcal{E}^I)$ of probability measures on $\mathcal{F}^I$ (that is, for any $F\subset\mathcal{F}^I$, the mapping $x\to\omega_x(F):\mathcal{E}^I\to[0,1]$ is Borel measurable). We shall use also the notation $\omega(.|x)$ for $\omega_x(.)$. All channels in this paper are assumed stationary, viz. $\omega(TF|Tx)=\omega(F|x)$; $F\subset\mathcal{F}^I$, $x\in\mathcal{E}^I$. If $X=(X_i; i\in I)$ is the process corresponding to a given source $[\mathcal{E},\lambda]$, we shall indicate this by writing $\mathrm{dist}(X)=\lambda$. If $[\mathcal{E},\omega,\mathcal{F}]$ is a channel, we let $Y|x$ denote the conditional output process of ω given the input sequence $x\in\mathcal{E}$. Consequently, $\mathrm{dist}(Y|x)=\omega_x$. Given λ and ω, let (X,Y) denote the joint input/output process. Then $\mathrm{dist}(X,Y)=\lambda\omega$, where $\lambda\omega$ is the probability measure on $\mathcal{E}^I\times\mathcal{F}^I$, uniquely determined by the properties that

$$\lambda\omega(E\times F)=\int_E \omega_x(F)\,\lambda(dx);\quad E\subset\mathcal{E}^I,\ F\subset\mathcal{F}^I. \qquad (6)$$

The output process of ω given the input process X will be denoted by $Y|X$. If $\mathrm{dist}(X)=\lambda$ then $\mathrm{dist}(Y|X)=\eta(\lambda,\omega)$ (we shall sometimes abbreviate this as η) is the output marginal of $\lambda\omega$, i.e.,

$$\eta(F)=\lambda\omega(\mathcal{E}^I\times F)=\int\omega_x(F)\lambda(dx);\quad F\subset\mathcal{F}^I. \qquad (7)$$

Let $i:\ \mathcal{E}^I\times\mathcal{F}^I\to[0,\infty)$ denote the mutual information density, that is, the (essentially) unique invariant bounded measurable function such that for any stationary source λ and for any stationary channel ω

$$\int i(x,y)\lambda\omega(dx,dy)=I(\lambda\omega)=h(\lambda)+h(\eta(\lambda,\omega))-h(\lambda\omega) \qquad (8)$$

(cf.[3]). For $\epsilon\in(0,1)$ put $C_\epsilon^*(\omega)=\sup_\lambda \sup\{r:\ \lambda\omega[i\le r]<\epsilon\}$, where the supremum is taken over all stationary input sources (or, equivalently, over all stationary and ergodic ones or, over all block stationary ones). The information quantile capacity $C^*(\omega)$ is defined as the limit

$$C^*(\omega)=\lim_{\epsilon\to 0+} C_\epsilon^*(\omega). \qquad (9)$$

An excellent review on block channel coding, in particular, an exhaustive study of relations among various concepts of capacity, may be found in[3].

Following[5], the channel entropy of a stationary channel $[\mathcal{E},\omega,\mathcal{F}]$ is defined as

$$h(\omega)=\sup\{h(\lambda\omega)-h(\lambda)\}, \qquad (10)$$

where the supremum is taken over all stationary input sources $[\mathcal{E},\lambda]$.

We assume that all the channels $\nu_{(1)}, \nu_{(2)},\dots,\nu_{(K)}$ (cf. the introduction) are $\bar{d}$-continuous, and all the channels $\nu_{(1)},\dots,\nu_{(K-1)}$ are conditionally almost block independent (CABI). These properties will not be used directly so that we refer to[4,5] for precise definitions. Roughly, a channel is $\bar{d}$-continuous if the statistics of a long segment of output depends, essentially, only on the corresponding amount of input, where "essentially" is measured in $\bar{d}$. A channel is CABI if long blocks of output are almost independent, "almost" again measured in $\bar{d}$ (cf.[3,4,5] for definition and basic properties of $\bar{d}$-distance).

3. THE MAIN RESULT

The main result of this paper is the following joint source and channel block coding theorem:

Theorem. Let $[\mathfrak{U},\mu]$ be a stationary source, and let $([\mathfrak{X}(k), \nu_{(k)}, \mathfrak{Y}(k)]; 1 \leq k \leq K)$ be an ordered K-tuple of stationary $\bar{d}$-continuous channels, where $\nu_{(1)}, \ldots, \nu_{(K-1)}$ are also conditionally almost block independent. Suppose that

$$H^*(\mu) < C^*(\nu_{(1)}) \tag{11}$$

and

$$H^*(\mu) + \sum_{i=1}^{k} h(\nu_{(i)}) < C^*(\nu_{(k+1)}), \quad 1 \leq k \leq K-1. \tag{12}$$

Then $[\mathfrak{U},\mu]$ is block transmissible over the K-tuple $(\nu_{(1)}, \ldots, \nu_{(K)})$ see the definition in Section 1.

Comments. (a) Suppose $[\mathfrak{X}(1),\lambda]$ is an ergodic source and $\nu_{(1)}$, $\nu_{(2)}$ (K=2) are ergodic channels. Hence $H^*(\lambda) = h(\lambda)$ (cf. (5)) and $C^*(\nu_{(k)}) = C(\nu_{(k)})$ (k = 1,2), where C stands for the Shannon (information rate) capacity [3]. We have

$$h(\lambda\nu_{(1)}) = h(\lambda\nu_{(1)}) - h(\lambda) + h(\lambda) \leq \sup\{h(\lambda\nu_{(1)}) - h(\lambda)\} + h(\lambda) = h(\nu_{(1)}) + h(\lambda);$$

in particular (cf. (7)), $h(\eta(\lambda,\nu_{(1)})) \leq h(\nu_{(1)}) + h(\lambda)$.

Now the output source $[\mathfrak{Y}(1),\eta_{(1)}]$ $(\eta_{(1)} = \eta(\lambda,\nu_{(1)}))$ is ergodic so that a sufficient condition for its block transmissibility over $[\mathfrak{X}(2),\nu_{(2)},\mathfrak{Y}(2)]$ is that $h(\eta_{(1)}) < C(\nu_{(2)})$. Since we do not know $h(\eta_{(1)})$ but do have the above bound, it is natural to impose the condition

$$h(\nu_{(1)}) + h(\lambda) < C(\nu_{(2)}). \tag{13}$$

In the non-ergodic case one has to use the asymptotic rate and the information quantile capacities. In this case (13) will assume on the form (12).

(b) In order to prove the theorem we wish to combine the block transmission theorems for separate channels. Here we meet the following difficulty. As we shall see, if (11) is satisfied, then $[\mathfrak{U},\mu]$ will be block transmissible over $\nu_{(1)}$ (see Lemma 1 below). In particular, for any N large enough, there exist good block coding functions $F_N: \mathfrak{U}^N \to \mathfrak{X}(1)^N$ and $H_N: \mathfrak{Y}(1)^N \to \mathfrak{U}^N$. But this means that we do not have a fixed source $[\mathfrak{X}(1),\lambda]$ at the input of $\nu_{(1)}$, but rather a sequence of different input sources $\{[\mathfrak{X}(1),\mu\bar{F}_N^{-1}]\}_{N=1}^{\infty}$. Consequently, instead of a fixed output source $[\mathfrak{Y}(1),\eta(\lambda,\nu_{(1)})]$ we have

a sequence $\{[\mathfrak{Y}(1),\eta(\mu\bar{F}_N^{-1},\nu_{(1)})]\}_{n=1}^{\infty}$ of different output sources. Hence, the reasoning from the preceding comment does not seem to apply. Fortunately, an analogue of block transmissibility (and namely one which fits in our definition from Section 1) is still true.

(c) The central step in our proof consists of proving that for any stationary source $[\mathfrak{U},\mu]$ and for any sequence $F_N: \mathfrak{U}^N \to \mathfrak{X}(1)^N$ the joint input/output rate is upper bounded by $H^*(\mu) + h(\nu_{(1)})$ (cf. Lemma 2). This will follow from a result of Neuhoff and Shields [5], according to which the "local" output rates (i.e., rates of processes $Y|\bar{F}_N(u), u \in \mathfrak{U}^I$) are uniformly upper bounded by the channel entropy, combined with Ziv's [9] universal source coding technique.

(d) The idea from the preceding comment allows one to assign a natural interpretation to the bound $H^*(\mu) + h(\nu_{(1)})$. If the block length N is sufficiently large, then there are approximately $\exp_2(NH^*(\mu\bar{F}_N^{-1})) \leq \exp_2(NH^*(\mu))$ different input N-tuples. Given any input N-tuple, there are at most $\exp_2(Nh(\nu_{(1)}))$ output N-tuples for the corresponding "local" output processes. Hence, the total number of output N-tuples should be no more than

$$\exp_2(NH^*(\mu)) \cdot \exp_2(Nh(\nu_{(1)})) = \exp_2[N(H^*(\mu) + h(\nu_{(1)}))]$$

(here, $\exp_2$ is the inverse function to $\log_2$).

4. THE PROOF

In order to keep clear the main steps we divide the proof into a series of assertions. We shall deal with the case K = 2; an extension to any K > 2 is elementary.

Lemma 1. Let $[\mathfrak{U},\mu]$ be a stationary source, and let $[\mathfrak{X},\nu,\mathfrak{Y}]$ be a stationary $\bar{d}$-continuous channel. If $H^*(\mu) < C^*(\nu)$, then $[\mathfrak{U},\mu]$ is block transmissible over $[\mathfrak{X},\nu,\mathfrak{Y}]$; that is, for any $\epsilon > 0$ there is an N_0 and for any $N \geq N_0$ there exist block coding functions $F_N: \mathfrak{U}^N \to \mathfrak{X}^N$ and $H_N: \mathfrak{Y}^N \to \mathfrak{U}^N$ such that

$$\mathrm{Prob}[U^N \neq H_N[(Y|\bar{F}_N U)^N]] \leq \epsilon, \tag{14}$$

where $\bar{F}_N: \mathfrak{U}^I \to \mathfrak{X}^I$ is the induced infinite string coder.

Proof. Choose $R > 0$ and then $\delta > 0$ so small that

$$H^*(\mu) < R - \delta < R < C^*(\nu). \tag{15}$$

Choose and fix an $\epsilon > 0$. Using the left-hand inequality from (15) and the definition of the asymptotic rate (cf. (4)) we find an

N_0 such that for any $N \geq N_0$ there exists a set $\Gamma_N \subset \mathfrak{U}^N$ with

$$R(\Gamma_N) = N^{-1} \log \|\Gamma_N\| \leq R - \delta \tag{16}$$

and

$$\mu^N(\Gamma_N) > 1 - \epsilon/2 \tag{17}$$

(cf. (1)). Next we use the right-hand inequality from (15). For any stationary $\bar{d}$-continuous channel ν, $C^*(\nu)$ coincides with the operational block channel coding capacity (see [3], Corollary 8). Since $R < C^*(\nu)$, it follows from the definition of the latter capacity that, given $\epsilon > 0$ as above, we find an N_1 and for any $N \geq N_1$ we find a $(<2^{NR}>, N, \epsilon/2)$ block channel code $\{(x_i, \Gamma_i) : 1 \leq i \leq <2^{NR}>\}$. Let $N \geq \max\{N_0, N_1\}$. Then it follows from (16) that there are more codewords in the channel code than in the source code. Let $\Gamma_N = \{u_i : 1 \leq i \leq \|\Gamma\|\}$. We define F_N: $\mathfrak{U}^N \to \mathfrak{X}^N$ by $F_N(u) = x_i$ if $u = u_i$ and $F_N(u) = x_0$, say if $u \notin \Gamma_N$. Let H_N: $\mathfrak{Y}^N \to \mathfrak{U}^N$ be defined by $H_N(y^N) = u_i$ if $y^N \in \Gamma_i$, otherwise, let H_N be defined arbitrarily. Let $E_i = \{u \in \mathfrak{U}^I : u^N = u_i\}$, $i = 1, 2, \ldots, <2^{NR}>$. Then (by denoting $M = <2^{NR}>$)

$$\mathrm{Prob}[U^N \notin H_N[(Y|\bar{F}_N U)^N]] \leq \mu^N(\mathfrak{U}^N \setminus \Gamma_N) + \sum_{i=1}^{M} \int_{E_i} \nu^N_{\bar{F}_N(u)}(\mathfrak{Y}^N \setminus \Gamma_i)\mu(du) \leq -$$

$$< (\epsilon/2) + (\epsilon/2) \sum_{i=1}^{M} \int_{E_i} \mu(du) = (\epsilon/2) + (\epsilon/2)\mu^N(\Gamma_N) \leq \epsilon$$

using (17) and the fact that $\{(x_i, \Gamma_i)\}_{i=1}^{M}$ was an $\epsilon/2$-code.

Remark 1. The above proof modifies the proof of Theorem 4 in [3], and extends that theorem to transmission of stationary non-ergodic sources. Gray and Ornstein used ergodicity of the source and the Shannon-McMillan theorem in order to choose the sets Γ_N as in (16) and (17) consisting of source N-tuples typical of the source in the sense of entropy. We see that this typicality does not play any role in the proof, in particular, the above proof does not make explicit use of the Shannon-McMillan theorem (see [6], pp.67-70 for more on this point).

Lemma 2. Let $[\mathfrak{U}, \mu]$ be a stationary source and $[\mathfrak{X}, \nu, \mathfrak{Y}]$ a stationary, $\bar{d}$-continuous, and CABI channel. Let $\{F_N\}_{N=1}^{\infty}$ be any sequence of block coding functions $F_N: \mathfrak{U}^N \to \mathfrak{X}^N$. Let $\mu \bar{F}_N^{-1} \nu$ denote the corresponding sequence of joint input/output distributions. Then

$$\sup_{0<\epsilon<1} \limsup_{N\to\infty} R(\mu \bar{F}_N^{-1}\nu; \epsilon, N) \leq H^*(\mu) + h(\nu). \tag{18}$$

Proof. It suffices to prove that for any $R > H^*(\mu) + h(\nu)$ the left-hand side of (18) is not larger than R. Choose $R_1, R_2 > 0$ so that

$$H^*(\mu) < R_1, \quad h(\nu) < R_2, \quad R_1 + R_2 < R. \tag{19}$$

It follows from the first inequality in (19) that there is a sequence $\{C_N\}_{N=1}^{\infty}$ of sets $C_N \subset \mathfrak{U}^N$ such that

$$\mathrm{Prob}[U^N \in C_N] \to 1 \tag{20a}$$

and

$$R(C_N) = N^{-1} \log \|C_N\| \leq R_1 \tag{20b}$$

when N is large enough. It follows from the second inequality in (19) that for any $x \in \mathfrak{X}^I$ there is a sequence $\{F_N^x\}_{N=1}^{\infty}$ of sets $F_N^x \subset \mathfrak{Y}^N$ such that

$$\mathrm{Prob}[(Y|x)^N \in F_N^x] \to 1 \tag{21a}$$

and, for N large enough,

$$R(F_N^x) = N^{-1} \log \|F_N^x\| \leq R_2, \tag{21b}$$

where this latter bound works uniformly in x (this follows from the facts that $\nu_{(1)}$ is $\bar{d}$-continuous and CABI; cf. Theorem 4 in [5], and is the only place where we employ the CABI property).

Let $E_N = F_N(C_N)$, $N = 1, 2, \ldots$, where the F_N's are the given block coding functions. Then it follows from (20a,b) that

$$\mathrm{Prob}[X^N \in E_N] = \mathrm{Prob}[(\bar{F}_N U)^N \in F_N(C_N)] \geq \mathrm{Prob}[U^N \in C_N],$$

hence

$$\mathrm{Prob}[X^N \in E_N] \to 1 \tag{22a}$$

and, when N is large enough,

$$R(E_N)(\leq R(C_N)) \leq R_1, \tag{22b}$$

where $X^N = (\bar{F}_N U)^N$.

Let $\gamma > 0$ be small enough in order we have also $R_1 + R_2 + \gamma < R$. It follows from (21a,b) and from Lemma 3 below that there is a sequence $\{\Gamma_N\}_{N=1}^{\infty}$ of sets $\Gamma_N \subset \mathfrak{Y}^N$ such that

$$R(\Gamma_N) \leq R_2 + \gamma, \quad N = 1, 2, \ldots \quad \text{and} \tag{23a}$$

$$\mathrm{Prob}[(Y|x)^N \in \Gamma_N] \to 1 \tag{23b}$$

as $N \to \infty$ for any $x \in \mathfrak{X}^I$. Now consider the sequence of direct products $\{E_N \times \Gamma_N\}_{N=1}^{\infty}$. By (23a) and (22b), and by the third inequality in (19),

$$\|E_N \times \Gamma_N\| \leq \exp_2[N(R_1 + R_2 + \gamma)] < 2^{NR}. \tag{24}$$

At the same time,

$$\mathrm{Prob}[(\bar{F}_N U, Y|\bar{F}_N U)^N \in E_N \times \Gamma_N] = \int_{[E_N]} \nu^N(\Gamma_N | x)\, \mu \bar{F}_N^{-1}(dx) =$$

$$= \int_{[F_N^{-1} E_N]} \nu^N(\Gamma_N | \bar{F}_N(u))\, \mu(du) \geq \int_{C_N} \nu^N(\Gamma_N | \bar{F}_N(u))\, \mu(du),$$

where for a set $E \subset \mathfrak{S}^n$, $[E] = \{x \in \mathfrak{S}^I : x^n \in E\}$. By (20a) and (23b) we have from this $\mathrm{Prob}[(\bar{F}_N U, Y|\bar{F}_N U)^N \in E_N \times \Gamma_N] \to 1$, i.e.,

$$(\mu \bar{F}_N^{-1} \nu)^N (E_N \times \Gamma_N) \to 1 \quad \text{as} \quad N \to \infty. \tag{25}$$

It is clear that the inequality

$$\limsup_{N\to\infty} R(\mu \bar{F}_N^{-1} \nu; \epsilon, N) < R \tag{26}$$

follows from (24) and (25). Since ϵ has been chosen arbitrarily, and since R has been chosen arbitrarily subject only to the condition $R > H^*(\mu) + h(\nu)$, (18) is in turn implied by (26).

<u>Lemma 3</u>. Suppose for any $x \in \mathfrak{X}^I$ there is a sequence $\{F_N^x\}_{N=1}^{\infty}$ such that for any N large enough $R(F_N^x) \leq R_2$ for all $x \in \mathfrak{X}^I$ and $\mathrm{Prob}[(\mathfrak{Y}|x)^N \in F_N^x] \to 1$ as $N \to \infty$. Let $\gamma > 0$ be arbitrary. Then there is a sequence $\{\Gamma_N\}_{N=1}^{\infty}$ such that $R(\Gamma_N) \leq R_2 + \gamma$ and $\mathrm{Prob}[(Y|x)^N \in \Gamma_N] \to 1$ for any $x \in \mathfrak{X}^I$.

<u>Proof</u>. For any $x \in \mathfrak{X}^I$ let R_x^* denote the infimum of rates R such that $\inf\{\mathrm{Prob}[(Y|x)^N \in \mathfrak{Y}^N \setminus G_N] : R(G_N) \leq R\} \to 0$ as $N \to \infty$. Our assumption entails that $\sup\{R_x^* : x \in \mathfrak{X}^I\} \leq R_2 < R_2 + \gamma$. As Ziv (cf. /9/, Section II) we construct a sequence $\{\Gamma_N\}_{N=1}^{\infty}$ such that $R(\Gamma_N) \leq R_2 + \gamma$. In fact, we can construct a sequence $\{B_N\}_{N=1}^{\infty}$ of binary block code books $B_N \subset \mathfrak{B}^{2N(R_2+\gamma)}$ ($\mathfrak{B} = \{0,1\}$), where each B_N contains the vector $(0,\ldots,0)$ of length $2^{N(R_2+\gamma)}$ indicating an encoding error. Let Φ_N denote the corresponding binary block coding function, and let $\Gamma_N = \{y \in \mathfrak{Y}^N : \Phi_N(y) \neq 0\}$. Then $R(\Gamma_N) \leq R(B_N)$ for each N. Since $R_2 + \gamma > \sup\{R_x^* : x \in \mathfrak{X}^I\}$, a slight modification of the proof of theorem 2 in /9/ yields that $\mathrm{Prob}[(Y|x)^N \in \mathfrak{Y}^N \setminus \Gamma_N] \to 0$ for any $x \in \mathfrak{X}^I$, and we are done.

<u>Corollary 4</u>. Let the hypotheses of Lemma 2 prevail. Let $\eta_N = \eta(\mu \bar{F}_N^{-1}, \nu)$ (cf. (7)). Then

$$\sup_{0<\epsilon<1} \limsup_{N\to\infty} R(\eta_N; \epsilon, N) \leq H^*(\mu) + h(\nu). \tag{27}$$

<u>Remark 3</u>. Note that (27) is not a bound to the output rate of the channel ν, for we do not have a fixed input process which is being transmitted across the channel. Nevertheless, subsequent considerations will show that (27) will entail an appropriate version of the block transmission theorem for the sequence $\{\eta_N\}_{N=1}^{\infty}$ and a second $\bar{d}$-continuous stationary channel.

Choose $\epsilon > 0$ and find N_0 so large that (14) is true with $\epsilon/4$. For $N \geq N_0$, the block coding functions F_N are defined as in Lemma 1, for $N \leq N_0$, let $F_N : \mathfrak{U}^N \to \mathfrak{X}(1)^N$ be an arbitrary mapping. Let $\eta_N = \eta(\mu \bar{F}_N^{-1}, \nu_{(1)})$. Then

$$\sup_{0<\delta<1} \limsup_{N\to\infty} R(\eta_N; \delta, N) \leq H^*(\mu) + h(\nu_{(1)}). \tag{28}$$

Let us choose $R > 0$ and then $\gamma > 0$ so small that

$$H^*(\mu) + h(\nu_{(1)}) < R - \gamma < R < C^*(\nu_{(2)}). \tag{29}$$

This is possible by (12) (for $K = 2$). Since $\nu_{(2)}$ is $\bar{d}$-continuous, and since $R < C^*(\nu_{(2)})$, as in the proof of Lemma 1 we find an N_1 such that for any $N \geq N_1$ there exists a $(<2^{NR}>, N, \epsilon/4)$ channel code $\{(x(2)_i, \Gamma(2)_i) : 1 \leq i \leq <2^{NR}>\}$ for the channel $\nu_{(2)}$. Using the first inequality in (29) and (28) we find an N_2 such that for any $N \geq N_2$ there is a set $\Gamma_N \subset \mathfrak{Y}(1)^N$ with

$$R(\Gamma_N) = N^{-1}\log\|\Gamma_N\| \leq R - \gamma \tag{30}$$

and

$$\eta_N^N(\Gamma_N) = [\eta(\mu \bar{F}_N^{-1}, \nu_{(1)})]^N(\Gamma_N) > 1 - \epsilon/4. \tag{31}$$

By (29) and (30), if $N \geq \max\{N_0, N_1, N_2\}$ then there are again more codewords in the channel code than in the source code. Proceeding exactly as in the proof of Lemma 1, for any such N we find block coding functions $J_N : \mathfrak{Y}(1)^N \to \mathfrak{X}(2)^N$, $L_N : \mathfrak{Y}(2)^N \to \mathfrak{Y}(1)^N$ such that

$$\mathrm{Prob}[(Y|\bar{F}_N U)^N \neq L_N[(Y|\bar{J}_N(Y|\bar{F}_N U))^N]] < \epsilon/2. \tag{32}$$

Recall that at the same time

$$\mathrm{Prob}[U^N \neq H_N[(Y|\bar{F}_N U)^N]] < \epsilon/2. \tag{33}$$

Lemma 5. Let A, B be finite sets, and let $F : A \to B$. Let X, $\hat{X} : \Omega \to A$ and $U : \Omega \to B$ be random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$. Suppose $P[X = \hat{X}] > 1 - \delta$ and $P[F \circ X = U] > 1 - \epsilon$. Then $P[F \circ \hat{X} = U] > 1 - (\epsilon + \delta)$.

Proof. Since $[U = F \circ X] \cap [F \circ X = F \circ \hat{X}] \subset [U = F \circ \hat{X}]$ and $[X = \hat{X}] \subset$ $\subset [F \circ X = F \circ \hat{X}]$, the union bound for the complements of events $[U = F \circ X]$ and $[X = \hat{X}]$ gives the desired result.

Let us define, for $N \geq \max\{N_0, N_1, N_2\}$.

$$F_N : \mathcal{U}^N \to \mathcal{X}(1)^N \quad \text{(cf. Lemma 1)};$$

$$K_N^{(1)} = J_N : \quad \mathcal{Y}(1)_N^N \to \mathcal{X}(2)^N ; \tag{34}$$

$$G_N = H_N \circ L_N : \quad \mathcal{Y}(2)^N \to \mathcal{U}^N ,$$

where H_N is the decoding function from Lemma 1, and J_N, L_N have been defined so that (32) is valid. Take $A = \mathcal{Y}(1)^N$, $B = \mathcal{U}^N$, $X = (Y|\bar{F}_N U)^N$, $\hat{X} = L_N[(Y|\bar{J}_N(Y|\bar{F}_N U))^N]$, $U = U^N$, $F = H_N$, and $\epsilon = \delta = \epsilon/2$ in Lemma 5. Then (32), (33), (34), and Lemma 5 imply that $\mathrm{Prob}[U^N \neq G_N[(Y|\bar{J}_N(Y|\bar{F}_N U))^N]] < \epsilon$, and the proof of our theorem is complete.

5. CONCLUSION

Let us close by pointing out several open problems. There is a general problem always met in Shannon theory, namely, to find calculable formulae for the information quantities involved in coding theorems. The channel entropy can be calculated directly from the definition only in rather special cases (e.g., for the binary symmetric channel).

In the paper we followed the traditional approach of separating source and channel coding problems. That is, a good joint source and channel code was constructed from a good source code and a good channel code. It is of interest to consider Kieffer's weak channel codes for that purpose.

Finally it would be desirable to have also sliding-block and zero-error stationary transmission theorems in our setup.

REFERENCES

1. Billingsley P. Ergodic Theory and Information. J.Wiley, New York, 1965.
2. Gallager R.G. Information Theory and Reliable Communication. J.Wiley, New York, 1968.
3. Gray R.M., Ornstein D.S. IEEE Trans.Inform.Theory, 1979, IT-25, p.292.

4. Neuhoff D.L., Shields P.C. IEEE Trans.Inform.Theory, 1979, IT-25, p.440.
5. Neuhoff D.L., Shields P.C. Ann.Probab., 1981, 10, p.188.
6. Pötschke D., Sobik F. Mathematische Informationstheorie. Akademie V., Berlin, 1980.
7. Šujan Š. Trans. 3rd CS-SU-H Seminar on Inform.Theory, UTIA CSAV, Prague, 1980, p.201.
8. Šujan Š. Kybernetika, 1983, 19, supplement, 68 pp.
9. Ziv J. IEEE Trans. Inform.Theory, 1973, IT-18, p.384.

Received by Publishing Department
on November 5, 1984.

Шуян Ш. E5-84-713

Блочное кодирование для переноса источника через серию дискретных стационарных каналов

Получены границы для глобальной скорости создания информации на выходе стационарного, $\bar{d}$-непрерывного и условно почти блочно-независимого канала. Эти границы применяются к доказательству теоремы совместного блочного кодирования для переноса стационарного дискретного источника через упорядоченную серию K-стационарных и $\bar{d}$-непрерывных каналов, первые (K-1) из которых являются также условно почти блочно-независимыми.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна 1984

Šujan Š. E5-84-713

Joint Source and Channel Block Coding for Series of Discrete Stationary Channels

Bounds on the global output rate of a stationary, $\bar{d}$-continuous and conditionally almost block-independent channel are obtained. Using them, a joint source and channel block coding theorem for transmission of a discrete stationary source over an ordered K-tuple of stationary $\bar{d}$-continuous channels, the first K-1 of which are also conditionally almost block independent, is derived.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna 1984