

Separation of Random Number Generation and Resolvability *

K. Visweswariah, S. Kulkarni, S. Verdú
Department of Electrical Engineering
Princeton University, Princeton, NJ 08544

Abstract

We consider the problem of when a given source can approximate the output due to any input to a given channel. We provide necessary and sufficient conditions for a general source and channel. For the special case of a full rank discrete memoryless channel we give a stronger necessary condition than for a general channel.

1 Introduction

The classical separation theorem essentially states that source and channel coding can be done separately without losing optimality. In [1] a Source-Channel separation theorem was shown that is more general than the classical separation theorem. Here we investigate an analogous separation theorem in the case of the resolvability and the random number generation. Random number generation involves finding a deterministic transformation to generate a sequence of equiprobable bits from a given source of randomness. Resolvability is a property of a channel which gives the amount of randomness required to simulate, at the output of the channel, any distribution that can be achieved by a random input to the channel. We consider the two problems together, i.e. when can a given source of randomness be used to simulate the output of a given channel due to any input to arbitrary accuracy. In Section 2 we formalize the problem we consider. Sections 3 and 4 give sufficient and necessary conditions for a source to approximate the output of a given channel due to any input. Section 5 deals with the special case when the channel is a discrete memoryless channel of full rank. For this special case we will be able to show a necessary condition which is stronger than the necessary conditions in Section 4 for a general channel.

2 Preliminaries and Problem Formulation

In this section we will give some basic definitions, give a precise statement of the problem we are looking at and give results which are already known for the problem of separation of resolvability and random number generation.

*This work was partially supported by the National Science Foundation under Grants NYI Award IRI-9457645 and NCR 9523805

This definition is as in [3] and we repeat it for the sake of completeness.

Definition 1 A channel \mathbf{W} with input alphabet A and output alphabet B is a sequence of conditional distributions

$$\mathbf{W} = \{W^n(y^n|x^n); (x^n, y^n) \in A^n \times B^n\}_{n=1}^{\infty}.$$

Definition 2 A source of randomness \mathbf{X} is a sequence of finite dimensional distributions $\{P_{X^n}\}_{n=1}^{\infty}$ with X^n taking values in A^n .

Note that there are no consistency requirements on the sequence of finite dimensional distributions, this is the difference between Definition 2 and the standard definition of a random process. Throughout we assume that the source and channel alphabets are finite.

We now give some notation that use throughout this paper. Let the output distribution, when the input is distributed according to Q^n be denoted by $Q^n W^n$, thus

$$Q^n W^n(y^n) = \sum_{x^n \in A^n} W^n(y^n|x^n) Q^n(x^n).$$

Also let

$$i_{X^n W^n}(a^n, b^n) \triangleq \log \frac{W^n(b^n|a^n)}{P_{Y^n}(b^n)}$$

(where P_{Y^n} is the output distribution when P_{X^n} is input to the channel W^n) and

$$h_{Z^n}(z^n) \triangleq \log \frac{1}{P_{Z^n}(z^n)}.$$

We will use l_1 distance to measure the difference between two distributions on the same alphabet. We will denote the distance between P and Q by $d(P, Q)$. We note that $d(P, Q) = 2 \sup_{E \subseteq A} |P(E) - Q(E)|$ where A is the alphabet on which the two distributions are defined.

We now define precisely what we mean by a source \mathbf{Z} being an approximating source for a channel \mathbf{W} .

Definition 3 For any $\epsilon > 0$, the source \mathbf{Z} with alphabet F is called an ϵ -approximating source for the channel \mathbf{W} if for any arbitrary input source $\tilde{\mathbf{X}}$, there exists a sequence of deterministic mappings $\{\phi_n : F^n \rightarrow A^n\}$ such that for sufficiently large n ,

$$d(Y^n, \tilde{Y}^n) < \epsilon$$

where Y^n and \tilde{Y}^n are the outputs of the channel due to $\phi(Z^n)$ and \tilde{X}^n respectively.

Using this definition of an ϵ -approximating source we can now define \mathbf{Z} to be an approximating source for \mathbf{W} if it is ϵ -approximating for \mathbf{W} for all $\epsilon > 0$.

The problem of resolvability of a channel was considered in [3]. The resolvability of a channel is the minimum number of random bits required per input sample to approximate arbitrarily well the output of the channel due to any input process (See [3] for a formal definition of resolvability). It is shown in [3] that the resolvability of a channel

is $\sup_{\mathbf{X}} \bar{I}(\mathbf{X}, \mathbf{Y})$. The problem we look at here is that of finding necessary and sufficient conditions for a given source \mathbf{Z} to be an approximating source for a channel \mathbf{W} . To use a given source to approximate the output of a channel due to another source we could use the source to generate random bits at the best possible rate and then use a deterministic transformation of the random bits at the input of the channel. The problem of random bit generation was considered in [2] where fundamental limits on the rate at which random bits can be generated from a given source were given. Using the results of [3] and [2] and a two step process outlined above we can easily show the following sufficient condition:

Theorem 1 *If $\underline{H}(\mathbf{Z}) > S$ then \mathbf{Z} is an approximating source for a channel \mathbf{W} , where S is the resolvability of the channel \mathbf{W} .*

The necessary condition below can also be derived using the results in [3]. A source \mathbf{Z} can be approximated using $\bar{H}(\mathbf{Z})$ random bits per sample (Theorem 3, [3]). This also means any process derived from \mathbf{Z} by a deterministic transformation can be approximated using $\bar{H}(\mathbf{Z})$ random bits per sample. Thus if a source \mathbf{Z} is an approximating source for a channel then we can approximate any output of the channel with $\bar{H}(\mathbf{Z})$ random bits per sample and so the resolvability S of the channel must be smaller than $\bar{H}(\mathbf{Z})$. Thus we have the following theorem:

Theorem 2 *If $\bar{H}(\mathbf{Z}) < S$ then \mathbf{Z} is not an approximating source for a channel \mathbf{W} .*

In the next two sections we will give stronger versions of the above two theorems.

3 A Sufficient Condition

In this section we give a sufficient condition for a source \mathbf{Z} to be an approximating source for a channel \mathbf{W} which implies the sufficient condition given by Theorem 1. The sufficient condition is analogous to sufficient condition derived in [1] for the Source-Channel separation problem. We will define first the notion of a source strictly dominating a channel analogous to the definition in [1].

Definition 4 *A source \mathbf{Z} is said to strictly dominate a channel \mathbf{W} if for every channel input process \mathbf{X} there exists a $\delta > 0$ such that*

$$\lim_{n \rightarrow \infty} \inf_{c_n \in \mathbb{R}} \left\{ P \left[\frac{1}{n} h_{Z^n}(Z^n) \leq c_n + \delta \right] + P \left[\frac{1}{n} i_{X^n W^n}(X^n; Y^n) \geq c_n \right] \right\} = 0.$$

We note that $\underline{H}(\mathbf{Z}) > S$ implies that \mathbf{Z} dominates \mathbf{W} since if we take $c_n = (\underline{H}(\mathbf{Z}) + S)/2$ the required condition is satisfied. The converse statement however, is not true.

Theorem 3 *If a source \mathbf{Z} strictly dominates a channel \mathbf{W} then \mathbf{Z} is an approximating source for the channel \mathbf{W} .*

Proof The source \mathbf{Z} strictly dominates the channel \mathbf{W} implies that for each input process \mathbf{X} there exists a sequence $\{c_n\}_{n=1}^{\infty}$ and a $\delta > 0$ such that

$$P \left[\frac{1}{n} h_{Z^n}(Z^n) \leq c_n + \delta \right] \leq \tau_n \tag{1}$$

and

$$P\left[\frac{1}{n}i_{X^nW^n}(X^n; Y^n) \geq c_n\right] \leq \tau_n \quad (2)$$

where $\tau_n \rightarrow 0$ as $n \rightarrow \infty$. Equation (1) implies that we can approximate any distribution with type less than $\exp n(c_n + \frac{2}{3}\delta)$ using the distribution Z^n . To show this we use the procedure in the Aggregation Lemma [2].

Define $S^{(n)} = \{z^n \in F^n : P_{Z^n}(z^n) \leq \exp^{-n(c_n+\delta)}\}$. Consider any M -type distribution P_M on $\{1, 2, \dots, M\}$. We will place elements of $S^{(n)}$ in bin $B_n(i)$ until we have

$$P_{Y^n}(B_n(i)) > P_M(i) - \exp^{-n(c_n+\delta)}.$$

We stop either when we complete this process for all $i = 1, 2, \dots, M$ or when we run out of sequences in the set $S^{(n)}$. All remaining sequences in F^n are placed in $B_n(1)$. At the end of this process we have

$$\sum_{i=1}^M |P_{Z^n}(B_n(i)) - P_M(i)| \leq \max(2M \exp^{-n(c_n+\delta)}, 2\tau_n + M \exp^{-n(c_n+\delta)})$$

For any $M \leq \exp n(c_n + \frac{2}{3}\delta)$ the right hand side of the last equation goes to zero, since $\tau_n \rightarrow 0$.

We will now state and use a lemma the proof of which readily follows from the argument used in the proof of Theorem 4, [3].

Lemma 1 *If*

$$P\left[\frac{1}{n}i_{X^nW^n}(X^n; Y^n) \geq c_n\right] \leq \tau_n$$

where $\tau_n \rightarrow 0$ as $n \rightarrow \infty$ then for any $\gamma > 0$ there exists a process \tilde{X} such that

$$\lim_{n \rightarrow \infty} d(Y^n, \tilde{Y}^n) = 0$$

and \tilde{X}^n is an M -type distribution with $M \leq \exp n(c_n + \gamma)$

Equation (2) and Lemma 1 imply that there exist a process \tilde{X}^n with type smaller than $\exp n(c_n + \frac{2}{3}\delta)$ which approximates the output due to X . We can approximate arbitrarily closely any distribution with type less than or equal to $\exp n(c_n + \frac{2}{3}\delta)$ (and hence \tilde{X}^n) using Z^n . This along with the fact that $d(PW, QW) \leq d(P, Q)$ for any channel W and distributions P, Q imply that Z can be used to approximate the output of the channel W when the input process is X . Since we can find a sequence $\{c_n\}_{n=1}^\infty$ and a $\delta > 0$ which satisfy equations (1) and (2) for any input process X the source Z can approximate the output due to any process X .

□

4 Necessary Condition

In this section we will give necessary conditions for a source Z to be an approximating source for a channel W . Again we would like to have a result completely analogous to the necessary condition in [1] for the Source-Channel separation problem. To state that necessary condition we will need the notion of a source Z dominating as opposed to strictly dominating a channel W .

Definition 5 A source \mathbf{Z} dominates the channel \mathbf{W} if for every process \mathbf{X} , for every $\delta > 0$ and for every sequence of non-negative numbers $\{c_n\}_{n=1}^{\infty}$

$$\lim_{n \rightarrow \infty} P \left[\frac{1}{n} h_{\mathbf{Z}^n} (Z^n) \leq c_n - \delta \right] P \left[\frac{1}{n} i_{\mathbf{X}^n \mathbf{W}^n} (X^n; Y^n) \geq c_n \right] = 0.$$

It can be verified that if a source strictly dominates a channel then it dominates the channel. Also note that $\bar{H}(\mathbf{Z}) < S$ implies that the source does not dominate the channel. The final result we would like to have would be: \mathbf{Z} is an approximating source for a channel \mathbf{W} implies that the source dominates the channel which would be a stronger statement than Theorem 2. We have not been able to show this statement but we give two necessary conditions, one of which gives a stronger statement than Theorem 2 and the other neither implies nor is implied by Theorem 2.

Theorem 4 Suppose that for some process \mathbf{X} , for some $\alpha, \delta > 0$ and for some sequence of non-negative numbers $\{c_n\}_{n=1}^{\infty}$

$$P \left[\frac{1}{n} h_{\mathbf{Z}^n} (Z^n) \leq c_n - \delta \right] > \beta_n \tag{3}$$

and

$$P \left[\frac{1}{n} i_{\mathbf{X}^n \mathbf{W}^n} (X^n; Y^n) \geq c_n \right] > \alpha \tag{4}$$

for all $n \in I$ where I is an infinite set of integers and $\beta_n \rightarrow 1$ as $n \rightarrow \infty$. Then the source \mathbf{Z} is not an approximating source for the channel \mathbf{W} .

Proof In this proof we will concentrate on $n \in I$. We will prove this theorem by contradiction. Assume Equations (3),(4) are satisfied and that \mathbf{Z} is an approximating source for the channel \mathbf{W} .

We note that if $V^n = \phi^n(Z^n)$ and if Z^n satisfies Equation (3) then V^n also satisfies the same property. This implies that for any deterministic transformation ϕ^n , $\phi^n(Z^n)$ can be approximated arbitrarily closely for sufficiently large $n \in I$ by an $M = \exp n \left(c_n - \frac{\delta}{2} \right)$ type distribution. To see this consider V^n satisfying

$$P \left[\frac{1}{n} h_{V^n} (V^n) \leq c_n - \delta \right] > \beta_n$$

where $\beta_n \rightarrow 1$ as $n \rightarrow \infty$. Let

$$B^n = \left\{ v^n : \frac{1}{n} h_{V^n} (V^n) \leq c_n - \delta \right\}.$$

Clearly $|B^n| \leq \exp n (c_n - \delta)$. We can approximate the distribution of V^n to within $2 \frac{|B^n|}{M} + (1 - P_{V^n}(B^n))$ with an M type distribution. Thus the distribution P_{V^n} can be approximated to within $\exp - \left(n \frac{\delta}{2} \right) + (1 - \beta_n)$ by an $M = \exp n \left(c_n - \frac{\delta}{2} \right)$ type distribution.

Now using a proof similar to the proofs of Lemmas 7 and 8 in [3] Equation (4) implies that for sufficiently small $\epsilon > 0$ we can find N distributions $\{Q_i\}_{i=1}^N$ with

$$\frac{1}{n} \log \log N > c_n - \frac{\delta}{10}$$

and

$$\min_{i \neq j} d(Q_i^n W^n, Q_j^n W^n) \geq 2\epsilon$$

for all sufficiently large $n \in I$.

Input distributions which are ϵ -apart result can result in output distributions which are at most ϵ -apart. We note that the Definition of an ϵ -approximating source in Definition 3 is the same as its uniform version where the “sufficiently large n ” for which the statement holds is independent of \tilde{X} . This follows from a proof similar to that in Lemma 6 of [3]. Thus if Z^n is an approximating source for the channel then we must have for sufficiently large $n \in I$ at least

$$\exp\left(\exp n \left(c_n - \frac{\delta}{10}\right)\right)$$

distributions derived by deterministic transformations from Z^n , no two of which are less than 2ϵ -apart.

We have already seen that for any $\epsilon > 0$, any distribution derived by a deterministic transformation from Z^n can be approximated to within ϵ by an $M = \exp n \left(c_n - \frac{\epsilon}{2}\right)$ type distribution. The number of these M type distributions is upper bounded by

$$|A^n|^{\exp n (c_n - \frac{\epsilon}{2})}.$$

Which for sufficiently large n is less than

$$\exp\left(\exp n \left(c_n - \frac{\delta}{10}\right)\right)$$

Thus the number of distributions derived by deterministic transformations from Z^n is not enough to approximate outputs of the channel to all possible inputs and so we have a contradiction. □

We note that the conditions in Equations (3), (4) imply the negation of the condition that the source Z dominates the channel W and thus the necessary condition we have is weaker than the necessary condition analogous to the separation for Source-Channel separation problem in [1]. The result is stronger than that in Theorem 2.

We now give a stronger definition of an approximating source and give another necessary condition for a source to be a strong approximating source.

Definition 6 For any $\epsilon > 0$, the source Z with alphabet F is called a strong ϵ -approximating source for the channel W if for any arbitrary input source \tilde{X} , there exists a sequence of deterministic mappings $\{\phi_n : F^n \mapsto A^n\}$ such that $P_{\phi_n(Z^n)} \ll P_{\tilde{X}^n}$ and such that for sufficiently large n ,

$$d(Y^n, \tilde{Y}^n) < \epsilon$$

where Y^n and \tilde{Y}^n are the outputs of the channel due to $\phi(Z^n)$ and \tilde{X}^n respectively.

We call Z a strong approximating source for W if it is a strong ϵ -approximating source for W for all $\epsilon > 0$.

Note that the only difference between Definitions 3 and 6 is that the latter places an extra constraint that the deterministic transformation can only map to those sequences which have non-zero probability under the source whose output we are trying to approximate.

The next theorem states precisely a necessary condition for the source to be a strong approximating source for a channel. Again the statement is in the contra-positive form.

Theorem 5 *Suppose that for some process X , for some $\alpha, \delta > 0$ and for some sequence of non-negative numbers $\{c_n\}_{n=1}^\infty$*

$$P \left[\frac{1}{n} h_{Z^n} (Z^n) \leq c_n - \delta \right] > \alpha \quad (5)$$

and

$$P \left[\frac{1}{n} i_{X^n W^n} (X^n; Y^n) \geq c_n \right] > \beta_n \quad (6)$$

for all $n \in I$ where I is an infinite set of integers and $\beta_n \rightarrow 1$ as $n \rightarrow \infty$. Then the source Z is not a strong approximating source for the channel W .

Proof Equation (6) along with Feinstein's Lemma [6] implies that there exists a code of length n with $\exp n \left(c_n - \frac{\delta}{10} \right)$ codewords and with maximal probability of error less than ϵ_n for $n \in I$ where $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$. Consider a good code with $M = \exp n \left(c_n - \frac{\delta}{10} \right)$ codewords. Let the codewords be $\{b_i\}_{i=1}^M$ and their corresponding decoding sets be $\{D_i\}_{i=1}^M$. Consider now the process \tilde{X}^n which places mass $\frac{1}{M}$ on each of the M codewords, $\{b_i\}_{i=1}^M$. We will show that the output due to this process cannot be approximated well with our source.

Consider the set $S^n = \{z^n : \frac{1}{n} h_{Z^n} (Z^n) \leq c_n - \delta\}$. By Equation (5) $P_{Z^n} (S^n) > \alpha$ for all $n \in I$. Consider the N codewords to which the sequences in S^n get mapped. Assume the codewords are numbered so that these codewords are $\{b_i\}_{i=1}^N$. Since $|S^n| \leq \exp n (c_n - \delta)$, we have $N \leq \exp n (c_n - \delta)$.

Let

$$B = \bigcup_{i=1}^N D_i$$

and \tilde{Y}^n be the output due to \tilde{X}^n .

$$\begin{aligned} P_{\tilde{Y}^n} (B) &= \frac{1}{M} \sum_{i=1}^N P (B|b_i) + \frac{1}{M} \sum_{i=N+1}^M P (B|b_i) \\ &\leq \frac{N}{M} + \frac{\epsilon_n}{M} (M - N) \\ &\leq 2\epsilon_n \end{aligned}$$

where the last inequality holds for sufficiently large $n \in I$. Now let X be any process derived by a deterministic transformation from Z and let Y be the output process due to the process with X as the input to the channel.

$$P_{Y^n} (B) = \sum_{i=1}^M P (B|b_i) P_{X^n} (b_i)$$

$$\begin{aligned}
&\geq \sum_{i=1}^N P(D_i|b_i) P_{X^n}(b_i) \\
&\geq (1 - \epsilon_n) P_{Z^n}(S^n) \\
&\geq (1 - \epsilon_n) \alpha.
\end{aligned}$$

Thus we have

$$P_{\bar{Y}^n}(B) - P_{Y^n}(B) \geq (1 - \epsilon_n) \alpha - 2\epsilon_n$$

for all sufficiently large $n \in I$. We note that the R.H.S of the equation above does not go to zero as n increases and so \mathbf{Z} cannot be an approximating source in the strong sense for the channel \mathbf{W} .

□

5 A Special Case

We will now look at the special case where the channel is a full rank, discrete, memoryless channel (FRDMC). A channel \mathbf{W} is of full rank if the transition vectors $\{W(\cdot|a)\}_{a \in A}$ are linearly independent. For the FRDMC we will be able to show a result analogous to [1].

Theorem 6 *For a FRDMC if the source \mathbf{Z} is an approximating source for the channel \mathbf{W} then the source \mathbf{Z} dominates the channel \mathbf{W} .*

Proof We will show that if the source \mathbf{Z} does not dominate the FRDMC \mathbf{W} then \mathbf{Z} is not an approximating source for the channel \mathbf{W} . If a source \mathbf{Z} does not dominate \mathbf{W} then there is a process $\bar{\mathbf{X}}$ such that for some $\alpha, \delta > 0$ and for some sequence of non-negative numbers $\{c_n\}_{n=1}^{\infty}$

$$P\left[\frac{1}{n} h_{Z^n}(Z^n) \leq c_n - \delta\right] > \alpha \quad (7)$$

and

$$P\left[\frac{1}{n} i_{\bar{\mathbf{X}}^n \mathbf{W}^n}(\bar{\mathbf{X}}^n; \bar{\mathbf{Y}}^n) \geq c_n\right] > \alpha \quad (8)$$

for all $n \in I$ where I is an infinite set of integers. From Corollary 2 in [4] we have that for a FRDMC $S = C$. But for a DMC, $C = \sup_X I(X; Y)$. Since $S = \sup_X \bar{I}(\mathbf{X}; \mathbf{Y})$ we have for a FRDMC $\sup_X \bar{I}(\mathbf{X}; \mathbf{Y}) = \sup_X I(X; Y)$. Thus for a FRDMC if there is a process $\bar{\mathbf{X}}$ such that Equations (7) and (8) are satisfied then there exists an i.i.d process \mathbf{X} such that

$$c_n \leq I(X; Y) + \frac{\delta}{10} \quad (9)$$

for sufficiently large $n \in I$. We will show that the source \mathbf{Z} cannot approximate the output due to process \mathbf{X} .

Define the set of sequences in A^n that are not γ typical ($\gamma > 0$) by

$$D_{X^n}(\gamma) = \left\{x^n : \left| \frac{1}{n} N(a|x^n) - P_X(a) \right| > \gamma \text{ for some } a \in A \right\}$$

where $N(a|x^n)$ denotes the number of times that a occurs in the sequence x^n .

Also define the set of sequences jointly typical with x^n by

$$T_W^n(x^n, \gamma) = \left\{ y^n : \left| \frac{1}{n} N(a, b|x^n, y^n) - \frac{1}{n} N(a|x^n)W(b|a) \right| \leq \gamma \text{ for all } (a, b) \in A \times B \right\}$$

where $N(a, b|x^n, y^n)$ denotes the number of times that (a, b) occurs in (x^n, y^n) .

Let W denote the matrix whose columns are the transition vectors $\{W(\cdot|a)\}_{a \in A}$. Let \mathbf{r} be an $|A|$ dimensional vector and let $\mathbf{s} = W\mathbf{r}$. Since W is of full rank if $|r_i| > \gamma$ for some $i \in \{1, 2, \dots, |A|\}$ then $|s_j| > k\gamma$ for some $j \in \{1, 2, \dots, |B|\}$ and some $k > 0$ independent of \mathbf{r} . Thus we have that if $x^n \in D_X^n(\gamma)$ then $WN(x^n)/n$ differs from WP_X by $k\gamma$ in at least one component, where $\mathbf{N}(x^n), P_X$ are column vectors consisting of $N(a|x^n), P_X(a)$ respectively for $a \in A$. Now if $y^n \in T_W^n(x^n, \gamma)$ then

$$\left| \frac{1}{n} N(b|y^n) - \sum_{a \in A} \frac{1}{n} N(a|x^n)W(b|a) \right| \leq |A|\gamma$$

for all $b \in B$. Thus we have that if $x^n \in D_X^n(\frac{\gamma}{k})$ and $y^n \in T_W^n(x^n, \frac{\gamma}{3|A|})$ then $\mathbf{N}(y^n)/n$ differs from P_Y in some component by at least $\frac{2\gamma}{3}$. Thus $y^n \in D_Y^n(\frac{2\gamma}{3})$. We have shown that if $x^n \in D_X^n(\frac{\gamma}{k})$ then $T_W^n(x^n, \frac{\gamma}{3|A|}) \subseteq D_Y^n(\frac{2\gamma}{3})$. But $P_{Y^n}(D_Y^n(\frac{2\gamma}{3})) \rightarrow 0$ as $n \rightarrow \infty$ and $W^n(T_W^n(x^n, \frac{\gamma}{3|A|})|x^n) \rightarrow 1$ as $n \rightarrow \infty$ at a rate independent of x^n (From Lemma 2.12 in [5]). So if the input $\phi^n(Z^n)$ approximates the output due to X^n then we must have $P_{Z^n}(\phi^n(Z^n) \in D_X^n(\frac{\gamma}{k})) \rightarrow 0$ as $n \rightarrow \infty$. Since if this were not so,

$$P_{\phi^n(Z^n)}W^n(D_Y^n(\frac{2\gamma}{3})) - P_{Y^n}(D_Y^n(\frac{2\gamma}{3})) > \beta$$

infinitely often, for some $\beta > 0$. This would imply that the input $\phi^n(Z^n)$ does not approximate the output due to X^n .

Now by a slight modification of Lemma 2.13 in [5]

$$\left| \frac{1}{n} \log \left| T_W^n \left(x^n, \frac{\gamma}{3|A|} \right) \right| - H(Y|X) \right| \leq \epsilon(\gamma)$$

for every $x^n \in D_{X^n}^c(\frac{\gamma}{k})$ where $\epsilon(\gamma)$ is continuous in γ , independent of x^n and $\epsilon(\gamma) \rightarrow 0$ as $\gamma \rightarrow 0$. Define $Q^n = \{z^n : \frac{1}{n} h_{Z^n}(z^n) \leq c_n - \delta\}$. Clearly $|Q^n| \leq \exp(n(c_n - \delta))$. Also define R^n as the image of Q^n under the mapping ϕ^n and

$$S^n = \bigcup_{x^n \in R^n \cap D_{X^n}^c(\frac{\gamma}{k})} T_W^n \left(x^n, \frac{\gamma}{3|A|} \right).$$

Then we have $|S^n| \leq \exp(n(c_n - \delta)) \exp(n(H(Y|X) + \epsilon(\gamma)))$ and $P_{\phi^n(Z^n)}W^n(S^n) \geq \frac{\alpha}{2}$ for sufficiently large $n \in I$. This is because

$$W^n \left(T_W^n \left(x^n, \frac{\gamma}{3|A|} \right) | x^n \right) \rightarrow 1$$

as $n \rightarrow \infty$ at a rate independent of x^n ,

$$P_{Z^n} \left[\phi^n(Z^n) \in D_{X^n}^c(\frac{\gamma}{k}) \right] \rightarrow 1$$

as $n \rightarrow \infty$ and

$$P_{Z^n} [\phi^n(Z^n) \in R^n] > \alpha$$

for all $n \in I$. Using Equation (9) we can upper-bound $|S^n|$ as

$$|S^n| \leq \exp n \left(H(Y) - \frac{9\delta}{10} + \epsilon(\gamma) \right) \leq \exp n \left(H(Y) - \frac{\delta}{2} \right)$$

for sufficiently large $n \in I$, where the second inequality holds if we chose $\gamma > 0$ sufficiently small. Thus we have $P_{Y^n}(S^n) \rightarrow 0$ as $n \rightarrow \infty$ (By the strong source coding theorem for i.i.d sources). Thus for any deterministic mapping ϕ^n we can find a set $S^n \subseteq B^n$ such that $P_{\phi^n(Z^n)} W^n(S^n) \geq \frac{\alpha}{2}$ for sufficiently large $n \in I$ and $P_{Y^n}(S^n) \rightarrow 0$ as $n \rightarrow \infty$. So Z cannot be an approximating source for the FRDMC W .

□

We note that Theorem 6 implies that if $\underline{H}(Z) < S$ then the source Z is not an approximating source for a FRDMC channel with resolvability S .

References

- [1] S. Vembu, S. Verdú, and Y. Steinberg, "The Source-Channel Separation Theorem Revisited," *IEEE Trans. Information Theory*, vol. 41, pp. 44-54, Jan 1995.
- [2] S. Vembu, S. Verdú, "Generating Random Bits from an Arbitrary source: Fundamental Limits," *IEEE Trans. Information Theory*, vol. 41, pp. 1322-1332, Sept. 1995.
- [3] T.S. Han, S. Verdú, "Approximation Theory of Output Statistics," *IEEE Trans. Information Theory*, vol. 39, pp. 752-772, May 1993.
- [4] T.S. Han, S. Verdú, "Spectrum invariancy under output approximation for full rank discrete memoryless channels," *Probl. Peredach. Inform.* (in Russian), no. 2, pp. 101-118, 1993.
- [5] I. Csiszár, Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981
- [6] A. Feinstein, "A new basic theorem in information theory," *IRE Trans. Information Theory*, vol. IT-4, pp. 2-22, 1954.