Chin. Ann. of Math.
4B (1) 1983

ON THE ERROR BOUND IN SLEPIAN-WOLF THEOREM

ZHANG ZHEN (张 箴)
(*Nankai University*)

Abstract

In this paper, we discuss the problem of estimation of the probability of error in Slepian-Wolf Theorem. We give both an upper bound and a lower bound of the error exponent of the best code $C(R_x, R_y)$. For a main part of the achievable rates, we have determined the error exponent completely, for the others, our estimation is accurate.

Slepian-Wolf theorem is considered as one of the most important results in the coding theory of correlated sources. In the original work[1], its proof was relatively tedious, therefore Ahlswede, R. F. and Körner, J. gave a better proof later. Recently, using random coding argument, Cover, T. M. has offered a still better proof with the advantage that, besides its simplicity, it also provides the possibility of estimating the error bound in Slepian-Wolf theorem.

In this paper, the problem relating to the estimation of the error bound in Slepian-wolf theorem is discussed, and the results obtained are tight.

This paper is prepared under the supervision of the research directors Professor Hu Guoding and Associate Professor Shen Shiyi. The auther is indebted to Prof. Shen for suggesting the use of the method employed in [3]; the thanks are also given to Mr. Wang Gungshu and Dai Changjun for their careful check-up on the paper.

§ 1. Introduction

The Slepian-Wolf coding problem can be described as follows: (Fig. 1)

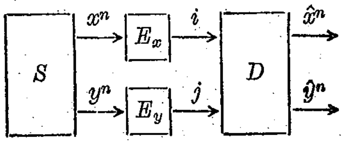


Fig. 1

In Fig. 1, $\mathscr{S}=[\mathscr{X}\otimes\mathscr{Y}, p(x, y)]$ is memoryless correlated sources, here $\mathscr{X}$, $\mathscr{Y}$ are finite sets, and $p(x, y)$ is a probability distribution over $\mathscr{X}\otimes\mathscr{Y}$. Let $\{(x_i, y_i)\}_{i=1}^n$ be the first n outputs of the sources, and $x^n=(x_1, \cdots, x_n)$, $y^n=(y_1, \cdots, y_n)$, then the

Manuscript received November 19, 1980.

joint distribution of x^n and y^n is

$$p^{(n)}(x^n, y^n)=\prod_{i=1}^{n} p(x_i, y_i). \tag{1.1}$$

Let $p(x|y)$, $p(y|x)$, $p(x)$ and $p(y)$ be the marginal and conditional distributions determined by $p(x, y)$. Similarly, in this paper, the denotations such as $p^{(n)}(x^n|y^n)$, $p^{(n)}(y^n)$ are used, the denotations $h(x|y)$, $h(x)$, $h(x^n|y^n)$ and $h(y^n|x^n)$ are also employed to express different kinds of entropy densities. It is assumed that all of the entropy densities take only finite values, for example, when $p(x|y)=0$, it is assumed that $h(x|y)=0$. Let $H(X, Y)$, $H(X|Y)$, $H(Y, X)$, $H(Y)$ and $H(X)$ be all kinds of the entropies of the random variables (X, Y) whose joint distribution is $p(x, y)$.

Encoders E_x, E_y. Observing x^n, y^n respectively, encode the messages by means of the following encoding functions.

$$f_x:\ \mathscr{X}^n\to I_{M_1}=\{1, 2, \cdots, M_1\}, \tag{1.2}$$

$$f_y:\ \mathscr{Y}^n\to I_{M_2}=\{1, 2, \cdots, M_2\}, \tag{1.3}$$

where $\mathscr{X}^n=\overbrace{\mathscr{X}\otimes\cdots\otimes\mathscr{X}}^{n}$, $\mathscr{Y}^n=\overbrace{\mathscr{Y}\otimes\cdots\otimes\mathscr{Y}}^{n}$. $r_x=\frac{1}{n}\ln M_1$, $r_y=\frac{1}{n}\ln M_2$ are called the rates of the code.

Decoder D. Observing the outputs of the encoders $i=f_x(x^n)$, $j=f_y(y^n)$ simultaneously, reproduce messages x^n. y^n. Its outputs are denoted by $\hat{x}^n$, $\hat{y}^n$. The decoding function is

$$g:\ I_{M_1}\otimes I_{M_2}\to\mathscr{X}^n\otimes\mathscr{Y}^n. \tag{1.4}$$

The probability of error is defined as

$$P_e^{(n)}=p^{(n)}\{(x^n, y^n):(\hat{x}^n, y^n)\neq(x^n, \hat{y}^n)\}. \tag{1.5}$$

Thus we have defined a triple (f_x, y_y, g) which is called a (n, r_x, r_y) code. A vector (R_x, R_y) is called an achievable vector, if there exists a series of $(n, r_x^{(n)}, r_y^{(n)})$ codes $\mathfrak{C}^{(n)}$ such that $r_x^{(n)}\leqslant R_x$, $r_y^{(n)}\leqslant R_y$ and

$$P_e^{(n)}(\mathfrak{C}^{(n)})\to 0 \quad (n\to\infty). \tag{1.6}$$

The set consisting of all of the achievable vectors is denoted by $\mathscr{R}$, which is determined by the following theorem.

Theorem 1. (Slepian-Wolf 1973)

$$\mathscr{R}=\{(R_x, R_y):R_x>H(X|Y),\ R_y>H(Y|X),\ R_x+R_y>H(X, Y)\}. \tag{1.7}$$

For each $(R_x, R_y)\in\mathscr{R}$, let $\mathfrak{C}^{(n)}(R_x, R_y)$ *be the set consisting of all of the* (n, r_x, r_y) *codes satisfying* $r_y\leqslant R_y$, $r_x\leqslant R_x$ and

$$m_e^n(R_x, R_y)=\inf\{P_e^{(n)}(\mathfrak{C}):\mathfrak{C}\in\mathfrak{C}^{(n)}(R_x, R_x)\}, \tag{1.8}$$

$$C(R_x, R_y)=\varliminf_{n\to\infty}\left\{-\frac{1}{n}\ln m_e^n(R_x, R_y)\right\}. \tag{1.9}$$

In this paper, it is proved that for each $((R_x, R_y)\in\mathscr{R}$, $C(R_x, R_y)>0$, moreover both an upper and a lower bounds of $C(R_x, R_y)$ are offered; in fact, for a main part of achievable vectors, $C(R_x, R_y)$ are determined completely.

§ 2. Lemmas

In this section, developing the method used in § 1.6 of [3], we estimate the numbers of the elements in some sets and the probabilities of some events. The results obtained will be useful not only in this paper, but also for some other problems in the coding theory of eorrelated sources.

$\|A\|$ is used to denote the numbers of the elements in set A. For eaeh $y^n\in\mathscr{Y}^n$, if $y^n=(y_1, \cdots, y_n)$, let $t(y)=\|\{i: y_i=y\}\|$. For each $(x^n, y^n)\in\mathscr{X}^n\otimes\mathscr{Y}^n$, if $(x^n, y^n)=((x_1, y_1), \cdots, (x_n, y_n))$, let $s(x, y)=\|\{i: (x_i, y_i)=(x, y)\}\|$, and $\bar{p}(y)=\frac{t(y)}{n}$, $\bar{q}(x|y)=\frac{s(x, y)}{t(y)}$, $\bar{p}(x, y)=\frac{s(x, y)}{n}$. It is easy to know that $\bar{p}(x, y)$ is a distribution over $\mathscr{X}\otimes\mathscr{Y}$, the marginal and conditional distributions of which are $p(y)$ and $q(x|y)$ respectively. All of these distributions are called frequency distributions or, simply, frequencies.

If $[\mathscr{X}\otimes\mathscr{Y}, p(x, y)]$ is a 2-dimentional finite probabilistic space, f: $\mathscr{X}\otimes\mathscr{Y}\to R'$ is a real valued function, then $f^{(n)}$: $\mathscr{X}'\otimes\mathscr{Y}^n\to R'$ can be defined as follows: for each $(x^n, y^n)\in\mathscr{X}^n\otimes\mathscr{Y}^n$, $(x^n, y^n)=((x_1, y_1), \cdots, (x_n, y_n))$, $f^{(n)}(x^n, y^n)=\frac{1}{n}\sum_{i=1}^{n}f(x_i, y_i)$. Let $E_{\bar{p}}f=\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}}\bar{p}(y)p(x|y)f(x, y)$, where $\bar{p}(y)$ is a distribution over $\mathscr{Y}$. If $\bar{q}(x|y)$ is a conditional distributions, let

$$H(\bar{p}, \bar{q}, p)=-\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}}\bar{p}(y)\bar{q}(x|y)\ln p(x|y), \tag{2.1}$$

$$H(\bar{p}, \bar{q})=-\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}}\bar{p}(y)\bar{q}(x|y)\ln \bar{q}(x|y), \tag{2.2}$$

$$E_{\bar{p},\bar{q}}f=\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}}\bar{p}(y)\bar{q}(x|y)f(x, y). \tag{2.3}$$

We define the following function

$$J^{+}(f(x, y), K|\bar{p})=\inf_{\bar{q}\in Q}\{H(\bar{p}; \bar{q}, p)-H(\bar{p}; \bar{q})\}, \tag{2.4}$$

where Q is the set of all of the conditional distributions $\bar{q}(x|y)$ satisfying the following conditions

$$E_{\bar{p},\bar{q}}f\geqslant K. \tag{2.5}$$

Function $J^{-}(f(x, y), K|\bar{p})$ can be defined in the same way as $J^{+}(f(x, y), K|\bar{p})$, but the following (2.6) is used instead of (2.5)

$$E_{\bar{p},\bar{q}}f\leqslant K. \tag{2.6}$$

Let

$$J(f(x, y), K|\bar{p})=J^{+}(f(x, y), K|\bar{p})+J^{-}(f(x, y), K|\bar{p}). \tag{2.7}$$

Lemma 1. *If $y^n\in\mathscr{Y}^n$ and the frequency of y^n is $\bar{p}(y)$, we have*

$$p^{(n)}\{x^n: f^{(n)}(x^n, y^n)\geqslant K|y^n\}=\exp\{-nJ^{+}(f(x, y), K|\bar{p})+O(\ln n)\} \tag{2.8}$$

$$p^{(n)}\{x^n: f^{(n)}(x^n, y^n)\leqslant K|y^n\}=\exp\{-nJ^{-}(f(x, y), K|\bar{p})+O(\ln n)\}. \tag{2.9}$$

Proof Given $y^n \in \mathscr{Y}^n$, the number of x^n, for which the conditional frequency is $\bar{q}(x|y) = \frac{s(x, y)}{t(y)}$, is

$$\prod_{y\in\mathscr{Y}} \frac{t(y)!}{\prod\limits_{x\in\mathscr{X}} s(x, y)!}. \tag{2.10}$$

For each above x^n, it can be obtained that

$$p^{(n)}(x^n|y^n) = \prod_{x\in\mathscr{X}} \prod_{y\in\mathscr{Y}} [p(x|y)]^{s(x,y)}, \tag{2.11}$$

therefore

$$p^{(n)}\left\{x^n\text{: the frequency of } (x^n, y^n) \text{ is } \bar{p}(y)\bar{q}(x|y) = \frac{s(x, y)}{n}\Big|y^n\right\}$$

$$= \left\{\prod_{y\in\mathscr{Y}} \frac{t(y)!}{\prod\limits_{x\in\mathscr{Y}} s(x, y)!}\right\} \cdot \left\{\prod_{x\in\mathscr{X}} \prod_{y\in\mathscr{Y}} [p(x|y)]^{s(x,y)}\right\}. \tag{2.12}$$

By using stirling formula, it can be deduced from (2.12) that the left side of (2.12) is

$$\exp\{-n[H(\bar{p}; \bar{q}, p) - H(\bar{p}; \bar{q})] + O(\ln n)\}. \tag{2.13}$$

Noticing that

$$f^{(n)}(x^n, y^n) = \sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{Y}} \bar{p}(y)\bar{q}(x|y)f(x, y), \tag{2.14}$$

we can obtain that

$$p^{(n)}\{x^n\text{: } f^{(n)}(x^n, y^n) \geqslant K|y^n\} \geqslant \exp\{-nJ^+(f(x, y), K|\bar{p}) + O(\ln n)\}. \tag{2.15}$$

Considering that the total number of different frequencies is $e^{O(\ln n)}$, we obtain

$$p^{(n)}\{x^n\text{: } f^{(n)}(x^n, y^n) \geqslant K|y^n\} \leqslant e^{O(\ln n)} \exp\{-nJ^+(f(x, y), K|\bar{p}) + O(\ln n)\}, \tag{2.16}$$

which proves (2.8). (2.9) can be proved similarly.

Lemma 2. *If K satisfies the following condition*

$$\sum_{y\in\mathscr{Y}} \bar{p}(y) \min_{x\in\mathscr{X}} f(x, y) \leqslant K \leqslant \sum_{y\in\mathscr{Y}} \bar{p}(y) \max_{x\in\mathscr{X}} f(x, y), \tag{2.17}$$

then

$$J(f(x, y), K|\bar{p}) = \lambda_0 K - F(\bar{p}, f, \lambda_0), \tag{2.18}$$

where λ_0 is the unique real root of the following equation

$$\sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{Y}} \bar{p}(y) \frac{p(x|y)e^{\lambda f(x,y)}}{\sum\limits_{x'\in\mathscr{X}} p(x'|y)e^{\lambda f(x',y)}} f(x, y) = K, \tag{2.19}$$

and

$$F(\bar{p}, f, \lambda) = \sum_{y\in\mathscr{Y}} \bar{p}(y) \ln \sum_{x\in\mathscr{X}} p(x|y)e^{\lambda f(x,y)}. \tag{2.20}$$

The $\bar{q}$ which gives the infimum in (2.4) (the same for $J^-(f(x, y), K|\bar{p})$) is

$$\bar{q}(x|y) = \frac{p(x|y)e^{\lambda_0 f(x,y)}}{\sum\limits_{x'\in\mathscr{X}} p(x'|y)e^{\lambda_0 f(x',y)}}. \tag{2.21}$$

If $K \leqslant E_{\bar{p}}f$, then

$$J^+(f(x, y), K|\bar{p}) = 0; \tag{2.22}$$

if $K \geqslant E_{\bar{p}}f$, then

$$J^-(f(x, y), K|\bar{p}) = 0; \tag{2.23}$$

if $K \in [\sum_{y\in\mathscr{Y}} \bar{p}(y) \min_{x\in\mathscr{X}} f(x, y), \sum_{y\in\mathscr{Y}} \bar{p}(y) \max_{x\in\mathscr{X}} f(x, y)]$.

Let

$$J(f(x, y), K|\bar{p}) = +\infty. \tag{2.24}$$

The proof of this lemma is almost the same as in § 1.6 of [3].

Lemma 3. *$F(\bar{p}, f, \lambda)$ has the following properties:*

(1) $F(\bar{p}, f, 0) = 0,$ (2.25)

(2) $\left.\frac{\partial F}{\partial \lambda}\right|_{\lambda=0} = E_{\bar{p}} f,$ (2.26)

(3) $\frac{\partial^2 F}{\partial \lambda^2} \geqslant 0,$ (2.27)

and the necessary and sufficient condition that the equality stands in (2.27) is that for all x satisfying $p(x|y) \neq 0$, $f(x, y) \equiv c(y)$.

Proof Because of the simplicity of the proof of (1), (2), we prove only (3) here. This can be done by considering that

$$\frac{\partial^2 F}{\partial \lambda^2} = \sum_{y \in \mathscr{Y}} \bar{p}(y) \left\{ \frac{\sum_{x \in \mathscr{X}} p(x|y) e^{\lambda f(x, y)} f^2(x, y)}{\sum_{x' \in \mathscr{X}} p(x'|y) e^{\lambda f(x', y)}} - \left(\frac{\sum_{x \in \mathscr{X}} p(x|y) e^{\lambda f(x, y)} f(x, y)}{\sum_{x' \in \mathscr{X}} p(x'|y) e^{\lambda f(x', y)}} \right)^2 \right\}$$

$$\geqslant 0 \tag{2.28}$$

and the necessary and sufficient conditions for (2.28) to become an equality are that for all of the x for which $p(x|y) \neq 0$, $f(x, y) \equiv c(y)$.

Lemma 4. *$J(f(x, y), K|\bar{p})$ has the following properties:*

(1) $J^+(f(x, y), K|\bar{p}) = \sup_{0 \leqslant \lambda < \infty} \{\lambda K - F(\bar{p}, f, \lambda)\},$ (2.29)

$J^-(f(x, y), K|\bar{p}) = \sup_{-\infty < \lambda \leqslant 0} \{\lambda K - F(\bar{p}, f, \lambda)\},$ (2.30)

$J(f(x, y), K|\bar{p}) = \sup_{-\infty < \lambda < +\infty} \{\lambda K - F(\bar{p}, f, \lambda)\},$ (2.31)

(2) $\frac{\partial J}{\partial K} = \lambda, \quad \left.\frac{\partial J}{\partial K}\right|_{K = E_{\bar{p}} f} = 0,$ (2.32)

(3) $\frac{\partial^2 J}{\partial K^2} > 0,$ (2.33)

(4) $J(f(x, y), K|\bar{p}) \geqslant 0,$ (2.34)

the equality stands in (2.34) if and only if $K = E_{\bar{p}} f$.

The proof of this lemma is omitted.

Remark. Assume that $\mathscr{Y}$ is a single-element set, then the results in § 1.6 of [3] can be easily deduced from Lemmas 1 and 2. In this case, the simple denotations $F(f, \lambda)$ and $J(f(x), K)$ are used.

Lemma 5.

$$F(h(x), 1) = \ln \|\mathscr{X}_0\|, \text{ where } \mathscr{X}_0 = \{x: x \in \mathscr{X}, p(x) \neq 0\}; \tag{2.35}$$

$$F(h(x, y), 1) = \ln \|\mathscr{Z}_0\|, \text{ where } \mathscr{Z}_0 = \{(x, y): x \in \mathscr{X}, y \in \mathscr{Y}, p(x, y) \neq 0\} \tag{2.36}$$

$$F(\bar{p}, h(x, y), 1) = \sum_{y \in \mathscr{Y}} \bar{p}(y) \ln \|\{x: p(x|y) \neq 0\}\|. \tag{2.37}$$

The proof of this lemma is omitted.

Now, we estimate the numbers of the elements in the following sets:

(1) $\{x^n: h(x^n) \leqslant nK\},$ (2.38)

(2) $\{(x^n, y^n): h(x^n, y^n) \leqslant nK\}$, (2.39)

(3) $\{x^n: h(x^n | y^n) \leqslant nK\}$. (2.40)

First we consider (1). According to Lemma 1, if $K' > H(X)$, then

$$p^{(n)}\{x^n: nK' \leqslant h(x^n) \leqslant nK\} = \exp\{-nJ(h(x), K') + O(\ln n)\} - \exp\{-nJ(h(x), K) + O(\ln n)\}, \tag{2.41}$$

if $K > K'$, it can be easily deduced from the strict monotonicity of $J(h(x), K)$ (when $K > H(x)$) that

$$p^{(n)}\{x^n: nK' \leqslant h(x^n) \leqslant nK\} = \exp\{-nJ(h(x), K') + O(\ln n)\} \tag{2.42}$$

and every element in set $\{x^n: nK' \leqslant h(x^n) \leqslant nK\}$ satisfies that

$$e^{-nK} \leqslant p^{(n)}(x^n) \leqslant e^{-nK'}. \tag{2.43}$$

Therefore

$$\exp\{n(K - J(h(x), K')) + O(\ln n)\} \geqslant \|\{x^n: nK' \leqslant h(x^n) \leqslant nK\}\| \geqslant \exp\{n(K' - J(h(x), K)) + O(\ln n)\}. \tag{2.44}$$

Hence, for every $K' < K$, we obtain

$$\|\{x^n: h(x^n) \leqslant nK\}\| \geqslant \exp\{n(K' - J(h(x), K')) + O(\ln n)\} \tag{2.45}$$

and therefore

$$\|\{x^n: h(x^n) \leqslant nK\}\| \geqslant \exp\{n(K - J(h(x), K)) + o(n)\}. \tag{2.46}$$

If $K \leqslant \left.\frac{\partial F}{\partial \lambda}\right|_{\lambda=1}$, that is $\frac{\partial J}{\partial K} \leqslant 1$, from (2.44) it is obtained that

$$\begin{aligned}\|\{x^n: h(x^n) \leqslant nK\}\| &= \left\| \bigcup_{i=0}^{n-1} \{x^n: iK \leqslant h(x^n) \leqslant (i+1)K\} \right\| \\ &\leqslant n \exp\left\{ \max_{0 \leqslant i \leqslant n-1} n \left\{ \frac{i+1}{n} K - J\left(h(x), \frac{i+1}{n} K\right)\right\} + O(\ln n)\right. \\ &\leqslant \exp\{n(K - J(h(x), K)) + O(\ln n)\}.\end{aligned} \tag{2.47}$$

Then, by using (2.46), it is reached that

$$\|\{x^n: h(x^n) \leqslant nK\}\| = \exp\{n(K - J(h(x), K)) + o(n)\}. \tag{2.48}$$

If $K \leqslant H(X)$, we can prove (2.48) in the same way, thus we obtain the following

Lemma 6. *If* $K \leqslant \left.\frac{\partial F}{\partial \lambda}\right|_{\lambda=1}$, *then*

$$\|\{x^n: h(x^n) \leqslant nK\}\| = \exp\{n(K - J(h(x), K)) + o(n)\}; \tag{2.49}$$

if $K > \left.\frac{\partial F}{\partial \lambda}\right|_{\lambda=1}$, *then*

$$\|\{x^n: h(x^n) \leqslant nK\}\| = \exp\{n \ln\|x_0\| + o(n)\}. \tag{2.50}$$

Proof Only (2.51) has to be proved. Using the fact: if $K = \left.\frac{\partial F}{\partial \lambda}\right|_{\lambda=1}$, then $K - J(h(x), K) = F(h(x), 1) = \ln\|\mathscr{X}_0\|$, we obtain, if $K > \left.\frac{\partial F}{\partial \lambda}\right|_{\lambda=1}$

$$\|\{x^n: h(x^n) \leqslant nK\}\| \geqslant \exp\{n \ln\|\mathscr{X}_0\| + o(n)\}, \tag{2.51}$$

and the oppositely directed inequality is obvious.

The following Lemmas, the proof of which is omitted here, are similar to Lemma 6.

Lemma 6′.

(1) *If* $K\leqslant\frac{\partial F(h(x,y),\lambda)}{\partial\lambda}\Big|_{\lambda=1}$, then

$$\|(x^n, y^n): h(x^n, y^n)\leqslant nK\}\| = \exp\{n(K-J(h(x,y),K))+o(n)\}. \tag{2.52}$$

If $K>\frac{\partial F(h(x,y),\lambda)}{\partial\lambda}\Big|_{\lambda=1}$, *then*

$$\|\{(x^n, y^n): h(x^n, y^n)\leqslant nK\}\| = \exp\{n\ln\|\mathscr{X}_0\|+o(n)\} \tag{2.53}$$

($\mathscr{X}_0$, *see Lemma* 5).

(2) *If* $K\leqslant\frac{\partial F(\bar{p}, h(x|y),\lambda)}{\partial\lambda}\Big|_{\lambda=1}$, *then*

$$\|\{x^n: h(x^n|y^n)\leqslant nK\}\| = \exp\{n(K-J(h(x|y), K|\bar{p}))+o(n)\} \tag{2.54}$$

where $\bar{p}(y)$ *is the frequency of* y^n; *if* $K>\frac{\partial F(\bar{p}, h(x|y),\lambda)}{\partial\lambda}\Big|_{\lambda=1}$, *then*

$$\|\{x^n: h(x^n|y^n)\leqslant nK\}\| = \exp\{n\sum_{y\in\mathscr{Y}}\bar{p}(y)\ln\|\{x: p(x|y)\neq 0\}\|\}. \tag{2.55}$$

If $K\leqslant\frac{\partial F}{\partial\lambda}\Big|_{\lambda=0}$, function $R=K-J(h(x), K)$ is a monotone function the inverse of which can be denoted by $K=K(h(x), R)$. Let

$$W(h(x), R)=J(h(x), K(h(x), R)), \tag{2.56}$$

$$W^+(h(x), R)=J^+(h(x), K(h(x), R)), \tag{2.57}$$

$$W^-(h(x), R)=J^-(h(x), K(h(x), R)), \tag{2.58}$$

where the domains of definition of K, W^+, W^- and W are $0\leqslant R\leqslant\ln\|\mathscr{X}_0\|$, and $W=W^+=\infty$ if $R>\ln\|\mathscr{X}_0\|$. We can define functions $K(h(x,y), R)$, $W(h(x,y), R)$, $K(h(x|y), R|\bar{p})$, $W(h(x|y), R|\bar{p})$, etc similarly to the definitions of $K(h(x), R)$ and $W(h(x), R)$, and their domains of definition are $0\leqslant R\leqslant\ln\|\mathscr{X}_0\|$ and

$$0\leqslant R\leqslant\sum_{y\in\mathscr{Y}}\bar{p}(y)\ln\|\{x: p(x|y)\neq 0\}\|$$

respectively.

The following lemma is a consequence of Lemma 6.

Lemma 7

(1) *If* $\mathscr{X}^{n*}$ *is a subset with* e^{nR} *elements of* $\mathscr{X}^n$, *then*

$$p^{(n)}(\overline{\mathscr{X}^{n*}})\geqslant\exp\{-nW^+(h(x), R)+o(n)\}, \tag{2.59}$$

$$p^{(n)}(\overline{\mathscr{X}^{n*}}|y^n)\geqslant\exp\{-nW^+(h(x|y), R|\bar{p})+o(n)\}. \tag{2.60}$$

(2) *If* $\mathscr{Z}^{n*}$ *is a subset of* $\mathscr{X}^n\otimes\mathscr{Y}^n$ *with* e^{nR} *elements, then*

$$p^{(n)}(\overline{\mathscr{Z}^{n*}})\geqslant\exp\{-nW^+(h(x,y), R)+o(n)\}. \tag{2.61}$$

§ 3. The Lower Bound of the Probability of Error

In this section, we estimate the probability of error in Slepian-Wolf theorem. For each code $\mathfrak{C}\in\mathfrak{C}^{(n)}(R_x, R_y)$, the number of the pairs (x^n, y^n), which can be decoded correctly, is fewer than $\exp\{n(R_x+R_y)\}$. From the consequence of Lemma 6 we know that

$$P_e^{(n)}(\mathfrak{C}) \geqslant \exp\{-nW^+(h(x, y), R_x+R_y)+o(n)\}. \tag{3.1}$$

If the message $x_0^n \in \mathscr{X}_0$ is fixed, the number of y^n, which can be decoded correctly, must be fewer than e^{nR_y}, so it is obtained that

$$P_e^{(n)}\{y^n: y^n \neq \hat{y}^n | x_0^n\} \geqslant \exp\{-W^+(h(y|x), R_y|\bar{p})+o(n)\}, \tag{3.2}$$

therefore

$$P_e^{(n)}(\mathfrak{C}) \geqslant \sum_{x_0^n \in \mathscr{X}^n} p^{(n)}(x_0^n) \exp\{-W^+(h(y|x), R_y|\bar{p})+o(n)\}, \tag{3.3}$$

where $\bar{p}$ is the frequency of x_0^n. It can be proved similarly that

$$P_e^{(n)}(\mathfrak{C}) \geqslant \sum_{y_0^n \in \mathscr{Y}^n} p^{(n)}(y_0^n) \exp\{-W^+(h(x|y), R_x|\bar{p})+o(n)\}. \tag{3.4}$$

Following the method used in the proof of Lemma 1, we can prove that the right side of (3.4) is

$$\exp\{-\inf_{\bar{p}} n[W^+(h(x|y), R_x|\bar{p})+H_{\bar{p}}(p)-H_{\bar{p}}(\bar{p})]+o(n)\}, \tag{3.5}$$

where

$$H_{\bar{p}}(p) = -\sum_{y \in \mathscr{Y}} \bar{p}(y) \ln p(y), \tag{3.6}$$

$$H_{\bar{p}}(\bar{p}) = -\sum_{y \in \mathscr{Y}} \bar{p}(y) \ln \bar{p}(y). \tag{3.7}$$

Let

$$S^+(h(x|y), R) = \inf_{\bar{p}} \{W^+(h(x|y), R|\bar{p})+H_{\bar{p}}(p)-H_{\bar{p}}(\bar{p})\}. \tag{3.8}$$

(similarly we can define $S(h(x|y), R)$ and $S^-(h(x|y), R)$), where the infimum is taken over the set of all possible $\bar{p}(y)$. Hence

$$P_e^{(n)}(\mathfrak{C}) \geqslant \exp\{-nS^+(h(x|y), R_x)+o(n)\}, \tag{3.9}$$

$$P_e^{(n)}(\mathfrak{C}) \geqslant \exp\{-nS^+(h(y|x), R_y)+o(n)\}, \tag{3.10}$$

so we obtain the following

Theorem 2. *For Slepian-Wolf coding problem*

$$C(R_x, R_y) \leqslant \min\{W^+(h(x, y), R_x+R_y), S^+(h(x|y), R_x), S^+(h(y|x), R_y)\}. \tag{3.11}$$

Now we discuss the properties of the S-functions.

Lemma 8.

(1) *If* $R \leqslant H(X|Y)$, *then*

$$S(h(x|y), R) = S^-(h(x|y), R), \tag{3.12}$$

$$S^+(h(x|y), R) = 0. \tag{3.13}$$

(2) *If* $R \geqslant H(X|Y)$, *then*

$$S(h(x|y), R) = S^+(h(x|y), R), \tag{3.14}$$

$$S^-(h(x|y); R) = 0. \tag{3.15}$$

The proof is omitted.

Let

$$\mathscr{L} = W(h(x|y), R|\bar{p})+H_{\bar{p}}(p)-H_{\bar{p}}(\bar{p})-\mu\Big(\sum_{y \in \mathscr{Y}} \bar{p}(y)-1\Big), \tag{3.16}$$

then

$$\frac{\partial \mathscr{L}}{\partial \bar{p}(y)}=\frac{\partial W}{\partial \bar{p}(y)}+\ln \bar{p}(y)+1-\ln p(y)-\mu. \tag{3.17}$$

From $R=K-W$, we obtain

$$0=\frac{\partial K}{\partial \bar{p}(y)}+\frac{\partial K}{\partial \lambda}\frac{\partial \lambda}{\partial \bar{p}(y)}-\frac{\partial J}{\partial \bar{p}(y)}-\frac{\partial J}{\partial \lambda}\frac{\partial \lambda}{\partial \bar{p}(y)} \tag{3.18}$$

and

$$\frac{\partial W}{\partial \bar{p}(y)}=\frac{\partial J}{\partial \bar{p}(y)}+\frac{\partial J}{\partial \lambda}\frac{\partial \lambda}{\partial \bar{p}(y)}. \tag{3.19}$$

Using (3.17), we obtain

$$\frac{\partial \lambda}{\partial p(y)}=\frac{\dfrac{\partial J}{\partial \bar{p}(y)}-\dfrac{\partial K}{\partial \bar{p}(y)}}{\dfrac{\partial K}{\partial \lambda}-\dfrac{\partial J}{\partial \lambda}}. \tag{3.20}$$

Substituting (3.20) into (3.19), we obtain

$$\frac{\partial W}{\partial \bar{p}(y)}=\frac{\dfrac{\partial K}{\partial \lambda}\dfrac{\partial J}{\partial \bar{p}(y)}-\dfrac{\partial K}{\partial \bar{p}(y)}\dfrac{\partial J}{\partial \lambda}}{\dfrac{\partial K}{\partial \lambda}-\dfrac{\partial J}{\partial \lambda}}. \tag{3.21}$$

From $J=\lambda K-F$ $\left(\text{and } \dfrac{\partial F}{\partial \lambda}=K\right)$,

$$\frac{\partial J}{\partial \lambda}=\lambda\frac{\partial K}{\partial \lambda}, \tag{3.22}$$

therefore

$$\frac{\partial W}{\partial \bar{p}(y)}=\frac{1}{1-\lambda}\left\{\frac{\partial J}{\partial \bar{p}(y)}-\lambda\frac{\partial K}{\partial \bar{p}(y)}\right\}. \tag{3.23}$$

From (2.18)

$$\frac{\partial J}{\partial \bar{p}(y)}=\lambda\frac{\partial K}{\partial \bar{p}(y)}-\frac{\partial F}{\partial \bar{p}(y)}, \tag{3.24}$$

and

$$\frac{\partial W}{\partial \bar{p}(y)}=-\frac{1}{1-\lambda}\frac{\partial F}{\partial \bar{p}(y)}=-\frac{1}{1-\lambda}\ln \sum_{x\in\mathscr{X}} p(x|y)^{1-\lambda}. \tag{3.25}$$

Then

$$\frac{\partial \mathscr{L}}{\partial \bar{p}(y)}=-\frac{1}{1-\lambda}\ln \sum_{x\in\mathscr{X}} p(x|y)^{1-\lambda}+\ln\frac{\bar{p}(y)}{p(y)}+1-\mu=0 \tag{3.26}$$

therefore

$$\bar{p}(y)=e^{\mu-1}p(y)\{\sum_{x\in\mathscr{X}} p(x|y)^{1-\lambda}\}^{\frac{1}{1-\lambda}}. \tag{3.27}$$

Using $\sum_{y\in\mathscr{Y}}\bar{p}(y)=1$, we obtain

$$\bar{p}(y)=\frac{p(y)\{\sum_{x\in\mathscr{X}} p(x|y)^{1-\lambda}\}^{\frac{1}{1-\lambda}}}{\sum_{y'\in\mathscr{Y}} p(y')\{\sum_{x'\in\mathscr{X}} p(x'|y')^{1-\lambda}\}^{\frac{1}{1-\lambda}}}, \tag{3.28}$$

and

$$K=\sum_{y\in\mathscr{Y}}\sum_{x\in\mathscr{X}}\bar{p}(y)\frac{p(x|y)^{1-\lambda}h(x|y)}{\sum_{x'\in\mathscr{X}} p(x'|y)^{1-\lambda}}, \tag{3.29}$$

$$W=\lambda K-F, \tag{3.30}$$

$$F=\sum_{y\in\mathscr{Y}}\bar{p}(y)\ln\sum_{x\in\mathscr{X}}p(x|y)^{1-\lambda}, \tag{3.31}$$

$$R=K-W, \tag{3.32}$$

$$S=W+\sum_{y\in\mathscr{Y}}\bar{p}(y)\ln\frac{\bar{p}(y)}{p(y)}, \tag{3.33}$$

thus we obtain a parametric representation of $S(h(x|y),\ R)$.

From (3.28), (3.33), it can be obtained that

$$\frac{\partial S}{\partial\lambda}=\frac{\lambda}{1-\lambda}\frac{\partial R}{\partial\lambda}, \tag{3.34}$$

therefore

$$\frac{\partial S}{\partial R}=\frac{\lambda}{1-\lambda}. \tag{3.35}$$

Moreover

$$\frac{\partial^2 S}{\partial R^2}=\frac{1}{(1-\lambda^2)}\left(\frac{\partial R}{\partial\lambda}\right)^{-1}. \tag{3.36}$$

From direct computation we know that

$$\frac{\partial R}{\partial\lambda}\geqslant 0, \tag{3.37}$$

therefore

$$\frac{\partial^2 S}{\partial R^2}>0. \tag{3.38}$$

Hence, we obtain the following

Lemma 9. *$S(h(x|y),\ R)$ has the following properties*

(1) *Its domain of definition is*

$$0\leqslant R\leqslant\max_{y\in\mathscr{Y}}\ln\|\{x\colon p(x|y)\neq 0\}\|; \tag{3.39}$$

if $R>\max_{y\in\mathscr{Y}}\ln\|\{x\colon p(x|y)\neq 0\}\|$, let $S=\infty$.

(2) (3.28)—(3.33) *are the parametric representation of the function* $S(h(x|y),\ R)$.

(3) $S(h(x|y),\ R)\geqslant 0$, *the equality stands if and only if* $R=H(X|Y)$.

(4) $\frac{\partial S}{\partial R}=\frac{\lambda}{1-\lambda},\ \frac{\partial^2 S}{\partial R^2}>0$.

§ 4. The Upper Bound of the Probability of Error

In this section, we estimate the probability of error in Slepian-Wolf theorem. First, we assume that

$$(R_x,\ R_y)\in\mathscr{R}, \tag{4.1}$$

$$R_x\leqslant\ln\|\mathscr{X}_0\|,\quad R_y\leqslant\ln\|\mathscr{Y}_0\| \tag{4.2}$$

($\mathscr{X}_0$, $\mathscr{Y}_0$ see Lemma 5) and construct a random code as follows.

Encoding. We encode any $x^n\in\mathscr{X}^n$ as one of the elements of the set $\{1, 2, \cdots, e^{nR_x}\}$ independently and with probability e^{-nR_x}. Similarly, we encode any $y^n\in\mathscr{Y}^n$.

Moreover, let the encoding of $y^n \in \mathcal{Y}^n$ be independent from the encoding of $x^n \in \mathcal{X}^n$ (without loss of generality, we assume that e^{nR_x}, e^{nR_y} are integers). Thus, all of the possible encodings are equiprobabilistic.

Decoding. For fixed encodings, if the received signals are (i, j), it can be deduced that the messages being sent are in the set $f_x^{-1}(i) \otimes f_y^{-1}(j)$. If there is a unique pair (x^n, y^n) in the set $f_x^{-1}(i) \otimes f_y^{-1}(j)$ satisfying that for each $(\bar{x}^n, \bar{y}^n) \in f_x^{-1}(i) \otimes f_y^{-1}(j)$

$$p^{(n)}(x^n, y^n) \geqslant p^{(n)}(\bar{x}^n, \bar{y}^n), \tag{4.3}$$

we decode (i, j) as (x^n, y^n), that is, $g(i, y) = (x^n, y^n)$; in other cases that there are at least two pairs in $f_x^{-1}(i) \otimes f_y^{-1}(j)$ satisfying (4.3), we decode (i, j) as one of the pairs arbitrarily.

The average probability of error of the random code. First, we compute the conditional probability of error under the condition that (x_0^n, y_0^n) is the message emitted from the sources. Let

$$P_e^{(n)}(x_0^n, y_0^n) = p^{(n)}\{(\hat{x}^n, \hat{y}^n) \neq \{(x_0^n, y_0^n) \mid (x_0^n, y_0^n) \text{ is the message emitted from the source}\}. \tag{4.4}$$

There are three possible cases for making mistakes, that is, if the received signals are (i, j),

(1) there is such a $y^n \in f_y^{-1}(j)$, $y^n \neq y_0^n$ that

$$p^{(n)}(x_0^n, y^n) \geqslant p^{(n)}(x_0^n, y_0^n), \tag{4.5}$$

(2) there is such an $x^n \in f_x^{-1}(i)$, $x^n \neq x_0^n$ that

$$p^{(n)}(x^n, y_0^n) \geqslant p^{(n)}(x_0^n, y_0^n), \tag{4.6}$$

(3) there is such a pair $(x^n, y^n) \in f_x^{-1}(i) \otimes f_y^{-1}(j)$, $x^n \neq x_0^n$, $y^n \neq y_0^n$ that

$$p^{(n)}(x^n, y^n) \geqslant p^{(n)}(x_0^n, y_0^n). \tag{4.7}$$

If we use $\bar{P}_t^{(n)}(x_0^n, y_0^n)$ to denote the conditional probability of error in the t th case, it is easy to know that

$$\bar{P}_e^{(n)}(x_0^n, y_0^n) = \sum_{t=1}^{3} \bar{P}_t^{(n)}(x_0^n, y_0^n). \tag{4.8}$$

It is obvious that

$$\begin{aligned} \bar{P}_1^{(n)}(x_0^n, y_0^n) &\leqslant e^{-nR_y} \| \{y^n: p^{(n)}(x_0^n, y^n) \geqslant p^{(n)}(x_0^n, y_0^n)\} \| \\ &= e^{-nR_y} \| \{y^n: h(y^n \mid x_0^n) \leqslant h(y_0^n \mid x_0^n)\} \|, \end{aligned} \tag{4.9}$$

$$\bar{P}_2^{(n)}(x_0^n, y_0^n) \leqslant e^{-nR_x} \| \{y^n: h(x^n \mid y_0^n) \leqslant h(x_0^n \mid y_0^n)\} \|, \tag{4.10}$$

$$\bar{P}_3^{(n)}(x_0^n, y_0^n) \leqslant e^{-n(R_x+R_y)} \| \{(x^n, y^n): h(x^n, y^n) \leqslant h(x_0^n, y_0^n)\} \|. \tag{4.11}$$

Let $\bar{P}_e^{(n)}$ be the average probability of error of random code, then

$$\begin{aligned} \bar{P}_e^{(n)} &= \sum_{x_0^n \in \mathcal{X}^n} \sum_{y_0^n \in \mathcal{Y}^n} p^{(n)}(x_0^n, y_0^n) \bar{P}_e^{(n)}(x_0^n, y_0^n) \\ &= \sum_{t=1}^{3} \{ \sum_{x_0^n \in \mathcal{X}^n} \sum_{y_0^n \in \mathcal{Y}^n} p^{(n)}(x_0^n, y_0^n) \bar{P}_t^{(n)}(x_0^n, y_0^n) \}. \end{aligned} \tag{4.12}$$

Let

$$\bar{P}_t^{(n)} = \sum_{x_0^n \in \mathscr{X}^n} \sum_{y_0^n \in \mathscr{Y}^n} p^{(n)}(x_0^n, y_0^n) \bar{P}_t^{(n)}(x_0^n, y_0^n), \tag{4.13}$$

then

$$\bar{P}_e^{(n)} = \sum_{t=1}^{3} \bar{P}_t^{(n)}. \tag{4.14}$$

Now, we estimate $\bar{P}_t^{(n)}$ $(t=1, 2, 3)$ respectively.

$$\bar{P}_3^{(n)} \leqslant \sum_{x_0^n \in \mathscr{X}^n} \sum_{y_0^n \in \mathscr{Y}^n} p^{(n)}(x_0^n, y_0^n) \min\{1, e^{-n(R_x+R_y)} \|\{(x^n, y^n): h(x^n, y^n) \leqslant h(x_0^n, y_0^n)\}\|\}$$

$$\leqslant \exp\{-nW^+(h(x, y), R_x+R_y)+o(n)\}$$

$$+ \sum_{h(x_0^n, y_0^n) < nK} p^{(n)}(x_0^n, y_0^n) e^{-n(R_x+R_y)} \|\{(x^n, y^n): h(x^n, y^n) \leqslant h(x_0^n, y_0^n)\}\|, \tag{4.15}$$

where $K = K(h(x, y), R_x+R_y)$. Now, we estimate the second term in (4.15).

$$\sum_{h(x_0^n, y_0^n) < nK} p^{(n)}(x_0^n, y_0^n) e^{-n(R_x+R_y)} \|\{(x^n, y^n): h(x^n, y^n) \leqslant h(x_0^n, y_0^n)\}\|$$

$$= \sum_{i=0}^{n-1} e^{-n(R_x+R_y)} \sum_{nK_i < h(x_0^n, y_0^n) < nK_{i+1}} p^{(n)}(x_0^n, y_0^n) \|\{(x^n, y^n): h(x^n, y^n) \leqslant h(x_0^n, y_0^n)\}\|$$

$$\leqslant \sum_{i=0}^{n-1} e^{-n(R_x+R_y)} p^{(n)}\{(x_0^n, y_0^n): nK_i \leqslant h(x_0^n, y_0^n)$$

$$\leqslant nK_{i+1}\} \|\{(x^n, y^n): h(x^n, y^n) \leqslant h(x_0^n, y_0^n)\}\|$$

$$\leqslant \sum_{i=0}^{n-1} \exp\left\{-n\left[R_x+R_y-\frac{i}{n}(R_x+R_y)+W^+\left(h(x, y), \frac{i}{n}(R_x+R_y)\right)\right]+o(n)\right\}$$

$$\leqslant \sum_{i=0}^{n-1} \exp\{-n[R_x+R_y-\sup_{R \leqslant R_x+R_y}(R-W^+(h(x, y), R))]+o(n)\}, \tag{4.16}$$

where $K_i = K\left(h(x, y), \frac{i}{n}(R_x+R_y)\right)$. There are two different cases, which we are to discuss.

Noticing that $\frac{\partial W}{\partial R} = \frac{\lambda}{1-\lambda}$, $\frac{\partial^2 W}{\partial R^2} > 0$ (which can be proved in the same way as Lemma 9), we can prove that

(1) if $\left.\frac{\partial W}{\partial R}\right|_{R=R_x+R_y} > 1$, that is, if $\lambda > \frac{1}{2}$, the supremum in (4.16) is achieved at R^* which satisfies that $\left.\frac{\partial W}{\partial R}\right|_{R=R^*} = 1$, that is, $\lambda = \frac{1}{2}$; denote this R^* by R_3, then

$$\bar{P}_3^{(n)} \leqslant \exp\{-n[R_x+R_y-R_3+W^+(h(x, y), R_3)]+o(n); \tag{4.17}$$

(2) if $\left.\frac{\partial W}{\partial R}\right|_{R=R_x+R_y} \leqslant 1$, that is, if $\lambda \leqslant \frac{1}{2}$, the supremum in (4.16) is achieved at $R = R_x+R_y$, then

$$\bar{P}_3^{(n)} \leqslant \exp\{-nW^+(h(x, y), R_x+R_y)+o(n)\}. \tag{4.18}$$

Let

$$\widetilde{W}^+(h(x, y), R) = W^+(h(x, y), R), \quad R \leqslant R_3;$$

$$\widetilde{W}^+(h(x, y), R) = R-R_3+W^+(h(x, y), R_3), \quad R > R_3. \tag{4.19}$$

Summing up the above results, we obtain

$$\bar{P}_3^{(n)} \leqslant \exp\{-n\widetilde{W}^+(h(x, y), R_x+R_y)+o(n)\}. \tag{4.20}$$

Now, we estimate $\bar{P}_1^{(n)}$

$$\begin{aligned}\bar{P}_1^{(n)} &\leqslant \sum_{x_0^n\in\mathscr{X}^n}\sum_{y_0^n\in\mathscr{Y}^n} p^{(n)}(x_0^n, y_0^n)\min\{1, e^{-nR_y}\|\{y^n: h(y^n|x_0^n)\leqslant h(y_0^n|x_0^n)\}\|\}\\ &\leqslant \sum_{x_0^n\in\mathscr{X}^n} p^{(n)}(x_0^n)\exp\{-nW^+(h(y|x), R_y|\bar{p})+o(n)\}\\ &+\sum_{x_0^n\in\mathscr{X}^n}\sum_{h(y_0^n|x_0^n)\leqslant nK(x_0^n)} p^{(n)}(x_0^n, y_0^n)e^{-nR_y}\|\{y^n: h(y^n|x_0^n)\leqslant h(y_0^n|x_0^n)\}\|,\end{aligned}\tag{4.21}$$

where $K(x_0^n)=K(h(y|x), R_y|\bar{p})$, $\bar{p}$ is the frequency of x_0^n. From (3.4), (3.5), (3.8), we know that the first term of the right side of (4.21)

$$\begin{aligned}&\sum_{x_0^n\in\mathscr{X}^n} p^{(n)}(x_0^n)\exp\{-nW^+(h(y|x), R_y|\bar{p})+o(n)\}\\ &\leqslant\exp\{-nS^+(h(y|x), R_y)+o(n)\}.\end{aligned}\tag{4.22}$$

The estimation of the second term is as follows

$$\begin{aligned}&\sum_{x_0^n\in\mathscr{X}^n}\sum_{h(y_0^n|x_0^n)\leqslant nK(x_0^n)} p^{(n)}(x_0^n, y_0^n)e^{-nR_y}\|\{y^n: h(y^n|x_0^n)\leqslant h(y_0^n|x_0^n)\}\|\\ &\leqslant\sum_{x_0^n\in\mathscr{X}^n} p^{(n)}(x_0^n)e^{-nR_y}\sum_{i=0}^{n-1}\exp\left\{iR_y-nW^+\left(h(y|x), \frac{iR_y}{n}\Big|\bar{p}\right)+o(n)\right\}\\ &=\sum_{i=0}^{n-1}e^{-nR_y}\sum_{x_0^n\in\mathscr{X}^n} p^{(n)}(x_0^n)\exp\left\{iR_y-nW^+\left(h(y|x), \frac{iR_y}{n}\Big|\bar{p}\right)+o(n)\right\}\\ &=\sum_{i=0}^{n-1}e^{-nR_y}\exp\left\{iR_y-nS^+\left(h(y|x), \frac{iR_y}{n}\right)+o(n)\right\}\\ &=ne^{-nR_y}\exp\{n\sup_{R<R_y}[R-S^+(h(y|x), R)]+o(n)\}.\end{aligned}\tag{4.23}$$

The supremum in (4.23) can be discussed in two different cases.

(1) if $\left.\frac{\partial S^+}{\partial R}\right|_{R=R_y}\leqslant 1$, then the supremum is achieved at $R=R_y$, and

$$\bar{P}_1^{(n)}\leqslant\exp\{-nS^+(h(y|x), R_y)+o(n)\}.\tag{4.24}$$

(2) if $\left.\frac{\partial S^+}{\partial R}\right|_{R=R_y}>1$, then the supremum is achieved at R^* (which is denoted by R_1) satisfying $\frac{\partial S^+}{\partial R}=1$, that is, $\lambda=\frac{1}{2}$, and (4.23) becomes (notice (4.21), (4.22))

$$\bar{P}_1^{(n)}\leqslant\exp\{-n[R_y-R_1+S^+(h(y|x), R_1)]+o(n)\}.\tag{4.25}$$

Let

$$\begin{aligned}&\tilde{S}^+(h(y|x), R)=S^+(h(y|x), R), \text{ if } R\leqslant R_1;\\ &\tilde{S}^+(h(y|x), R)=R-R_1+S^+(h(y|x), R_1), \text{ if } R>R_1.\end{aligned}\tag{4.26}$$

Then

$$\bar{P}_1^{(n)}\leqslant\exp\{-n\tilde{S}^+(h(y|x), R_y)+o(n)\}.\tag{4.27}$$

Discussing $\bar{P}_2^{(n)}$ in the same way, we can define function $\tilde{S}^+(h(x|y), R)$, and obtain

$$\bar{P}_2^{(n)}\leqslant\exp\{-n\tilde{S}^+(h(x|y), R_x)+o(n)\}.\tag{4.28}$$

Thus we obtain the following

Theorem 3. *For the Slepian-Wolf coding problem, it is true that*

(1) *if* $R_x\leqslant\ln\|\mathscr{X}_0\|$, $R_y\leqslant\ln\|\mathscr{Y}_0\|$, then

$$C(R_x, R_y)\geqslant\min\{\tilde{W}^+(h(x, y), R_x+R_y), \tilde{S}^+(h(y|x), R_y), \tilde{S}^+(h(x|y)R_x\},\tag{4.29}$$

(2) *if* $R_x>\ln\|\mathscr{X}_0\|$, $R_y>\ln\|\mathscr{Y}_0\|$, *then*

$$C(R_x, R_y)=\infty,\tag{4.30}$$

(3) *if* $R_x > \ln\|\mathscr{X}_0\|$, $R_y \leqslant \ln\|\mathscr{Y}_0\|$, *then*

$$C(R_x, R_y) \geqslant \tilde{S}^+(h(y|x), R_y), \tag{4.31}$$

(4) *if* $R_x \leqslant \ln\|\mathscr{X}_0\|$, $R_y > \ln\|\mathscr{Y}_0\|$, *then*

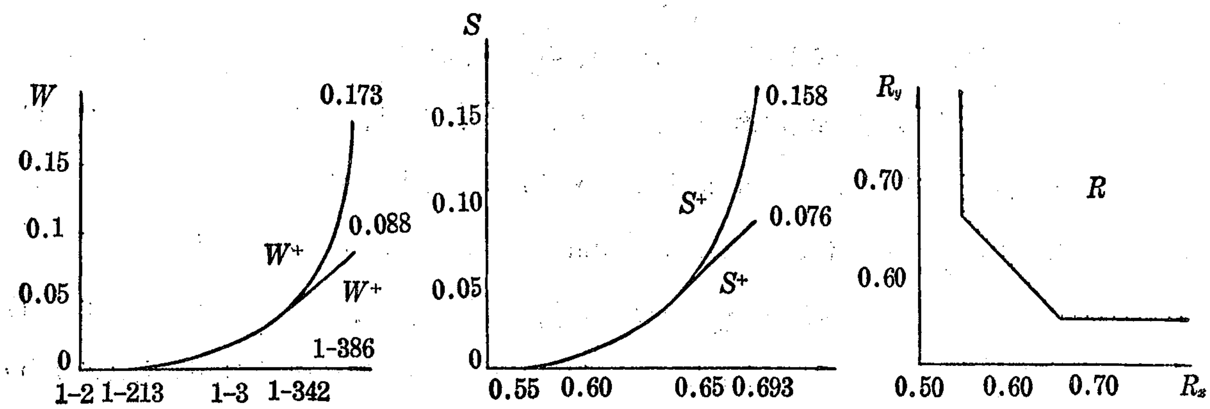
$$C(R_x, R_y) \geqslant \tilde{S}^+(h(x|y), R_y). \tag{4.32}$$

Proof We have to prove only (3). In this case, all of the $x^n \in \mathscr{X}^n$ can be transpoted without error, so we have to construct the random code at only the encoder E_y. Then, following the example of the proof of (1), we can prove (3) step by step.

Example Let $\mathscr{X} = \mathscr{Y} = \{0, 1\}$, $p(x, y)$ be given by the following matrix

$$\begin{pmatrix} \frac{1}{8}, & \frac{1}{2} \\ \frac{1}{4}, & \frac{1}{8} \end{pmatrix}.$$

Then $H(X, Y) = 1.2130$, $H(X|Y) = H(Y|X) = 0.5514$, and the curves of W^+, $\widetilde{W}^+$, S^+, $\tilde{S}^+$ and region $\mathscr{R}$ are shown in the following figures.



References

[1] Slepian, D. and Wolf, J. K., Noiseless Coding of Correlated Information Sources, *IEEEE Trans Information Theory*, IT-**19** (1973), 477—480.

[2] Ahlswede, R. F. and Körner, J., Source Coding with Side Information and a Converse for Degraded Broadcast Channels, *IEEE Trans Information Theory*, IT-**21** (1975), 629—637.

[3] Флейшман. Б. С., Конструктивные методы оптимального кодирования для каналов с шумами, Издательство Академии Наук, СССР. москва, (1963).