THE TRANSACTIONS OF THE IECE OF JAPAN, VOL. E 63, NO. 10 OCTOBER 1980

Source Coding Theory for Multiterminal Communication Systems with a Remote Source

Hirosuke YAMAMOTO* *and* Kohji ITOH**, *Regular Members*

UDC 519.728.2

SUMMARY The source coding problems are studied on the Slepian-Wolf-type system with a remote source (Fig. 1) and the Wyner-Ziv-type system with a remote source (Fig. 4). For the former, inner and outer bounds are obtained on the admissible rate region to attain a prescribed distortion tolerance. For the latter, the rate-distortion function is derived. As examples, a Gaussian remote source and a binary remote source are analyzed.

1. Introduction

In most cases the source coding problem assumes that a source output can be directly encoded for transmission over a channel. In some practical situations, however, uncoded noisy transmission may intervene between a source and an encoder. For example, in telephone networks, any sophisticated encoding and decoding equipments could not be located at terminals while central offices could be equipped with encoders and decoders of considerable complexity. Also, there are some cases where information to be transmitted is measured data corrupted by measurement errors or a source output is quantized for digital processing. The source whose output may be distorted prior to encoding is called remote source[(1)]. The source coding problem relating to the communication system comprising a single encoder connected to a remote source, a single channel and a single decoder has been studied in the literatures[(1)-(3)]. On the other hand, a practical communication system is often multiterminal. So the source coding problems for certain multiterminal communication systems also have been studied by various authors[(4)-(7)].

In this paper we consider source coding problems for two types of multiterminal communication systems with a remote source. Section 2 deals with the system in which a single remote source is connected to two separate encoders, their outputs being supplied to a decoder via individual channels (Fig. 1). The system resembles the one due to Slepian and Wolf[(5)] differing in that the decoder estimates the output of the original source and not the inputs to the encoders. Inner and outer bounds are obtained to the region of admissible coding rates to attain a prescribed distortion tolerance in this system. The case of Gaussian remote source is analyzed as an example. In Section 3, the rate-distortion function is obtained for the Wyner-Ziv[(8),(9)]-type system consisting of one encoder connected to a remote source and one decoder with side information. This is a special case of the Slepian-Wolf-type system in Section 2. As examples, the rate-distortion functions for a binary and a Gaussian remote sources are determined.

2. Slepian-Wolf-Type System with a Remote Source

Let us consider first the system depicted in Fig. 1, where $\{X_k\}_{k=1}^{\infty}$ and $\{Y_{1k}, Y_{2k}\}_{k=1}^{\infty}$ are sequences of random variables representing the source output and the noisy channel outputs, respectively. Let the source and the noisy channel be memoryless: that is, let (X_k, Y_{1k}, Y_{2k}), $k=1, 2, \cdots$ be generated by repeated independent drawings of a triplet of random variables (X, Y_1, Y_2) which take values in $\mathcal{X}, \mathcal{Y}_1$ and $\mathcal{Y}_2$, respectively. Encoder i, $i=1,2$ can receive only the sequence $\{Y_{ik}\}$ and encodes it at respective rate, R_i bits per input symbol, for transmission to the common decoder, which in its turn emits a reproduction sequence $\{\hat{X}_k\}$ corresponding to the source output sequence $\{X_k\}$. $\hat{X}_k$ takes values in $\hat{\mathcal{X}}$. $\mathcal{X}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\mathcal{X}$ are either discrete sets, the reals or arbitrary measurable spaces.

Encoding and decoding are done in blocks of length n, and the fidelity criterion is given by

$$E\frac{1}{n}\sum_{k=1}^{n} D(X_k, \hat{X}_k) \tag{1}$$

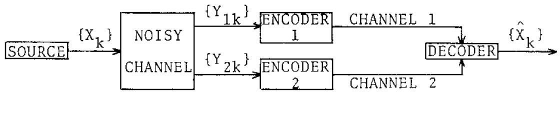


Fig. 1 Slepian-Wolf-type system with a remote source.

Manuscript received March 31, 1980.

*The author is now with the Department of Electronic Engineering, Tokushima University, Tokushima, 770 Japan.

**The author is with the Institute of Space and Aeronautical Science, The University of Tokyo, Tokyo, 153 Japan.

55-60 [P-36]

where $D:\mathcal{X}\times\hat{\mathcal{X}}\to[0,\infty)$ is a given distortion function and E denotes expectation. A code (n, M_1, M_2, Δ) is defined by three mappings F_{E1}, F_{E2} and F_D, which correspond to encoder 1, 2 and the decoder, respectively;

$$\begin{cases} F_{E1}: \mathcal{Y}_1^n \to I_{M_1} & \text{(2a)} \\ F_{E2}: \mathcal{Y}_2^n \to I_{M_2} & \text{(2b)} \\ F_D: I_{M_1}\times I_{M_2}\to\hat{\mathcal{X}}^n & \text{(2c)} \end{cases}$$

where $I_{M_i}\triangleq\{0, 1, \cdots, M_i-1\}$. The parameter Δ is defined by

$$\Delta \triangleq E\frac{1}{n}\sum_{k=1}^{n} D(X_k, \hat{X}_k) \tag{3}$$

where $\hat{\boldsymbol{X}}^n=F_D(F_{E1}(\boldsymbol{Y}_1^n), F_{E2}(\boldsymbol{Y}_2^n))$, bold face letters representing vectors with n components.

A rate-pairs (R_1, R_2) is said to be d-admissible if, for any given $\varepsilon>0$ and n sufficiently large, there exists a code (n, M_1, M_2, Δ) such that

$$M_i \leq e^{n(R_i+\varepsilon)},\ i=1, 2 \tag{4}$$

and

$$\Delta \leq d+\varepsilon\ . \tag{5}$$

We define $\mathcal{R}^*(d)$ as the set of d-admissible rate-paris.

Our main problem is to determine the region $\mathcal{R}^*(d)$. However, this problem is very difficult to solve exactly and bounds on $\mathcal{R}^*(d)$ are derived in the following.

Let $\hat{Y}_1$ and $\hat{Y}_2$ be random variables distributed jointly with (X, Y_1, Y_2) and take values in certain sets $\hat{\mathcal{Y}}_1$ and $\hat{\mathcal{Y}}_2$, respectively. Define subset $\mathcal{R}_{\hat{Y}_1\hat{Y}_2}$ in two dimensional Euclidean space by

$$\begin{aligned} \mathcal{R}_{\hat{Y}_1\hat{Y}_2} \triangleq \{(R_1, R_2):\ & R_1 \geq I(Y_1Y_2\,;\hat{Y}_1|\hat{Y}_2), \\ & R_2 \geq I(Y_1Y_2\,;\hat{Y}_2|\hat{Y}_1), \\ & R_1+R_2 \geq I(Y_1Y_2\,;\hat{Y}_1\hat{Y}_2)\}. \end{aligned} \tag{6}$$

Let $\mathcal{P}^i(d)$ be the set of pairs of random variables $(\hat{Y}_1\hat{Y}_2)$ that satisfy properties (i) and (iii) below, and let $\mathcal{P}^o(d)$ be the set of those which satisfy properties (ii) and (iii).

(i) $I(\hat{Y}_i\,;XY_j\hat{Y}_j|Y_i)=0,\ (i,j)=(1,2),(2,1);$ (7)
that is, $\hat{Y}_1-Y_1-Y_2-\hat{Y}_2$ and $X-(Y_1Y_2)-(\hat{Y}_1\hat{Y}_2)$ form Markov chains in that order.

(ii) $I(\hat{Y}_i\,;XY_j|Y_i)=0,\ (i,j)=(1,2),(2,1),$ (8)
which means $\hat{Y}_1-Y_1-(Y_2X)$ and $\hat{Y}_2-Y_2-(Y_1X)$ form Markov chains in that order.

(iii) There exists a function $f:\hat{\mathcal{Y}}_1\times\hat{\mathcal{Y}}_2\to\hat{\mathcal{X}}$ such that

$$E[D(X, f(\hat{Y}_1, \hat{Y}_2))] \leq d. \tag{9}$$

Now define two regions, $\mathcal{R}^i(d)$ and $\mathcal{R}^o(d)$, as follows.

$$\mathcal{R}^o(d) \triangleq \left[\bigcup_{\hat{Y}_1\hat{Y}_2\in\mathcal{P}^o(d)} \mathcal{R}_{\hat{Y}_1\hat{Y}_2}\right]^c \tag{10}$$

$$\mathcal{R}^i(d) \triangleq \left[\bigcup_{\hat{Y}_1\hat{Y}_2\in\mathcal{P}^i(d)} \mathcal{R}_{\hat{Y}_1\hat{Y}_2}\right]^c \tag{11}$$

where $[\quad]^c$ denotes set closure. Then the following theorem holds.

[Theorem 1]

Assume that a source satisfies the following conditions†

(1) If $\mathcal{X}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$ and $\hat{\mathcal{X}}$ are finite discrete sets, it holds that for $\forall x\in\mathcal{X}$ and $\forall\hat{x}\in\hat{\mathcal{X}}$

$$D(x, \hat{x})<\infty \tag{12}$$

(2) Otherwise††, it holds that for all $\hat{x}\in\hat{\mathcal{X}}$

$$E[D(X, \hat{x})]<\infty. \tag{13}$$

Furthermore, for the random variable $\hat{X}$ satisfying $ED(X,\hat{X})<\infty$ and for arbitrary $\varepsilon>0$, there exists a finite subsets $\{\hat{x}_j\}_{j=1}^J\subseteq\hat{\mathcal{X}}$ and a quantization mapping $f_Q:\hat{\mathcal{X}}\to\{\hat{x}_j\}_{j=1}^J$ such that

$$E[D(X, f_Q(\hat{X}))] \leq (1+\varepsilon)E[D(X,\hat{X})]\,. \tag{14}$$

Then,

$$\mathcal{R}^i(d)\subseteq\mathcal{R}^*(d)\subseteq\mathcal{R}^o(d). \tag{15}$$

Theorem 1 can be proved in the similar way as Housewright[(6),(10)] or Tung[(7),(11)] did. However, we omit it because it is too lengthy. (The proof is given in Ref. (12).)

As an example, let us consider a Gaussian remote source shown in Fig. 2, where X, W_1, W_2 are independent Gaussian random variables, and let the distortion measure be $D(x,\hat{x})=(x-\hat{x})^2$. Even under these specifications, it has been impossible to delimit the set $\mathcal{P}^i(d)$. Accordingly, we make assumption here that $(Y_1, Y_2, \hat{Y}_1, \hat{Y}_2)$ are jointly Gaussian. $\mathcal{R}^i(d)$ obtained by this contraction of $\mathcal{P}^i(d)$ is an innerbound anyhow, though it may be slightly narrower††† than the exact one. Owing to the Gaussian assumption and the condition (i),

$$\hat{Y}_i=Y_i+V_i,\ i=1, 2 \tag{16}$$

where V_i is the Gaussian random variable $N(0, \sigma_{V_i}^2)$ independent with X, W_1 and W_2, and let the function

† It is quite natural to assume that $D(x,\hat{x})\geq 0$ and for each $x\in\mathcal{X}$ there exists at least one $\hat{x}$ such that $D(x,\hat{x})=0$.

†† As shown in Ref. (9), if $\mathcal{X}$ and $\hat{\mathcal{X}}$ are the reals, $D(x,\hat{x})=|x-\hat{x}|^r$, $r>0$ and $E|X|^r<\infty$, then condition (2) is satisfied.

††† There is an indication in Ref. (13) that the jointly Gaussian assumption is insufficient.

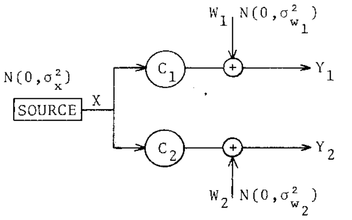


Fig. 2 Gaussian remote source.

f be the minimum mean square estimator of X given $\hat{Y}_1$ and $\hat{Y}_2$.

It is well-known that the differential entropy of a jointly Gaussian L-vector $\boldsymbol{Z}^L$ can be expressed in terms of the determinant of the covariance matrix of $\boldsymbol{Z}^L$, $|\Phi_{\boldsymbol{Z}^L}|$.

$$H_d(\boldsymbol{Z}^L)=\frac{L}{2}\log(2\pi e|\Phi_{\boldsymbol{Z}^L}|) \quad (17)$$

By using this relation, we can show that the region $\mathcal{R}^i(d)$ defined by Eq. (11) is represented as follows.

$$\mathcal{R}^i(d_n)=\bigcup_{\substack{b_1,b_2\\ \text{satisfy (22)}}}\{(R_1,R_2):R_1\geq r_1,\ R_2\geq r_2,\ R_1+R_2\geq r\} \quad (18)$$

$$r_1=\frac{1}{2}\log\frac{k}{b_1^2(1+a_2^2+b_2^2)} \quad (19)$$

$$r_2=\frac{1}{2}\log\frac{k}{b_2^2(1+a_1^2+b_1^2)} \quad (20)$$

$$r=\frac{1}{2}\log\frac{k}{b_1^2\,b_2^2} \quad (21)$$

$$(a_1^2+b_1^2)+(a_2^2+b_1^2)+(a_1^2+b_1^2)(a_2^2+b_2^2)=\frac{(a_1^2+b_1^2)(a_2^2+b_2^2)}{d_n}\triangleq k \quad (22)$$

where b_i, $i=1,2$ are parameters and the following normalizations

$$d_n\triangleq\frac{d}{\sigma_X^2},\qquad a_i^2=\frac{\sigma_{W_i}^2}{c_i^2\,\sigma_X^2},\qquad i=1,2 \quad (23)$$

are performed.

Figure 3 illustrates $\mathcal{R}^i(d_n)$. For the purpose of comparison, also drawn in the figure by dotted lines are the regions for the case where the two encoders can receive both Y_1 and Y_2 , and the broken lines indicate for the case where the independent encoders make time-shared use of their respective input. The former are outerbounds on $\mathcal{R}^*(d_n)$'s and the latter are innerbounds. The boundaries of $\mathcal{R}^i(d_n)$'s obtained are seen to lie between the boundaries of the corresponding bounds.

3. Wyner-Ziv-Type System with a Remote Source

If the decoder in Fig. 1 can receive the sequence $\{Y_1\}$ directly or equivalently, in case of finite $\mathcal{Y}_1$, $R_1\geq H(Y_1)$, the system reduces to the Wyner-Ziv-type system with a remote source, which is illustrated in Fig. 4.

Let us denote the rate-distortion function for this system by $R_{wz}^{r*}(d)$, and define $R_{wz}^{r}(d)$ as follows.

$$R_{wz}^{r}(d)\triangleq\inf_{\hat{Y}_2\in\mathcal{P}_{wz}^r(d)} I(Y_2;\hat{Y}_2|Y_1) \quad (24)$$

where $\mathcal{P}_{wz}^r(d)^\dagger$ is the set of the random variables $\hat{Y}_2$ jointly distributed with (X,Y_1,Y_2) satisfying the conditions,

(iv) $I(\hat{Y}_2;XY_1|Y_2)=0$ (25)

(v) There exists a function $f_{wz}^r:\ \hat{\mathcal{Y}}_1\times\hat{\mathcal{Y}}_2\to\hat{\mathcal{X}}$ such that

$$E[D(X,f_{wz}^r(Y_1,\hat{Y}_2))]\leq d. \quad (26)$$

Then, the following theorem is obtained from Theorem 1 by noting that $R_{wz}^{r*}(d)=\inf_{(R_1,R_2)\in\mathcal{R}^*(d)|_{\hat{Y}_1=Y_1}} R_2$ and $\mathcal{R}^i(d)=\mathcal{R}^*(d)=\mathcal{R}^o(d)$ if $\hat{Y}_1=Y_1$.

[Theorem 2]

If a source satisfies conditions (1) and (2) in section 2, then

$$R_{wz}^{r*}(d)=R_{wz}^{r}(d) \quad (27)$$

As the first example, consider again the Gaussian remote source, Fig. 2. By letting $b_1^2\to 0$ in Eqs. (20) and (21), we obtain

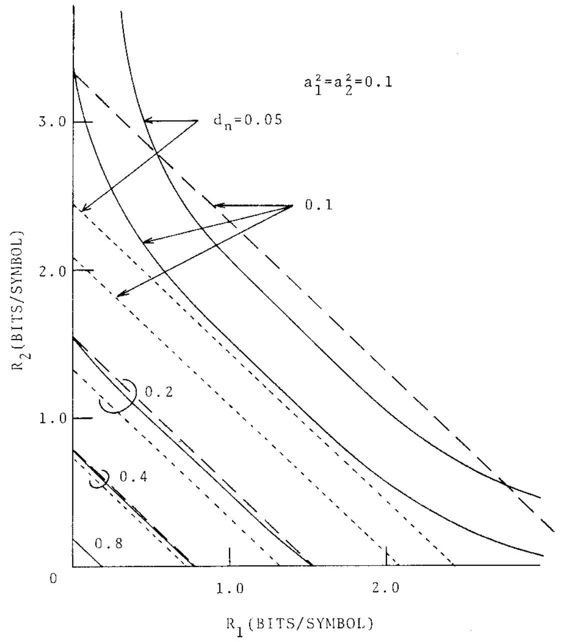
$$r_2=\frac{1}{2}\log\frac{1}{b_2^2(1+a_1^2)}\frac{a_1^2(a_2^2+b_2^2)}{d_n} \quad (28)$$

$$a_1^2+a_2^2+b_2^2+a_1^2(a_2^2+b_2^2)=\frac{a_1^2(a_2^2+b_2^2)}{d_n} \quad (29)$$

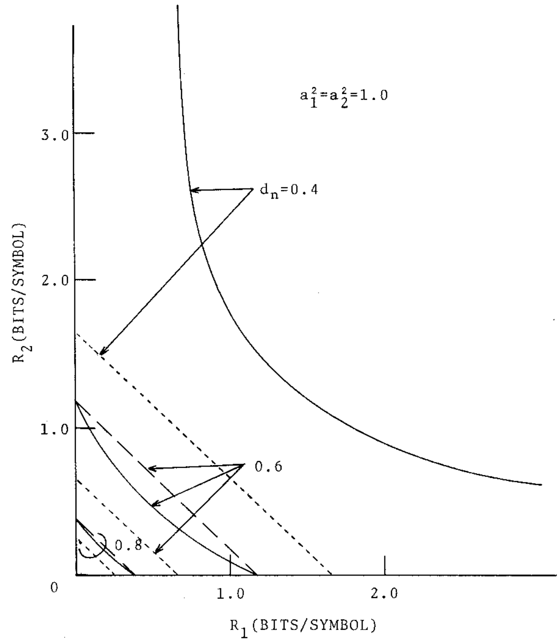
By eliminating b_2^2 from these equations, we obtain

$$R_{wz}^{r*}(d_n)=\begin{cases}\dfrac{1}{2}\log\dfrac{1}{a_2^2\left(1+\dfrac{1}{a_1^2}\right)\left(1+\dfrac{1}{a_1^2}+\dfrac{1}{a_2^2}\right)\left(d_n-\dfrac{1}{1+\dfrac{1}{a_1^2}+\dfrac{1}{a_2^2}}\right)}, \\[2ex] \qquad \dfrac{1}{1+\dfrac{1}{a_1^2}+\dfrac{1}{a_2^2}}<d_n\leq\dfrac{1}{1+\dfrac{1}{a_1^2}} \end{cases}$$

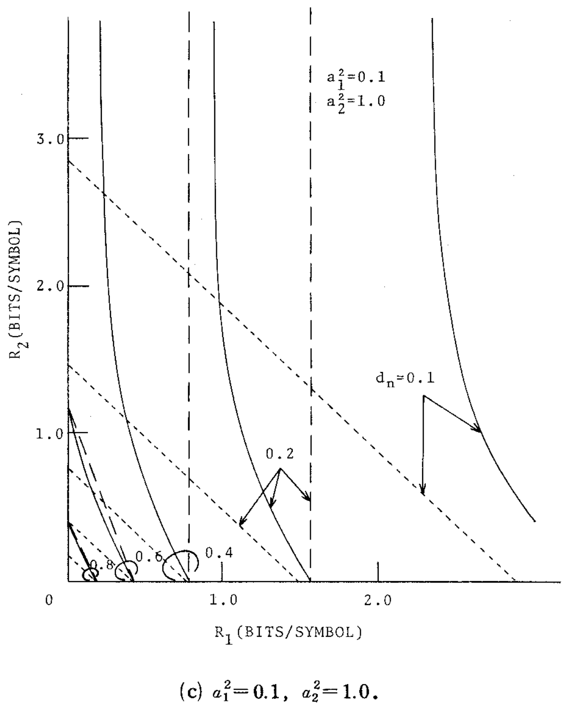
† When $\hat{Y}_1=Y_1$, both $\mathcal{P}^i(d)$ and $\mathcal{P}^o(d)$ equal to $\mathcal{P}_{wz}^r(d)$.



(a) $a_1^2=a_2^2=0.1$,



(b) $a_1^2=a_2^2=1.0$,



(c) $a_1^2=0.1,\ a_2^2=1.0$.

Fig. 3 $\mathcal{R}^i(d_n)$ for the Gaussian remote source.

$$\left| \; 0, \qquad \frac{1}{1+\frac{1}{a_1^2}} \leqq d_n. \right. \tag{30}$$

In this example, even if the encoder could know $\{Y_1\}$, the rate-distortion function would not decrease just by the same reason as in the jointly Gaussian example for the Wyner-Ziv system(9, section3). The reason can be explained by means of Figs. 5(a) and (b) which represent the optimal test channels for each situations, respectively, where α is an arbitrary constant and f' is the minimum mean square estimator of X given Y_1 and $\hat{Y}_3$. Two configurations can be transformed to each other and attain the same values of the mean square error and the conditional mutual information as follows.

$$d=E[(X-\hat{X})^2]=\frac{1}{\dfrac{1}{\sigma_X^2}+\dfrac{c_1^2}{\sigma_{W1}^2}+\dfrac{c_2^2}{\sigma_{W2}^2+\sigma_{V2}^2}} \tag{31}$$

$$I(Y_2;\hat{Y}_2|Y_1)=I(Y_3;\hat{Y}_3|Y_1)$$
$$=\frac{1}{2}\log\frac{\dfrac{\sigma_{W1}^2}{c_1^2\sigma_X^2+\sigma_{W1}^2}\,c_2^2\sigma_X^2+\sigma_{W2}^2+\sigma_{V2}^2}{\sigma_{V2}^2} \tag{32}$$

As the second example, let us consider a binary remote source, Fig. 6, where $\mathcal{X}=\mathcal{Y}_1=\mathcal{Y}_2=\hat{\mathcal{X}}=\{0, 1\}$,

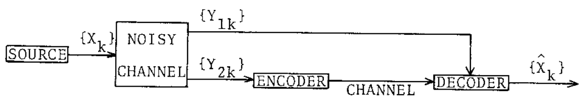


Fig. 4 Wyner-Ziv-type system with a remote source.

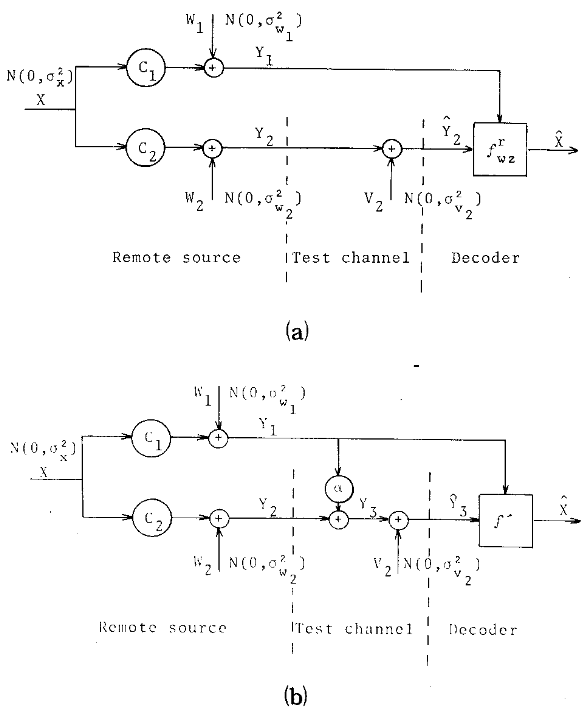


Fig. 5 Optimal test channels for the Gaussian remote source (a) in case the encoder cannot observe Y_1, (b) in case the encoder can observe Y_1.

and the noisy channel consists of two binary symmetric channels. The input probability is $Q_X(0)=Q_X(1)=\frac{1}{2}$ and the bit error probabilities are P_1 and P_2. Let the distortion measure be the Hamming metric.

$$D(x,\hat{x})=\begin{cases}0 & \text{if } x=\hat{x}\\ 1 & \text{if } x\neq\hat{x}\end{cases} \tag{33}$$

Then, the following theorem is obtained.

[Theorem 3]

The rate-distortion function for the Wyner-Ziv -type system with the binary remote source is given as follows.

(i) Average distortion such that $d<\min(P_1, P_2)$ is not attainable.

(ii) For $d\geqq\min(P_1,P_2)$,

if $P_1\geqq P_2$

$$R_{wz}^{r*}(d)=\begin{cases}g(d), & P_2\leqq d\leqq P_1\\ 0, & P_1\leqq d\end{cases} \tag{34}$$

if $P_1\leqq P_2$

$$R_{wz}^{r*}(d)=0,\quad d\geqq P_1 \tag{35}$$

where $g(d)$ is defined by

$$g(d)\triangleq \inf_{\substack{0\leqq\theta\leqq 1\\ 0\leqq\beta*P_2\leqq P_1\\ d=\theta(P_2*\beta)+(1-\theta)P_1}} [\theta\{h(P_1*P_2*\beta)-h(\beta)\}], \tag{36}$$

and the following notations are employed.

$$x*y\triangleq x(1-y)+(1-x)y \tag{37}$$

$$h(x)\triangleq -x\log x-(1-x)\log(1-x) \tag{38}$$

The proof is given in the appendix.

$R_{wz}^{r*}(d)$ versus d for the Gaussian and binary remote sources are depicted in Figs. 7 and 8, respectively.

If Y_2 equals X in Fig. 4, the system reduces to the Wyner-Ziv system[8],[9]. This situation is attained by letting $a_2^2=0$ in Eq. (30) or in $P_2=0$ in Eq. (34), and $R_{wz}^{r*}(d)$ is observed to coincide with the rate -distortion function for the Wyner-Ziv system[8],[9]. If, on the other hand, Y_1 is independent of X and Y_2 in Fig. 4, the system reduces to the Shannon system with a remote source. This situation is achieved by letting $a_1^2=\infty$ in Eq. (30) or $P_1=\frac{1}{2}$ in Eq. (34), and it is easy to see that $R_{wz}^{r*}(d)$ becomes equal to the rate -distortion function for the Shannon system with a remote source[1].

4. Concluding Remarks

We have discussed the source coding problems for the Slepian-Wolf-type system with a remote source and the Wyner-Ziv-type system with a remote source. For the former system the upper and lower bounds were obtained on the admissible rate region, and for the latter system the rate-distortion function was derived. The result can be easily extended to a more general multi-terminal system having several remote sources, encoders and decoders[6],[10],[12].

In some situations, inputs to the separate encoders, Y_1 and Y_2, may be outputs of correlated sources and, at the decoder, it is wanted to reproduce some

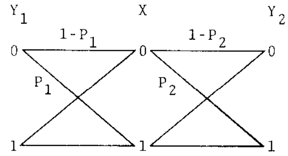


Fig. 6 Binary remote source.

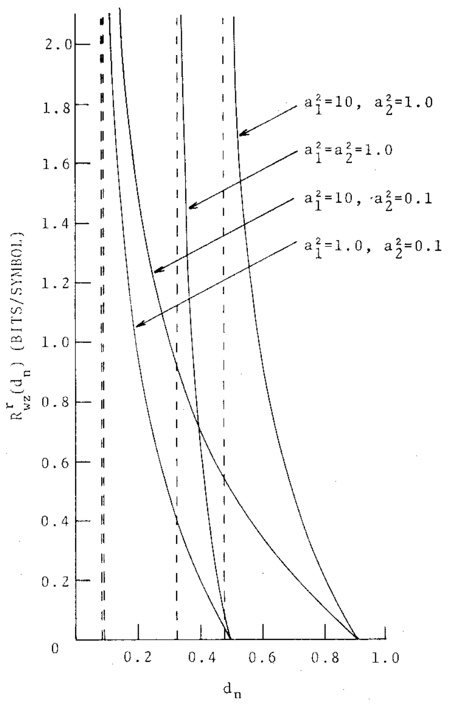


Fig. 7 $R^r_{wz}(d_n)$ versus d_n for the Guassian remote source.

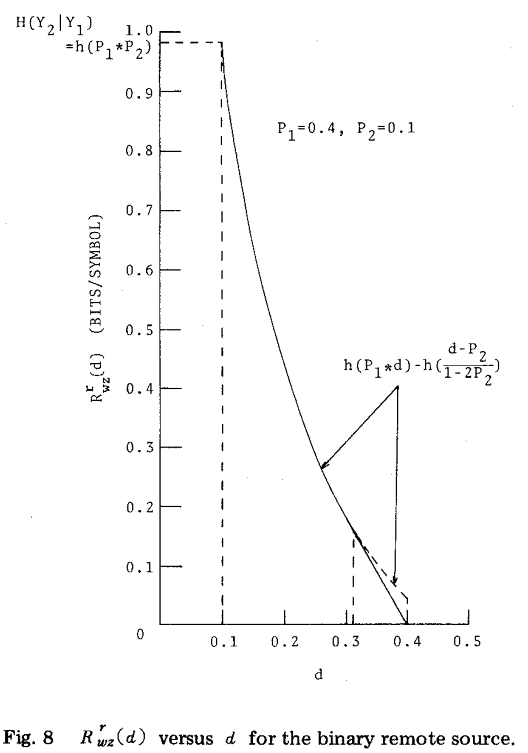


Fig. 8 $R^r_{wz}(d)$ versus d for the binary remote source.

function of Y_1 and Y_2, $X \triangleq F(Y_1, Y_2)$, within a prescribed distortion tolerance. For example, $X = Y_1 - Y_2$ for continuous amplitude sources, or $X = Y_1 + Y_2$ (mod 2) for binary sources. It is easily noticed that theorems 1 and 2 hold for such situations.

Acknowledgement

The authors wish to thank Dr. T.S. Han for providing the literatures and for his helpful suggestion.

References

(1) Berger, T.: "Rate distortion theory, a mathematical basis for data compression", Prentice-Hall, Inc. (1971).
(2) Dobrushin, R.L. and Tsybakov, B.S.: "Information transmission with additional noise", IRE Trans. Inform. Theory, IT-8, pp. 293—304 (Sept. 1962).
(3) Wolf, J.K. and Ziv, J.: "Transmission of noisy information to a noisy receiver with minimum distortion", IEEE Trans. Inform. Theory, IT-16, pp. 406—411 (July 1970).
(4) Berger, T.: "Multiterminal source coding", information theory approach to communications (G. Long, ed.), CISM Cources and Lectures #229, Springer-Verlag (1978).
(5) Slepiano, D. and Wolf, J.K.: "Noiseless coding of correlated information sources", IEEE Trans. Inform. Theory, IT-21, pp. 226—228 (March 1975).
(6) Omura, J.K. and Housewright, K.B.: "Source coding studies for information networks", Proc. of IEEE 1977 Int. Conf. on Communication, IEEE Press, Chicago, III, pp. 237—240 (June 1977).
(7) Berger, T. and Tung, S.Y.: "Multiterminal source coding", IEEE Int. Symp. Inform. Theory, Ithaca, N.Y. (Oct. 1977).
(8) Wyner, A.D. and Ziv, J.: "The rate-distortion function for source coding with side information at the decoder", IEEE Trans. Inform. Theory, IT-22, pp. 1—10 (Jan. 1976).
(9) Wyner, A.D.: "The rate-distortion function for source coding with side information at the decoder-II: General sources", Inf. & Control, 38, pp. 60—80 (1978).
(10) Housewright, K.B.: "Source coding studies for multiterminal system", Ph.D. dissertation, Systems Science Dept., Univ. of California, Los Angels (1977).
(11) Tung, S.Y.: "Multiterminal source coding", Ph. D. thesis, Cornell Univ. (May 1978).
(12) Yamamoto, H.: "Source coding theory for multiterminal communication systems", Dr. Eng. degree thesis, Dept. Elec. Eng., Univ. of Tokyo, Japan (1980).
(13) Kawabata, T.: "Gaussian multiterminal source coding", M.E. thesis, Dept. Math. Eng. and Instrumen. Phys., Univ. of Tokyo, Japan (1980).

Appendix: Proof of Theorem 3

It is not difficult to see that average distortion d less than min (P_1, P_2) cannot be attained even if we estimate X from Y_1 and Y_2 according to the maximum a posteriori probability criterion. So that $d < \min(P_1, P_2)$ cannot be realized at the decoder no matter what code is used. On the other hand, the decoder can obviously estimate X with $d=P_1$ from Y_1 only. Thus $R_{wz}^{r*}(d)=0$ for $d \geq P_1$.

The most important part of theorem 3, Eq. (34), can be proved in a similar way as in Ref. (8), Section II except the following points. In proving the inequality $R_{wz}^{r*}(d) \leq g(d)$, let $\hat{X}=f_{wz}^{r}(Y_1,\hat{Y}_2)=\hat{Y}_2$ and $\hat{X}=f_{wz}^{r}(Y_1,\hat{Y}_2)=Y_1$ instead of letting $\hat{X}=f(Y,Z)=Z$ and $\hat{X}=f(Y,Z)=Y$ at paragraphs a) and b) in Ref. (8), Section II, respectively.

In order to prove the inequality $R_{wz}^{r*}(d) \geq g(d)$, define

$$\mathcal{A} \triangleq \{\hat{y}_2 : f_{wz}^{r}(0,\hat{y}_2)=f_{wz}^{r}(1,\hat{y}_2)\} \qquad \text{(A·1)}$$

$$\mathcal{A}^c \triangleq \hat{\mathcal{Y}}_2-\mathcal{A}=\{\hat{y}_2 : f_{wz}^{r}(0,\hat{y}_2)\neq f_{wz}^{r}(1,\hat{y}_2)\} \qquad \text{(A·2)}$$

$$\theta \triangleq P_r\{\hat{Y}_2 \in \mathcal{A}\} \qquad \text{(A·3)}$$

$$\lambda_{\hat{y}_2} \triangleq \frac{P_r\{\hat{Y}_2=\hat{y}_2\}}{P_r\{\hat{Y}_2 \in \mathcal{A}\}} \qquad \text{(A·4)}$$

$$d_{\hat{y}_2} \triangleq E[D(X,\hat{X})\,|\,\hat{Y}_2=\hat{y}_2] \qquad \text{(A·5)}$$

$$d' \triangleq \theta \sum_{\hat{y}_2 \in \mathcal{A}} \lambda_{\hat{y}_2} d_{\hat{y}_2}+(1-\theta)P_1, \qquad \text{(A·6)}$$

then we can show

$$d' \leq d \qquad \text{(A·7)}$$

$$I(Y_2;\hat{Y}_2|Y_1) \geq \theta \sum_{\hat{y}_2 \in \mathcal{A}} \lambda_{\hat{y}_2}[H(Y_1|\hat{Y}_2=\hat{y}_2)-H(Y_2|\hat{Y}_2=\hat{y}_2)] \qquad \text{(A·8)}$$

in the similar way as in Ref. (8), Eqs. (39) and 40). Furthermore, define

$$\gamma(\hat{y}_2) \triangleq f_{wz}^{r}(0,\hat{y}_2)=f_{wz}^{r}(1,\hat{y}_2) \quad \text{for } \hat{y}_2 \in \mathcal{A} \qquad \text{(A·9)}$$

$$\alpha_{\hat{y}_2} \triangleq P_r\{Y_2 \neq \gamma(\hat{y}_2)\,|\,\hat{Y}_2=\hat{y}_2\} \qquad \text{(A·10)}$$

$$G(u) \triangleq h(P_1 * P_2 * u)-h(u) \qquad \text{(A·11)}$$

then we obtain

$$H(Y_2|\hat{Y}_2=\hat{y}_2)=h(\alpha_{\hat{y}_2}) \qquad \text{(A·12)}$$

$$H(Y_1|\hat{Y}_2=\hat{y}_2)=h(\alpha_{\hat{y}_2} * P_1) \qquad \text{(A·13)}$$

$$d_{\hat{y}_2}=\alpha_{\hat{y}_2} * P_2 \qquad \text{(A·14)}$$

From Eqs. (A·3), (A·4), (A·8), (A·11)–(A·13) and convexity of G (8, Lemma A),

$$\begin{aligned} I(Y_2;\hat{Y}_2|Y_1) &\geq \theta \sum_{\hat{y}_2 \in \mathcal{A}} \lambda_{\hat{y}_2} G(\alpha_{\hat{y}_2}) \\ &\geq \theta G\Big(\sum_{\hat{y}_2 \in \mathcal{A}} \lambda_{\hat{y}_2}\alpha_{\hat{y}_2}\Big) \\ &= \theta[h(P_1 * P_2 * \beta)-h(\beta)] \end{aligned} \qquad \text{(A·15)}$$

where

$$\beta \triangleq \sum_{\hat{y}_2 \in \mathcal{A}} \lambda_{\hat{y}_2}\alpha_{\hat{y}_2} \qquad \text{(A·16)}$$

On the other hand, from Eqs. (A·6) and (A·4),

$$\begin{aligned} d' &= \theta \sum_{\hat{y}_2 \in \mathcal{A}} \lambda_{\hat{y}_2}(\alpha_{\hat{y}_2} * P_2)+(1-\theta)P_1 \\ &= \theta\Big\{\Big(\sum_{\hat{y}_2 \in \mathcal{A}} \lambda_{\hat{y}_2}\alpha_{\hat{y}_2}\Big) * P_2+(1-\theta)P_1\Big\} \\ &= \theta(\beta * P_2)+(1-\theta)P_1 \end{aligned} \qquad \text{(A·17)}$$

where the second equality follows from Eq. (A·20). Thus, Eqs. (A·15), (A·17) and (34)–(36) yield

$$I(Y_2;\hat{Y}_2|Y_1) \geq g(d') \qquad \text{(A·18)}$$

Now, because of (A·7) and since $g(d)$ is nonincreasing in d (8), we have

$$I(Y_2;\hat{Y}_2|Y_1) \geq g(d) \qquad \text{(A·19)}$$

Note:

Let $\sum_k a_k=1$. Then,

$$\begin{aligned} \sum_k a_k(x_k * y) &= \sum_k a_k\{x_k(1-y)+(1-x_k)y\} \\ &= \sum_k a_k x_k(1-y)+\sum_k a_k(1-x_k)y \\ &= \Big(\sum_k a_k x_k\Big)(1-y)+\Big(1-\sum_k a_k x_k\Big)y \\ &= \Big(\sum_k a_k x_k\Big) * y. \end{aligned} \qquad \text{(A·20)}$$