1st Semester 2012/13

Homework Set #3

source coding

1. **Slepian-Wolf for deterministically related sources.** Find and sketch the Slepian-Wolf rate region for the simultaneous data compression of (X, Y), where $y = f(x)$ is some deterministic function of x.

   Solutions

   The quantities defining the Slepian Wolf rate region are $H(X, Y) = H(X)$, $H(Y|X) = 0$ and $H(X|Y) \geq 0$. Hence the rate region is as shown in the Figure 1.

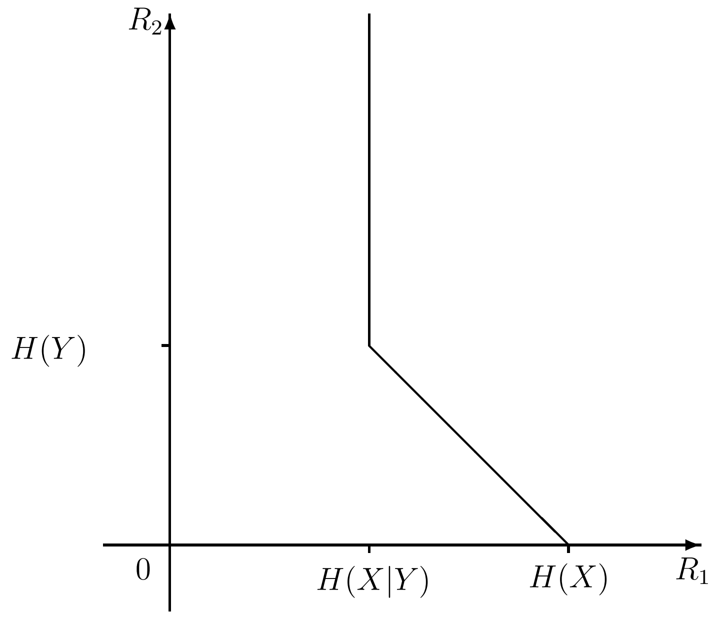


Figure 1: Slepian Wolf rate region for $Y = f(X)$.

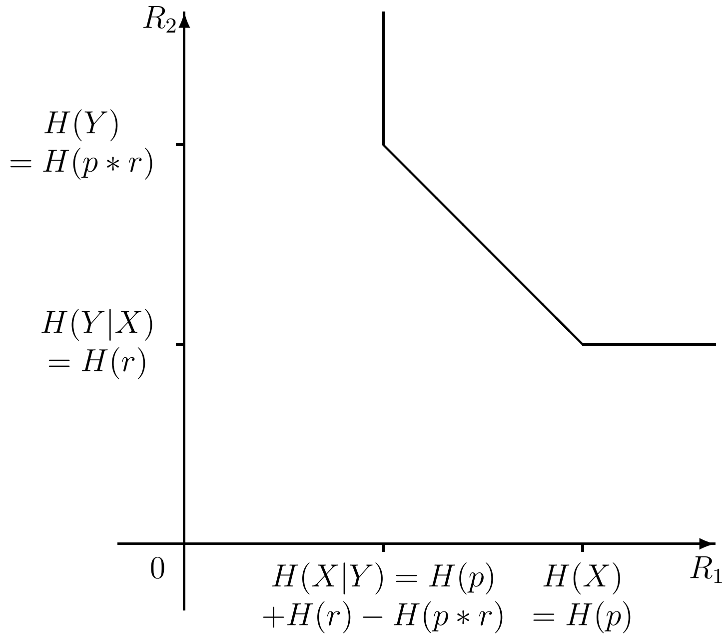


Figure 2: Slepian Wolf region for binary sources

2. **Slepian Wolf for binary sources.** Let X_i be i.i.d. Bernoulli(p). Let Z_i be i.i.d. $\sim$ Bernoulli(r), and let $\mathbf{Z}$ be independent of $\mathbf{X}$. Finally, let $\mathbf{Y} = \mathbf{X} \oplus \mathbf{Z}$ (mod 2 addition). Let $\mathbf{X}$ be described at rate R_1 and $\mathbf{Y}$ be described at rate R_2. What region of rates allows recovery of $\mathbf{X}, \mathbf{Y}$ with probability of error tending to zero?

   Solutions

   $X \sim \text{Bern}(p)$. $Y = X \oplus Z$, $Z \sim \text{Bern}(r)$. Then $Y \sim \text{Bern}(p * r)$, where $p*r = p(1-r)+r(1-p)$. $H(X) = H(p)$. $H(Y) = H(p*r)$, $H(X, Y) = H(X, Z) = H(X) + H(Z) = H(p) + H(r)$. Hence $H(Y|X) = H(r)$ and $H(X|Y) = H(p) + H(r) - H(p * r)$.

   The Slepian Wolf region in this case is shown in Figure 2.

3. **Computing simple example of Slepian Wolf**

   Let (X, Y) have the joint pmf $p(x, y)$

p(x,y)	1	2	3
1	α	β	β
2	β	α	β
3	β	β	α

where $\beta = \frac{1}{6} - \frac{\alpha}{2}$. (Note: This is a joint, not a conditional, probability mass function.)

(a) Find the Slepian Wolf rate region for this source.

(b) What is $\Pr\{X = Y\}$ in terms of α?

(c) What is the rate region if $\alpha = \frac{1}{3}$?

(d) What is the rate region if $\alpha = \frac{1}{9}$?

Solutions

(a) $H(X,Y) = -\sum p(x,y) \log p(x,y) = -3\alpha \log \alpha - 6\beta \log \beta$. Since X and Y are uniformly distributed

$$H(X) = H(Y) = \log 3 \tag{1}$$

and

$$H(X|Y) = H(Y|X) = H(3\alpha, 3\beta, 3\beta) \tag{2}$$

Hence the Slepian Wolf rate region is

$$R_1 \geq H(X|Y) = H(3\alpha, 3\beta, 3\beta) \tag{3}$$

$$R_2 \geq H(Y|X) = H(3\alpha, 3\beta, 3\beta) \tag{4}$$

$$R_1 + R_2 \geq H(X,Y) = H(3\alpha, 3\beta, 3\beta) + \log 3 \tag{5}$$

(b) From the joint distribution, $\Pr(X = Y) = 3\alpha$.

(c) If $\alpha = \frac{1}{3}$, $\beta = 0$, and $H(X|Y) = H(Y|X) = 0$. The rate region then becomes

$$R_1 \geq 0 \tag{6}$$

$$R_2 \geq 0 \tag{7}$$

$$R_1 + R_2 \geq \log 3 \tag{8}$$

(d) If $\alpha = \frac{1}{9}$, $\beta = \frac{1}{9}$, and $H(X|Y) = H(Y|X) = \log 3$. X and Y are independent, and the rate region then becomes

$$\begin{aligned} R_1 &\geq \log 3 &(9)\\ R_2 &\geq \log 3 &(10)\\ R_1 + R_2 &\geq 2\log 3 &(11) \end{aligned}$$

4. **An example of Slepian-Wolf**: Two senders know random variables U_1 and U_2 respectively. Let the random variables (U_1, U_2) have the following joint distribution:

$U_1 \backslash U_2$	0	1	2	$\cdots$	$m-1$
0	α	$\frac{\beta}{m-1}$	$\frac{\beta}{m-1}$	$\cdots$	$\frac{\beta}{m-1}$
1	$\frac{\gamma}{m-1}$	0	0	$\cdots$	0
2	$\frac{\gamma}{m-1}$	0	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$m-1$	$\frac{\gamma}{m-1}$	0	0	$\cdots$	0

where $\alpha+\beta+\gamma = 1$. Find the region of rates (R_1, R_2) that would allow a common receiver to decode both random variables reliably.

Solutions For this joint distribution,

$$H(U_1) = H(\alpha+\beta, \frac{\gamma}{m-1}, \ldots, \frac{\gamma}{m-1}) = H(\alpha+\beta, \gamma) + \gamma \log(m-1) \tag{12}$$

$$H(U_2) = H(\alpha+\gamma, \frac{\beta}{m-1}, \ldots, \frac{\beta}{m-1}) = H(\alpha+\gamma, \beta) + \beta \log(m-1) \tag{13}$$

$$\begin{aligned} H(U_1, U_2) &= H\left(\alpha, \frac{\beta}{m-1}, \ldots, \frac{\beta}{m-1}, \frac{\gamma}{m-1}, \ldots, \frac{\gamma}{m-1}\right) &(14)\\ &= H(\alpha, \beta, \gamma) + \beta\log(m-1) + \gamma\log(m-1) &(15) \end{aligned}$$

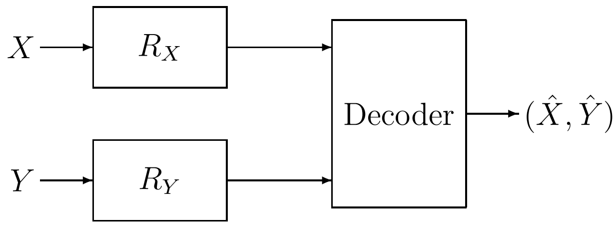
$$\begin{aligned} H(U_1|U_2) &= H(\alpha, \beta, \gamma) - H(\alpha+\gamma, \beta) + \gamma\log(m-1) &(16)\\ H(U_2|U_1) &= H(\alpha, \beta, \gamma) - H(\alpha+\beta, \gamma) + \beta\log(m-1) &(17) \end{aligned}$$

and hence the Slepian Wolf region is

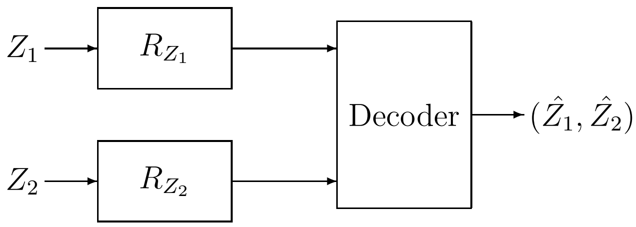
$$R_1 \geq H(\alpha, \beta, \gamma) - H(\alpha + \gamma, \beta) + \gamma \log(m-1) \quad (18)$$
$$R_2 \geq H(\alpha, \beta, \gamma) - H(\alpha + \beta, \gamma) + \beta \log(m-1) \quad (19)$$
$$R_1 + R_2 \geq H(\alpha, \beta, \gamma) + \beta \log(m-1) + \gamma \log(m-1) \quad (20)$$

5. **Stereo.** The sum and the difference of the right and left ear signals are to be individually compressed for a common receiver. Let Z_1 be Bernoulli (p_1) and Z_2 be Bernoulli (p_2) and suppose Z_1 and Z_2 are independent. Let $X = Z_1 + Z_2$, and $Y = Z_1 - Z_2$.

   (a) What is the Slepian Wolf rate region of achievable (R_X, R_Y)?



   (b) Is this larger or smaller than the rate region of (R_{Z_1}, R_{Z_2})? Why?



   There is a simple way to do this part.

Solutions

The joint distribution of X and Y is shown in following table

Z_1	Z_2	X	Y	probability
0	0	0	0	$(1-p_1)(1-p_2)$
0	1	1	-1	$(1-p_1)p_2$
1	0	1	1	$p_1(1-p_2)$
1	1	2	0	$p_1 p_2$

and hence we can calculate

$$H(X) = H(p_1p_2, p_1 + p_2 - 2p_1p_2, (1-p_1)(1-p_2)) \tag{21}$$

$$H(Y) = H(p_1p_2 + (1-p_1)(1-p_2), p_1 - p_1p_2, p_2 - p_1p_2) \tag{22}$$

and

$$H(X,Y) = H(Z_1, Z_2) = H(p_1) + H(p_2) \tag{23}$$

and therefore

$$\begin{aligned} H(X|Y) &= H(p_1) + H(p_2) - H(p_1p_2 + (1-p_1)(1-p_2), p_1 - p_1p_2, p_2 - p_1p_2) \\ H(Y|X) &= H(p_1) + H(p_2) - H(p_1p_2, p_1 + p_2 - 2p_1p_2, (1-p_1)(1-p_2)) \quad (24) \end{aligned}$$

The Slepian Wolf region in this case is

$$\begin{aligned} R_1 &\geq H(X|Y) = H(p_1) + H(p_2) - H(p_1p_2 + (1-p_1)(1-p_2), p_1 - p_1p_2, p_2 - p_1p_2) \\ R_2 &\geq H(Y|X) = H(p_1) + H(p_2) - H(p_1p_2, p_1 + p_2 - 2p_1p_2, (1-p_1)(1-p_2)) \\ R_1 + R_2 &\geq H(p_1) + H(p_2) \quad (25) \end{aligned}$$

The Slepian Wolf region for (Z_1, Z_2) is

$$R_1 \geq H(Z_1|Z_2) = H(p_1) \tag{26}$$

$$R_2 \geq H(Z_2|Z_1) = H(p_2) \tag{27}$$

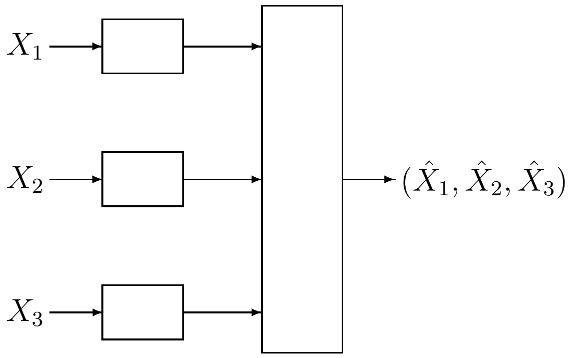
$$R_1 + R_2 \geq H(Z_1, Z_2) = H(p_1) + H(p_2) \tag{28}$$

which is a rectangular region.

The minimum sum of rates is the same in both cases, since if we knew both X and Y, we could find Z_1 and Z_2 and vice versa. However, the region in part (a) is usually pentagonal in shape, and is larger than the region in (b).

6. **Distributed data compression.** Let Z_1, Z_2, Z_3 be independent Bernoulli(p). Find the Slepian-Wolf rate region for the description of (X_1, X_2, X_3) where

$$\begin{aligned} X_1 &= Z_1 \\ X_2 &= Z_1 + Z_2 \\ X_3 &= Z_1 + Z_2 + Z_3 \ . \end{aligned}$$



Solutions.

To establish the rate region, appeal to Theorem 14.4.2 in the text, which generalizes the case with two encoders. The inequalities defining the rate region are given by

$$R(S) > H(X(S)|X(S^c))$$

for all $S \subseteq \{1, 2, 3\}$, and $R(S) = \sum_{i \in S} R_i$.

The rest is calculating entropies $H(X(S)|X(S^c))$ for each S. We have

$$\begin{aligned}
H_1 &= H(X_1) = H(Z_1) = H(p), \\
H_2 &= H(X_2) = H(Z_1 + Z_2) = H(p^2, 2p(1-p), (1-p)^2), \\
H_3 &= H(X_3) = H(Z_1 + Z_2 + Z_3) \\
&= H(p^3, 3p^2(1-p), 3p(1-p)^2, (1-p)^3), \\
H_{12} &= H(X_1, X_2) = H(Z_1, Z_2) = 2H(p), \\
H_{13} &= H(X_1, X_3) = H(X_1) + H(X_3|X_1) = H(X_1) + H(Z_2 + Z_3) \\
&= H(p^2, 2p(1-p), (1-p)^2) + H(p), \\
H_{23} &= H(X_2, X_3) = H(X_2) + H(X_3|X_2) = H(X_2) + H(Z_3) \\
&= H(p^2, 2p(1-p), (1-p)^2) + H(p), \quad \text{and} \\
H_{123} &= H(X_1, X_2, X_3) = H(Z_1, Z_2, Z_3) = 3H(p).
\end{aligned}$$

Using the above identities and chain rule, we obtain the rate region as

$$
\begin{aligned}
R_1 &> H(X_1|X_2,X_3) = H_{123} - H_{23} \\
&= 2H(p) - H(p^2, 2p(1-p), (1-p)^2) = 2p(1-p), \\
R_2 &> H(X_2|X_1,X_3) = H_{123} - H_{13} = 2p(1-p), \\
R_3 &> H(X_3|X_1,X_2) = H_{123} - H_{12} = H(p), \\
R_1 + R_2 &> H(X_1,X_2|X_3) = H_{123} - H_3 \\
&= 3H(p) - H(p^3, 3p^2(1-p), 3p(1-p)^2, (1-p)^3) = 3p(1-p)\log(3), \\
R_1 + R_3 &> H(X_1,X_3|X_2) = H_{123} - H_2 \\
&= 3H(p) - H(p^2, 2p(1-p), (1-p)^2) = H(p) + 2p(1-p), \\
R_2 + R_3 &> H(X_2,X_3|X_1) = H_{123} - H_1 = 2H(p), \quad \text{and} \\
R_1 + R_2 + R_3 &> H_{123} = 3H(p).
\end{aligned}
$$

7. **Gaussian Wyner-Ziv** Show that for a mean square distortion where the source X and Y are jointly Gaussian, the best auxiliary random variable is Gaussian and the distortion is as Y is known to the Encoder.

8. **Wyner ziv with two decoders and one message** Consider the coordination Wyner-Ziv problem, but where the message at rate R reaches two decoders. Assume that there is a degradation of the side information at the decoders, and the first decoder has a side information Y and the second decoder does not have any side information at all. Let $\hat{X}_1$ be the action taken by Decode 1 and $\hat{X}_2$ the action taken by Decode 2. Show that in order to achieve a (weak) coordination $P_0(x,y)P(\hat{x}_2|x)P(\hat{x}_1|x,y,\hat{x}_2)$ then

$$R \geq I(X;\hat{X}_2) + I(X;U|X_2,Y) \tag{29}$$

where the joint distribution is of the form $P_0(x,y)P(\hat{x}_2|x)P(\hat{u}|x,y,\hat{x}_2)P(\hat{x}_1|u,x,y,\hat{x}_2)$ with a marginal (summing over u) is the coordination pmf $P_0(x,y)P(\hat{x}_2|x)P(\hat{x}_1|x,y,\hat{x}_2)$. Furthermore, show that if (29) holds for some auxiliary U then there exists a code that achieves a coordination $P_0(x,y)P(\hat{x}_2|x)P(\hat{x}_1|x,y,\hat{x}_2)$.