ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 23 Information Theory and Coding
Solution 10 December 7, 2010, SG1 – 15:15pm-17:00

Problem 1. (i) Notice that the capacity of C_2 can be computed via the expression

$$\text{Max}I(Y;X) = \text{Max}(H(Y) - H(Y|X)) = \text{Max}(H(Y) - H(Z))$$

where the maximum is taken over all the possible distributions on the input alphabet. Therefore we only need to find the maximum of $H(Y)$ over all possible distribution on the set $\{-1, 1\}$. To do this job we need to compute a non-trivial integral.

(ii) , Notice that this channel is in fact a BSC channel with error probability equal to $1 - \Phi(\sqrt{P})$. So the capacity of this channel is equal to $1 - H_2(1 - \Phi(\sqrt{P}))$

Problem 2. (i) Since among the random variables of a given variance, the Gaussian random variable has the largest entropy, we can argue that

$$h(X|Y = i) \leq \frac{1}{2}log(2\pi e\text{Var}(X|Y = y)$$

for every value of y. Therefore we can conclude that

$$h(X|Y) \leq E_Y(\frac{1}{2}log(2\pi e\text{Var}(X|Y)).$$

On the other hand, since log is a concave function, we can use the Jensen's inequality to conclude that $E_Y(\frac{1}{2}log(2\pi e\text{Var}(X|Y))) \leq \frac{1}{2}\log(2\pi eE_Y(\text{Var}(X|Y)))$. This completes the proof of the first part.

(ii) Since the minimum mean square error estimator of os X, based on the observation $Y = y$ is $\hat{X}_{MMSE}(y) = E(X|Y = y)$, for every arbitrary estimator $\hat{X}(y)$ we have $E(E_X((\hat{X}_{MMSE} - X)^2)) = E(E_X((E(X|Y) - X)^2)) \leq E((\hat{X}(y) - X)^2)$. But notice that $E_X((E(X|Y) - X)^2) = Var(X|Y)$. So, $E(E_X((E(X|Y) - X)^2) = E_Y((Var(X|Y)))$ and therefore, for every arbitrary estimator $\hat{X}(y)$ of X we have $E_Y(Var(X|Y) \leq E(E_X((\hat{X} - X)^2))$. In particular, let $\hat{X}_Y(y) = \frac{a}{a+b}y$. For this particular $\hat{X}(y)$ we will find $E(E_X((\hat{X} - X)^2))$. Since $Y = X + Z$ and X and Z are independent random variables, we have:

$$\begin{aligned}
E(E_X((\hat{X} - X)^2)) &= E(E_X((\frac{a}{a+b}Y - X)^2)) \\
&= E(E_X((\frac{a}{a+b}(X+Z) - X)^2)) \\
&= E(E_X((\frac{a}{a+b}Z - \frac{b}{a+b}X)^2)) \\
&= E(\frac{a^2}{(a+b)^2}E_X(Z^2) + \frac{b^2}{(a+b)^2}E_X(X^2)) \\
&= \frac{a^2}{(a+b)^2}E(E_X(Z^2)) + \frac{b^2}{(a+b)^2}E_X(X^2) \\
&= \frac{a^2b}{(a+b)^2} + \frac{ab^2}{(a+b)^2} \\
&= \frac{ab}{a+b}.
\end{aligned}$$

Thus, $E_Y(\text{Var}(X|Y)) \leq \frac{ab}{a+b}$.

(iii) Since the logarithm is an increasing function,

$$\frac{1}{2}\log(2\pi e E_Y(\text{Var}(X|Y))) \leq \frac{1}{2}\log(2\pi e \frac{ab}{a+b}).$$

This inequality together with the result of the part (i) completes the proof.

Problem 3. We have

$$\begin{aligned} C &= \sup_{X:E[X^2]\leq P} I(X;Y_1,Y_2) \\ &= h(Y_1,Y_2) - h(Y_1,Y_2|X) \\ &= h(X+Y_1,X+Y_2) - h(Z_1,Z_2|X) \\ &= h(X+Y_1,X+Y_2)_h(Z_1,Z_2), \end{aligned}$$

where we have used the fact that X and (Z_1, Z_2) are independent. Now since

$$(Z_1,Z_2) \sim \mathcal{N}((0,0), \begin{bmatrix} M & M\sigma \\ M\sigma & M \end{bmatrix}),$$

we have

$$h(Z_1,Z_2) = \frac{1}{2}\log((2\pi e)^2|K|) = \frac{1}{2}\log((2\pi e)^2 M^2(1-\sigma^2)).$$

Further, we have

$$h(X+Z_1,X+Z_2) \leq \frac{1}{2}\log((2\pi e)^2|\tilde{K}|) = \frac{1}{2}\log((2\pi e)^2(M(M-P(-2+\sigma)) - (P+M\sigma)\sigma M)).$$

where $\tilde{K}$ is given by

$$\begin{bmatrix} P+M & P+M\sigma \\ P+M\sigma & P+M \end{bmatrix}).$$

Note that we get equality by assuming that $X \sim \mathcal{N}(0,P)$. Hence the capacity is:

$$C = h(X+Z_1,X+Z_2) - h(Z_1,Z_2) = \frac{1}{2}(\log(1+\frac{2P}{M(1+\sigma)})).$$

So, we only need to substitute $\sigma = 1, \frac{1}{2}$ and -1 to find the solution of each part.

Problem 4. (i) All rates less than $\frac{1}{2}\log(1+\frac{P}{\sigma_1^2})$ are achievable.

(ii) The new noise $Z_1 - \rho Z_2$ has zero mean and variance $E((Z_1-Z_2)^2) = \sigma_1^2 + \rho^2\sigma_2^2 - 2\rho\sigma_3$. Therefore, all rates less than $\frac{1}{2}\log(1+\frac{P}{\sigma_1^2+\rho^2\sigma_2^2-2\rho\sigma_3}))$ are achievable.

(iii) The capacity is $C = \max I(X;Y_1,Y_2) = \max(h(Y_1,Y_2) - h(Z_1,Z_2))$. We can easily see that the capacity achieving distribution on the input is $x \sim \mathcal{N}(0,P)$ and therefore Y_1, Y_2 are jointly Gaussian random variables and the determinant of their covariance matrix will be equal to $(\sigma_1^2+P)\sigma_2^2 - \sigma_3^2$. Hence, the capacity of the channel 3 will be equal to $\frac{1}{2}\log(\frac{(\sigma_1^2+P)\sigma_2^2-\sigma_3^2}{\sigma_1^2\sigma_2^2-\sigma_3^2})$.