Solutions: Homework Set # 10

Problem 1

Parallel channels and waterfilling. By the result of parallel channels that you have seen in the course, it follows that we will put all the signal power into the channel with less noise until the total power of noise + signal in that channel equals the noise power in the other channel. After that, we will split any additional power evenly between the two channels. Thus the combined channel begins to behave like a pair of parallel channels when the signal power is equal to the difference of the two noise powers, i.e., when $2P = \sigma_1^2 - \sigma_2^2$.

Problem 2

Normally one would water-fill over the eigenvalues of the noise covariance matrix. Here we have the degenerate case (i.e., one of the eigenvalue is zero), which we can exploit easily.
Musing upon the structure of the noise covariance matrix, one can see $Z_1 + Z_2 = Z_3$. Thus, by processing the output vector as $Y_1+Y_2-Y_3 = (X_1+Z_1)+(X_2+Z_2)-(X_3+Z_3) = X_1+X_2-X_3$; we can get rid of the noise completely. Therefore, we have infnite capacity.
Note that we can reach the conclusion by water-filling on the zero eigenvalue.

Problem 3

(a) By data processing inequality we can write

$$I(S^k; \hat{S}^k) \leq I(X^m; Y^m).$$

Then we can write

$$\min_{p(\hat{s}|s):\ \mathbb{E}[d(s,\hat{s})]\leq D} I(S^k; \hat{S}^k) \leq I(X^m; Y^m).$$

Let $p^*(\hat{s}|s)$ be the minimizer for $I(S^k; \hat{S}^k)$. This means that the encoder and decoder are chosen such that minimize $I(S^k; \hat{S}^k)$. So these particular choise for the encoder and decoder impose some input distribution for the Gaussian channel *i.e.*, impose some distribution on the sequence X^m which can be different from the distribution which maximize $I(X^m; Y^m)$. So we can write

$$R(D) = \min_{p(\hat{s}|s):\ \mathbb{E}[d(s,\hat{s})]\leq D} I(S^k; \hat{S}^k) \leq I(X^m; Y^m) \leq \max_{p(x):\ \mathbb{E}[\rho(X)]\leq P} I(X^m; Y^m) = C(P).$$

(b) We want to find an expression for $R(D)$ where

$$R(D) = \min_{p(\hat{s}|s):\ \mathbb{E}[d(s,\hat{s})]\leq D} I(S^k; \hat{S}^k).$$

Let us write

$$\begin{aligned}
I(S;\hat{S}) &= H(S) - H(S|\hat{S}) \\
&= 1/2\log(2\pi e Q) - H(S|\hat{S}) \\
&= 1/2\log(2\pi e Q) - H(S-\hat{S}|\hat{S}) \\
&\geq 1/2\log(2\pi e Q) - H(S-\hat{S}) \\
&\geq 1/2\log(2\pi e Q) - H(\mathcal{N}(0,\mathbb{E}(S-\hat{S})^2)) \\
&= 1/2\log(2\pi e Q) - 1/2\log(2\pi e\mathbb{E}(S-\hat{S})^2) \\
&\geq 1/2\log(2\pi e Q) - 1/2\log(2\pi e D) \\
&= 1/2\log\frac{Q}{D},
\end{aligned}$$

so $I(S;\hat{S}) \geq \left(\frac{1}{2}\log\frac{Q}{D}\right)^+$, and remember that it is shown in the course that in fact this rate is acheivable so finally we have

$$R(D) = \left[\frac{1}{2}\log\frac{Q}{D}\right]^+ .$$

Now let us compute $C(P)$. We have

$$C(P) = \max_{p(x):\ \mathbb{E}[X^2]\leq P} .$$

Then we can write

$$\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&= H(Y) - H(X+Z|X) \\
&= H(Y) - H(Z) \\
&= H(Y) - 1/2\log(2\pi e N).
\end{aligned}$$

We know that X and Z are independent so $\mathbb{E}Y^2 = \mathbb{E}X^2 + N \leq P + N$. Then we have $I(X;Y) \leq 1/2\log(1+P/N)$ and this rate is acheivable by choosing the input distribution to be Gaussian, so we have

$$C(P) = \frac{1}{2}\log\left(1+\frac{P}{N}\right).$$

From part (a) we conclude that

$$\frac{NQ}{N+P} \leq D.$$

(c) We have $X_\ell = \alpha S_\ell$ so we can write $\mathbb{E}X_\ell^2 = \alpha^2\mathbb{E}S_\ell^2$. Then we have $P = \alpha^2 Q$ which means

$$\alpha = \sqrt{\frac{P}{Q}}.$$

For the second part we have $Y_\ell = \alpha S_\ell + Z_\ell$ and $\hat{S}_\ell = \beta Y_\ell$. Then we can write

$$\mathbb{E}\Big[(S_\ell - \hat{S}_\ell)^2\Big] = \mathbb{E}\big[(S_\ell - \alpha\beta S_\ell - \beta Z_\ell)^2\big].$$

Taking the derivative with respect to β and setting it to zero we have

$$\mathbb{E}[(S_\ell - \alpha\beta S_\ell - \beta Z_\ell)(\alpha S_\ell + Z_\ell)] = 0,$$

then

$$\alpha(1-\alpha\beta)\mathbb{E}S_\ell^2 - \beta\mathbb{E}Z_\ell^2 = 0,$$

so we have $\alpha(1-\alpha\beta)\frac{P}{\alpha^2} = \beta N$. Finally, we can write

$$\beta = \frac{\sqrt{PQ}}{P+N}.$$

(d) Here we have $\rho(x) = x^2$ and $x = f(s) = \alpha' s$ and we want to check that the condition (ii) is satisfied. We have

$$P_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi N}} \exp\left\{-\frac{(y-x)^2}{2N}\right\}.$$

Moreover we have $Y = X + Z = \alpha' S + Z$ so $Y \sim \mathcal{N}(0, \underbrace{\alpha'^2 Q + N}_{\sigma_Y^2})$ then we can write

$$P_Y(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} \exp\left\{-\frac{y^2}{\sigma_Y^2}\right\}.$$

Now let us find $D(P_{Y|X}||P_Y)$. We can write

$$\begin{aligned}
D(P_{Y|X}||P_Y) &= \int P_{Y|X} \log\frac{P_{Y|X}}{P_Y} d_y \\
&= \int \frac{1}{\sqrt{2\pi N}} \exp\left[-\frac{(y-x)^2}{2N}\right] \log\left(\frac{\sigma_Y}{\sqrt{N}} e^{-\frac{(y-x)^2}{2N} + \frac{y^2}{2\sigma_Y^2}}\right) d_y \\
&= \log\sqrt{\frac{\sigma_Y^2}{N}} + \left(\frac{N}{2\sigma_Y^2} - 1/2\right)\log e + \frac{\log e}{2\sigma_Y^2} x^2,
\end{aligned}$$

so choosing $a = \frac{2\sigma_Y^2}{\log e}$ and $b = -\left(\log\sqrt{\frac{\sigma_Y^2}{N}} + (\frac{N}{2\sigma_Y^2} - 1/2)\log e\right)$ we would obtain $\rho(x) = aD(\cdot||\cdot) + b$.

Problem 4

(a) Let us write

$$\begin{aligned}
I(X;\hat{X}|Y) &= h(X|Y) - h(X|\hat{X}, Y) \\
&= h(X|Y) - h(X - \hat{X}|\hat{X}, Y) \\
&\geq h(X|Y) - h(X - \hat{X}|Y) \\
&\geq h(X|Y) - h(X - \hat{X}) \\
&\geq h(X|Y) - 1/2\log(2\pi e D).
\end{aligned}$$

Now for $h(X|Y)$ we can write

$$
\begin{aligned}
h(X|Y) &= h(X,Y) - h(Y) \\
&= 1/2 \log \left[(2\pi e)^2 \det \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right] - 1/2 \log \left[2\pi e \times 1 \right] \\
&= 1/2 \log \left[(2\pi e)(1-\rho^2) \right],
\end{aligned}
$$

so

$$
R(D) \geq 1/2 \log \left[(2\pi e)(1-\rho^2) \right] - 1/2 \log 2\pi e D = -1/2 \log \frac{1-\rho^2}{D}.
$$

(b) Consider the test channel depicted in Fig. 1. Let us choose

$$
\hat{X}|Y \sim \mathcal{N}(0, 1-\rho^2-D),
$$

and

$$
Z \sim \mathcal{N}(0, D),
$$

and Z is independent of $\hat{X}$ and Y. Then we have

$$
\tilde{X}|Y \sim \mathcal{N}(0, 1-\rho^2).
$$

Let us choose $X = \tilde{X}$ then we can write

$$
\begin{aligned}
I(X;\hat{X}|Y) &= h(X|Y) - h(X|\hat{X}, Y) \\
&= h(X|Y) - h(Z) \\
&= 1/2 \log \frac{1-\rho^2}{D},
\end{aligned}
$$

so we have shown that the lower bound is achieved.

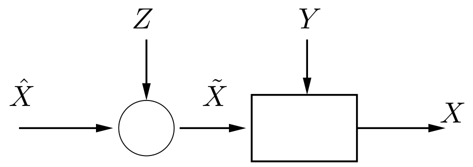

Figure 1: Test channel for the problem 4 part b.

Problem 5

Hamming distortion. X is uniformly distributed on the set $\{1, \cdots, m\}$. The distortion measure is

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

Consider any joint distribution that satisfies the distortion constraint D. Since $D = Pr(X \neq \hat{X})$, we have by Fano's inequality

$$
H(X|\hat{X}) \leq H(D) + D \log(m-1)
$$

and hence

$$I(X;\hat{X}) = H(X) - H(X|\hat{X}) \geq \log m - H(D) - D\log(m-1).$$

We can achieve this lower bound by choosing $p(\hat{x})$to be the uniform distribution, and the conditional distribution of $p(x|\hat{x})$ to be

$$p(x|\hat{x}) = \begin{cases} 1-D & \text{if } x = \hat{x} \\ \frac{D}{m-1} & \text{if } x \neq \hat{x} \end{cases} \tag{2}$$

It is easy to verify that this gives the right distribution on X and satisfies the bound with equality for $D < 1 - \frac{1}{m}$. Hence

$$R(D) = \begin{cases} 1\log(m) - H(D) - D\log(m-1) & \text{if } 0 \leq D \leq 1 - \frac{1}{m} \\ 0 & \text{if } D > 1 - \frac{1}{m} \end{cases} \tag{3}$$