Homework Set #10

Problem 1 (Parallel Channels and Water-filling)

Consider a pair of parallel Gaussian channels:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}, \tag{1}$$

where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right), \tag{2}$$

and there is a power constraint $\mathbb{E}(X_1^2 + X_2^2) \leq 2P$. Assume that $\sigma_1 > \sigma_2$. At what power does the channel stop behaving like a single channel with noise power σ_2^2, and begins behaving like a pair of channels?

Problem 2 (Vector Gaussian channel)

Consider the vector Gaussian noise channel $\mathbf{Y} = \mathbf{X} + \mathbf{Z}$, where $\mathbf{X} = (X_1, X_2, X_3)$, $\mathbf{Z} = (Z_1, Z_2, Z_3)$, and $\mathbf{Y} = (Y_1, Y_2, Y_3)$, $\mathbb{E}\|\mathbf{X}\|^2 \leq P$, and

$$\mathbf{Z} \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\right). \tag{3}$$

Find the capacity. The answer might be surprising.

Problem 3 (Source-channel coding)

Consider the Gaussian channel shown in Fig. 1 with

$$Y = X + Z,$$

where $Z \sim \mathcal{N}(0, N)$ is a Gaussian noise independent of X, and $X \in \mathbb{R}$ and $Y \in \mathbb{R}$ are the input and the output of the channel, respectively. Let $\rho : \mathcal{X} \to \mathbb{R}^+$ be the cost function for the channel input, and define the channel capacity for a given cost P as

$$C(P) = \max_{p(x):\mathbb{E}[\rho(X)] \leq P} I(X;Y).$$

The source in Fig. 2 produces Gaussian symbols $S \sim \mathcal{N}(0, Q)$ with zero mean and variance Q. The encoder maps a sequence of length k of the source symbols to a sequence of length m of the channel input X^m, using the encoding function $X^m = f(S^k)$ and X^m is fed to the channel. The decoder uses the channel output Y^m to estimate the source sequence $\hat{S}^k = g(Y^m)$. The

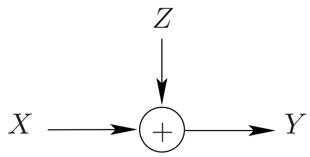

Figure 1: The Gaussian channel

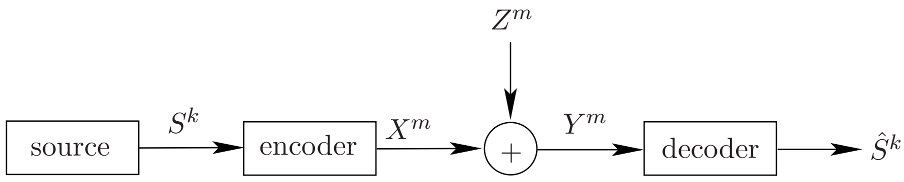

Figure 2: Transmission with distortion

quality of this reconstruction is measured by a distortion function $d(\cdot,\cdot) : \mathcal{S} \times \hat{\mathcal{S}} \to \mathbb{R}^+$. The rate-distortion function for a given distortion D is defined as

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

(a) Prove that $R(D) \leq C(P)$.

(b) Evaluate the functions $R(D)$ and $C(P)$ for $d(s,\hat{s}) = (s - \hat{s})^2$ and $\rho(x) = x^2$. Use the inequality $R(D) \leq C(P)$ to obtain a bound on D in terms of P, Q, and N.

(c) Suppose we use $m = k = 1$ and take the source symbol S_ℓ and scale to obtain the channel input which satisfies the power constraint, *i.e.*, $X_\ell = \alpha S_\ell$, such that $\mathbb{E}[X_\ell^2] = P$. Find the value of α.

Given $Y_\ell = \alpha S_\ell + Z_\ell$, the decoder finds $\hat{S}_\ell = \beta Y_\ell$, such that $\mathbb{E}[(S_\ell - \hat{S}_\ell)^2]$ is minimized. Find the β, and show that $\mathbb{E}[(S_\ell - \hat{S}_\ell)^2] \leq D$.
Hint: Compute $\mathbb{E}[(S_\ell - \hat{S}_\ell)^2]$ in terms of P, Q, N, and β. Take the derivative with respect to β.

The solution of part (b) shows that one can achieve optimal performance through uncoded transmission. In class, we have proved a source-channel separation result which showed that it is optimal to separately encode the source and do channel coding on it. In parts (c), and (d) we show that even when separation holds, one can get very simple scheme if we combine source and channel coding, if the conditions of the *matching theorem* are satisfied.

The matching theorem states that using encoding and decoding functions of block length 1 ($m = k = 1$) is optimal if and only if the following conditions hold.

(i) $I(S;\hat{S}) = I(X;Y)$,

(ii) $\rho(x) = aD(p_{Y|X=x}(y|x) \parallel p_Y(y)) + b$, where a and b are constant, and $D(\cdot \parallel \cdot)$ denotes the Kullback-Leibler divergence,

(iii) $d(s,\hat{s}) = -c \log p_{S|\hat{S}}(s|\hat{s}) + d(s)$, where c is a constant and $d(s)$ is an arbitrary function of s (does not depend on $\hat{s}$).

Using this theorem, we seek conditions to have optimal code block length 1 for this problem.

(d) Let $\rho(x) = x^2$ be the cost function and $x = f(s) = \alpha' s$ be the encoding function. Show that condition (ii) is satisfied.

(e) Let the distortion function be $d(s, \hat{s}) = (s - \hat{s})^2$ and the decoding function be of the form $\hat{s} = g(y) = \beta' y$. For given $\alpha' = \sqrt{P/Q}$ find the value of β' such that condition (iii) is satisfied. Compare β' to β you have found in part (c). How can you explain it?
Hint: You can use the facts that $\hat{S} \sim \mathcal{N}(0, \beta'^2(P+N))$, *i.e.*,

$$p_{\hat{S}}(\hat{s}) = \frac{1}{\sqrt{2\pi(P+N)}\beta'} \exp\left[-\frac{\hat{s}^2}{2\beta'^2(P+N)}\right],$$

and $\hat{S}|s \sim \mathcal{N}(\alpha'\beta' s, \beta'^2 N)$, *i.e*,

$$p_{\hat{S}|S}(\hat{s}|s) = \frac{1}{\sqrt{2\pi N}\beta'} \exp\left[-\frac{(\hat{s} - \alpha'\beta' s)^2}{2\beta'^2 N}\right].$$

Use the Bayes' rule

$$p_{S|\hat{S}}(s|\hat{s}) = \frac{p_{\hat{S}|S}(\hat{s}|s) p_S(s)}{p_{\hat{S}}(\hat{s})}.$$

to find $p_{S|\hat{S}}(s|\hat{s})$.

Problem 4 (Lossy Compression with Side-Information)

Suppose we want to encode an i.i.d. Gaussian source $\{X_i\}$ for a squared error distortion criterion. That is, for a given distortion constraint D and block length n, we want to find $\hat{X}^n = f(X^n)$ such that

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} |X_i - \hat{X}_i|^2\right] \leq D. \tag{4}$$

Now, suppose that $\{Y_i\}$ is also an i.i.d. Gaussian sequence with

$$\mathbb{E}\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \mathbb{E}\left[\begin{bmatrix} X_i \\ Y_i \end{bmatrix} \begin{bmatrix} X_i & Y_i \end{bmatrix}\right] = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

Assume Y^n is given to both encoder and decoder as shown in Fig. 3 and we want to compress X^n given knowledge of Y^n at both the encoder and decoder.

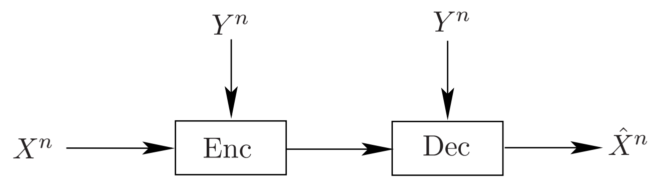

Figure 3: Lossy Gaussian data compression with side information.

(a) One can show that the minimum required rate of compression in order to satisfy the distortion constraint in (4) is given by

$$R(D) = \min_{p(\hat{x}|x,y):\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}|X_i-\hat{X}_i|^2\right]\leq D.} I(X;\hat{X}|Y).$$

Show that $R(D) \geq \frac{1}{2}\log\frac{\sigma^2_{x|y}}{D}$ where $\sigma^2_{x|y} = 1-\rho^2$.

(b) Demonstrate a conditional density $p(\hat{x}|x,y)$ that achieves this lower bound.
Hint: In class we studied the encoding scheme and the optimal density for the case when there was no side-information. There, we derived the optimal $p(\hat{x}|x)$ along with an equivalent "test channel" summarizing this conditional distribution, as $X = \hat{X} + Z$ where all variables are Gaussian and Z and $\hat{X}$ are independent. You can build on this scheme.

Problem 5 (Rate distortion for uniform source with hamming distortion)

Consider a source X uniformly distributed on the set $\{1,\cdots,m\}$. Find the rate distortion function for this source with Hamming distortion; i.e.,

$$d(x,\hat{x}) = \begin{cases} 0 & \text{if } x = \hat{x} \\ 1 & \text{if } x \neq \hat{x} \end{cases} \tag{5}$$