Sample problems for the 11th lecture (May 24)

1. Assume $\rho : \mathcal{Z}^3 \to \hat{\mathcal{Z}}^3$ is the additive Hamming distortion measure, where $\mathcal{Z} = \hat{\mathcal{Z}} = \{0, 1\}$. Suppose Z^3 is uniform distributed over $\mathcal{Z}^3$.

   (a) Let
   $$h(z^3) = \begin{cases} 000, & \text{if } z^3 \in \{000, 001, 010, 011\} \\ 111, & \text{if } z^3 \in \{011, 101, 110, 111\} \end{cases}.$$
   Find the average distortion of this lossy data compressor per source letter.

   (b) Find the mutual information $I(Z^3; \hat{Z}^3)$, where $\hat{Z}^3 = h(Z^3)$.

   Solution.

   (a)
   $$\begin{aligned} \frac{1}{3}E[\rho(Z^3, \hat{Z}^3)] &= \frac{1}{3}E[\rho(Z^3, h(Z^3))] \\ &= \frac{1}{3}\left(\frac{1}{8}\rho(000, 000) + \frac{1}{8}\rho(001, 000) + \frac{1}{8}\rho(010, 000)\right. \\ &\quad +\frac{1}{8}\rho(100, 000) + \frac{1}{8}\rho(011, 111) + \frac{1}{8}\rho(101, 111) \\ &\quad \left. +\frac{1}{8}\rho(110, 111) + \frac{1}{8}\rho(111, 111)\right) \\ &= \frac{1}{3}\left(0 + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + 0\right) \\ &= \frac{1}{4}. \end{aligned}$$

   (b) Noting that $\hat{Z}^3$ is uniformly distributed over $\{000, 111\}$, we obtain
   $$\begin{aligned} I(Z^3; \hat{Z}^3) &= H(\hat{Z}^3) - \underbrace{H(\hat{Z}^3|Z^3)}_{=0} \\ &= H(\hat{Z}^3) \\ &= \frac{1}{2}\log_2\frac{1}{1/2} + \frac{1}{2}\log_2\frac{1}{1/2} \\ &= 1. \end{aligned}$$

2. Suppose $0 \le \delta, x \le 1$, and $n \ge 2$ is a positive integer.

(a) Let $f(x) = (1 - \delta x)^n$ and $g(x) = 1 - x$. Let A satisfy $f(1) = g(1) + A$. Argue that $f(x) \le g(x) + A$ for $0 \le 1 \le x$.

(b) Prove that $(1 - \delta)^n \le e^{-n\delta}$ using the fundamental inequality.

Solution.

(a) Apparently, $A = (1 - \delta)^n$.

Note that $f(x)$ is a convex function over $x \in [0, 1]$ because

$$f''(x) = \delta^2 n(n - 1)(1 - \delta x)^{n-2} > 0$$

and hence the straight line connecting

$$(0, f(0)) = (0, 1) \quad \text{and} \quad (1, f(1)) = (1, (1 - \delta)^n)$$

must be above $f(x)$ but this straight line is below $1-x+(1-\delta)^n$. Hence, $f(x) \le g(x) + (1 - \delta)^n$.

(b) We first note that $(1 - \delta)^n \le e^{-n\delta}$ iff $(1 - \delta) \le e^{-\delta}$. Leting $u = 1-\delta$ yields $u \le e^{-(1-u)}$. Taking logarithm on both sides produces $\ln(u) \le u - 1$, which is exactly the fundamental inequality. In summary, we have

$$\begin{aligned} & (1 - \delta)^n \le e^{-n\delta} \\ \text{iff} \quad & (1 - \delta) \le e^{-\delta} \\ \text{iff} \quad & u \le e^{-(1-u)} \\ \text{iff} \quad & \ln(u) \le u - 1. \end{aligned}$$

3. The *distortion δ-typical set* with respect to the memoryless (product) distribution $P_{Z,\hat{Z}}$ on $\mathcal{Z}^n \times \hat{\mathcal{Z}}^n$ and a bounded additive distortion measure $\rho_n(\cdot, \cdot)$ is defined by

$$\begin{aligned} \mathcal{D}_n(\delta) \; := \; & \Big\{ (z^n, \hat{z}^n) \in \mathcal{Z}^n \times \hat{\mathcal{Z}}^n : \\ & \left| -\frac{1}{n} \log_2 P_{Z^n}(z^n) - H(Z) \right| < \delta, \\ & \left| -\frac{1}{n} \log_2 P_{\hat{Z}^n}(\hat{z}^n) - H(\hat{Z}) \right| < \delta, \\ & \left| -\frac{1}{n} \log_2 P_{Z^n,\hat{Z}^n}(z^n, \hat{z}^n) - H(Z, \hat{Z}) \right| < \delta, \\ & \text{and} \left| \frac{1}{n} \rho_n(z^n, \hat{z}^n) - E[\rho(Z, \hat{Z})] \right| < \delta \Big\}. \end{aligned}$$

(a) Rewrite the distortion typical set in the form of the normalized sum of n quantities.

(b) A researcher provides an alternative distortion typical set as follows.

$$\begin{aligned}\mathcal{D}_n^{(I)}(\delta) \;:=\; & \Big\{(z^n, \hat{z}^n) \in \mathcal{Z}^n \times \hat{\mathcal{Z}}^n : \\ & \left|\frac{1}{n}\log_2 \frac{P_{Z^n,\hat{Z}^n}(z^n,\hat{z}^n)}{P_{Z^n}(z^n)P_{\hat{Z}^n}(\hat{z}^n)} - I(Z;\hat{Z})\right| < \delta, \\ & \text{and } \left|\frac{1}{n}\rho_n(z^n,\hat{z}^n) - E[\rho(Z,\hat{Z})]\right| < \delta\Big\}.\end{aligned}$$

Does the alternative distortion typical set also satisfy Theorem 6.18 (by replacing 3δ with δ for Property 2)? Justify your answer.

Solution.

(a)

$$\begin{aligned}\mathcal{D}_n(\delta) := \Big\{ & (z^n, \hat{z}^n) \in \mathcal{Z}^n \times \hat{\mathcal{Z}}^n : \\ & \left|\frac{\log_2 \frac{1}{P_Z(z_1)} + \log_2 \frac{1}{P_Z(z_2)} + \cdots + \log_2 \frac{1}{P_Z(z_n)}}{n} - H(Z)\right| < \delta, \\ & \left|\frac{\log_2 \frac{1}{P_{\hat{Z}}(\hat{z}_1)} + \log_2 \frac{1}{P_{\hat{Z}}(\hat{z}_2)} + \cdots + \log_2 \frac{1}{P_{\hat{Z}}(\hat{z}_n)}}{n} - H(\hat{Z})\right| < \delta, \\ & \left|\frac{\log_2 \frac{1}{P_{Z,\hat{Z}}(z_1,\hat{z}_1)} + \log_2 \frac{1}{P_{Z,\hat{Z}}(z_2,\hat{z}_2)} + \cdots + \log_2 \frac{1}{P_{Z,\hat{Z}}(z_n,\hat{z}_n)}}{n} - H(Z,\hat{Z})\right| < \delta, \\ & \text{and } \left|\frac{\rho(z_1,\hat{z}_1) + \rho(z_2,\hat{z}_2) + \cdots + \rho(z_n,\hat{z}_n)}{n} - E[\rho(Z,\hat{Z})]\right| < \delta\Big\}\end{aligned}$$

(b) We can rewrite $\mathcal{D}_n^{(I)}(\delta)$ as

$$\begin{aligned}\mathcal{D}_n^{(I)}(\delta) := \Big\{ & (z^n, \hat{z}^n) \in \mathcal{Z}^n \times \hat{\mathcal{Z}}^n : \\ & \left|\frac{\log_2 \frac{P_{Z,\hat{Z}}(z_1,\hat{z}_1)}{P_Z(z_1)P_{\hat{Z}}(\hat{z}_1)} + \cdots + \log_2 \frac{P_{Z,\hat{Z}}(z_n,\hat{z}_n)}{P_Z(z_n)P_{\hat{Z}}(\hat{z}_n)}}{n} - I(Z;\hat{Z})\right| < \delta, \\ & \text{and } \left|\frac{\rho(z_1,\hat{z}_1) + \rho(z_2,\hat{z}_2) + \cdots + \rho(z_n,\hat{z}_n)}{n} - E[\rho(Z,\hat{Z})]\right| < \delta\Big\}.\end{aligned}$$

Thus, by the memorylessness of $\left\{\log_2 \frac{P_{Z,\hat{Z}}(Z_i,\hat{Z}_i)}{P_Z(Z_i)P_{\hat{Z}}(\hat{Z}_i)}\right\}_{i=1}^n$ and $\left\{\rho(Z_i,\hat{Z}_i)\right\}_{i=1}^n$ as well as the boundedness of distortion measure $\rho(\cdot,\cdot)$, the law of large numbers hold, i.e.,

$$\frac{1}{n}\log_2 \frac{P_{Z^n,\hat{Z}^n}(Z^n,\hat{Z}^n)}{P_{Z^n}(Z^n)P_{\hat{Z}^n}(\hat{Z}^n)} \to I(Z;\hat{Z}) \text{ in probability}$$

and

$$\frac{1}{n}\rho_n(Z^n,\hat{Z}^n) \to E[\rho(Z,\hat{Z})] \text{ in probability.}$$

Property 1 of Theorem 6.18 accordingly holds.

For Property 2, we observe

$$I(Z;\hat{Z}) - \delta \le \frac{1}{n}\log_2 \frac{P_{Z^n,\hat{Z}^n}(z^n,\hat{z}^n)}{P_{Z^n}(z^n)P_{\hat{Z}^n}(\hat{z}^n)} = \frac{1}{n}\log_2 \frac{P_{\hat{Z}^n|Z^n}(\hat{z}^n|z^n)}{P_{\hat{Z}^n}(\hat{z}^n)} \le I(Z;\hat{Z}) + \delta$$

which implies

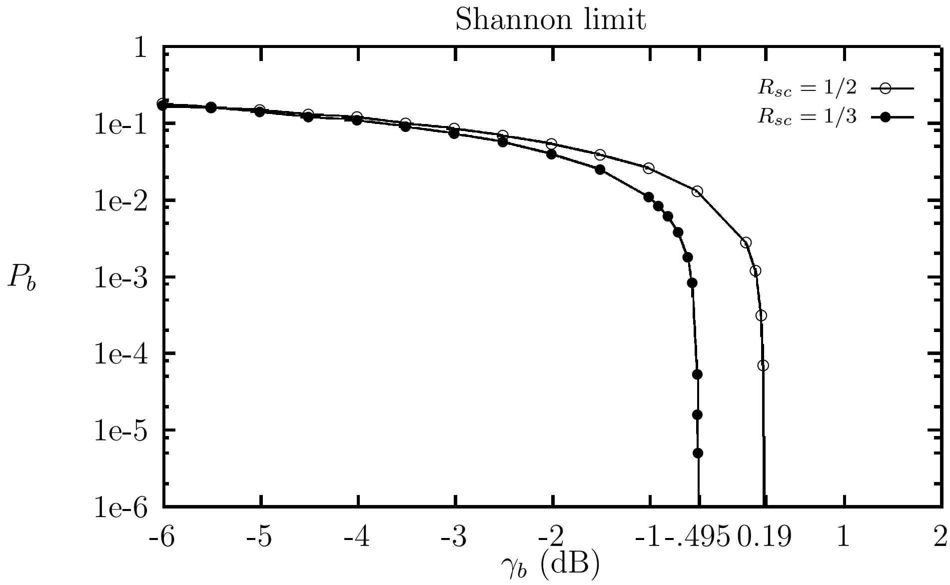
$$P_{\hat{Z}^n}(\hat{z}^n) \ge P_{\hat{Z}^n|Z^n}(\hat{z}^n|z^n)2^{-n[I(Z;\hat{Z})+\delta]}.$$

4. (a) Consider to transmit binary memoryless source $\{Z_i\}_{i=1}^n$ with uniform marginal distribution over the BEC with erasure probability $\epsilon = \frac{1}{2}$. Suppose the bit error rate (BER) is concerned. Let the source be generated at a rate of $1/T_s$ symbol/second, and the channel is used at a rate of $1/T_c$ channel-usage/second. According to lossy joint source-channel coding theorem, please give the condition on T_s and T_c such that the source can be transmitted with BER less than 0.1.

   Hint: For Hamming distortion measure, the average distortion is exactly the bit error rate, i.e., $D = \text{BER}$, and the rate distortion function is given by $R(D) = 1 - h_b(D)$ bits/source symbol, where $h_b(D) = D\log_2(D) + (1-D)\log_2(1-D)$.

   (b) Suppose a binary memoryless sequence is transmitted over the BPSK-input AWGN channel. A design is targeted to transmit the binary memoryless source at code rate 1/2 with bit error rate (BER) 10^{-5} under $\gamma_b = E_b/N_0 = -3$ dB, but fails. Can we complete the design by relaxing the BER requirement to 10^{-3} but still operate under $\gamma_b = E_b/N_0 = -3$ dB? Justify your answer.

   Hint: See the figure of Shannon limit below.



The Shannon limits for $(2, 1)$ and $(3, 1)$ codes under binary-input AWGN channel.

Solution.

(a) The condition is

$$[1 - h_b(D)] \frac{\text{bits}}{\text{source symbol}} \times \frac{1}{T_s} \frac{\text{source symbol}}{\text{second}}$$

$$< (1 - \epsilon) \frac{\text{bits}}{\text{channel usage}} \times \frac{1}{T_c} \frac{\text{channel usage}}{\text{second}}.$$

In other words,

$$\frac{T_s}{T_c} > 2(1 - h_b(0.1)).$$

(b) According to the lossy joint source-channel coding theorem (converse part), for any sequence of $(m = 1)$-to-$(n_m = 2)$ lossy source-channel codes $(f^{(sc)}, g^{(sc)})$ satisfying the average distortion fidelity criterion BER $\leq 10^{-3}$, the channel must have

$$\gamma_b \geq 0.18 \text{ dB}$$

as shown in the figure of Shannon limit. Thus, it would be theoretically infeasible to have such a system spec.