Exam

TA: Heng-Chien Liou

Solution:

a) $Z = X + Y$, where $X \sim f, Y \sim g$ and $X \perp Y$. Verify that $f * g \geq 0$ and $\int f * g = 1$

b) Following the warm-up, let $X \sim f, Y \sim g$ independently and $Z = X + Y$

(1) always hold. $\mathrm{h}(Z) \geq \mathrm{h}(Z|Y) = \mathrm{h}(X+Y|Y) = \mathrm{h}(X)$. Similarly, $\mathrm{h}(Z) \geq \mathrm{h}(Y)$.

(2) doesn't always hold. For example, take $X, Y \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, \sigma^2)$, then $Z \sim \mathrm{N}(0, 2 \times \sigma^2)$. And $\mathrm{h}(X) = \mathrm{h}(Y) = \frac{1}{2}\log(2\pi e\sigma^2)$, while $\mathrm{h}(Z) = \frac{1}{2}\log(2\pi e\sigma^2) + \log\sqrt{2}$. Then for $\pi e\sigma^2 < 1$, $\mathrm{h}(X) + \mathrm{h}(Y) < \mathrm{h}(Z)$

c) Similarly, let $X \sim f, Y \sim g, Z \sim h$ and $Z \perp (X, Y)$, then by the data processing inequality

$$\mathrm{D}(f \| g) = \mathrm{D}(X \| Y) \geq \mathrm{D}(X + Z \| Y + Z) = \mathrm{D}(f * h \| g * h)$$

Solution:

a) For a stationary ergodic Bernoulli-q source, by Theorem 8 of Lecture 1, $\mathcal{H}(\{S_i\}) = \widetilde{\mathcal{H}}(\{S_i\})$, hence it suffices to consider $\widetilde{\mathcal{H}}(\{S_i\}) = \lim_{n\to\infty} \mathrm{H}(S_n|S^{n-1})$. And note that $\forall n, \mathrm{H}(S_n|S^{n-1}) \leq \mathrm{H}(S_n)$ by the property of conditioning reduce entropy. Hence $\lim_{n\to\infty} \mathrm{H}(S_n|S^{n-1}) \leq \lim_{n\to\infty} \mathrm{H}(S_n) = \mathrm{H}(\mathrm{Ber}(q))$

$S_i \overset{\text{i.i.d.}}{\sim} \mathrm{Ber}(q)$ attains the maximum.

Considering the result in L1, p.77, solving $\left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right) = (q, 1-q)$ yields $\beta = \frac{1-q}{q}\alpha$.

$$\mathcal{H}(\{S_i\}) = (1-q)\mathsf{H_b}(\alpha) + q\mathsf{H_b}\left(\frac{1-q}{q}\alpha\right)$$

Extend the source-channel separation theorem to source with memory, if R is achievable

$$\mathrm{R} < \frac{\mathrm{C}(\mathsf{P}_{Y|X})}{\mathcal{H}(\{S_i\})} = \frac{1 - \mathsf{H_b}(p)}{(1-q)\mathsf{H_b}(\alpha) + q\mathsf{H_b}\left(\frac{1-q}{q}\alpha\right)}$$

Then R is achievable.

b) Extend the source-channel separation theorem to lossy source coding and channel coding with input power constraint, for fixed D, if R is achievable, then

$$\mathrm{R} \le \frac{\mathrm{C}(\mathsf{B})}{\mathrm{R}_s(\mathsf{D})}$$

And note that $\mathsf{D}_{\max} = \sigma_S^2$ and $\mathsf{D}_{\min} = 0$. Therefore, if a pair (R, D) is achievable, we can ensure that

$$\begin{aligned}
\mathrm{R} \le \frac{\mathrm{C}(\mathsf{B})}{\mathrm{R}_s(\mathsf{D})} &= \frac{\frac{1}{2}\log\left(1+\frac{\mathsf{B}}{\sigma_Z^2}\right)}{\frac{1}{2}\log\frac{\sigma_S^2}{\mathsf{D}}} \\
\mathrm{R}\log\mathsf{D} &\ge \mathrm{R}\log\sigma_S^2 - \log\left(1+\frac{\mathsf{B}}{\sigma_Z^2}\right) \\
\log\mathsf{D} &\ge \log\sigma_S^2 - \frac{\log\left(1+\frac{\mathsf{B}}{\sigma_Z^2}\right)}{\mathrm{R}} \\
\mathsf{D} &\ge \frac{\sigma_s^2}{\left(1+\frac{\mathsf{B}}{\sigma_Z^2}\right)^{\frac{1}{\mathrm{R}}}}
\end{aligned}$$

We can also make the observation that $\mathsf{D}_{\min} < \sigma_s^2 / \left(1+\frac{\mathsf{B}}{\sigma_Z^2}\right)^{\frac{1}{\mathrm{R}}} < \mathsf{D}_{\max}, \forall \mathrm{R} > 0$

Solution:

a) Observe that permutation channel is a special form of sum channel from Homework 4

$$\begin{aligned}
\mathrm{C} &= \max_{\mathsf{P}_{\boldsymbol{X}}}\{\mathrm{I}(\boldsymbol{X};\boldsymbol{Y})\} \\
&= \max_{\mathsf{P}_{\boldsymbol{X},L}}\{\mathrm{I}(\boldsymbol{X};\boldsymbol{Y},L)\} \\
&= \max_{\mathsf{P}_{\boldsymbol{X},L}}\{\mathrm{I}(\boldsymbol{X};L)+\mathrm{I}(\boldsymbol{X};\boldsymbol{Y}|L)\} \\
&= \max_{\mathsf{P}_{\boldsymbol{X},L}}\{\mathrm{H}(L)+\mathrm{I}(\boldsymbol{X};\boldsymbol{Y}|L)\} \\
&= \max_{\mathsf{P}_L}\left\{\mathrm{H}(L)+\max_{\mathsf{P}_{\boldsymbol{X}|L}}\mathrm{I}(\boldsymbol{X};\boldsymbol{Y}|L)\right\}
\end{aligned}$$

And

$$\max_{\mathsf{P}_{\boldsymbol{X}|L=l}} \mathrm{I}(\boldsymbol{X};\boldsymbol{Y}|L=l) = 0$$

Any P_X such that $L \sim \mathrm{Unif}\{0,\ldots,d\}$ is capacity achieving.

b)

$$\begin{aligned}
\mathsf{C}(\mathsf{B}) &= \max_{\mathsf{P}_{\boldsymbol{X}}:\mathsf{E}[b(\boldsymbol{X})]\le \mathsf{B}} \mathrm{I}(\boldsymbol{X};\boldsymbol{Y}) \\
&= \max_{\mathsf{P}_L:\mathsf{E}[L]\le \mathsf{B}} \left\{ \mathrm{H}(L) + \max_{\mathsf{P}_{\boldsymbol{X}|L}} \mathrm{I}(\boldsymbol{X};\boldsymbol{Y}|L) \right\} \\
&= \max_{\mathsf{P}_L:\mathsf{E}[L]\le \mathsf{B}} \mathrm{H}(L) \\
&= \max_{0\le\mu\le \mathsf{B}} \left\{ \max_{\mathsf{P}_L:\mathsf{E}[L]=\mu} \mathrm{H}(L) \right\}
\end{aligned}$$

where we reformulate it into a two-step optimization problem. First we need to solve

$$\begin{aligned}
\text{maximize} \quad & \mathrm{H}(L) \\
\text{subject to} \quad & \mathsf{E}[L] = \mu
\end{aligned}$$

Denote $\mathsf{P}_L(i) = p_i, \forall i \in \{0, ..., d\}$, and consider any other sequence $\{q_i\}_{i=0}^d$ where $\sum_{i=0}^d q_i = 1$ and $\sum_{i=0}^d iq_i = \mu$, we have

$$\begin{aligned}
0 \le \mathrm{D}(p\|q) = \sum_{i=0}^d p_i \log \frac{p_i}{q_i} = -\mathrm{H}(L) - \sum_{i=0}^d p_i \log q_i \\
\mathrm{H}(L) \le -\sum_{i=0}^d p_i \log q_i
\end{aligned}$$

Inspired by HW2, Problem 1, now we limit $\{q_i\}_{i=0}^d$ to the form that $\log q_i = -\alpha i - \beta$ then

$$\mathrm{H}(L) \le -\sum_{i=0}^d p_i \log q_i = \sum_{i=0}^d p_i (\alpha i + \beta) = \alpha\mu + \beta$$

where α, β satisfy $\sum_{i=0}^d 2^{-\alpha i - \beta} = 1$ and $\sum_{i=0}^d i 2^{-\alpha i - \beta} = \mu$, hence we can consider them as function of μ. Let $f(\mu) = \alpha\mu + \beta$, the problem then becomes

$$\mathsf{C}(\mathsf{B}) = \max_{0\le\mu\le \mathsf{B}} f(\mu)$$

Now we want to show that $f(\mu)$ is monotonically increasing for $\mu \le d/2$ and monotonically decreasing for $\mu > d/2$

$$\frac{\mathrm{d}}{\mathrm{d}\mu} f(\mu) = \mu \frac{\mathrm{d}}{\mathrm{d}\mu}\alpha + \alpha + \frac{\mathrm{d}}{\mathrm{d}\mu}\beta = \alpha$$

where

$$0 = \sum_{i=0}^d 2^{-\alpha i - \beta} \left(-i \frac{\mathrm{d}}{\mathrm{d}\mu}\alpha - \frac{\mathrm{d}}{\mathrm{d}\mu}\beta \right) \ln 2 = \left(\mu \frac{\mathrm{d}}{\mathrm{d}\mu}\alpha + \frac{\mathrm{d}}{\mathrm{d}\mu}\beta \right) \ln 2$$

And $\alpha \geq 0$ iff $\mu \leq d/2$.

If $\mathsf{B} \geq d/2$, the capacity achieving input distribution is $L \sim \text{Unif}\{0, ..., d\}$, and $\mathrm{C}(\mathsf{B}) = \log(d+1)$.

Else, the capacity achieving input distribution is the form of $\mathsf{P}_L(i) = 2^{-\alpha i - \beta}$, where (α, β) is determined so that $\sum_{i=0}^{d} \mathsf{P}_L(i) = 1$ and $\sum_{i=0}^{d} i\mathsf{P}_L(i) = \mathsf{B}$. And $\mathrm{C}(\mathsf{B}) = \alpha\mathsf{B} + \beta$.

c) We can use the result of sum channel from Homework 4

$$\mathrm{C} = \log \sum_{l=0}^{d} 2^{\mathrm{C}_l}$$

$$\mathrm{C}_l = \log \binom{d}{l} + \left(1 - \alpha + \frac{\alpha}{\binom{d}{l}}\right) \log \left(1 - \alpha + \frac{\alpha}{\binom{d}{l}}\right) + \left(\binom{d}{l} - 1\right) \frac{\alpha}{\binom{d}{l}} \log \frac{\alpha}{\binom{d}{l}}$$

$$\mathsf{P}_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{2^{\mathrm{C}_{\|\boldsymbol{x}\|_1}}}{\sum_{i=0}^{d} 2^{\mathrm{C}_i}} \frac{1}{\binom{d}{\|\boldsymbol{x}\|_1}}.$$

Solution:

a)

$$\begin{aligned}
\mathsf{D}_{\text{alt}}^{(n)} &= \mathsf{E}_{(X^n, \hat{S}^n)} \left[\tilde{d}(X^n, \hat{S}^n)\right] = \mathsf{E}_{(X^n, \hat{S}^n)} \left[\frac{1}{n} \sum_{i=1}^{n} \tilde{d}(X_i, \hat{S}_i)\right] \\
&= \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}_{(X^n, \hat{S}^n)} \left[\mathsf{E}_{S_i \sim \mathsf{P}_{S|X}} \left[d(S_i, \hat{S}_i) \middle| X_i\right]\right] \\
&= \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}_{(S^n, \hat{S}^n)} \left[d(S_i, \hat{S}_i)\right] = \mathsf{E}_{(S^n, \hat{S}^n)} \left[d(S^n, \hat{S}^n)\right] \\
&= \mathsf{D}_{\text{remote}}^{(n)}
\end{aligned}$$

A pair (R, D) is achievable in the alternative problem if exist a sequence of $(n, \lfloor n\mathrm{R} \rfloor)$ codes such that $\limsup_{n\to\infty} \mathsf{D}_{\text{alt}}^{(n)} \leq \mathsf{D}$. By the above derivation we know that $\mathsf{D}_{\text{alt}}^{(n)} = \mathsf{D}_{\text{remote}}^{(n)}$, hence this sequence of code also have $\limsup_{n\to\infty} \mathsf{D}_{\text{remote}}^{(n)} \leq \mathsf{D}$, (R, D) is achievable in the remote problem. Same argument for the other directions.

b)

$$\begin{aligned}
\mathrm{R}(\mathsf{D}) &= \min_{\mathsf{P}_{\hat{S}|X} : \mathsf{E}\left[d(S, \hat{S})\right] \leq \mathsf{D}} \mathrm{I}\left(X; \hat{S}\right) \\
\mathsf{D}_{\max} &= \min_{\hat{s} \in \hat{\mathcal{S}}} \mathsf{E}\left[d(S, \hat{s})\right] \\
\mathsf{D}_{\min} &= \min_{\hat{s} : \mathcal{X} \to \hat{\mathcal{S}}} \mathsf{E}\left[d(S, \hat{s}(X))\right]
\end{aligned}$$