Boulder Theoretical Biophysics 2019

Neuroscience Mini-course: Exercise Set 3 **solutions**

1. Data processing inequality: This question based on problem 8.9 from MacKay. There, they use the mutual information chain rule, which is, for any ensemble XYZ,

$$I(X;Y,Z) = I(X;Y) + I(X;Z|Y)$$

. We will not appeal to that directly here, but it may be of general interest.

We want to show that

$$I(S;R) \leq I(S;T)$$

given that

$$P(s,t,r) = P(s)P(t|s)P(r|t)$$

So,

$$\begin{aligned} H(S|R,T) &= H(S|T) \\ H(S|R) &\geq H(S|T) \\ H(S) - H(S|R) &\leq H(S) - H(S|T) \\ I(S;R) &\leq I(S;T) \end{aligned}$$

From lecture: Capacity of a binary symmetric channel. Given some binary source distribution ($p(x)$) and bit-flip probability (f), we want to find the capacity of a binary symmetric channel. This is taken from problem 9.2 in MacKay.

That is, we want to find:

$$\begin{aligned} C &= \max_{p(x)} I(X;Y) \\ &= \max_{p(x)} H(Y) - H(Y|X) \end{aligned}$$

and we can note now that $H(Y)$ is maximized by $p(y) = 1/2$ and that, from the symmetry of the channel, this can be achieved by setting $p(x) = 1/2$.

However, we can also observe that,

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H_2(p_0 f + p_1(1-f)) - H_2(f) \end{aligned}$$

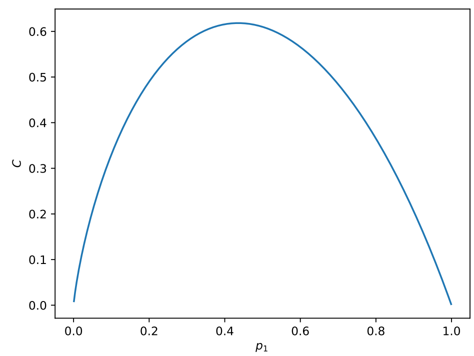

Figure 1: The capacity of a Z-channel with $f = .2$ as a function of p_1.

and we know that $H_2(a)$ is maximal for $a = 1/2$. We can choose $p_0 = 1/2$ to ensure that the argument to the first H_2 will always be $1/2$. So,

$$\begin{aligned} &= H_2\left(\frac{1}{2}f + \frac{1}{2}(1-f)\right) - H_2(f) \\ &= H_2(1/2) - H_2(f) \\ C = \max_{p(x)} I(X;Y) &= 1 - H_2(f) \end{aligned}$$

2. The Z channel. This is similar to exercise 9.15 in MacKay.

We want to find, for $f = .2$,

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H_2(p_1(1-f)) - p_1 H_2(f) \end{aligned}$$

Now, we can maximize this expression with respect to p_1 using the fact that

$$H_2' = -\log_2 \frac{p}{1-p}$$

$$\begin{aligned} 0 = \frac{\partial I}{\partial p_1} &= -(1-f)\log_2 \frac{p_1(1-f)}{1-p_1(1-f)} - H_2(f) \\ H_2(f) &= (1-f)\log_2 \frac{1-p_1(1-f)}{p_1(1-f)} \\ 2^{\frac{H_2(f)}{1-f}} &= \frac{1-p_1(1-f)}{p_1(1-f)} \\ 1 + 2^{\frac{H_2(f)}{1-f}} &= \frac{1}{p_1(1-f)} \\ p_1 &= \frac{1}{1-f}\left[1 + 2^{\frac{H_2(f)}{1-f}}\right]^{-1} \end{aligned}$$

and with $f = .2$, we can see that $p_1^* = .436$ and $C = .618$.

The Z channel continued. This is similar to exercise 9.15 in MacKay.

There is no noise for the 0 symbol, and there is noise for the 1. Thus, $p_1 < p_0$ because while we are sacrificing some source entropy we are increasing our overall transmission reliability (we are injecting less noise entropy).

One must take the limit of the expression for $p_1^*(f)$ as f approaches 1. Using L'Hospital's rule, one can show that this becomes $\frac{1}{e}$.

3. The Gaussian channel. This is similar to exercise 11.5 in MacKay. As shown previously in the class, the capacity of a Gaussian channel (given that the variance of the source is constrained to be σ_s^2) is

$$C = \frac{1}{2} \log \left(1 + \frac{\sigma_s^2}{\sigma_n^2} \right)$$

4. If the input is binary, the capacity of the channel will be achieved by using both symbols with equal probability. Then,

$$\begin{aligned} C' = I(X;Y) &= H(Y) - H(Y|X) \\ &= -\int_{-\infty}^{\infty} dy Q(y) \log Q(y) + \int_{-\infty}^{\infty} dy N(y; 0, \sigma_n) \log N(y; 0, \sigma_n) \\ &= -\int_{-\infty}^{\infty} dy Q(y) \log Q(y) + \frac{1}{2} \ln \left(2\pi e \sigma_n^2 \right) \end{aligned}$$

where

$$Q(y) = \frac{1}{2} \left[N(y; -\sigma_s, \sigma_n) + N(y; \sigma_s, \sigma_n) \right]$$

and

$$N(y; x, s) = \frac{1}{\sqrt{2s^2\pi}} \exp \left[(y-x)^2 / 2s^2 \right]$$

5. If the output is thresholded, then the channel becomes equivalent to a binary symmetric channel with a transition probability determined by the level of noise. We can write this using the error function,

$$\phi(z) = \int_{-\infty}^{z} dz \frac{1}{\sqrt{2\pi}} \exp \left[-z^2/2 \right]$$

So, now we have a binary symmetric channel with transition probability $f = \phi(\sigma_s/\sigma_n)$ and

$$C'' = 1 - H_2(f)$$

6. See Figure 2.

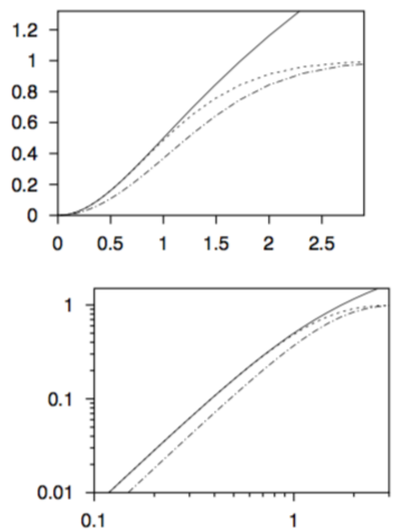

Figure 11.9. Capacities (from top to bottom in each graph) C, C', and C'', versus the signal-to-noise ratio ($\sqrt{v}/\sigma$). The lower graph is a log–log plot.

Figure 2: Taken from MacKay.