Small Project 1

Due: Thursday, February 12

In this project, we will compute the capacity of the meta-channel obtained by concatenating a repetition encoder with the Binary Symmetric Channel with crossover probability p (BSC_p). That is, let R_k denote the encoder that merely repeats its input k times:

$$R_3(\texttt{01}) = \texttt{000111}.$$

Let $C_{k,p}$ denote the meta-channel obtained by passing data first through R_k and then through BSC_p:

$$C_{k,p}(x) = BSC_p(R_k(w)).$$

In this project, we will experimentally estimate the capacity of this channel for varying p and k.

Before I describe how we will do this, I should point out that it is possible to exactly compute the capacity of this channel, and that this computation can be performed quickly. We will perform this computation experimentally for three reasons:

- Because it is very rare that one can compute these capacities exactly. In fact, this is almost the only example in which one can.

- Because it is a good idea to test our methodology on an example for which we know the answer before we apply it to other situations.

- To give you practice at setting up computer experiments.

To estimate the capacity of $C_{k,p}$, you will repeatedly:

1. generate a random $w \in 0, 1$.

2. encode x (*i.e.*, compute $x = R_k(w)$).

3. obtain the output of passing x through the channel (*i.e.*, compute $y = BSC_p(x)$).

4. compute $\Pr[y]$ and $\Pr[y|w]$, and use these to compute

$$i(w; y) \stackrel{\text{def}}{=} \log_2\left(\frac{\Pr[w \text{ and } y]}{\Pr[w]\Pr[y]}\right) = \log_2\left(\frac{\Pr[y|w]}{\Pr[y]}\right)$$

Finally, we will estimate the capacity of $C_{k,p}$ to be the average of all the $i(w; y)$ values computed.

For each k and p for which you perform the experiment, you will have to decide how many times to sample before you consider your estimate to be good. In a later lecture, we will find out how many samples we actually need. For now, you could just do enough samples to make your plots look reasonable (see below), or as many as you can in the time you have.

You should:

- Perform this experiment for $k = 2, 3, 5, 8, 10, 20, 40, 80$.

- For each k, you should vary p between 0 and $1/2$. The region near $1/2$ will be much more interesting, so be sure to include at least the points $p = 1/2 - 2^{-i}$ for $i = 2, \ldots, 8$.

- Produce plots in which the y-axis is the capacity of the meta-channel, the x-axis is the capacity of the BSC_p, and a separate curve appears for each value of k.

- Produce the same plots, but with the y and x axes being the *logarithms* of this data. Even better, try to produce these with logarithmic axes. If you are using matlab, you can do this by just replacing statement your `plot` from the previous example with `loglog`. (see the discussion in "Plotting Data" under "Programming Tips"))

- Submit your plots and code.

- If you can figure out how to compute these capacities exactly, then compute these and put them on the plot as well.

- Record for my reference: how long did this take you? How was your time spent?

Collaboration: You are free to discuss how do this project with others in the class (and especially to get technical help). But, you should write your own code. Also, list those you have collaborated with.

Note: If computation time is a barrier, just do as much as you can.