ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 22 Information Theory and Coding

Graded Homework, due Friday December 14th, noon December 3rd, 2012

- You can work in teams of up to 4 people. You should submit one report per team.
- Your work must be submitted electronically to marc.desgroseilliers@epfl.ch. The report should be typed in Latex. Attach suitably commented source code (we should be able to run your code).
- Your code should be written in Matlab, C, C++ or Python. If you want to use another language, discuss with a teaching assistant first.
- This week's exercise session will be devoted to this homework.
- No late submissions will be accepted.

PROBLEM 1. For a given value $0 \le z_0 \le 1$, define the following random process:

$$Z_0 = z_0, \quad Z_{i+1} = \begin{cases} Z_i^2 & \text{with probability } 1/2 \\ 2Z_i - Z_i^2 & \text{with probability } 1/2 \end{cases} \qquad i \ge 0,$$

with the sequence of random choices made independently. Observe that the Z process keeps track of the polarization of a Binary Erasure Channel with erasure probability z_0 as it is tranformed by the polar transform: $\Pr(Z_i = z)$ is exactly the fraction of 2^i Binary Erasure Channels the polar transform synthesizes at the ith level that have erasure probability z. In the class last week, by a numerical experiment we observed that the fraction of unpolarized channels seem to vanish as we keep applying the polar transform. In other words, for any $\delta > 0$, $\Pr\big(Z_i \in (\delta, 1-\delta)\big) \to 0$ as i gets large. The aim of this problem is to prove this rigorously.

(a) Define $Q_i = \sqrt{Z_i(1-Z_i)}$. Find $f_1(z)$ and $f_2(z)$ so that

$$Q_{i+1} = Q_i \times \begin{cases} f_1(Z_i) & \text{with probability } 1/2 \\ f_2(Z_i) & \text{with probability } 1/2 \end{cases}$$

(b) Show that $f_1(z) + f_2(z) \le \sqrt{3}$. Based on this, find a $\rho < 1$ so that

$$E\big[Q_{i+1} \bigm| Q_0, \ldots, Q_i\big] \le \rho Q_i.$$

(c) Show that, for the ρ you found in (b), $E[Q_i] \le \frac{1}{2}\rho^i$.

(d) Show that

$$\Pr\big(Z_i \in (\delta, 1-\delta)\big) = \Pr\big(Q_i > \sqrt{\delta(1-\delta)}\big) \le \frac{\rho^i}{2\sqrt{\delta(1-\delta)}}$$

PROBLEM 2. In this question, you will be asked to evaluate different schemes to pick the channels on which to communicate when using polar coding.

First, recall that after one step of the polar transform, a binary erasure channel W with erasure probability ϵ is turned into two channels which are again BECs: W^+ is a BEC(ϵ^2) and W^- is a BEC($2\epsilon - \epsilon^2$).

(a) Consider sending 93 bits of information in a 256 bit codeword (hence the polar transform is applied 8 times) over a BEC(0.5). One way to pick over which channels to transmit information would be to take the channels indexed with the most + signs. Use your favorite programming language to identify the erasure probability associated with these channels.

(b) Redo (a) by picking the 93 channels that have the smallest erasure probabilities. List their indices and their associated erasure probabilities (starting from 0). What is the threshold, i.e. what are the probabilities associated with the 93rd and 94th channel.

(c) What is the sum of the erasure probabities of the channels that are picked by method in (a)? What is this sum for method (b)? Which indices are picked by both methods?

PROBLEM 3. The method used in part (b) of the previous problem chooses the channels with the lowest error (erasure) probabilities and utilizes them for data transmission. That the erasure probabilities of the polarized channels can be efficiently computed when we start from a BEC made the method computationally feasible. However, if we start with a channel that is *not* a BEC, no efficient procedure exists to find the exact error probabilities of the polarized channels. In this problem we will explore a way to estimate these error probabilities.

Recall that the channels W^- and W^+ are given by

$$W^-(y_1y_2|u_1) = \sum_{u_2} \tfrac{1}{2} W(y_1|u_1 \oplus u_2)W(y_2|u_2) \tag{1}$$

$$W^+(y_1y_2u_1|u_2) = \tfrac{1}{2} W(y_1|u_1 \oplus u_2)W(y_2|u_2). \tag{2}$$

Recall also that W^+ describes a genie-aided receiver who has access to the true value of u_1. Based on this, we can get a numerical estimate of the quality of W^- and W^+ by the following procedure:

(i) Pick u_1 and u_2 randomly. Pick y_1 randomly according to the distribution $W(y_1|u_1 \oplus u_2)$, pick y_2 randomly according to the distribution $W(y_2|u_2)$.

(ii) Compute the likelihoods $[W(y_1|0), W(y_1|1)]$ and $[W(y_2|0), W(y_2|1)]$.

(iii) Compute $[W^-(y_1y_2|0), W^-(y_1y_2|1)]$ and $[W^+(y_1y_2u_1|0), W^+(y_1y_2u_1|1)]$ based on (1) and (2)

(iv) Compute the maximum likelihood estimates of u_1 and u_2 based on the values computed in (iii).

(v) Repeat (i)-(iv) a large number of times to estimate the probability of making a wrong decision for $\hat{u}_1$ (the error probability for W^-) and the probability of making a wrong decision for $\hat{u}_2$.

This procedure can be somewhat simplified if the channel W is symmetric. Instead of generating u_1, u_2 randomly in (i) we can set $u_1 = u_2 = 0$ and generate the y_i's according to $W(y|0)$; step (iii) is also simplified as we can use $u_1 = 0$ in (2).

Henceforth we will assume that W is symmetric, to fix ideas set W to be a binary symmetric channel.

The procedure generalizes to multiple, say ℓ, levels of polarization: in step (i) we generate $n = 2^\ell$ independent $y_1, \ldots, y_n$ according to the distribution $W(y|0)$, in step (iii) we would compute $[W^{-\cdots-}(y^n|0), W^{-\cdots-}(y^n|1)], \ldots, [W^{+\cdots+}(y^n0^{n-1}|0), W^{+\cdots+}(y^n0^{n-1}|1)]$ by recursively applying equations (1) and 2. (Two likelihood values for each of the n synthesized channels.)

(a) Implement the procedure descibed for ℓ level polarization above for symmetric channels in your favorite programming language.

(b) For $\ell = 1, 2, 4, 8$, use this procedure to estimate the error probabilities of each of the 2^ℓ channels created by an ℓ-fold application of the polarization transform to the binary symmetric channel with crossover probability 0.11. (For $\ell = 1$ make sure you check the computational results against an exact hand computation of the two error probabilities.) For the $l = 8$ case, sort and plot these error probabilities.

(c) For the $\ell = 8$ case how many channels can we pick and keep the union bound on error probability $\sum_i p_i$ below 10^{-3}? (The sum is over the picked channels, p_i denotes the error probability of channel i.)