Introduction to Information and Coding Theory September 21, 2014
(Prof. Permuter Haim, Mr. Tal Kopetz and Mr. Oron Sabag)

Final Exam

Total time for the exam: 3 hours!

1) **True or False (20 Points)**:
Copy each relation to your notebook and write **true** or **false**. Then, if it's true, prove it. If it is false give a counterexample or prove that the opposite is true.
   a) Let X, Y be a pair of random variables jointly distributed according to $P_{X,Y}$. For every $y \in \mathcal{Y}$ we have
   $$H(X|Y = y) \leq H(X).$$
   b) Consider the channel in Fig. 1 where Z is a Gaussian noise with power N and a is a deterministic constant. There is a power constraint on X such that $E[X^2] \leq P$. The capacity between X and Y is denoted by C. Is $C = \frac{1}{2}\log(1 + \frac{P}{N})$?

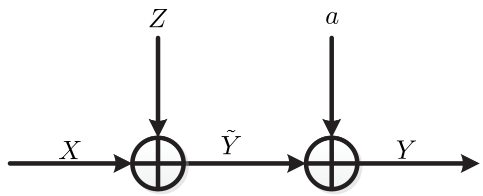

Fig. 1: Cascaded AWGN channel.

   c) Consider the channel in Fig. 2 where Z is a Gaussian noise with power N. There is a power constraint on X such that $E[X^2] \leq P$. The capacity between X and Y is denoted by C. Is $C = \frac{1}{2}\log(1 + \frac{P}{N})$?

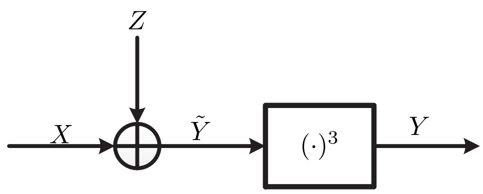

Fig. 2: AWGN channel.

2) **Cascaded additive modulo-2 channels with dependent noise (20 Points)**
Consider the additive modulo-2 channel in Fig. 3.

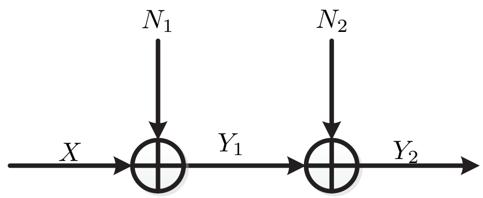

Fig. 3: Cascaded additive modulo-2 channels.

The input to the channel, X, is binary. The noise N_1 is distributed with $Bernoulli(\epsilon), \epsilon < \frac{1}{2}$. If $N_1 = 1$, the noise N_2 is distributed $Bernoulli(\alpha_1)$. If $N_1 = 0$, the noise N_2 is distributed with $Bernoulli(\alpha_2)$. N_1 and N_2 are independent of X. User 1 observes the output $Y_1 = X + N_1$. The capacity of the channel between X and Y_1 is denoted by C_1. User 2 observes the output $Y_2 = X + N_1 + N_2$. The capacity of the channel between X and Y_2 is denoted by C_2.
   a) Find the capacities C_1 and C_2 as functions of $\epsilon, \alpha_1, \alpha_2$.
   b) Find the constraints on $\epsilon, \alpha_1, \alpha_2$ such that:
      i) $C_1 = C_2$.
      ii) $C_1 < C_2$
      iii) $C_1 > C_2$
   c) We now introduce a 3rd user which obtains both (Y_1, Y_2). The capacity of the channel between X and (Y_1, Y_2) is C_3. Is it true that $C_1 + C_2 \geq C_3$? If not, give a counterexample.

3) **Network Coding for Broadcast Channel (30 Points)**
Consider the network in Fig. 4 where a source S wants to send a file consisting of M parts from the set $\mathcal{M} = \{1, \ldots, M)$ to the receivers $\tau = \{t_1, t_2\}$. Each receiver $t_i, i \in \{1, 2\}$ knows a subset of messages $\mathcal{M}_i \subseteq \mathcal{M}$. The source S knows the subsets $\mathcal{M}_i$. At each time slot the source S is able to transmit one packet.
   a) What is the minimum number of transmissions required by S such that all receivers can decode the file? prove it.
   b) Suggest a simple coding scheme that achieves the upper bound from (a).
   c) We now assume that $|\tau| > 2$ (there are more than two receivers in the network), what is the minimum number of transmissions required by S?
   d) Provide a coding scheme that achieves the upper bound from (c).

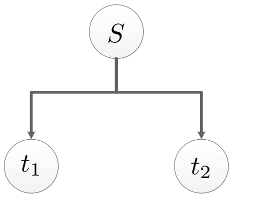

Fig. 4: A network that describes a broadcast channel.

4) **Conditional Information Divergence (30 Points)**
   a) Let X, Z be random variables jointly distributed according to $P_{X,Z}$. We define the conditional informational divergence as follows:
$$D\big(P_{X|Z}||Q_{X|Z}|P_Z\big) = \sum_{(x,z)\in\mathcal{X}\times\mathcal{Z}} P_{X,Z}(x,z)\log\left(\frac{P_{X|Z}(x|z)}{Q_{X|Z}(x|z)}\right).$$
   With respect to this definition, write each of the following relations to your notebook and write **true** or **false**:
   For any pair of random variables A, B that are jointly distributed according to $P_{A,B}$,
      i)
$$D\big(P_{A,B}||Q_{A,B}\big) = D\big(P_A||Q_A\big) + D\big(P_{B|A}||Q_{B|A}|P_A\big).$$
      ii)
$$D\big(P_{A,B}||P_A P_B\big) = D\big(P_{B|A}||P_B|P_A\big).$$
      iii)
$$I(A;B) = D\big(P_{B|A}||P_B|P_A\big).$$
      iv)
$$D\big(P_{A|B}||Q_{A|B}|P_B\big) = \sum_{b\in\mathcal{B}} P_B(b) D\big(P_{A|B=b}||Q_{A|B=b}\big).$$
   b) Consider the setting in Fig. 5.

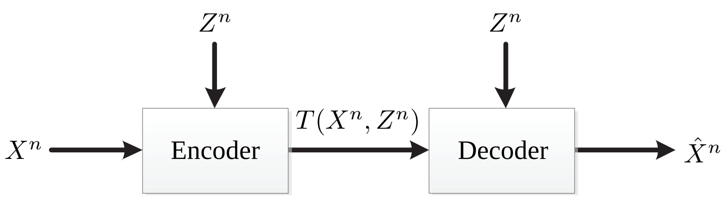

Fig. 5: Source coding with side information.

   We would like to compress the source sequence X^n losslessly using a prefix code with side information Z^n which is available to the encoder and the decoder. The sources (X^n, Z^n) are distributed i.i.d. according to $P_{X,Z}$ and that all the distribution and conditional distributions are dyadic (i.e., P_X is dyadic if $P_X(x) = 2^{-i}$, for some i, for all $x \in \mathcal{X}$). We denote the average number of bits per symbol needed to compress the source X^n as L.
      i) What is the minimal L?
      ii) Although the distribution of (X^n, Z^n) is $P_{X,Z}$, the distribution that is used design the optimal prefix code is $Q_{X|Z}P_Z$. What is the actual L (average bits per symbol) of this code?
      iii) Now, the distribution that is used to design the prefix code is $Q_{X,Z}$. What is the actual L now?

Good Luck!