Problem Set #3
Due: Tuesday, Nov. 29, 2016

1. *Strong Secrecy is Too Weak:* Consider a communication system that sends a packet of 256 bits and guarantees "strong secrecy," according to a given, very small ϵ. Without reading the small-print disclaimer on the communication device, an end user decides to use the system in the following way. They have packets of size 192 bits (3/4 of the device packet). They choose to pad their packets with 0's to fill the missing bits in the communication device. Aside from wasting efficiency, we will show that they may have severely crippled the security of the communication.

   Let the packet size (for the communication device) be N bits. For any sequence of N transmitted bits, suppose that the eavesdropper observes the symbol $*$ as long as the last 1/4 of the packet is not all-zero (i.e. the eavesdropper cannot distinguish between this set of transmissions). On the other hand, if the last 1/4 of the bits in the packet are all zeros, the eavesdropper sees exactly what was transmitted.

   How secure is this system from the point of view of the end-user described in the first paragraph?

   What is the mutual information (also give a simple approximation for large N) between the input to the communication device and the eavesdropper's observation under the assumption that the message is uniformly distributed?

2. *Converse for Gelfand-Pinsker:* The capacity of a memoryless channel $P_{Y|X,S}$, where the state variables is i.i.d. according to P_S and is known non-causally to the encoder, is given as follows:

$$C = \max_{P_{U,X|S}} \left(I(U;Y) - I(U;S) \right). \tag{1}$$

   Prove the converse for this claim, which is that the capacity C is less than or equal to the right side of (1). Feel free to refer to the converse proof of the wiretap channel from the course notes (which should be very similar), but do not refer to other references. As an extra challenge (optional), can you also claim the restriction that X is a function of S and U. This will likely fall immediately out of your converse proof if you look carefully.

3. *Excess distortion criterion:* Prove achievability of the rate-distortion theorem under the excess-distortion criterion using the likelihood encoder (do not refer to the literature). In class we proved achievability under the expected-distortion criterion. Many of the steps will be the same, and you may refer to those notes.

   In other words, prove that if

$$R > \min_{P_{\hat{S}|S} \,:\, \mathbb{E}[d(\hat{S},S)] \le D} I(S;\hat{S}), \tag{2}$$

   then for all $\epsilon > 0$ there exists a blocklength n and an encoder and decoder at rate R such that

$$\mathbb{P}\left[\frac{1}{n} \sum_{i=1}^{n} d(S_i, \hat{S}_i) > D \right] < \epsilon. \tag{3}$$