Spring 2016
6.441 - Information Theory
Homework 9
Due: Thur, Apr 28, 2016 (in class)
`Prof. Y. Polyanskiy`

1 Reading (optional)

1. Read [1, Chapter 8,9]

2 Exercises

NOTE: Each exercise is 10 points. Only 3 exercises per assignment will be graded. If you submit more than 3 solved exercises please indicate which ones you want to be graded.

1 Let $\{Z_j, j = 1, 2, \ldots\}$ be a stationary Gaussian process with variance 1 such that Z_j form a Markov chain $Z_1 \to \ldots \to Z_n \to \ldots$ Consider an additive channel

$$Y^n = X^n + Z^n$$

with power constraint $\sum_{j=1}^n |x_j|^2 \le nP$. Suppose that $I(Z_1; Z_2) = \epsilon \ll 1$, then capacity-cost function

$$C(P) = \frac{1}{2}\log(1+P) + B\epsilon + o(\epsilon)$$

as $\epsilon \to 0$. Compute B and interpret your answer.

How does the frequency spectrum of optimal signal change with increasing ϵ?

2 A semiconductor company offers a random number generator that outputs a block of random n bits $Y_1, \ldots, Y_n$. The company wants to secretly embed a signature in every chip. To that end, it decides to encode the k-bit signature in n real numbers $X_j \in [0, 1]$. To each individual signature a chip is manufactured that produces the outputs $Y_j \sim \mathrm{Bern}(X_j)$. In order for the embedding to be inconspicuous the average bias P should be small:

$$\frac{1}{n}\sum_{j=1}^n \left| X_j - \frac{1}{2} \right| \le P\,.$$

As a function of P how many signature bits per output (k/n) can be reliably embedded in this fashion? Is there a simple coding scheme achieving this performance?

3 (Strong converse for BSC) In this exercise we give a combinatorial proof of the strong converse for the binary symmetric channel. For BSC(δ) with $0 < \delta < \frac{1}{2}$,

1. Given any $(n, M, \epsilon)_{\max}$-code with deterministic encoder f and decoder g, recall that the decoding regions $\{D_i = g^{-1}(i)\}_{i=1}^M$ form a partition of the output space. Prove that for all $i \in [M]$,

$$|D_i| \ge \sum_{j=0}^{L} \binom{n}{j}$$

where L is the smallest integer such that $\mathbb{P}\left[\mathrm{Binomial}(n, \delta) \ge L\right] \ge 1 - \epsilon$.

2. Conclude that
$$M \le 2^{n(1-h(\delta))+o(n)}. \tag{1}$$
3. Show that (1) holds for average probability of error. (Hint: how to go from maximal to averange probability of error?)
4. Conclude that strong converse holds for BSC. (Hint: argue that requiring deterministic encoder/decoder does not change the asymptotics.)

4 Recall that AWGN is specified by

$$Y^n = X^n + Z^n\,, \qquad Z^n \sim \mathcal{N}(0, I_n)\,, \quad \mathsf{c}(x^n) = \frac{1}{n}||x^n||^2$$

Prove the strong converse for the AWGN via the following steps:

1. Let $c_i = f(i)$ and $D_i = g^{-1}(i), i = 1, \ldots, M$ be the codewords and the decoding regions of an $(n, M, P, \epsilon)_{max}$ code. Let
$$Q_{Y^n} = \mathcal{N}(0, (1+P)I_n)\,.$$
Show that there must exist a codeword c and a decoding region D such that
$$P_{Y^n|X^n=c}[D] \quad \ge \quad 1-\epsilon \tag{2}$$
$$Q_{Y^n}[D] \le \frac{1}{M}\,. \tag{3}$$
2. Show that then
$$\beta_{1-\epsilon}(P_{Y^n|X^n=c}, Q_{Y^n}) \le \frac{1}{M}\,. \tag{4}$$
3. Show that hypothesis testing problem
$$P_{Y^n|X^n=c} \quad \text{vs.} \quad Q_{Y^n}$$
is equivalent to
$$P_{Y^n|X^n=Uc} \quad \text{vs.} \quad Q_{Y^n}$$
where $U \in \mathbb{R}^{n\times n}$ is an orthogonal matrix. (Hint: use spherical symmetry of white Gaussian distributions.)
4. Choose U such that
$$P_{Y^n|X^n=Uc} = P^n\,,$$
where P^n is an iid Gaussian distribution of mean that depends on $||c||^2$.
5. Apply Stein's lemma to show:
$$\beta_{1-\epsilon}(P^n, Q_{Y^n}) = \exp\{-nE + o(n)\}$$
6. Conclude via (4) that
$$\log M \le nE + o(n) \quad \Longrightarrow \quad C_\epsilon \le \frac{1}{2}\log(1+P)\,.$$

5 *Mixtures of DMCs.* Consider two DMCs $U_{Y|X}$ and $V_{Y|X}$ with a common capacity achieving input distribution and capacities $C_U < C_V$. Let $T = \{0, 1\}$ be uniform and consider a channel $P_{Y^n|X^n}$ that uses U if $T = 0$ and V if $T = 1$, or more formally:

$$P_{Y^n|X^n}(y^n|x^n) = \frac{1}{2}U^n_{Y|X}(y^n|x^n) + \frac{1}{2}V^n_{Y|X}(y^n|x^n)\,. \tag{5}$$

Show:

1. Is this channel $\{P_{Y^n|X^n}\}_{n\geq 1}$ stationary? Memoryless?
2. Show that the Shannon capacity C of this channel is not greater than C_U (Hint: use strong converse. In fact the capacity equals C_U.)
3. the maximal mutual information is

$$\lim_{n\to\infty} \frac{1}{n} \sup_{X^n} I(X^n; Y^n) = \frac{C_U + C_V}{2}$$

4. conclude that

$$C < \lim_{n\to\infty} \frac{1}{n} \sup_{X^n} I(X^n; Y^n)\,.$$

6 Routers A and B are setting up a covert communication channel in which the data is encoded in the ordering of packets. Formally: router A receives n packets, each of type A or D (for Ack/Data), where type is i.i.d. Bernoulli(p) with $p \approx 0.9$. It encodes k bits of secret data by reordering these packets. The network between A and B delivers packets in-order with loss rate $\delta \approx 5\%$ (Note: packets have sequence numbers, so each loss is detected by B).

What is the maximum rate $\frac{k}{n}$ of reliable communication achievable for large n? Justify your answer!

References

[1] T. Cover and J. Thomas, *Elements of Information Theory,* Second Edition, Wiley, 2006