TUTORIAL VII

1 Expurgation

Let C be a M-code with error probability $P_{err}(C) = \delta$.

1. Show that you can build a $\lfloor M/2 \rfloor$-code with maximal error probability $\leq 2\delta$.

2 Channel coding problems

1. Consider the channel W with input alphabet $\mathcal{X} = \{a, b, c\}$ and output alphabet $\{0, 1\}$, with $W(0|a) = 1, W(0|b) = \frac{1}{2}, W(1|b) = \frac{1}{2}$ and $W(1|c) = 1$. Then, let $W^{\times n}$ be n independent copies of W.
   (a) For any M, determine the optimal (i.e., smallest possible) error probability for an M-code for $W^{\times n}$, as a function of M and n.
   Note that this bound is achievable, by losslessly sending messages $m \in \{1, 2, \ldots, 2^n\}$ (so that they are correctly decodable with probability 1).

   We encode all other message as b^n (note that they are never decoded correctly, as they are not in the range of the decoding function).

   Thus, we can correctly decode 2^n out of M messages, so the error probability is $1 - \frac{2^n}{M}$.
   (b) Compute $\mathrm{C}(W)$.
2. Let $a \in \{1, 2\}$. Consider the additive noise channel with input alphabet $\mathcal{X} = \{0, 1\}$ and output alphabet $\mathcal{Y} = \{0, 1, 2, 3\}$, where the output Y is given by $x + Z$ when x is the input symbol and Z is a random variable with distribution $\mathbf{P}\{Z = 0\} = \mathbf{P}\{Z = a\} = \frac{1}{2}$. Compute the information capacity of this channel.

3 Algorithmic approach to the channel coding problem

The main objective here is to take an algorithmic approach for the channel coding problem. The input to our algorithmic problem is the specification of a noisy channel W from an input set $\mathcal{X}$ to an output set $\mathcal{Y}$. We would like to send M messages and we ask what is the minimum error probability that we can achieve.

This will be a good opportunity to introduce *submodular* functions which is an interesting property to keep in mind and a rich area of study in optimization and approximation algorithms.

1. (Maximization of submodular functions)
   A function $f : 2^{\mathcal{X}} \to \mathbb{R}_+$ (with $\mathcal{X}$ is finite) taking as input a subset $S \subseteq X$ that has the following property.

$$f(S \cup T) + f(S \cap T) \leq f(S) + f(T) \,. \tag{4}$$

   It is said to be monotone if $f(S) \leq f(T)$ whenever $S \subseteq T$.
   (a) Show that an equivalent definition for submodular function is that $f(T \cup \{j\}) - f(T) \leq f(S \cup \{j\}) - f(S)$ for any $S \subseteq T$ and any $j \in \mathcal{X} - T$. This can be interpreted as a "diminishing returns" property.
   (b) (Remark: this question is independent of the following questions) Let $Z_1, \ldots, Z_n$ be a family of random variables. For a subset $S \subseteq \{1, \ldots, n\}$, let Z_S be the collection of random variables $\{Z_i\}_{i \in S}$. Show that $f(S) = H(Z_S)$ is a submodular and monotone function.
   (c) Let f be a submodular, monotone and nonnegative function and consider the following optimization problem $\max_{S \subseteq \mathcal{X}, |S| = M} f(S)$. Let S^* of size M be such that $f(S^*) = \max_{S \subseteq \mathcal{X}, |S| = M} f(S)$. Computing such an S^* is computationally hard in general. But there is a natural greedy algorithm for this problem: start with $S_0 = \emptyset$, then choose $S_{i+1} = S_i \cup \arg\max\{f(S_i \cup \{j\}) : j \in \mathcal{X} - S_i\}$. Show that $f(S^*) \leq f(S_i) + M(f(S_{i+1}) - f(S_i))$.
   (d) Prove that $f(S^*) - f(S_{i+1}) \leq (1 - \frac{1}{M})(f(S^*) - f(S_i))$.
   (e) Conclude that the greedy algorithm gives a constant factor approximation for this problem (and say what the constant is).

2. (Channel coding as a submodular optimization problem) Let $\mathrm{S}(W, M)$ be the largest average success probability of a code for M messages:

$$\mathrm{S}(W, M) = \max_{e,d} \frac{1}{M} \sum_{i=1}^{M} \sum_{y \in \mathcal{Y}: d(y)=i} W(y|e(i)) \,, \tag{15}$$

where the maximization is over functions $e : \{1, \ldots, M\} \to \mathcal{X}$ and $d : \mathcal{Y} \to \{1, \ldots, M\}$.

   (a) Show that $\mathrm{S}(W, M)$ can be written as maximizing some function f over all subsets of $\mathcal{X}$ of size M. Then show that f is submodular and monotone.

   (b) Conclude that it is possible to efficiently (here efficiently means polynomial in the description of the channel W and of M) find a code that achieves a success probability that is at least $(1 - 1/e) \cdot \mathrm{S}(W, M)$.