Homework 8-9: On channel coding

CR07: Selected Topics Information Theory (Fall 2019). ENS de Lyon
jean-marie.gorce@inria.fr – Deadline: December, 9, 2019 at 11:59 p.m.

1 Fundamental limits of channel codes

1.1 Preliminaries

Consider a general channel (may be discrete or continuous), defined as

$$\left(\mathcal{X}, \mathcal{Y}, P_{Y|X}\right).$$

The source is characaterized by a random variable W, uniformly distributed on $\mathcal{W} = \{1, \dots M\}$, with $|\mathcal{W}| = M$.

A $M-$code for $P_{Y|X}$ is an encoder-decoder pair (f, g) of functions :

$$\begin{array}{rcccc} f & : & \{1 \dots M\} & \longrightarrow & \mathcal{X} \\ g & : & \mathcal{Y} & \longrightarrow & \{1 \dots M\} \end{array}$$

Remind the two definitions :

- A code (f, g) is an (M, ϵ) code for $P_{Y|X}$ if the average error of this code noted $P_e = \mathbb{P}(w \neq \hat{w})$, verifies $P_e \leq \epsilon$.
- A code (f, g) is an $(M, \epsilon)_{max}$ code for $P_{Y|X}$ if the max error of this code noted $P_{e,max} = \max_{m \in \mathcal{W}} \mathbb{P}(\hat{W} \neq m | W = m)$, verifies $P_{e,max} \leq \epsilon$.

The following rules (partially proved in the course) are known:

Theorem 1 (general weak converse) *Any M-code for $P_{Y|X}$ satisfies*

$$\log M(\epsilon) \leq \frac{\sup_{P_X} I(X;Y) + h(P_e)}{1 - P_e}.$$

We note $i_{XY}(x;y) = \frac{dP_{Y|X=x}}{dP_Y}(y)$. where $\frac{dP_{Y|X=x}}{dP_Y}$ is the Radon-Nikodym derivative. For the sake of simplicity, $i_{XY}(x;y)$ will be shortly written when no confusion is possible as $i(x;y)$ or $i(X;Y)$ when considered as a random variable.

Theorem 2 (Shannon's bound) *For a given $P_{Y|X}$, $\forall P_X$, $\forall \tau > 0$, there exists an (M,ϵ)-code such that :*

$$\epsilon(M) \leq \mathbb{P}\left[i(X;Y) \leq \log M + \tau\right] + \exp(-\tau).$$

Theorem 3 (Dependence testing bound) *For a given $P_{Y|X}$, $\forall P_X$, there exists an (M,ϵ)-code such that :*

$$\epsilon(M) \leq \mathbb{E}\left[e^{-\left\{i(X;Y)-\log\left(\frac{M-1}{2}\right)\right\}^+}\right].$$

Theorem 4 (Feinstein's lemma) *For a given $P_{Y|X}$, $\forall P_X$, $\forall \gamma > 0$, there exists an $(M,\epsilon)_{max}(M)$-code such that :*

$$\epsilon(M) \leq \mathbb{P}\left[i(X;Y) < \log \gamma\right] + \frac{M}{\gamma}.$$

Theorem 5 (Gallager's bound) *For a given $P_{Y|X}$, $\forall P_X$, $\forall \lambda \in [0,1]$, there exists an $(M,\epsilon)_{max}$-code such that :*

$$\epsilon(M) \leq (M-1)^{\lambda}\mathbb{E}_{\bar{Y}}\left[\left\{\mathbb{E}_{X,Y}\left[exp\left(\frac{i_{XY}(X;\bar{Y}}{1+\lambda}\right)\right]\right\}^{1+\lambda}\right]$$

Theorem 6 (ML error of random coding) *For a given $P_{Y|X}$, $\forall P_X$, the average error associated to random coding is given by :*

$$\epsilon = 1 - \sum_{l=0}^{M-1}\binom{M-1}{l}\frac{1}{l+1}\mathbb{E}\left[U^l V^{M-1-l}\right],$$

with $U = \mathbb{P}\left[i_{XY}(\bar{X};Y) = i_{XY}(\bar{X};Y)|XY\right]$ *and* $V = \mathbb{P}\left[i_{XY}(\bar{X};Y) < i_{XY}(\bar{X};Y)|XY\right]$.

1.2 Questions

1. Write down the proof of the 6 theorems above. Clearly identify the different assumptions used to derive each proof.

2. Derive for each of them the relation between $\epsilon^*(M)$ or $M^*(\epsilon)$.

3. For a discrete memoryless symmetric channel (with M states), compute and compare the resulting function $M^*(\epsilon)$.

4. Compare these bounds in terms of tightness and computational complexity.

5. Evaluate the asymptotic behavior of at least three of these expressions when they are used for a stationarty discrete meomryless channel with n channel uses. first and second order approximations are expected.

6. Comments on the results are welcome.

2 Channel coding with input constraints

2.1 Preliminaries

A stationaty memoryless channel with a separable cost constraint is a channel $\left(\mathcal{X}, \mathcal{Y}, p_{\boldsymbol{Y}|\boldsymbol{X}}\right)$ where $\mathcal{X} = \mathbb{A}^n$, $\mathcal{Y} = \mathcal{B}^n$ and where :

- $\mathcal{A}, \mathcal{B}$ are input/output spaces associated to the samples
- $p_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{k=1}^{n} p_{Y|X}(y_k|x_k)$,
- Cost $c : \mathcal{A} \to \mathbb{R}^+$
- Average cost constraint: $c(\boldsymbol{x}) = \frac{1}{n} \cdot \sum_{k=1}^{n} c(x_k) \leq P_M$, where P_M is a positive constant.

The additive white gaussian noise (AWGN)under average power constraint is a channel with cost constraint with $\mathcal{A} = \mathcal{B} = \mathbb{R}$ and $c(x) = x^2$.
The vector output of the channel is given by :

$$\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{Z}.$$

where $\boldsymbol{Z} \sim \mathcal{N}(0, 1)$ is a normalized received noise.

2.2 Questions

1. Give the conditional probability density function : $p_{Y|X}$.

2. Proof that the capacity of this channel is given by:

$$C(P) = \frac{1}{2} \log \left(1 + P_M\right),$$

   and show that the optimal input distribution follows a normal distribution.

3. Show that, under a unitary variance constraint, the gaussian noise is the worst case.

4. From one of the achievability expressed in the first part, establish a second order approximation of the Gaussian channel ϵ-capacity (you can use a power-shell input distribution and explore the litterature on this topic).