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

\((X, Y, P_{Y|X})\).

The source is characterized by a random variable \(W\), uniformly distributed on \(\mathcal{W} = \{1, \ldots, M\}\), with \(|\mathcal{W}| = M\).

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

\[
\begin{align*}
  f & : \{1 \ldots M\} \to \mathcal{X} \\
  g & : \mathcal{Y} \to \{1 \ldots M\}
\end{align*}
\]

Remind the two definitions:

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

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

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

\[
\log M(\epsilon) \leq \sup_{P_X} I(X; Y) + h(P_e) \frac{1}{1 - P_e}.
\]
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 $(M, \epsilon)$-code such that:

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

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

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

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

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

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

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

**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:

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

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

### 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 $M^*(\epsilon)$. 

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

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

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

The vector output of the channel is given by:

$$Y = X + Z,$$

where $Z \sim \mathcal{N}(0, 1)$ 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:

\[ C(P) = \frac{1}{2} \log (1 + P_M), \]

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 \( \epsilon \)-capacity (you can use a power-shell input distribution and explore the litterature on this topic).