CIE6020/MAT3350
Selected Topics in Information Theory

Lecture 13: Achievability of Channel Coding Theorem

14 March 2019

The Chinese University of Hong Kong, Shenzhen

Achievability

Outline of Achievability Proof

1. Generate a random code with rate close to $I(X;Y)$.
2. Define a jointly typical decoding algorithm.
3. Evaluate the expected P_e of all the codes in the ensemble.
4. Last enhance the code so that $\lambda_{\max} < \epsilon$.

Random Code Generation

- Fix $p(x)$ and $\epsilon > 0$.
- Let M be an *even* integer such that

$$I(X;Y) - \frac{\epsilon}{2} < \frac{\log M}{n} < I(X;Y) - \frac{\epsilon}{4},$$

  where $(X,Y) \sim p \cdot W$.
- Generate a codebook $\mathcal{C}$ of M codewords independently according to the distribution $p(\mathbf{x}) = \prod_i p(x_i)$. Let

$$\mathcal{C} = \{\mathbf{X}_1, \ldots, \mathbf{X}_M\}.$$

- $\mathbf{X}_i \sim p(\mathbf{x})$ are independent
- Assume that both the sender and the receiver know the instance of $\mathcal{C}$ to use.

Jointly Typical Decoding

- The first message $\mathbf{X}_1$ is transmitted, and the channel output is $\mathbf{Y}$.
- $(\mathbf{X}_1, \mathbf{Y}) \sim p(\mathbf{x})W_n(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(x_i)W(y_i|x_i)$.
- Jointly typical decoding:
  - The sequence $\mathbf{Y}$ is decoded to the message k if $(\mathbf{X}_k, \mathbf{Y}) \in T^n_{[XY]\delta}$, where δ is a positive quantity to be specified later.
  - If no such message exists or if there is more than one such a message, an error is declared.

Average Error Probability I

- λ_1 is a function of the code $\mathcal{C}$.
- Let $E(\mathbf{x}) = \{(\mathbf{x}, \mathbf{Y}) \in T^n_{[XY]\delta}\}$.
- We have

$$\begin{aligned}\mathbb{E}[\lambda_1] &= \Pr\{E^c(\mathbf{X}_1) \cup E(\mathbf{X}_2) \cup \cdots \cup E(\mathbf{X}_M)\} \\ &\leq \Pr\{E^c(\mathbf{X}_1)\} + \sum_{k=2}^{M} \Pr\{E(\mathbf{X}_k)\}.\end{aligned}$$

- Since $(\mathbf{X}_1, \mathbf{Y}) \sim \prod_{i=1}^n p(x_i)W(y_i|x_i)$, by the property of (strongly) typical sets, for sufficiently large n,

$$\Pr\{E^c(\mathbf{X}_1)\} = \Pr\{(\mathbf{X}_1, \mathbf{Y}) \notin T^n_{[XY]\delta}\} \leq \delta.$$

Average Error Probability II

- For $k > 1$, $\mathbf{X}_k$ and $\mathbf{Y}$ are independent.
- We have that for $k > 1$,

$$\begin{aligned}
\Pr\{E(\mathbf{X}_k)\} &= \Pr\left\{(\mathbf{X}_k, \mathbf{Y}) \in T^n_{[XY]\delta}\right\} \\
&= \sum_{(\mathbf{x},\mathbf{y}) \in T^n_{[XY]\delta}} p(\mathbf{x}) p_{\mathbf{Y}}(\mathbf{y}) \\
&\leq 2^{-n(H(X)-\eta_1)} 2^{-n(H(Y)-\eta_2)} \left| T^n_{[XY]\delta} \right| \\
&\leq 2^{-n(H(X)-\eta_1)} 2^{-n(H(Y)-\eta_2)} 2^{n(H(X,Y)+\eta_3)} \\
&= 2^{-n(I(X;Y)-\tau)},
\end{aligned}$$

where $\tau = \eta_1 + \eta_2 + \eta_3 \to 0$ as $\delta \to 0$.

Average Error Probability II

- For sufficiently large n

$$\begin{aligned}\mathbb{E}[\lambda_1] &\leq \delta + \sum_{k=2}^{M} 2^{-n(I(X;Y)-\tau)} \\ &\leq \delta + M2^{-n(I(X;Y)-\tau)} \\ &< \delta + 2^{-n(\epsilon/4-\tau)}.\end{aligned}$$

- Note that $\tau \to 0$ as $\delta \to 0$.
- Let δ be sufficiently small such that $\delta < \epsilon/3$ and $\tau < \epsilon/4$.
- For n sufficiently large,

$$\mathbb{E}[\lambda_1] < \epsilon/2.$$

Existence: Average Error Probability

- Since the error probabilities of all codewords follow the same calculation, we have

$$\mathbb{E}P_e(\mathcal{C}) = \mathbb{E}\frac{1}{M}\sum_i \lambda_i(\mathcal{C}) = \frac{1}{M}\sum_i \mathbb{E}\lambda_i(\mathcal{C}) < \epsilon/2.$$

- Therefore, there exits at least one codebook $\mathbb{C}$ such that

$$P_e(\mathbb{C}) < \epsilon/2.$$

Existence: Maximal Error Probability

- Let $\mathbb{C}^*$ be the subset of $\mathbb{C}$ with the best half codewords (in terms of λ_i).
- The maximal error probability of the codewords in $\mathbb{C}^*$ is less than ϵ.
- The rate of $\mathbb{C}^*$ is

$$\frac{1}{n} \log \frac{M}{2} > I(X;Y) - \frac{\epsilon}{2} - \frac{1}{n} > I(X;Y) - \epsilon$$

when n is sufficiently large.