Distinguishing Codes from Noise

Student: Da Wang
Advisor: Gregory Wornell

Massachusetts Institute of Technology

Motivation

- Communication: **synchronization** and **coding**
- Synchronization: mostly done by training:

x_c	x_c	...	x_c	capacity achieving code

$\underbrace{\qquad\qquad}_{(1-\frac{R}{C})n \text{ symbols}}$ $\underbrace{\qquad}_{\frac{R}{C}n \text{ symbols}}$

Is that always good? Can we do better?

Distinguishing Codes from Noise

Asynchronous channel

$W_{Y|X}(\cdot|\cdot)$

$\mathcal{X} \quad \mathcal{Y}$

$\star$

- $\star$: input symbol to model that nothing is sent.

$$\underset{X_1 \quad X_{t_1} \qquad\qquad X_{t_2}}{\star \cdots \star \overset{\text{codeword 1}}{c_1 \cdots c_n} \star \star \overset{\text{codeword 2}}{c_1 \cdots c_n} \star} \qquad t$$

Slotted simplification

Communicate in pre-defined timeslots:

$$\star^n \quad \star^n \quad \star^n \quad \mathbf{c}(i) \quad \star^n \quad \star^n \quad \mathbf{c}(j) \quad \star^n$$

$\longrightarrow t$

detect & locate $\overset{\text{slotting}}{\longrightarrow}$ detection only

Mathematical Setup

For a channel code $\mathcal{C} = \{X^n(k)\}$ with rate R, we have the following hypothesis testing problem:

$$\begin{cases} H_0: & Y_i \overset{i.i.d.}{\sim} W(\cdot|\star) \quad i = 1, 2, \cdots n \\ H_1: & Y^n \sim W(\cdot|X^n(k)) \ k \in \{1, 2, \cdots, M\} \end{cases}$$

Define

$$P_m \triangleq \mathbb{P}[\{H_1 \to H_0\}] \doteq \exp(-nE_m)$$
$$P_f \triangleq \mathbb{P}[\{H_0 \to H_1\}] \doteq \exp(-nE_f)$$

Analysis objectives

Characterize the $E_m - E_f$ trade-off at rate R.

Special case: $E_m = 0$

Optimal $E_f(R)$ when $E_m = 0$

Given $E_m = 0$ and hence $P_m \to 0$,

$$E_f(R) = \max_{P_X: I(P_X, W) = R} D(P_Y \| Q_\star)$$

- i.i.d. codebook with distribution P_X
- noise output distribution $Q_\star = W(\cdot|\star)$.
- ✓ use rate-achieving i.i.d. codebook rather than capacity-achiving codebook.

BSC Example

For a binary symmetric channel (BSC) with $W(\cdot|\star) = \text{Bern}(u)$ ($u \le 0.5$), we can achieve

$$E_f(R) = D(\underbrace{\text{Bern}(s^*)}_{\text{c.w. output dist.}} \| \underbrace{\text{Bern}(u)}_{\text{noise output dist.}})$$

where $H_b(s^*) - H_b(\varepsilon) = R$.

Figure: $E_f(R)$ for BSC with $\varepsilon = 0.01$

Comparison with training

(a) $u = 0.1$ $\qquad$ (b) $u = 0.5$

Figure: Comparison with training for BSC with $\varepsilon = 0.01$.

✓ Large gain over training at high rate.

General case: $E_m \ge 0$

Achievable $E_f(R)$ when $E_m \ge 0$

Given $P_m \overset{\cdot}{\le} \exp(-nE_m)$,

$$E_f(R, E_m) = \max_{P_X: I(P_X, W) \ge R} \min_{V: D(V\|W|P_X) \le E_m} \left[D(Q_V \| Q_\star) + \{I(P_X, V) - R\}^+ \right]$$

- achieved by constant composition codebook with maximizing distribution P_X^*.
- i.i.d. codebook is suboptimal in general.
- non-trivial converse for DMC is unknown.

Comparison with i.i.d. codebook and training for BSC channel

(a) $R = 0.405$ $\qquad$ (b) $R = 0.708$

Figure: Performance comparison between constant composition codebook, i.i.d. codebook, and training for BSC with $\varepsilon = 0.05$ and $u = 0.5$.

✓ Again, large gain over training at high rate.

Extensions & Connections

AWGN channel, unequal error protection (UEP), ...

Conclusion

For certain communication scenarios, designing codes for both **detection and information transmission jointly** achieves significantly larger detection error exponents than the traditional separate sync–coding approach.