Lecture 14: Proof of channel coding theorem

- Achievability: when $R < C$, exists zero error code
- Converse: zero error code must have $R < C$

Channel coding theorem

Theorem. *(Shannon, 1948)*
For a DMC

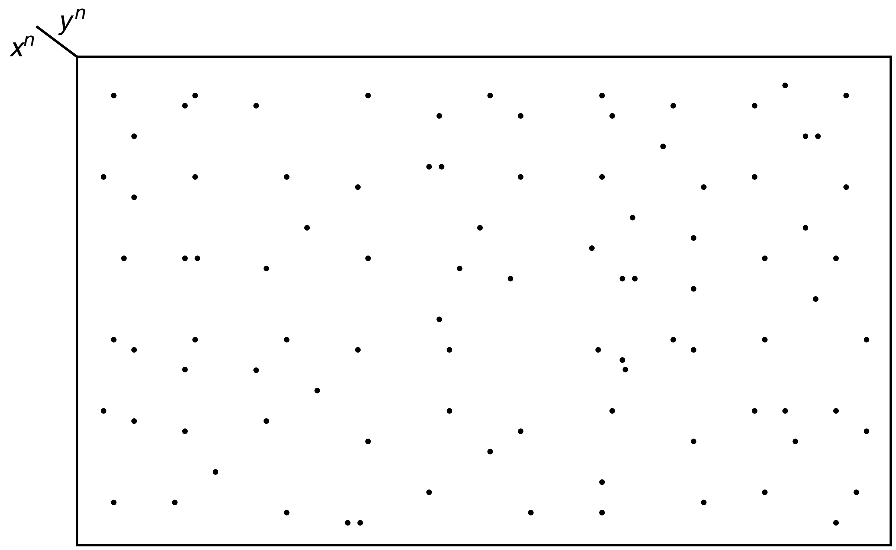
1. *all rates below capacity* $R < C$ *are achievable.*

2. *Converse: any sequence of* $(2^{nR}, n)$ *codes with* $\lambda^{(n)} \to 0$ *must have* $R \leq C$.

Joint typical decoding

- Decoder find $\hat{W}$ if $(X^n(\hat{W}), Y^n)$ is jointly typical

- No confusion: no more than $X^n(\hat{W})$ jointly typical with Y^n



Proof for achievability

- calculate the probability of error averaged over all codes randomly generated according to $p(x)$
- Average P_e does not depend on which index was sent
- For typical X^n, two type of errors

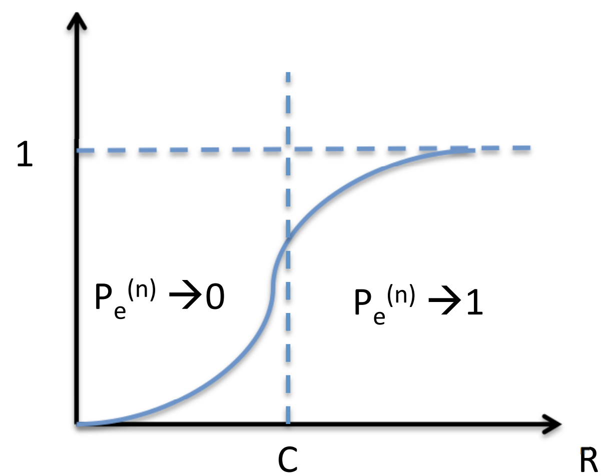
(a) (X^n, Y^n) not jointly typical
(b) $(\tilde{X}^n, Y^n)$ is typical, but $\tilde{X}^n \neq X^n$

- Use AEP to bound (a) and (b)
- Conditional probability of error

$$\lambda_i = P\{g(Y^n) \neq i | X^n = x^n(i)\}$$

Proof for converse

- Use Fano's inequality to lower bound P_e



Implications of the theorem

- It shows that there exist good codes with exponentially small probability of error for long block length
- it does not provide a way to construct the best codes
- random code, without structure, very difficult to code (look-up table)
- property of capacity achieving codes
- example of capacity achieving: noisy typewriter
- new capacity achieving code: polar codes (2009)