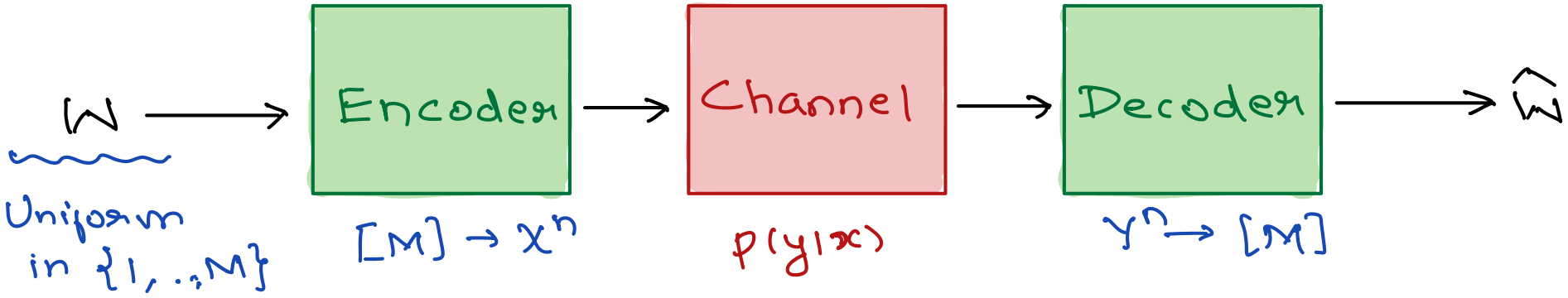
Recap



Rate: $R = \frac{\log M}{\log |X|^n} = \frac{\log M}{n \cdot \log |X|}$

Capacity: $C = \max_{P(x)} I(X;Y)$

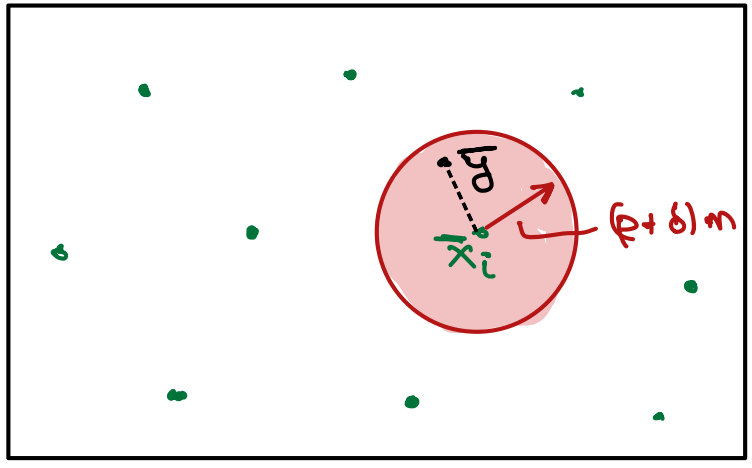
- Sup {achievable R} $= R^* = C$

Achieving Capacity (Binary Symmetric Channel)

$M = 2^{nR}$. $C = \{\bar{x}_1, \ldots, \bar{x}_M\} \subseteq \{0,1\}^n$

$\bar{y} = \bar{x}_i + \bar{z} \pmod 2$

$\forall j \in [n]$ $z_j = \begin{cases} 1 & \text{wp } p \\ 0 & \text{wp } 1-p \end{cases}$



$$Dec(\bar{y}) = \begin{cases} \bar{x}_i & \text{if } \bar{x}_i \text{ is unique codeword with } \Delta(\bar{x}_i, \bar{y}) \leq (p+\delta)\cdot n \\ \text{garbage} & \text{o.w.} \end{cases}$$

Δ: Hamming distance

$$\mathop{\mathbb{E}}_{C}[p_e] = \mathop{\mathbb{E}}_{C}[\mathbb{P}[\hat{W} \neq W]]$$

$$= \mathop{\mathbb{E}}_{C}\left[\frac{1}{M}\sum_{i} \mathbb{P}[\hat{W} \neq i \mid W = i]\right]$$

$$= \mathop{\mathbb{E}}_{C}\left[\mathbb{P}[\hat{W} \neq 1 \mid W = 1]\right]$$

$$\leq \mathop{\mathbb{P}}_{C,\bar{y}}\left[\Delta(\bar{x}_1, \bar{y}) > (p+\delta)\cdot n \mid W=1\right] + \underbrace{\sum_{i=2}^{M} \mathop{\mathbb{P}}_{C,\bar{y}}\left[\Delta(\bar{x}_i, \bar{y}) \leq (p+\delta)\cdot n \mid W=1\right]}_{\text{(M-1)}\cdot \mathbb{P}[\Delta(\bar{x}_2, \bar{y}) \leq (p+\delta)\cdot n \mid W=1]}$$

$$= \mathbb{P}\left[\sum Z_j \geq (p+\delta)\cdot n\right]$$

$$\bar{Z} \sim \underbrace{(\mathrm{Bern}(p))^n}_{Q^n}$$

$$\leq (n+1)\cdot 2^{-n\cdot D(p+\delta \,\|\, p)}$$

$$= \mathbb{P}\left[\sum Z_j' \leq (p+\delta)\cdot n\right]$$

$$\bar{Z}' \sim (\mathrm{Bern}(1/2))^n$$

$$\leq M\cdot(n+1)\cdot 2^{-nD(p+\delta\|1/2)}$$

Bounding expected error

$$\mathbb{E}[p_e] \leq \underbrace{n \cdot 2^{-n \cdot D(p+\delta \| p)}}_{\to 0 \text{ as } n \to \infty} + \underbrace{M \cdot n \cdot 2^{-n \cdot D(p+\delta \| 1/2)}}$$

$$\leq 2^{nR} \cdot n \cdot 2^{-n(1 - H_2(p+\delta))}$$

$$\to 0$$

if $R < 1 - H_2(p+\delta) - \epsilon$ (say)

if δ s.t. $H_2(p \star \delta) \leq H_2(p) + \epsilon$

then $R \geq 1 - H_2(p) - 2\epsilon$ suffices.

Ex: Suffices to take $n \gg 1/\epsilon^2$ when $R \leq 1 - H_2(p) - \epsilon$.

Are we done with codes then?

Random Codes

- Rate ✓
- $P_e \to 0$ ✓
- Storage ✗
- Encoding/Decoding ✗
- Structural Properties ✗
- Other channels/ Hamming model ✓

Linear Codes over Finite Fields

$\mathbb{F}_2$ $+, \cdot$ mod 2

$\mathbb{F}_p$ $+, \cdot$ mod p

$[M] \simeq \mathbb{F}_2^k$

$G \in \mathbb{F}_2^{n \times k}$

$C = \{ Gw \mid w \in \mathbb{F}_2^k \}$

$= \text{im}(G)$

- C is a subspace $\mathbb{F}_2^n$
- Succinct representation
- Encoding is easy
- Decoding without error is easy

Generators and Parity Checks

$G \in \mathbb{F}_2^{n \times k}$

$C = \{ Gw \mid w \in \mathbb{F}_2^k \} = \text{im}(G) = \ker(H)$. $H \in \mathbb{F}_2^{(n-k) \times n}$

$= \{ x \in \mathbb{F}_2^n \mid Hx = 0 \}$

e.g. $H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \in \mathbb{F}_2^{3 \times 7}$

$x \longrightarrow y = x + e_i$

bit flip, unknown i

Find x

$Hy = \cancel{Hx}^{\,0} + He_i$