DATA TRANSMISSION (CHANNEL CODING THEOREM)

$$m \xrightarrow{\text{encoding}} \underline{x} \xrightarrow[Q(\underline{y}|\underline{x})]{\text{channel}} \underbrace{\underline{y} \longrightarrow \underline{x}^{d} \longrightarrow m'}_{\text{decoding}}$$

$m \in \{0,1\}^M$; $\underline{x} \in \{0,1\}^N$; BSC $\rightarrow$ $\underline{y} \in \{0,1\}^N$

p = bin crossover probability (channel noise)

posterior probability

$$\mu_{\underline{y}}(\underline{x}) = \frac{1}{Z(\underline{y})} \prod_{i=1}^{N} Q(y_i|x_i)\, \underbrace{\mu_0(\underline{x})}_{\text{prior probability}}$$

uniform prior $\mu_0(\underline{x}) = \frac{1}{|\mathcal{C}_N|} \mathbb{1}(\underline{x} \in \mathcal{C}_N)$ — codebook

$|\mathcal{C}_N| = 2^M$

Word MAP decoding

$$\mu_y(\underline{x}) = \frac{1}{Z(\underline{y})} \exp[-2B\, d(\underline{x},\underline{y})]\, \mathbb{1}(\underline{x} \in \mathcal{C}_N)$$

$d(\underline{x},\underline{y})$ = Hamming distance ; $B = \frac{1}{2} \log\left(\frac{1-p}{p}\right) > 0$ $(p < 1/2)$

maximize $\mu_y(\underline{x})$ $\rightarrow$ find the closest codeword to channel output $\underline{y}$

decoding successful if $\underbrace{p < \delta_{GV}}$ $\quad \delta = d/N$

Symbol MAP decoding: $\mu_y^{(i)}(x_i) = \sum_{\underline{x} \setminus i} \mu_y(\underline{x})$

successful decoding

$\longrightarrow \quad p < \delta_{GV}$

FINITE TEMPERATURE DECODING

$$\mu_{\underline{y},\beta}(\underline{x}) = \frac{1}{Z(\beta)} \exp\left[-2\beta B\, d(\underline{x},\underline{y})\right] \mathbb{I}(\underline{x} \in \mathcal{C}_N)$$

$$0 \le \beta < +\infty$$

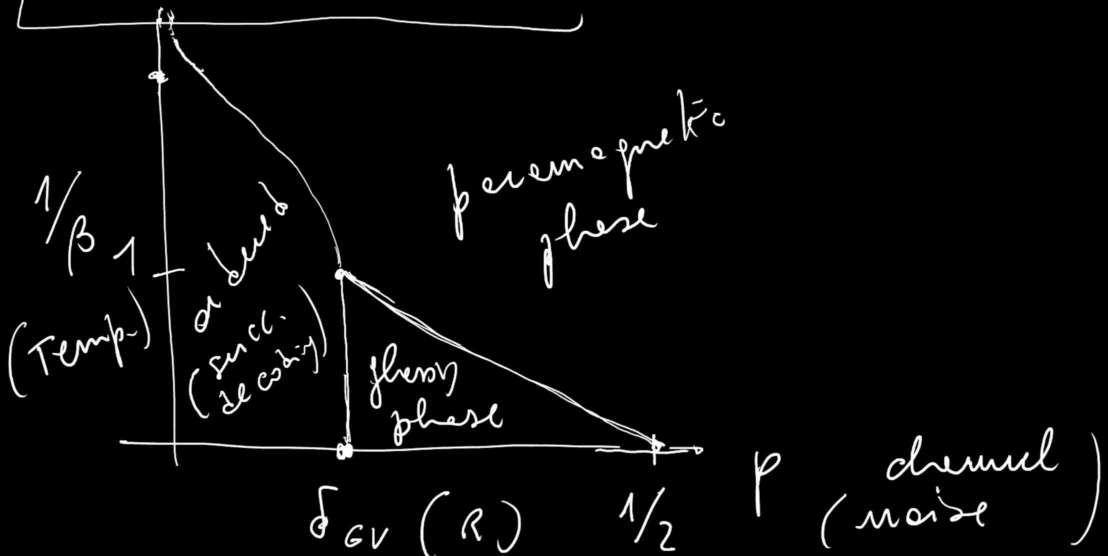
decoding: find x_i that maximizes $\mu^{(i)}_{\underline{y},\beta}(x_i) \ \forall i$

$\beta = 1 \Rightarrow$ symbol MAP

$\beta = \infty \Rightarrow$ word MAP (the minimum $d(\underline{x},\underline{y})$ $\rightarrow$ leading term in $Z(\beta)$)

$\beta = 0 \Rightarrow$ uniform posterior $\rightarrow$ complete noise $\rightarrow$ no decoding for any p

PHASE DIAGRAM



$R = \frac{M}{N}$

code rate

1) "ordered" phase $\left(\begin{matrix}\text{low noise}\\ \text{low } T\end{matrix}\right) \to$ successful decoding

2) "Glassy" phase $\left(\begin{matrix}\text{large noise}\\ \text{low } T\end{matrix}\right) \to$ unsuccessful decoding

measure ("posterior") $\mu_{\underline{y},\beta}(\underline{x})$ dominated by codewords CLOSEST to $\underline{y}$ $(\delta^* = \delta_{GV})$

$\to$ "MEASURE CONDENSATION"

Related Shannon Entropy $H(\mu_{y,\beta})$ sublinear in N

$$\lim_{N\to\infty} \frac{H(\mu_{\underline{y},\beta})}{N} = 0 \quad \text{"zero" entropy phase}$$

3) "paramagnetic" phase $\left(\begin{matrix}\text{large noise}\\ \text{high } T\end{matrix}\right) \to$ unsuccessful decoding

measure $\mu_{\beta,\underline{y}}(\underline{x})$ dominated by codewords at a distance

$$\delta^* = \frac{p^\beta}{p^\beta + (1-p)^\beta}$$

$\left(\begin{matrix}\beta = 1 \to \delta^* = p > \delta_{GV} \\ \text{(symbol MAP)} \\ \beta = 0 \to \delta^* = 1/2 \\ d = N/2 \to \text{complete noise}\end{matrix}\right)$

$H(\mu_{\underline{y},\beta})$ linear in N

$\to$ non zero Entropy <u>NO</u> measure condensation

$\#$ of codewords "dominating" $\mu \sim \exp(NH)$

LARGE DEVIATION THEORY

(review by H. Touchette 2009)

Large deviation principle

→ scaling law $P_n \overset{n \gg 1}{\approx} \exp(-n I)$

$n \gg 1$ P_n prob. ; I = rate function
"decay rate"

random variable A_n

(A_n is "intensive" : $\langle A_n \rangle$ converging for $n \to \infty$)

$$P(A_n \in B) \underset{n \gg 1}{\approx} \exp(-n I_B)$$

$$\lim_{n \to \infty} -\frac{1}{n} \ln\left[P(A_n \in B)\right] = I_B$$

A_n continuous variable ; $p(A_n = a)$

$$\lim_{n \to \infty} -\frac{1}{n} \ln p(A_n = a) = I(a)$$ — rate function

$$P(A_n \in [a, a+da]) = p(A_n = a)\, da$$

$$\begin{array}{c} \parallel \\ P(A_n \in da) \end{array} \overset{n \gg 1}{\approx} \exp(-n I(a))\, da$$

Examples of large deviations (rate function can be computed directly)

RANDOM BITS

sequence $b = (b_1, \dots, b_n)$ n bits (independent random bits)

$$R_n = \frac{1}{n}\sum_{i=1}^{n} b_i \qquad 0 \le R_n \le 1 \qquad \text{(fair coins)}$$

$$(n \to \infty \Rightarrow R_n \in [0,1])$$

$$P(R_n = r) = \frac{1}{2^n}\underbrace{\binom{n}{rn}} = \frac{1}{2^n}\frac{n!}{(rn)!\,[(1-r)n]!}$$

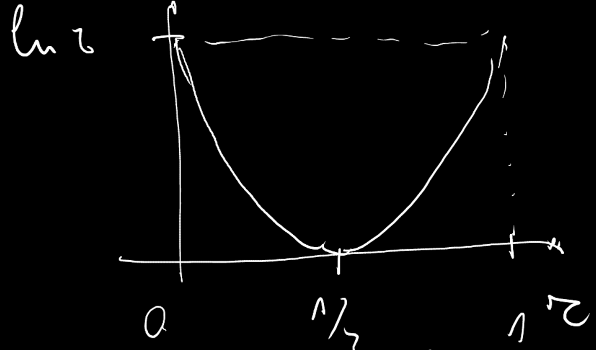
# of ways I can choose rn bits $= 1$ out of n bits

$$P(R_n = r) \underset{n \gg 1}{\approx} \exp(-n I(r))$$

$I(r) = -$ "growth" rate (+ additive constants)

$$I(r) = \ln 2 + r \ln r + (1-r)\ln(1-r)$$

$$\overset{!}{=} \ln 2 - \mathcal{H}(r)$$ ~ Shannon entropy of Bernoulli process



typical value of R_n $(n \to \infty) \to r = 1/2$

$0 \le r \le 1$

$I(r) \ge 0 \quad \forall\, 0 \le r \le 1$

$I(r)$ convex

$I(r)$ has a unique minimum (for $r = 1/2$) / $I(r = 1/2) = 0$

"Gaussian Sample Mean"

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad X_i \text{ i.i.d Gaussian variables}$$

$$p(X_i = x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mu - x_i)^2}{2\sigma^2}\right) \qquad \text{mean } \mu, \ \text{variance } \sigma^2$$

$$p(S_n = s) = \int_{\mathbb{R}^n} \delta(S_n(x) - s)\, p(x)\, dx \qquad p(x) = \prod_{i=1}^{n} p(x_i)$$

x = vector

$$\delta(S_n(x) - s) = \int \frac{dw}{2\pi} \exp\left(iw\left(s - \frac{\sum_i x_i}{n}\right)\right) \qquad (x_1, \dots, x_n)$$

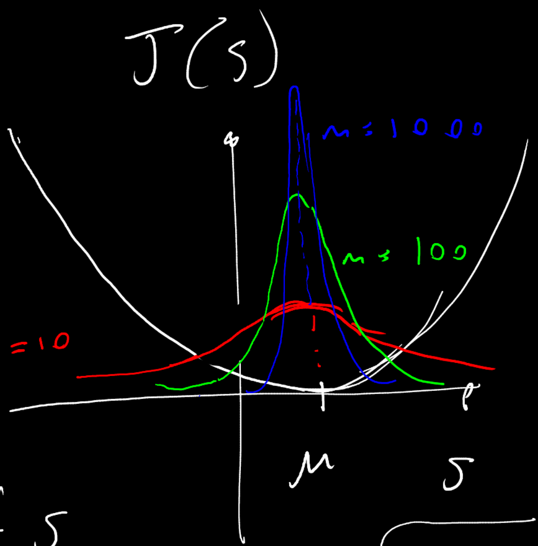
$$p(S_n = s) = \sqrt{\frac{n}{2\pi\sigma^2}} \exp\left(-n\frac{(s-\mu)^2}{2\sigma^2}\right)$$

gaussian with mean μ ; variance σ^2/n

Large deviation principle:

$$\boxed{P(S_n = s) \approx \exp(-n J(s))}$$

$$J(s) = \frac{(s-\mu)^2}{2\sigma^2} \qquad s \in \mathbb{R}$$



$J(s)$ convex ; $J(s) \geq 0 \ \forall s$

$J(s)$ single minimum $(s = \mu)$ / $J(\mu) = 0$

"Law of large numbers" $\rightarrow \lim_{n\to\infty} P(S_n \in [\mu - \delta, \mu + \delta]) = 1 \quad \forall \delta > 0$

"Symbol frequencies" $w = (w_1, w_2, \dots, w_n)$

sequence of i.i.d random variables drawn from the set

$\Lambda = \{1, \dots, q\}$ $\quad P(w_i = j) = \rho_j > 0$

$$\sum_{1}^{q} \rho_j = 1$$

$$L_{n,j}(w) = \frac{1}{n} \sum_{i=1}^{n} \delta_{w_i, j}$$

vector of random variables $\longrightarrow$ $L_n(w) = (L_{n,1}(w), \dots, L_{n,q}(w))$

empirical frequency

"empirical vector"

$e = (e_1, \dots, e_q)$

$$P(L_n = e) = \frac{n!}{\prod_{j=1}^{q} (n e_j)!} \prod_{j=1}^{q} \rho_j^{n e_j}$$

multinomial distributions

$$\underset{n \gg 1}{\approx} \exp[-n\, I_\rho(e)]$$

$$I_\rho(e) = \sum_{j=1}^{q} e_j \ln[e_j/\rho_j]$$

K-L divergence between e, ρ

$I_\rho(e) \geq 0$; $I_\rho(e)$ convex

$I_\rho(e) = 0 \iff e = \rho$