$$H(Z|X) = \underbrace{P(X=b)}_{1/2}\, H(Z|X=b) + \underbrace{P(X=2b)}_{1/2}\, H(Z|X=2b)$$

$p(z|x)$

Z \ X	1	2
1	$1-\gamma(x)$	$\gamma(x)$
2	$\gamma(x)$	$1-\gamma(x)$

$$H(Z|X) = \frac{1}{2}\left[H\left(\frac{e^{-b}}{e^{-b}+e^{b}}\right) + H\left(\frac{e^{-2b}}{e^{-2b}+e^{2b}}\right)\right]$$

$$= \frac{1}{2}\left[H(\alpha) + H(\beta)\right]$$

$$\alpha = \frac{e^{-b}}{e^{-b}+e^{b}} \qquad \beta = \frac{e^{-2b}}{e^{-2b}+e^{2b}}$$

$$P(Z=x_2) = P(X=x_1)\cdot P(Z=x_2 | X=x_1) + P(X=x_2)\, P(Z=x_2 | X=x_2)$$

$$= \frac{1}{2}\frac{e^{b}}{e^{-b}+e^{b}} + \frac{1}{2}\left(1 - \frac{e^{-2b}}{e^{-2b}+e^{2b}}\right)$$

$$P(Z=Y) = P(X)\, P(Z=Y|X) + P(Y)\, P(Z=Y|Y)$$

$$= \frac{1}{2}(1-\gamma(x)) + \frac{1}{2}\gamma(x)$$

$$\mu_1 = \frac{\frac{e^{-2b}}{e^{-2b}+e^{2b}}}{\frac{e^{-b}}{e^{-b}+e^{b}} + \frac{e^{-2b}}{e^{-2b}+e^{2b}}} = \frac{1}{2}$$

$$\frac{e^{-2b}}{e^{-2b}+e^{2b}} = \frac{e^{-b}}{2(e^{-b}+e^{b})} + \frac{e^{-2b}}{2(e^{-2b}+e^{2b})} \qquad \frac{e^{-b}}{2(e^{-b}+e^{b})} = \frac{e^{-2b}}{2(e^{-2b}+e^{2b})}$$

$$\frac{1}{e^{-b}+e^{+b}} = \frac{e^{-b}}{e^{-2b}+e^{2b}} \qquad e^{-2b}+e^{2b} = e^{-2b}+1$$

$$e^{2b} = 1 \qquad 2b \ln e = 0 \Rightarrow \boxed{b=0}$$

$$W(b,\gamma) = \sum_{x=1}^{2} \gamma(x)\cdot \mathrm{ld}\, o(x) + \sum_{x=1}^{2} \gamma(x)\cdot \mathrm{ld}\, b(x) =$$

$$= \underbrace{\left(\frac{1}{2}\mathrm{ld}\, 1 + \frac{1}{2}\mathrm{ld}\, 2\right)}_{1/2} + \frac{1}{2}\mathrm{ld}\frac{e^{-b}}{e^{-b}+e^{b}} + \frac{1}{2}\mathrm{ld}\frac{e^{-2b}}{e^{-2b}+e^{2b}}$$

$$\frac{1}{n}\mathrm{ld}\, S_n = \frac{1}{n}\mathrm{ld}\prod S(x) = \frac{1}{n}\sum \mathrm{ld}\, S(x) \to E[\mathrm{ld}\, S(x)] = W(\gamma,b)$$

$$\boxed{S_n = 2^{n\cdot W(\gamma,b)}}$$

$b = 1$

$$W(b,Y) = \frac{1}{2} + \frac{1}{2}\, ld \frac{e^{-1}}{e^{-1}+e^{1}} + \frac{1}{2}\, ld \frac{e^{-2}}{e^{-2}+e^{2}} =$$

$$= \frac{1}{2} + \frac{1}{2}\, ld\left(\frac{1}{1+e^2}\cdot\frac{1}{1+e^4}\right) =$$

$$\frac{e^{-b}}{e^{-b}+e^{b}} \cdot \frac{e^{-2b}}{e^{-2b}+e^{2b}} = \frac{1}{1+e^{2b}} \cdot \frac{1}{1+e^{4b}} =$$

$$\boxed{a = e^{2}}$$

$$= \frac{1}{(1+x^{b})} \cdot \frac{1}{(1+x^{2b})}$$

$$= \frac{1}{1+a^{2b}+a^{b}+a^{3b}} = \frac{1}{1+a^{b}+a^{2b}+a^{3b}} > \frac{1}{\sum_{n=0}^{\infty} e^{2b\cdot n}}$$

$$= \frac{1}{\frac{1}{1-e^{2b}}} = \underline{1-e^{2b}}$$

$$W(b,q) = \frac{1}{2} + \frac{1}{2}\, ld \frac{e^{-x}}{e^{-x}+e^{x}} + \frac{1}{2}\, ld\left(1-\frac{e^{-x}}{e^{-x}+e^{x}}\right)$$

$$= \frac{1}{2} + \frac{1}{2}\, ld \frac{e^{-x}}{e^{-x}+e^{x}} + \frac{1}{2}\, ld \frac{e^{x}}{e^{-x}+e^{x}} =$$

$$= \frac{1}{2} + \frac{1}{2}\, ld \frac{1}{(e^{-x}+e^{x})^2} = \frac{1}{2} + ld \frac{1}{e^{-x}+e^{x}}$$

$$= \frac{1}{2} - ld\left(e^{-x}+e^{x}\right) = \frac{1}{2} - ld\, e^{x}\left(1+e^{2x}\right) =$$

$$= \frac{1}{2} + xV_{a(e)} - ld\left(1+e^{2x}\right)$$

• EDITION 2 SOLUTIONS

(b) $E[Y|X] = P(X=x)\cdot E[Y|X=x] + P(X=x)\cdot E[Y|X=x$

$$= \frac{1}{2}\underbrace{2x} + \frac{1}{2}\cdot\underbrace{\frac{x}{2}} = x + \frac{x}{4} = \frac{5x}{4}$$

$Y \in \{2x \ldots \frac{x}{2}\}$

$$E[Y|X=x] = \sum Y\, P(Y|X=x) = 2x\cdot\frac{1}{2} + \frac{x}{2}\cdot\frac{1}{2} =$$

$$= x + \frac{x}{4} = \frac{5x}{4}$$

(c) CONDITIONED ON $J=1$ $X=2\theta$, PROBABILITY THAT $Z=X$ i.e. $J'=1$ is $p(2\theta)$

– SIMILARLY CONDITIONED ON $J=1$ $X=\theta$ ⟹ PROB. $J'=1$ is $1-p(\theta)$

$$P(J'=1|J=1) = P(X=\theta|J=1)\,P(J'=1|J=1, X=\theta) + P(X=2\theta|J=1)\,P(J'=1|J=1, X=2\theta) =$$

$$= \frac{1}{2}[1-p(\theta)] + \frac{1}{2}\cdot p(2\theta) = \frac{1}{2} + \frac{1}{2}[p(2\theta)-p(\theta)]$$

$$> \frac{1}{2} \Rightarrow J \,\&\, J' \text{ ARE NOT INDEPENDENT}$$

SINCE CONDITIONAL DISTRIBUTION IS NOT EQUAL TO THE UNCONDITIONAL DISTRIB.

$$P(J') = \left\{\frac{1}{2}, \frac{1}{2}\right\}$$

VLE OSNOV ZA OVA IZRESUVA DEKA $p(J'=1|J=1) < \frac{1}{2}$

$p(\theta) \geq p(2\theta)$ $\forall \theta > 0$ VIDI MAPLE

(d) $E[Z|J=1] = P(X=\theta|J=1)\,E[Z|X=\theta, J=1] +$

$+ P(X=2\theta|J=1)\cdot E[Z|X=2\theta, J=1] = \frac{1}{2} E[Z|X=\theta, J=1] +$

$+ \frac{1}{2} E[Z|X=2\theta, J=1] =$

$E[Z|X=\theta, J=1] = Y\,p(X) + X\cdot(1-p(X)) =$

$= 2\theta\,p(2\theta) + \theta[1-p(\theta)]$

$E[Z|X=2\theta, J=1] = X\cdot(1-p(X)) + Y\cdot p(X) =$

$= 2\theta[1-p(2\theta)] + \theta\,p(\theta)$

$= \frac{1}{2} P(J'=1|X=\theta, J=1)\cdot E[Z|J'=1, X=\theta, J=1] +$

$+ \frac{1}{2} P(J'=2|X=\theta, J=1)\cdot E[Z|J'=2, X=\theta, J=1] +$

$+ \frac{1}{2} P(J'=1|X=2\theta, J=1)\cdot E[Z|J'=1, X=2\theta, J=1] +$

$+ \frac{1}{2} P(J'=2|X=2\theta, J=1)\cdot E[Z|J'=2, X=2\theta, J=1] =$

$$= \frac{1}{2}\left[1-\gamma(b)\right] E[Z \mid J'=1, X=b, J=1] + \frac{1}{2}[\gamma(b)] \cdot E[Z \mid J'=2, X=b, J=1]$$

$$+ \frac{1}{2}\gamma(2b) \cdot E[Z \mid J'=1, X=2b, J=1] + \frac{1}{2}[1-\gamma(b)] \cdot E[Z \mid J'=2, X=2b, J=1]$$

ОДОБРАТНО ТРЕБА

$$= \frac{1}{2}(1-\gamma(b)) \cdot 2b + \frac{1}{2}\gamma(b) \cdot b + \frac{1}{2}\gamma(2b) \cdot 2b + \frac{1}{2}[1-(2b)] \cdot b$$

ЈАС МИСЛАМ ДЕКА ТРЕБА ВАКА: (АХА ... ЗАБОРАВИ ОВА ШТО ГО РАЗБИРАМ !!!)

$$= \frac{1}{2}[1+\gamma(b)] \cdot b + \frac{1}{2}\gamma(b) 2b + \frac{1}{2}[1-\gamma(2b)] \cdot 2b + \frac{1}{2}\gamma(2b) \cdot b$$

$$= \frac{b}{2} - \frac{\gamma(b) \cdot b}{2} + b \cdot \gamma(b) + \frac{2b}{2} - 2b \cdot \gamma(2b) + \frac{b}{2}\gamma(2b) =$$

$$= \frac{3b}{2} + \frac{b\,\gamma(b)}{2} - \frac{3b}{2}\gamma(2b) = \frac{3b}{2} + \frac{b}{2}\left[\gamma(b) - 3\gamma(2b)\right]$$

→ REVISITED (СЕГА ЗАКЛУЧУВАМ ДЕКА J ПОКАЖУВА НА ПОГОЛЕМОТО ПЛИКО И ДЕКА X ЗА СОРПИ ВРЕДНОСТА ОД ИЗБОРОТ (ИЗБОР))

$$E[Z \mid J=1] = \frac{1}{2}P(J'=1 \mid X=b, J=1) \cdot E[Z \mid J'=1, X=b, J=1] +$$

$$+ \frac{1}{2}P(J'=2 \mid X=b, J=1) \cdot E[Z \mid J'=2, X=b, J=1] + \frac{1}{2}P(J'=1 \mid X=2b, J=1) \cdot$$

$$\cdot E[Z \mid J'=1, X=2b, J=1] + \frac{1}{2}P(J'=2 \mid X=2b, J=1) \cdot E[Z \mid J'=2, X=2b, J=1]$$

$$= \frac{1}{2}[1-\gamma(b)] \cdot 2b + \frac{1}{2}\gamma(b) \cdot b + \frac{1}{2}(1-\gamma(2b)) \cdot 2b + \frac{1}{2}\gamma(2b) \cdot b =$$

ОСТАНУВА НА X И ДОБИВА — SWITCH-ува НА ... СЕ ЗАБОРАВА — ОСТАНУВА ДОБИВА — SWITCH-ува ГУБИ

$$= b - \gamma(b) \cdot b + \frac{1}{2}\gamma(b) \cdot b + 2b - 2b\,\gamma(2b) + \frac{b}{2}\gamma(2b) =$$

$$= 2b - \frac{1}{2}\gamma(b) \cdot b - \frac{1}{2}\gamma(2b) b = 2b - \frac{b}{2}\left[\gamma(b) + \gamma(2b)\right]$$

< 1 (ВИДИ ГОРЕ)

$$> 2b - \frac{b}{2} = \frac{3b}{2} \qquad \left(E[Z \mid J=1] > \frac{3b}{2} = E[X]\right)$$

(c) REVISITED: $P(J'=1 \mid J=1) = P(X=b \mid J=1) \cdot P(J'=1 \mid X=b, J=1)$

$$+ P(X=2b \mid J=1) \cdot P(J'=1 \mid X=2b, J=1) = \frac{1}{2} \cdot [1-\gamma(b)] +$$

$$+ \frac{1}{2}[1-\gamma(2b)] = 1 - \frac{1}{2}\left[\gamma(b) + \gamma(2b)\right] > \frac{1}{2} \quad (\text{ЗА } b \neq 0)$$

< 1 (ВИД ГОРЕ)

→ НЕМА ЗАВИСНОСТ ПОМЕЃУ ИЗБОРОТ ИЗБОР И МОМЕНТАЛНАТА ПОЗИЦИЈА (ИНДЕКС) НА ПОГОЛЕМОТО ПЛИКО.

$$P(J'=1 \mid J=1) > \frac{1}{2} = P(J') = P(J') \Rightarrow J \,\&\, J' \text{ ARE NOT INDEPENDENT}$$

PROBLEM 2 — DAMAGING. FIND HOUSE WIN PROBABILITIES $q_1, q_2, \ldots, p_m$

(a) MAXIMIZE THE DOUBLING RATE FOR GIVEN FIXED KNOWN ODDS.

(b) MINIMIZE THE DOUBLING RATE FOR GIVEN FIXED ODDS.

(a) $W(b,q) = \sum_x q(x) \cdot \operatorname{ld} b(x) o(x) =$

$= \sum_x q(x) \cdot \operatorname{ld} \frac{1}{\frac{1}{o(x)}} - \sum_x q(x) \cdot \operatorname{ld} \frac{1}{b(x)} =$

$= \sum_x q(x) \cdot \operatorname{ld} \frac{1}{r(x)} \cdot \frac{q(x)}{q(x)} \cdot b(x) = \sum_x q(x) \cdot \operatorname{ld} \frac{q(x)}{r(x)} - \sum_x q \operatorname{ld} \frac{q}{b}$

$W(b,q) = D(q\|r) - D(q\|b)$

↳ BOOKIE ESTIMATE

$\max [W(b,q)] = W^*(b,q) = D(q\|r)$ i.e $q = b$

(b) $\min [W(b,q)] = 0$ $b = r$ THE DOUBLING RATE IS MINIMAL i.e 0 IF YOU BET ACCORDING BOOKIE ESTIMATE

- DA IMA MIN I NEGATIVNA DA E DOUBLING RATE TREBA BUKMEJKERO DA POGODI VISTINSKATA DISTRIBUCIJA T.E. AKO

$q = r$ ie $q_i = \frac{1}{o_i}$

W_0 ZA SLUČAJ

$$W(b,q) = D(q\|p) - D(q\|b) = -D(q\|b)$$

$D(q\|b)$ e KONVEKSNA FUNKCIJA

$D(q\|b) = \sum q \operatorname{ld} \frac{q}{b}$ } DA SE MAKSIMIZIRA

$\frac{d D(q\|b)}{d b_i} = 0$ $\quad \nabla D(q\|b) + \lambda \nabla \sum b = 0$

$\frac{d}{d b_i}\left(q \operatorname{ld} \frac{q}{b_i}\right) + \lambda = 0$

$$\frac{d}{db_i}\left(p\, ld\frac{p}{b_i}\right) = -\frac{d}{db_i}\left(p_i\frac{\ln b_i}{\ln 2}\right) = \frac{p_i}{\ln 2}\frac{1}{b_i}$$

$$\frac{p_i}{\ln 2\, b_i} = \lambda \qquad \boxed{b_i = \frac{p_i}{\lambda \cdot \ln 2}}$$

$$\sum_i \frac{p_i}{\lambda \cdot \ln 2} = 1 \qquad \frac{1}{\lambda \cdot \ln 2} = 1 \quad \boxed{\lambda = \frac{1}{\ln 2}}$$

$$\Rightarrow \boxed{b_i = p_i} \; ?$$

$$W_*(b, p) = -D(p \| b) = -\sum p\, ld\frac{p}{b} =$$

$$= -\sum p\, ld\frac{1}{b} - \sum p\, ld\, p = -\sum p\, ld\frac{1}{b} + H(p) =$$

$$\left|\boxed{b = \frac{1}{m}}\right| = -ld\, m + H(p) < -ld\, m + ld\, m = \emptyset$$

$$H(p) < ld\, m \quad \Rightarrow \quad \boxed{W_*(p, p)} < \emptyset$$

MINIMAR DOVOLIKUS ...!!

- ZNAČI, AKO BOOKIE-TO ZA PORODI REZULTATA p(T) A TI SE KLADIŠ UNIFORMO TOGAŠ ĆE BIDEŠ VO NAJLOVEVA ZAGUBA. VO NAJDO DOBRO SLUČAJ PRIDOBA NA SE KLADIŠ KAKO BOOKIE-TO ZA DA BIDEŠ NA ~0.

- EDITION 2 SOLUTIONS $\quad W^*(p, b) = \sum_i p_i\, ld\, O_i p_i$

$$= \sum_i p_i\, ld\, O_i - \sum_i p_i\, ld\frac{1}{p_i} = \sum_i p_i\, ld\, O_i - H(p)$$

$$= \sum_i p_i\, ld\frac{p_i}{\frac{1}{O_i}} = \sum_i p_i\, ld\frac{p_i}{\frac{\frac{1}{O_i}}{\sum \frac{1}{O_i}}} \cdot \frac{1}{\sum \frac{1}{O_i}} =$$

$$= \sum_i p_i\, ld\frac{p_i}{q_i} - \sum_i p_i\, ld\sum_i\frac{1}{O_i} =$$

$$\left|\boxed{q_i = \frac{\frac{1}{O_i}}{\sum \frac{1}{O_i}}}\right| = \sum_i p_i\, ld\frac{p_i}{q_i} - ld\sum_i\frac{1}{O_i}$$

$$\min W^*(p, b) = -ld\sum_i\frac{1}{O_i}$$

$$\boxed{p_i = q_i}$$

• Maximum growth rate occurs when horse with maximum odds win in all races i.e. $p_i = 1$ for horse that provides maximum odds.

Ako celo vreme se vlozi so $b=1$ na konjot so najgolema ..., se zgolemuva.

Problem 6.13 Dutch Book. Consider the horse race with $m = 2$ horses:

$x = 1, 2$ $\quad p(x) = \{\frac{1}{2}, \frac{1}{2}\}$ $\quad o(x) = \{10, 30\}$

$b(x) = \{b, 1-b\}$

The odds are superfair.

(a) There is a bet b that guarantees the same payoff regardless of which horse wins. Such bet is called Dutch book. Find this bet and associated wealth factor $S(x)$.

(b) What is the maximum growth rate of the wealth for optimal choice of b? Compare it to the Dutch book.

a) $S(x) = b(x) \cdot o(x)$ $\qquad b \cdot 10 = (1-b) \cdot 30$

$10b = 30 - 30b$ $\qquad 40b = 30$ $\qquad \boxed{b = \frac{3}{4}}$

$S(x) = 10 \cdot \frac{3}{4} = \frac{30}{4} = 30\left(1 - \frac{3}{4}\right) = 30 \cdot \frac{1}{4} = \frac{30}{4}$

$\boxed{S(x) = 30/4}$ $\qquad S_n = 2^{2.9} = 7.5$ $\qquad W_D = \frac{1}{2} \cdot ld \frac{3}{4} \cdot 10 + \frac{1}{2} ld \frac{1}{4} \cdot 30 = ld \frac{30}{4} = 2.9$

b) $b(x) = p(x)$ $\qquad \boxed{b = \frac{1}{2}}$

$W^*(b, p) = \sum_x p(x) \cdot ld\, o(x) - H(\frac{1}{2}) =$

$= \frac{1}{2} \cdot ld 10 + \frac{1}{2} ld 30 - 1 = 4.11 - 1 = 3.11$

$W^* = \{E[ld S(x)]\}$ $\qquad S_n = 2^{n S(x)}$

$ld \frac{30}{4} = 2.9$ $\qquad \boxed{W_D = 2.9 < 3.11 = W_*}$

— KRAJNIOT PODELEN SE ODNESUVA SO MAKSIMUM ... MOŽE.

$\frac{1}{n} ld S_n = \frac{1}{n} ld \prod S(x) = \frac{1}{n} \sum ld S(x) \rightarrow$ (i.i.d.)

$E[ld S(x)] = W(x)$ $\qquad W(X, b) = E[ld S(x)]$

$\boxed{S_n = 2^{n \cdot W(X, b)}}$ $\qquad S_n = 2^{3.11} = 8.66$

PROBLEM 6.14 HORSE RACE. $p_1, p_2, \dots p_m$ WIN PROBABILITIES OF m HORSES. WHEN $(O_1, O_2, \dots O_m)$ YIELD HIGHER DOUBLING RATE THEN $(O_1', O_2', \dots O_m')$?

$$W(p,b) = \sum p_i \, ld\, b_i O_i \qquad W'(p,b) = \sum p_i \, ld\, b_i O_i'$$

$$W(p,b) - W'(p,b) = \sum_i p_i \, ld\, b_i O_i - \sum_i p_i \, ld\, b_i O_i' =$$

$$= \boxed{\sum_i p_i \, ld O_i - \sum_i p_i \, ld O_i'} \; (*) = \sum_i p_i \, ld \frac{O_i}{O_i'} \cdot \frac{p}{p} =$$

$$= \sum_i p_i \, ld \frac{p_i}{r_i} - \sum p_i \, ld \frac{p}{r_i'} = D(p\|r_i) - D(p\|r_i') > 0$$

$$D(p\|r_i) > D(p\|r_i') \quad \text{IFF} \quad (*) \; \boxed{E[\log O_i] > E[\log O_i']}$$

PROBLEM 6.15 ENTROPY OF A FAIR HORSE RACE

LET $X \sim p(x)$, $x = 1, 2, \dots m$, DENOTE THE WINNER OF A HORSE RACE. SUPPOSE THAT THE ODDS $O(x)$ ARE FAIR WITH RESPECT TO $p(x)$ [i.e. $O(x) = \frac{1}{p(x)}$]. LET $b(x)$ BE THE AMOUNT BET ON HORSE x: GIVEN $\sum_1^m b(x) = 1$. THEN THE RESULTING WEALTH FACTOR IS $S(x) = O(x) \cdot b(x)$ WITH PROBAB.: $p(x)$.

(a) FIND $E[S(x)]$ (EXPECTED WEALTH)

(b) FIND W^* OPTIMAL GROWTH RATE OF WEALTH.

(c) SUPPOSE: $Y = \begin{cases} 1 & x = 1 \text{ OR } 2 \\ 0 & \text{OTHERWISE} \end{cases}$

IF THIS SIDE INFORMATION IS AVAILABLE BEFORE THE BET. HOW MUCH DOES IT INCREASE THE GROWTH RATE W^*?

(d) FIND $I(X;Y)$

(a) $E[S(x)] = \sum_x p(x) \cdot S(x) = \sum_x p(x) \cdot b(x) \cdot O(x) =$

$$= \sum_x p(x) \cdot b(x) \cdot \frac{1}{p(x)} = \sum_x b(x) = 1$$

(b) $$W^* = \sum_x p(x) \cdot ld\, O(x) \cdot b(x) = \sum_x p(x) \cdot ld \frac{b(x)}{p(x)}$$

$$\nabla W^*(b_1) - \lambda \nabla b(x) = 0 \qquad \frac{d}{db_i}\left(p_i \, ld \frac{b_i}{p_i}\right) = p_i \cdot \frac{1}{b_i / p_i}$$

$$\frac{p_i^2}{b_i} - \lambda = 0 \qquad p_i^2 = \lambda \cdot b_i \qquad b_i = \frac{p_i^2}{\lambda}$$

$$\sum_i b_i = \sum_i \frac{p_i^2}{\lambda} = \frac{1}{\lambda} \sum p_i^2 = 1$$

$$p_i^2 = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\} \cdot \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\} = \left\{\frac{1}{4}, \frac{1}{16}, \frac{1}{16}\right\}$$

$$\sum p_i^2 = \frac{1}{4} + 2\frac{1}{16} = \frac{1}{4} + \frac{1}{8} = \frac{3}{8}$$

$$\boxed{\lambda = \sum p_i^2} \qquad \sum p_i^2 = E[p_i] \qquad \boxed{b_i = \frac{p_i^2}{\sum p_i^2}}$$

$$W = \sum_x p(x) \cdot ld \frac{\frac{p^2(x)}{\sum p^2(x)}}{p(x)} = \sum_x p(x)\, ld \frac{p(x)}{\sum_x p^2(x)} =$$

$$= \sum p(x) \cdot ld \frac{1}{\sum_x p^2(x)} - H(x)$$

$$W = \sum_i p_i\, ld \frac{b_i}{p_i} = H(p) + \sum_i p_i\, ld\, b_i$$

$$W = -\sum p_i\, ld \frac{p_i}{b_i} = -D(p\|b) \Rightarrow p = b \Rightarrow 0$$

$$\geq 0 \qquad \boxed{W^* = 0} \text{ MAXIMAL (OPTIMAL)}$$

GROWTH RATE.

ISTOVO MOZE SO LAGRANGE MULTIPLIERS:

$$\nabla W(b,p) = \lambda \nabla b(x) \qquad W(b,p) = \sum p(x)\, ld\, b(x) + H(x)$$

$$\nabla W(b,p) \;\forall b_i \qquad \frac{d\, p(x)\, ld\, b_i}{d b_i} - \lambda = 0$$

$$\boxed{p_i \cdot \frac{1}{b_i} = \lambda} \qquad b_i = \frac{p_i}{\lambda} \qquad \sum b_i = \sum \frac{p_i}{\lambda} = 1 \Rightarrow \boxed{\lambda = 1}$$

$$\Rightarrow p_i = b_i \quad \} \text{ OPTIMAL SOLUTION.}$$

(c) $$Y = \begin{cases} 1 & x = 1 \text{ or } 2 \\ 0 & \text{OTHERWISE} \end{cases}$$

$$W(X|Y) = \sum p(x) \cdot ld\, o(x) \cdot b(x|y) \Rightarrow$$

$$W^*(X|Y) = \sum_{x,y} p(x,y) \cdot ld\, o(x) \cdot p(x|y) = \sum p(x)\, ld\, o(x) + \sum p(x,y)\, ld\, p(x|y)$$

$$= \sum p(x)\, ld\, o(x) - \sum_{x,y} p(x,y)\, ld \frac{1}{p(x|y)} = H(x) - H(x|y)$$

$$W^*(X) = \sum p(x) \cdot ld\, o(x) p(x) = \sum_x p(x)\, ld\, o(x) - H(X)$$

$$W^*(X|Y) - W^*(X) = \sum p(x)\, ld\, o(x) - H(X|Y) -$$

$$- \sum_x p(x)\, ld\, o(x) + H(X) = H(X) - H(X|Y) = \underline{I(X;Y)}$$

(b) $H(X) = \sum_{i=1}^{n} p(x)\, ld \frac{1}{p(x)}$

$$H(X|Y) = \underline{p(Y=1)} \cdot H(X|Y=1) + p(Y=0) \cdot H(X|Y=0)$$

$$= p(Y=1) \, H(x_1, x_2) + p(Y=0) \cdot H(X_3^n)$$

$$= p(Y=1) \, H(x_1^2) + [1 - p(Y=1)] \, H(x_3^n)$$

$$I(X;Y) = H(X_1^n) - p(Y=1) H(X_1^2) - H(X_3^n) + p(Y=1) H(X_3^n)$$

$$I(X;Y) = \sum_{i=1}^{n} p_i\, ld \frac{1}{p_i} - p(Y=1) \sum_{i=1}^{2} p_i\, ld \frac{1}{p_i} - \sum_{i=3}^{n} p_i\, ld \frac{1}{p_i}$$

$$+ p(Y=1) \sum_{i=3}^{n} = \sum_{i=1}^{2} p_i\, ld \frac{1}{p_i} - p(Y=1) \sum_{i=1}^{2} p_i\, ld \frac{1}{p_i}$$

$$+ p(Y=1) \sum_{i=3}^{n} p_i\, ld \frac{1}{p_i} = \underline{p(Y=0) \sum_{i=1}^{2} p_i\, ld \frac{1}{p_i}} +$$

$$\underline{+ p(Y=1) \sum_{i=3}^{n} p_i\, ld \frac{1}{p_i}} = p(Y=0) \sum_{i=1}^{2} p_i\, ld \frac{1}{p_i} +$$

$$+ p(Y=1) \left[H(X) - \sum_{i=1}^{2} p_i\, ld \frac{1}{p_i} \right] =$$

$$= \sum_{i=1}^{2} p_i\, ld \frac{1}{p_i} \left[p(Y=0) - p(Y=1) \right] + p(Y=1)\, H(X) =$$

$$p(Y=1) = p(X=x_1) + p(X=x_2)$$

$$= \left[1 - p(Y=1) - p(Y=1) \right] \sum_{i=1}^{2} p_i\, ld\, p_i^{-1} + p(Y=1)\, H(X) =$$

$$= \left[1 - 2(p_1 + p_2) \right] \sum_{i=1}^{2} p_i\, ld\, p_i^{-1} + (p_1 + p_2)\, H(X)$$

$$= \left[1 - (p_1 + p_2) \right] \sum_{i=1}^{2} p_i\, ld\, 1/p_i + (p_1 + p_2) \sum_{i=3}^{n} p_i\, ld \frac{1}{p_i}$$

$$= \sum_{i=3}^{m} p_i \cdot \sum_{i=1}^{2} p_i \, ld \frac{1}{p_i} + \sum_{i=1}^{2} p_i \cdot \sum_{i=3}^{m} p_i \, ld \frac{1}{p_i}$$

$[p_1, p_2, p_3]$

$$p_3 \cdot \sum_{i=1}^{2} p_i \, ld \frac{1}{p_i} + (p_1 + p_2)\, p_3 \, ld \frac{1}{p_3} =$$

$$= p_1 \cdot p_3 \cdot ld \frac{1}{p_1} + p_3 \cdot p_2 \, ld \frac{1}{p_2} + p_1 \cdot p_3 \, ld \frac{1}{p_3} + p_2 p_3 \, ld \frac{1}{p_3}$$

$$= p_1 \cdot p_3 \left(ld \frac{1}{p_1} + ld \frac{1}{p_3} \right) + p_2 p_3 \left(ld \frac{1}{p_2} + ld \frac{1}{p_3} \right)$$

$$I(X;Y) = p \cdot \sum_{i=1}^{2} p_i \, ld \frac{1}{p_i} + (1-p) \sum_{i=3}^{m} p_i \, ld \frac{1}{p_i}$$

$p = p(Y=1)$ IF $p = \frac{1}{2}$

$$I(X;Y) = \frac{1}{2} \sum_{i=1}^{m} p_i \, ld \frac{1}{p_i} = \frac{1}{2} H(X)$$

EDITION 2 SOLUTION

(c) (d) $q = P(Y=1) = p(x_1) + p(x_2)$

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

(Y is a function of X and so zero)

$$I(X;Y) = H(Y) = q \, ld \frac{1}{q} + (1-q) \, ld \frac{1}{1-q} = H(q)$$

PROBLEM 6.16 NEGATIVE HORSE RACE. CONSIDER A HORSE RACE WITH m HORSES WITH WIN PROBABILITIES $p_1, p_2, \ldots p_m$. HERE THE GAMBLER HOPES THAT A GIVEN HORSE WILL LOSE. HE PLACES BETS $(b_1, b_2, \ldots b_m)$, $\sum_i b_i = 1$ ON THE HORSES, LOSES HIS BET b_i IF HORSE i WINS AND RETAINS THE REST OF HIS BETS. (NO ODDS). THUS, $S = \sum b_j$ WITH PROBABILITY p_i, AND ONE WISHES TO MAXIMIZE $\sum p_i \, ld (1 - b_i)$ SUBJECT

TO CONSTRAINT: $\sum b_i = 1$.

(a) FIND THE GROWTH RATE OPTIMAL INVESTMENT STRATEGY b^*. DO NOT CONSTRAIN THE BETS TO BE POSITIVE, BUT DO CONSTRAIN THE BETS TO SUM TO 1. (THIS EFFECTIVELY ALLOWS SHORT SELLING AND MARGIN.)

(b) WHAT IS THE OPTIMAL GROWTH RATE?

(a) $S(x) = 1 - b(x)$ } DOBITOČNOTO ŠTO GO DOBIVAŠ E OVA NAMALENO ZA OVA NA KONTOT ŠTO IZLOŽE

$$W(b,p) = \sum p(x) \cdot \operatorname{ld} S(x) = \sum_{i=1}^{m} p_i \operatorname{ld}(1-b_i)$$

$\sum b_i = 1$ $\qquad \dfrac{\partial\, p_i \operatorname{ld}(1-b_i)}{\partial b_i} + \lambda = 0$

$$-\frac{p_i}{\ln 2} \cdot \frac{1}{(1-b_i)} + \lambda = 0 \qquad \boxed{\lambda = \frac{p_i/\ln 2}{1-b_i}}$$

$\sum b_i = 1$ $\qquad 1 - b_i = \dfrac{p_i/\ln 2}{\lambda}$ $\qquad b_i = \left(1 - \dfrac{p_i/\ln 2}{\lambda}\right)$

$$\sum_{i=1}^{m}\left(1 - \frac{p_i/\ln 2}{\lambda}\right) = 1 \Rightarrow m - \frac{1}{\ln 2\,\lambda}\sum_{i=1}^{m} p_i = 1$$

$m - \dfrac{1}{\lambda \ln 2} = 1$ $\qquad m - 1 = \dfrac{1}{\lambda \ln 2}$ $\qquad \boxed{\lambda \ln 2 = \dfrac{1}{m-1}}$

$$b_i = 1 - \frac{p_i/\ln 2}{1/\ln 2 \cdot (m-1)} \leftarrow \lambda = \frac{1}{\ln 2 \cdot (m-1)}$$

$$\boxed{b_i = 1 - p_i(m-1)}$$

(b) $W^*(p,b) = \sum p_i \operatorname{ld}(1-b_i) = \sum p_i \operatorname{ld}\left[1 - 1 + p_i(m-1)\right]$

$$W^*(p,b) = \sum p_i \operatorname{ld} p_i(m-1) = \operatorname{ld}(m-1) - H(p)$$

EDITION 2 SOLUTION $\qquad b_i' = 1 - b_i \qquad \sum_{i=1} b_i' = m-1$

$$q_i = \frac{b_i'}{\sum_i b_i'} = \frac{b_i'}{m-1} \qquad W = \sum_i p_i \operatorname{ld}(1-b_i) =$$

$$= \sum p_i \operatorname{ld} b_i' = \sum_i p_i \operatorname{ld}\left[(m-1)q_i\right] = \operatorname{ld} m + \sum p_i \operatorname{ld} q_i$$

$$= \operatorname{ld} m + \sum p_i \operatorname{ld} \frac{p_i q_i}{p_i} = \operatorname{ld} m - D(p\|q) - H(p)$$

$\boxed{p = q} \Rightarrow$ MAXIMAL W $\qquad \dfrac{b_i'}{m-1} = p \qquad \dfrac{1-b_i}{m-1} = p$

$$b = 1 - q(m-1) \Rightarrow D(q\|?) = 0 \Rightarrow$$

$$W^*(b,q) = \log m - H(q)$$

PROBLEM 6.17 ST. PETERSBURG PARADOX. MANY CENTURIES AGO IN ANCIENT ST. PETERSBURG THE FOLLOWING GAMBLING PROPOSITION CAUSED GREAT CONSTERNATION. FOR AN ENTRY FEE OF "c" UNITS A GAMBLER RECEIVES A PAYOFF OF "2^k" UNITS WITH PROBABILITY 2^{-k}, $k = 1, 2, \ldots$

(a) SHOW THAT EXPECTED PAYOFF FOR THIS GAME IS INFINITE. FOR THIS REASON, IT WAS ARGUED THAT $c = \infty$ WAS A FAIR PRICE TO PAY TO PLAY THIS GAME. MOST PEOPLE FIND THIS ANSWER ABSURD.

(b) SUPPOSE THAT GAMBLER CAN BUY A SHARE OF THE GAME. FOR EXAMPLE, IF HE INVESTS $c/2$ UNITS IN THE GAME, HE RECEIVES 1/2 A SHARE AND A RETURN $X/2$, WHERE $Pr(X = 2^k) = 2^{-k}$, $k = 1, 2, \ldots$ SUPPOSE THAT $X_1, X_2, \ldots$ ARE I.I.D. ACCORDING TO THIS DISTRIBUTION, AND THAT THE GAMBLER REINVESTS ALL HIS WEALTH EACH TIME. THUS HIS WEALTH S_n AT TIME "n" IS GIVEN BY:

$$S_n = \prod_{i=1}^{n} \frac{X_i}{c}$$

SHOW THAT THIS … PROBABILITY 1 ACCORD… $c > c^*$. IDENTIFY THE … MORE REALISTICALLY, THE … ALLOWED TO KEEP …

$$\bar{b} = 1 - b$$

OF HIS MONEY IN HIS POCKET AND INVEST THE REST IN ST. PETERSBURG, ETC. HIS WEALTH AT TIME "n" IS THEN:

$$S_n = \prod_{i=1}^{n} \left(\bar{b} + \frac{b X_i}{c} \right)$$

LET:

$$W(b,c) = \sum_{k=1}^{\infty} 2^{-k} \log\left(1 - b + \frac{b\, 2^k}{c}\right)$$

WE HAVE: $S_n \doteq 2^{nW(b,c)}$, LET: $W^*(c) = \max_{0 \le b \le 1} W(b,c)$

- HERE ARE SOME QUESTIONS ABOUT $W^*(c)$:

(a) FOR WHAT VALUE OF THE ENTRY FEE "c" DOES THE OPTIMIZING VALUE b^* DROP BELOW "1"?

(b) HOW DOES b^* VARY WITH c?

(c) HOW DOES $W^*(c)$ FALL OFF WITH c?

NOTE THAT SINCE $W^*(c) > 0$ FOR $\forall c$, WE CAN CONCLUDE THAT ANY ENTRY FEE "c" IS FAIR.

(a) $W = \sum p(x) \cdot ld\, o(x) \cdot 6(x)$ $\quad o(x) = \frac{1}{2^k} \quad k = 1, 2, \ldots$

$W = \sum p(x) \cdot ld\, o(x) = H(p)$

$W = \sum \frac{1}{2^k} \cdot ld\, 2^k - H = \sum_{k=1}^{\infty} \frac{k}{2^k} - H = \frac{q}{(1-q)^2} - H \,\Big|\, q = \frac{1}{2}$

$s = \sum_{x=1}^{\infty} \frac{x}{2^x} = \sum_{x=1}^{\infty} x q^x \quad \Big| \int a^x dx = \frac{a^{x+1}}{x+1} \Big|$

$= q \sum_{x=1}^{\infty} x \cdot q^{x-1}$

$S = \sum q^x \qquad \frac{ds}{dq} = \sum_{x=1}^{\infty} x q^{x-1}$

$\frac{ds}{dq} = q \sum_{x=1}^{\infty} x q^x \qquad S = \frac{1}{1-q}$

$\frac{ds}{dq} = \left(\frac{1}{1-q}\right)' = \frac{1}{(1-q)^2} \qquad \frac{1}{(1-q)^2} = q \sum_{x=1}^{\infty} x q^x$

$S_1 = \sum_{n=1}^{\infty} n \cdot x^n \qquad \boxed{S = \sum_{n=1}^{\infty} x^n = \frac{x}{1-x}}$

$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 = \frac{1}{1-x}$

$\sum_{n=1}^{\infty} x^n = \Big| n = n-1 \Big| = \sum_{n=0}^{\infty} x^{n+1} = x \sum_{n=0}^{\infty} x^n$

$\frac{ds}{dx} = \frac{1}{(1-x)} + \frac{x}{(1-x)^2} = \frac{1-x+x}{(1-x)^2} = \frac{1}{(1-x)^2}$

$\frac{d}{dx} \sum_{n=1}^{\infty} x^n = \sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{x} \sum_{n=1}^{\infty} n x^n = \frac{1}{(1-x)^2}$

$\boxed{\sum_{n=1}^{\infty} n x^n = \frac{x}{(1-x)^2}}$

$W = \frac{\frac{1}{2}}{\left(\frac{1}{2}\right)^2} - H(p) = \frac{\frac{1}{2}}{\frac{1}{4}} = 2 - H(p)$

$S_n = 2^{n \cdot W}$ $\quad \max\left[\overline{H(Y)}\right] = n \cdot \frac{1}{n}\, ld\, n = ld\, n$

$S(X) = 2^k = b(X) \cdot o(X)$ ← ОВА Е ПОЛНО ПЛУСТАЈ !!!!

$$E[ld\, S(X)] = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot ld\, 2^k = \sum \frac{k}{2^k} = 2 = W$$

$S_n = 2^{n \cdot W} = 2^{2n}$ $\quad \lim_{n\to\infty} S_n = \infty$

– ВАЖНОСТА НА ТЕМА Е! МОЖЕ ДА ИМА

$$E[X] = E[S(X)] = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot 2^k = \sum_{k=1}^{\infty} 1 = \infty$$

($S(X) \triangleq X$ RETURN; $\frac{1}{2^k} = p(X)$, $2^k = S(X)$) ВАЖНА Е ДЕФИНИЦИЈАТА ВО МАТЕ!!!

(b) GAMBLER INVESTS $\frac{c}{2}$ UNITS IN THE GAME AND RECEIVES $\frac{1}{2}$ SHARE AND A RETURN $\frac{X}{2}$

$Pr(X = 2^k) = 2^{-k}$ $\quad k = 1, 2, \ldots$

$X_1, X_2, \ldots$ ARE I.I.D $\quad S_n = \prod_{i=1}^{n} \frac{X_i}{c}$

$S(X_1) = \frac{X_1}{2}$ $\quad S(X_2) = \frac{X_2}{2} = \frac{1}{2} \cdot \frac{X_1}{2}$

$S(X_3) = \frac{X_3}{2} = \frac{1}{2} \cdot \frac{X_2}{2} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{X_1}{2}$

$S(X_n) = \frac{X_1}{2^n}$

$$(1) \rightarrow \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2}\sum_{i=1}^{\infty} \frac{1}{2^{i-1}} = \frac{1}{2}\sum_{i=0}^{\infty} \frac{1}{2^i} = \frac{1}{2} \frac{1}{1-\frac{1}{2}} = \frac{1}{2} \cdot 2 = 1$$

$S_n = \prod_{i=1}^{n} \frac{X_i}{c} = \prod_{i=1}^{n} S(X_i)$

$$\frac{1}{n}\, ld\, S_n = \frac{1}{n}\, ld \prod_{i=1}^{n} S(X_i) = \frac{1}{n}\, ld \prod_{i=1}^{n} \frac{X_i}{c} = \frac{1}{n} \sum_{i=1}^{n} ld\left(\frac{X_i}{c}\right) \xrightarrow[n\to\infty]{w.p.}$$

$$\sum_{i=1}^{n} p(X_i) \cdot ld\left(\frac{X_i}{c}\right) = \ldots = \sum_{i=1}^{n\to\infty} \frac{1}{2^i} \cdot ld\, \frac{2^i}{c} =$$

$$= \sum_{i=1}^{n\to\infty} \frac{i}{2^i}\left[1 + ld\, c^{-1}\right] = \frac{\frac{1}{2}}{\left(1-\frac{1}{2}\right)^2} - \sum_{i=1}^{\infty} \frac{1}{2^i}\, ld\, c =$$

$$= 2 - ld\, c \underbrace{\sum_{i=1}^{\infty} \frac{1}{2^i}}_{(1)} = 2 - (ld\, c) \cdot 1 = \left| c = 2 \right| = 0$$

$1 \le c \le 2$ $\quad W = 2 - ld\, c. \rightarrow \frac{1}{n}\, ld\, S_n$

$S_n = 2^{n \cdot W} = 2^{n \cdot (2 - ld c)}$ $\boxed{c^* = 4}$

$\boxed{c > c_* = 4} \Rightarrow S_n = 2^{-n \cdot |ld c - 2|}$

$n \to \infty \quad S_n \to 0$ $\boxed{W = 2 - ld c}$

$1 \leq \boxed{c < c_* = 4} \quad S_n = 2^{n \cdot |2 - ld c|} \to \infty$

- MORE REALISTICALY $\bar{b} = 1 - b$ } KEEPING PROPORT.

$$S_n = \prod_{i=1}^{n} \left(\bar{b} + \frac{b x_i}{c}\right)$$

$$\boxed{W(b,c) = \sum_{k=1}^{\infty} 2^{-k}\, ld \left(1 - b + \frac{b x_i}{c}\right)} =$$

$$= \boxed{\sum_{k=1}^{\infty} 2^{-k}\, ld \left(1 - b + \frac{b 2^k}{c}\right)} = (*)$$

[redacted]

$$W^*(c) = \max_{0 \leq b \leq 1} W(b, c)$$

NEMA λ ZOŠTO NE DEFINIRAN CONSTR. INT $\sum b = 1$

$$\frac{d}{db}\left[2^{-k}\, ld \left(1 - b + \frac{b 2^k}{c}\right)\right] + \lambda = 0$$

$$\frac{2^k}{ln 2} \cdot \frac{1}{\left(1 - b + \frac{b \cdot 2^k}{c}\right)} \cdot \left(-1 + \frac{2^k}{c}\right) = 0$$

$$\frac{2^k \cdot c}{ln 2 (c - bc + b \cdot 2^k)} \left(\frac{2^k}{c} - 1\right) = 0 \qquad \frac{1}{c - bc + b 2^k} \cdot \left(\frac{2^k}{c} - 1\right) = 0$$

$$\frac{2^k}{c} - 1 = 0 \qquad \frac{2^k}{c} = 1 \qquad \boxed{c = 2^k} \Rightarrow W(b, 2^k) = 0$$

$$(*) = \sum_{k=1}^{\infty} 2^{-k}\, ld \left[\left(1 - b + \frac{b 2^k}{c}\right) \frac{2^{-k}}{2^{-k}}\right] = \sum_{k=1}^{\infty} 2^{-k}\, ld \frac{2^{-k}}{\left(\frac{1}{1 - b + \frac{b 2^k}{c}}\right)} +$$

$$+ \sum_{k=1}^{\infty} 2^{-k}\, ld \frac{1}{2^{-k}} = \sum_{k=1}^{\infty} 2^{-k}\, ld \frac{2^{-k}}{\left(\frac{1}{1 - b + \frac{b \cdot 2^k}{c}}\right)} -$$

$$+ \sum_{k=1}^{\infty} 2^{-k} \cdot k = \sum_{k=1}^{\infty} 2^{-k}\, ld \frac{2^{-k}}{\left(1 - b + \frac{b \cdot 2^k}{c}\right)} - \frac{\frac{1}{2}}{\left(1 - \frac{1}{2}\right)^2}$$

$$= \sum_{k=1}^{\infty} 2^{-k}\, ld \frac{2^{-k}}{\left(1 - b + \frac{b \cdot 2^k}{c}\right)} - 2 = D(p \| q) - 2$$

$$\frac{d}{dc}\left[2^{-k} ld\left(1-b+\frac{b2^k}{c}\right)\right] = \frac{2^{-k}}{ln 2} \frac{1}{1-b+\frac{b2^k}{c}} \cdot \frac{b \cdot 2^k (-1)}{c^2} = 0$$

$$A = \frac{b}{\frac{b}{c}2^k - b + 1} = 0 \qquad A = \frac{b \cdot c}{b(2^k-1)+c} = 0$$

$$\frac{b \cdot 2^k}{b(2^k-1)+2^k} = 0 \qquad \frac{b}{b(1-2^{-k})+1} = 0$$

$$\lim_{c\to\infty} A = \lim_{c\to\infty} \frac{b \cdot c}{b(c-1)+c} = \lim_{c\to\infty} \frac{bc}{bc+c-b} =$$

$$\lim_{c\to\infty} \frac{bc}{c(b+1)-b} = \frac{b}{b+1} = 0 \ ?$$

$$1-b+\frac{b \cdot 2^k}{c} \to \infty$$

- KUHN – TUCKER CONDITIONS ?
- GRITTON 2 SOLUTIONS [PART (C)]

$$W(b,c) = \sum_{k=1}^{\infty} 2^{-k} ld\left(1-b+\frac{b2^k}{c}\right)$$

$\boxed{b=0}$ $$W(b,c) = \sum_{k=1}^{\infty} 2^{-k} ld(1) = 0$$

$\boxed{b=1}$ $$W(b,c) = \sum_{k=1}^{\infty} 2^{-k} ld \frac{2^k}{c} = \sum_{k=1}^{\infty} \frac{k}{2^k} - \sum_{k=1}^{\infty} \frac{ld c}{2^k} =$$

$$\boxed{W = 2 - ld c} \qquad \boxed{\frac{\partial W(b,c)}{\partial b} = \sum_{k=1}^{\infty} 2^{-k} \frac{1}{1-b+\frac{b2^k}{c}} \left(\frac{b2^k}{c}-1\right)}$$

$\boxed{b=1}$ $$\frac{\partial W(b,c)}{\partial b} = \sum_{k=1}^{\infty} 2^{-k} \frac{1}{\frac{2^k}{c}} \cdot \left(\frac{2^k}{c}-1\right) = \sum_{k=1}^{\infty} 2^{-2k} c \left(\frac{2^k}{c}-1\right)$$

$$= \sum_{k=1}^{\infty} \left(2^{-k} - \frac{c}{2^{2k}}\right) = \sum_{k=1}^{\infty} 2^{-k} - \sum_{k=1}^{\infty} \frac{c}{2^{2k}} = \frac{1}{2}\frac{1}{1-\frac{1}{2}} - c\frac{1}{4}\frac{1}{1-\frac{1}{4}}$$

$$= \frac{1}{2} \cdot 2 - c \cdot \frac{1}{4} \cdot \frac{4}{3} = 1 - \frac{c}{3} \geq 0 \text{ FOR } \boxed{c \leq 3}$$

$$\sum_{k=1}^{\infty} 2^{-k} \frac{\frac{b2^k}{c}}{1-b+\frac{b2^k}{c}} - \sum_{k=1}^{\infty} \frac{2^{-k}}{1-b+\frac{b2^k}{c}} = 0$$

$$\sum_{k=1}^{\infty} \frac{\frac{b}{c}}{1-b+\frac{b2^k}{c}} = \sum_{k=1}^{\infty} \frac{2^{-k}}{1-b+\frac{b2^k}{c}}$$

$$\sum_{k=1}^{\infty} \frac{2^{-k}}{1-b+\frac{b\cdot 2^k}{c}} \left(\frac{b\cdot 2^k}{c} - 1\right) = \sum_{k=1}^{\infty} \frac{\left(\frac{b}{c} - 2^{-k}\right)}{1-b+\frac{b\cdot 2^k}{c}} =$$

$$= \sum_{k=1}^{\infty} \frac{\left(b - c\cdot 2^{-k}\right)}{(1-b)\cdot c + b\cdot 2^k} = \sum_{k=1}^{\infty} \frac{\left(b - \frac{c}{2^k}\right)}{(1-b)\cdot c + b\cdot 2^k}$$

$$c=0 \quad dW = \sum_{k=1}^{\infty} \frac{b}{b\cdot 2^k} = \frac{1}{2}\cdot\frac{1}{1-\frac{1}{2}} = 1$$

07.08.2017

PROBLEM 6.18 SUPER ST. PETERSBURG. FINALLY WE HAVE SUPER ST. PETERSBURG PARADOX, WHERE $Pr(X = 2^{2^k}) = 2^{-k}$, $k = 1, 2, \dots$ HERE THE EXPECTED LOG-WEALTH IS INFINITE FOR ALL $b > 0$, FOR ALL c, AND THE GAMBLERS WEALTH GROWS TO INFINITY FASTER THAN EXPONENTIALLY FOR ANY $b > 0$. BUT THAT DOESN'T MEAN THAT ALL INVESTMENT RATIOS b ARE EQUALLY GOOD. TO SEE THIS WE WISH TO MAXIMIZE THE RELATIVE GROWTH RATE WITH RESPECT TO SOME OTHER PORTFOLIO, SAY, $b = \left(\frac{1}{2}, \frac{1}{2}\right)$. SHOW THAT THERE EXIST A UNIQUE b MAXIMIZING:

$$E\left[\ln \frac{\bar{b} + b\,X/c}{\frac{1}{2} + \frac{1}{2}\frac{X}{c}}\right]$$

uyqifay
xunlgi

AND INTERPRET THE ANSWER.

$$E[\operatorname{ld} S(X)] = \sum_{k=1}^{\infty} \frac{1}{2^{-k}} \operatorname{ld} 2^{2^k} = \sum_{k=1}^{\infty} \frac{2^k}{2^{-k}} = \sum_{k=1}^{\infty} \frac{1}{2^{-2k}}$$

$$= \frac{1}{4}\sum_{k=0}^{\infty} \frac{1}{4^k} = \frac{1}{4}\cdot\frac{1}{1-\frac{1}{4}} = \frac{1}{4}\cdot\frac{4}{3} = \frac{1}{3}$$

$$S(X) = \frac{2^{2^k}}{c} \Rightarrow E[\operatorname{ld} S(X)] = \sum_{k=1}^{\infty} \frac{1}{2^{-k}} \cdot \operatorname{ld} \frac{2^{2^k}}{c} =$$

$$= \frac{1}{3} - \left(\frac{1}{2}\sum_{k=0}^{\infty} \frac{1}{2^k}\right) \operatorname{ld} c = \frac{1}{3} - \frac{1}{2}\cdot 2\cdot c = \frac{1}{3} - \operatorname{ld} c$$

$$W = \sum_{k=1}^{\infty} \frac{1}{2^k} \ln\left(\frac{1-b+\frac{b\,(X)}{c}}{\frac{1}{2}+\frac{X}{2c}}\right) = \sum_{k=1}^{\infty} \frac{1}{2^k} \ln\left(\frac{1-b+\frac{b\cdot 2^{2^k}}{c}}{\frac{1}{2}+\frac{2^{2^k}}{2c}}\right)$$

$$\frac{dW}{db} = \frac{1}{2^k} \frac{\frac{1}{2}+\frac{2^{2^k}}{2c}}{1-b+\frac{b\cdot 2^{2^k}}{c}} \cdot \left(-1 + \frac{2^{2^k}}{c}\right)$$

$$\frac{2^{2^k}}{c} = 1 \qquad \boxed{c = 2^{2^k}} \; (*)$$

18

$$\frac{dW}{dc} = \frac{1/2 + \frac{2^{2^k}}{2c}}{1-b+\frac{b\cdot 2^{2^k}}{2c}} \cdot \frac{\left(-\frac{b2^{2^k}}{c^2}\right)\left(\frac{1}{2}+\frac{2^{2^k}}{2c}\right) + \left(1-b+\frac{b\cdot 2^{2^k}}{c}\right)\frac{2^{2^k}}{c^2}}{\left(\frac{1}{2}+\frac{2^{2^k}}{2c}\right)} = 0$$

$$\left(\frac{1}{2}+\frac{2^{2^k}}{2c}\right) = 0 \Rightarrow \frac{2^{2^k}}{2c} = -\frac{1}{2} \quad c = -2^{2^k}$$

$$-\frac{b\cdot 2^{2^k}}{2c^2} - \frac{b\cdot 2^{2^k}\cdot 2^{2^k}}{2c^3} + \frac{2^{2^k}}{c^2} - \frac{b\cdot 2^{2^k}}{c^2} + \frac{b\cdot 2^{2^k\cdot 2}}{c^3} = 0$$

$$-\frac{3b2^{2^k}}{2c^2} - \frac{b\cdot 4^{2^k}}{2c^3} + \frac{2^{2^k}}{c^2} = 0$$

$$\frac{2^{2^k}}{c^2}\left(1-\frac{3b}{2}-\frac{b\cdot 2^{2^k}}{2c}\right) = 0$$

(★) →

$$\left(1-\frac{3b}{2}-\frac{b}{2}\right) = 0 \quad \left(1-\frac{4b}{2}\right) = 0 \quad (1-2b) = 0 \quad \boxed{b=\frac{1}{2}}$$

$$W(b,c) \quad \boxed{W\left(\frac{1}{2},c\right) = 0}$$

Problem 6.17 revisited

$$W(b,c) = \sum_{k=1}^{\infty} 2^{-k} \ln\left(1-b+\frac{b2^k}{c}\right)$$

$$\frac{dW(b,c)}{db} = \sum_{k=1}^{\infty} 2^{-k}\left(\frac{-1+\frac{2^k}{c}}{1-b+\frac{b2^k}{c}}\right) = 0$$

$$b=1 \quad dW = \sum_{k=1}^{\infty} 2^{-k}\frac{\frac{2^k}{c}-1}{\frac{2^k}{c}} = \sum_{k=1}^{\infty} 2^{-k}\frac{2^k-c}{2^k} = \sum_{k=1}^{\infty}\frac{1-c2^{-k}}{2^k} =$$

$$= \sum_{k=1}^{\infty}\frac{1}{2^k} - c\sum_{k=1}^{\infty}\frac{1}{4^k} = \frac{1}{2}\frac{1}{1-1/2} - c\,\frac{1}{4}\frac{1}{1-1/4} = 1 - c\,\frac{1}{4}\cdot\frac{4}{3}$$

$$\boxed{dW = 1-\frac{c}{3}}$$

$$dW = -\underbrace{\sum_{k=1}^{\infty}\frac{2^{-k}}{1-b+\frac{b2^k}{c}}}_{S_1} + \underbrace{\sum_{k=1}^{\infty}\frac{\frac{1}{c}}{1-b+\frac{b2^k}{c}}}_{S_2}$$

$$S_1 = \sum_{k=1}^{\infty}\frac{c}{c2^k(1-b)+b} = -c\sum_{k=1}^{\infty}\frac{1}{2^k\cdot c(1-b)+b} = -c\sum_{k=1}^{\infty}\ldots$$

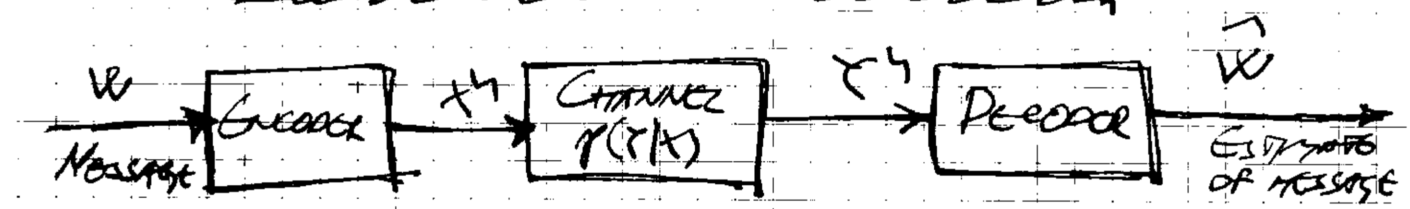
$$S_2 = \sum_{k=1}^{\infty}\frac{1}{c(1-b)+b2^k} = \sum_{k=1}^{\infty}\frac{1}{a+b\cdot 2^k}$$

$$E[U] = \sum_{k=1}^{\infty} \frac{\left[ld(w + 2^{k-1} - c) - ld(w)\right]}{2^k} < \infty$$

↑ expected utility of the lottery

Chapter 7: Channel Capacity

X5TECH



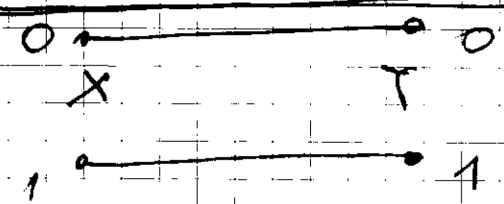
Definition: We define "information" channel capacity of the discrete memoryless channel as:

$$C = \max_{p(x)} I(X;Y)$$

where the maximum is taken over all possible input distributions.

Operational definition:
Channel capacity is the highest bit rate at which information can be sent with arbitrarily low probability of error.

7.1 Examples of Channel Capacity
7.1.1 Noiseless Binary Channel



$C = 1$

Information capacity: $C = \max I(X:Y) = 1$

$H(Y|X) = 0$ we [illegible] retrieved

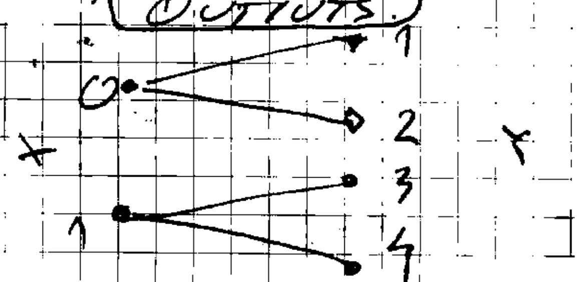
$Y = f(X)$

$$I(X;Y) = H(X) - \underbrace{H(X|Y)}_{0} = H(Y) - H(Y|X)$$

} one-to-one mapping

$\rightarrow H(X) = ld\, m = ld\, 2 = \underline{1}$ for $p(x) = \left\{\frac{1}{2}, \frac{1}{2}\right\}$

7.1.2 Noisy Channel with Nonoverlapping Outputs

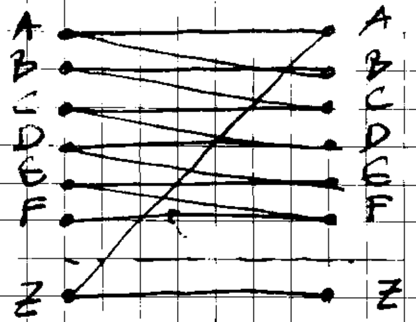


$H(X|Y) = P(Y=1) \cdot H(X|Y=1) + P(Y=2) \cdot H(X|Y=2) + P(Y=3) \cdot H(X|Y=3) + P(Y=4) \cdot H(X|Y=4)$

(each term $= 0$: we know X from Y)

$= 0 \cdot (P(Y=1) + P(Y=2) + P(Y=3) + P(Y=4)) = 0$

7.1.3 Noisy typewriter:



In this case channel input is either received unchanged at the output with probability 1/2 or is transformed into the next letter with probability 1/2.

$$C = \max I(X;Y) = \max [H(Y) - H(Y|X)] =$$

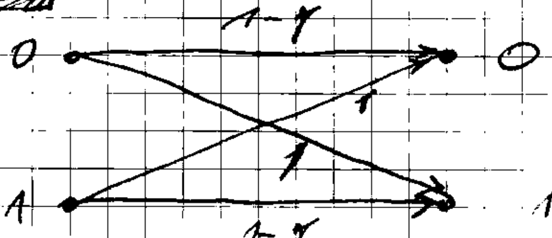
$$H(Y|X) = P(X=A)\, H(Y|X=A) + P(X=B)\, H(Y|X=B) + \dots + P(X=Z)\, H(Y|X=Z)$$

$$H(Y|X=A) = H(Y|X=B) = \dots = H(Y|X=Z) = \tfrac{1}{2}\,\mathrm{ld}\,2 + \tfrac{1}{2}\,\mathrm{ld}\,2 = 1$$

$$C = \max[H(Y) - 1] = \max[H(Y)] - 1 = \mathrm{ld}\,m - 1$$ for $p(Y) = \frac{1}{m}$ uniform.

$$\boxed{C = \mathrm{ld}\,26 - 1 = \mathrm{ld}\,26 - \mathrm{ld}\,2 = \mathrm{ld}\,13}$$

7.1.4 Binary symmetric channel



$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \sum_x p(x) \cdot H(Y|X) = H(Y) - p(X=0) \cdot H(Y|X=0) -$$

$$- p(X=1)\, H(Y|X=1) = H(Y) - p(X=0) \cdot p(Y=0|X=0) \cdot \mathrm{ld}\frac{1}{p(Y=0|X=0)} -$$

$$- p(X=0) \cdot p(Y=1|X=0) \cdot \mathrm{ld}\frac{1}{p(Y=1|X=0)} - p(X=1) \cdot p(Y=0|X=1) \cdot \mathrm{ld}\frac{1}{p(Y=0|X=1)} - p(X=1) \cdot p(Y=1|X=1) \cdot \mathrm{ld}\frac{1}{p(Y=1|X=1)} =$$

$$= H(Y) - p(X=0)\left[(1-p)\,\mathrm{ld}\frac{1}{1-p} + p\,\mathrm{ld}\frac{1}{p}\right] -$$

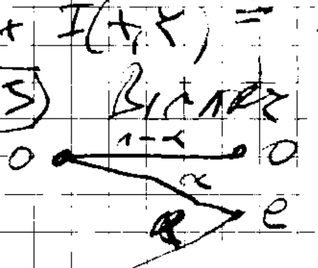
$$- p(X=1)\left[p\,\mathrm{ld}\,p^{-1} + (1-p)\,\mathrm{ld}\frac{1}{1-p}\right] = H(Y) - p(X=0) \cdot H(p)$$

$$- p(X=1) \cdot H(p) = H(Y) - \sum p(x) \cdot H(p) = H(Y) - H(p)$$

$$\leq \mathrm{ld}\,m - H(p) = \mathrm{ld}\,2 - H(p) = 1 - H(p)$$

$$\max I(X,Y) = 1 - H(p) \qquad \boxed{C = 1 - H(p)} \text{ bit}$$

7.1.5 Binary erasure channel



$$H(Y|X) = \sum_{x=0} p(x) \cdot H(Y|X) = \quad (1/1-\alpha)$$

$$= \sum p(x)\left[p(Y=0|X_i=0) \cdot \mathrm{ld}\,p(Y=0|X_i=0) - \mathrm{ld}\,p(Y=e|X_i=0)\right] = \sum p(x) \cdot H(\alpha)$$

$\alpha \leftarrow p(Y_i=e|X_i=0)$

$H(Y|X) = H(\alpha)$ $\qquad H(Y) = P(Y=0)\cdot ld\frac{1}{P(Y=0)} +$

$+ P(Y=1)\cdot ld\frac{1}{P(Y=1)} + P(Y=e)\cdot ld\frac{1}{P(Y=e)} =$

~~$= 2(1-\alpha)\, ld\frac{1}{(1-\alpha)} + 2\alpha\, ld\frac{1}{2\alpha}$~~

~~$2\cdot(1-\alpha) + 2\alpha = 2 - 2\alpha + 2\alpha = 2$?~~

$= \frac{(1-\alpha)}{2}\cdot ld\frac{2}{1-\alpha} + \alpha\cdot ld\frac{1}{\alpha} + \frac{1-\alpha}{2}\, ld\frac{2}{1-\alpha}$

$2\cdot\frac{1-\alpha}{2} + \alpha = 1-\alpha+\alpha = 1$

$= (1-\alpha)\cdot ld\frac{2}{1-\alpha} + \alpha\cdot ld\frac{1}{\alpha} = (1-\alpha) + H(\alpha)$

$I(X,Y) = H(Y) - H(Y|X) = 1-\alpha + H(\alpha) - H(\alpha) = 1-\alpha$

Cover Proof: $H(Y) = H(Y,E) = H(Y) + H(E|Y) =$

$= H(Y) + 0 = H(E) + H(Y|E)$

$E \rightarrow$ event $Y = e$

$H(Y) = ?$ $\quad H(E) = ?$ $\quad \boxed{P_r(X=1) = \pi}$

$H(Y,E) = H(Y) = H\big(\underbrace{(1-\pi)}_{P(X=1)}\cdot(1-\alpha),\ \alpha,\ \pi\cdot(1-\alpha)\big) =$

(*) — vidi pp.23

$= -\big[(1-\pi)(1-\alpha)\, ld\,(1-\pi)(1-\alpha) + \alpha\, ld\,\alpha + \pi(1-\alpha)\, ld\,\pi(1-\alpha)\big]$

$= -\big[(1-\pi)(1-\alpha)\cdot ld(1-\pi) + (1-\pi)(1-\alpha)\, ld(1-\alpha) + \alpha\, ld\,\alpha + \pi(1-\alpha)\, ld(1-\alpha) + \pi(1-\alpha)\, ld\,\pi$

$= -\big[(1-\pi)(1-\alpha)\, ld(1-\pi) + (1-\alpha)\, ld(1-\alpha) + \alpha\, ld\,\alpha + \pi(1-\alpha)\, ld\,\pi\big]$

$= -\big[(1-\alpha)\big[(1-\pi)\, ld(1-\pi) + \pi\, ld\,\pi\big]\big] + H(\alpha) =$

$= (1-\alpha)\cdot H(\pi) + H(\alpha)$

$H(Y,E) = \sum_{Y,E} P(Y,E)\, ld\frac{1}{P(Y,E)} = -\big[P(Y=1,\bar{E})\, ld\, P(Y=1,\bar{E})$

$+ P(Y=0,\bar{E})\, ld\, P(Y=0,\bar{E}) + P(E)\, ld\, P(E)\big] =$

$= H(\pi(1-\alpha), (1-\pi)\cdot(1-\alpha), \alpha)$

$\boxed{P(E) = P(E|X=1)\cdot P(X=1) + P(E|X=0)\cdot P(X=0) = \pi\cdot\alpha + \alpha\cdot(1-\pi) = \pi\alpha + \alpha - \alpha\pi = \alpha}$

$I(X;Y) = H(Y) - H(Y|X) = (1-\alpha)H(\pi) + H(\alpha) - H(\alpha)$

$I(X;Y) = (1-\alpha)H(\pi)$ $\quad C = \max_{p(x)} (1-\alpha)H(\pi)$

$\pi = p(x) = \frac{1}{2}$ $\quad$ $C = (1-\alpha)\cdot H(\frac{1}{2}) = (1-\alpha)$

(*) $P(Y,E)$

$E \backslash Y$	0	e	1	$P(E)$
E	0	α	$\emptyset$	α
$\bar{E}$	$(1-\pi)(1-\alpha)$	$\emptyset$	$\pi(1-\alpha)$	$(1-\alpha)$
$P(Y)$	$(1-\pi)(1-\alpha)$	α	$\pi(1-\alpha)$	

$H(Y,E) = -P(0,E)\,ld\,P(0,E) + P(e,E)\,ld\,P(e,E) + P(1,\bar{E})\,ld\,P(1,\bar{E}) + P(0,\bar{E})\,ld\,P(0,\bar{E}) + P(e,\bar{E})\,ld\,P(e,\bar{E}) + P(1,\bar{E})\dots$

... MOŽE SE ... = 0 ... SISTEMA.

THE EXPRESSION FOR THE CAPACITY HAS SOME INTUITIVE MEANING: SINCE THE PROPORTION OF α OF THE BITS ARE LOST IN THE CHANNEL, WE CAN RECOVER (AT MOST) A PROPORTION $(1-\alpha)$ OF THE BITS.

HENCE THE CAPACITY IS AT MOST $(1-\alpha)$.

• ...

$P = 10^{-8}$ $\quad$ $R = (1-p)\cdot T$

... $1MB = 10^6$ bit

$R = (1-10^{-8})\cdot 10^6 = 10^6 - 10^{-2} = 1.000000 - 0.01$

$= 999.999,99$

ZNAČI ... $= 999,9$ k...

• FEEDBACK DOESN'T INCREASE THE CAPACITY OF THE DISCRETE MEMORYLESS CHANNEL ...

7.2 SYMMETRIC CHANNELS

$p(y|x) = x\downarrow \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.5 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$

EACH COLUMN AND ROW ARE PERMUTATIONS OF EACH OTHER → SYMMETRIC CHANNEL.

$Y = X + Z \pmod{c}$

$Z \in \{0, 1, 2, \dots, c-1\}$

... SAME ALPHABET AS Z AND INDEPENDENT.

$$\begin{bmatrix} p(y_1) \\ p(y_2) \\ p(y_3) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}^T \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \end{bmatrix} \qquad y_1 = p_{11}p(x_1) + p_{12}p(x_2) + p_{13}p(x_3)$$

$$\begin{pmatrix} p(y_1) \\ p(y_2) \\ p(y_3) \end{pmatrix}^T = [p(x_1)\, p(x_2)\, p(x_3)] \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

OVAA FORMULACIJA KORISTI ZA (p_{ij})

$$p(y_j) = \sum_{i=1}^{m} p(x_i)\, p(y_j|x_i) \qquad p(y_j|x_i) = p_{ij}$$

r - ROW OF TRANSMISION MATRIX

$$I(X;Y) = H(Y) - H(Y|X)$$

$$H(Y|X) = \sum_x p(x) H(Y|X=x) = \sum_{x,r} p(x) \sum_{c=1}^{n} p_{rc}\, ld \frac{1}{p_{rc}} = H(r)$$

($\to H(r)$)

ZA SEKOJA REDICA E ISTO ZA IZZGLAVA PRED SUMATA

$$I(X;Y) = H(Y) - H(r) \leq ld|Y| - H(r)$$

- $p(x) = \frac{1}{|X|}$ $\Rightarrow$ ACHIEVES UNIFORM DISTRIBUTION FOR Y

$$p(y) = \sum_{x \in X} p(x) \cdot p(y|x) = \frac{1}{|X|} \sum_x p(y|x) = \frac{c}{|X|} = \frac{1}{|Y|}$$

MV

c = SUM OF PROBABILITIES IN ONE COLUMN OF THE PROBABILITY TRANSITION MATRIX.

THUS THE CHANNEL IN 7.17 HAS CAPACITY:

$$C = \max_{p(x)} I(X;Y) = ld\,m - H(0.5, 0.3, 0.2)$$

DEFINITION: A CHANNEL IS SAID TO BE SYMETRIC IF THE ROWS OF THE CHANNEL TRANSITION MATRIX $p(y|x)$ ARE PERMUTATIONS OF EACH OTHER AND THE COLUMNS ARE PERMUTATIONS OF EACH OTHER. THE CHANNEL IS SAID TO BE WEAKLY SYMETRIC IF EVERY ROW OF THE TRANSMISION MATRIX $p(\cdot|x)$ IS A PERMUTATION OF EVERY OTHER ROW AND ALL COLUMNS SUMS $\sum_x p(y|x)$ ARE EQUAL

$$p(y|x) = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \end{pmatrix} \Rightarrow \text{WEAKLY SIMETRIC CHANNEL.}$$

THEOREM 7.2.1 FOR A WEAKLY SYMETRIC CHANNEL

$$C = ld|Y| - H(\text{ROW OF TRANSITION MATRIX})$$

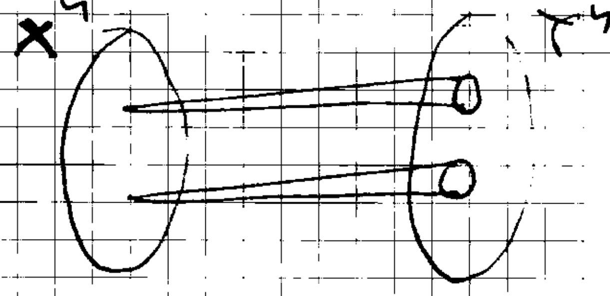
THAT IS ACHIEVED FOR THE UNIFORM DISTRO OF INPUT ALPHABET.

7.3 Properties of channel capacity

1. $C \geq 0$ since $I(X;Y) \geq 0$
2. $C \leq \log|X|$ since $C = \max I(X;Y) \leq \max H(X) = \log|X|$
3. $C \leq \log|Y|$ for the same reason
4. $I(X;Y)$ is continuous function of $p(x)$
5. $I(X;Y)$ is concave function of $p(x)$. (Theorem 2.7.4)

7.4 Preview of the channel coding theorem

- For each (typical) input sequence there are $2^{nH(Y|X)}$ possible Y sequences, all of them equally likely (Fig. 7.7).

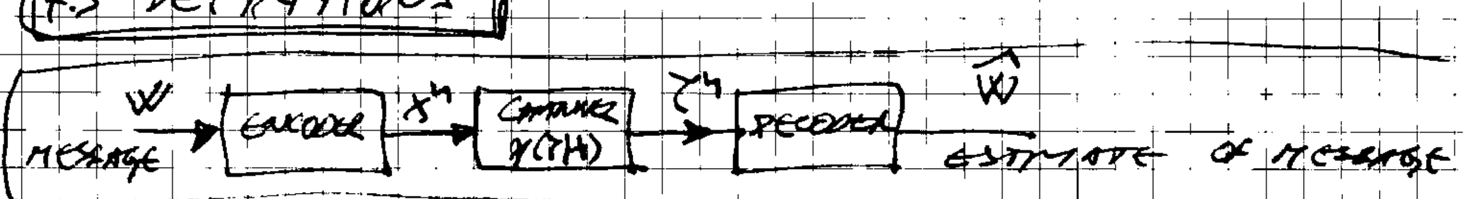


- The total number of possible (typical) Y sequences is $\doteq 2^{n \cdot H(Y)}$. This set has to be divided in sets of size $2^{nH(Y|X)}$ corresponding to the different input X sequences.
- Total number of disjoint sets is less than or equal to:

$$2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$$

Hence we can sent at most $\doteq 2^{nI(X;Y)}$ distinguishable sequences of length n.

7.5 Definitions



- A message W drawn from the index $\{1,2,..M\}$ results in the signal $X^n(W)$ which is received at the receiver as random sequence:

$$Y^n \sim p(y^n|x^n)$$

The receiver then guesses the index W by appropriate decoding rule: $\hat{W} = g(Y^n)$

THE RECEIVER MAKES AN ERROR IF $\hat{W}$ IS NOT THE SAME AS THE INDEX W THAT WAS TRANSMITTED.

DEFINITION A discrete channel, DENOTED BY $(\mathcal{X}, p(y|x), \mathcal{Y})$ CONSISTS OF TWO FINITE SETS $\mathcal{X}$ AND $\mathcal{Y}$ AND A COLLECTION OF PROBABILITY MASS FUNCTIONS $p_x(y|x)$, ONE FOR EACH $x \in \mathcal{X}$, SUCH FOR EVERY x AND y, $p(y|x) \geq 0$, AND FOR EVERY x, $\sum_y p(y|x) = 1$, WITH THE INTERPRETATION THAT X IS THE INPUT AND Y IS THE OUTPUT OF THE CHANNEL.

DEFINITION THE n-TH EXTENSION OF THE DISCRETE MEMORYLESS CHANNEL (DMC) IS THE CHANNEL $(\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n)$ WHERE

$$p(y_k|x^k, y^{k-1}) = p(y_k|x_k) \qquad k = 1, 2, \dots n$$

REMARK IF THE CHANNEL IS USED WITHOUT FEEDBACK [I.E. IF THE INPUT SYMBOLS DO NOT DEPEND ON THE PAST SYMBOLS, NAMELY $p(x_k|x^{k-1}, y^{k-1}) = p(x_k|x^{k-1})$] THE CHANNEL TRANSITION MATRIX REDUCES TO:

$$p(y^n|x^n) = \prod_{i=1}^{n} p(y_i|x_i)$$

DEFINITION: AN (M, n) **CODE** FOR CHANNEL $(\mathcal{X}, p(y|x), \mathcal{Y})$ CONSIST OF FOLLOWING:

1. AN INDEX SET $\{1, 2, \dots M\}$
2. AN ENCODING FUNCTION $X^n: \{1, 2, \dots M\} \to \mathcal{X}^n$ YIELDING CODEWORDS $x^n(1), x^n(2), \dots, x^n(M)$. THE SET OF CODEWORDS IS CALLED **CODEBOOK**.
3. A DECODING FUNCTION: $g: \mathcal{Y}^n \to \{1, 2, \dots M\}$ IS DETERMINISTIC RULE THAT ASSIGNS A GUESS TO EACH POSSIBLE RECEIVED VECTOR.

DEFINITION: (CONDITIONAL PROBABILITY OF ERROR) LET

$$\lambda_i = \Pr(g(Y^n) \neq i \mid X^n = x^n(i)) = \sum_{y^n} p(y^n|x^n(i)) I(g(y^n) \neq i)$$

BE THE CONDITIONAL PROBABILITY OF ERROR GIVEN THAT i WAS SENT WHERE $I(\cdot)$ IS THE INDICATOR FUNCTION. (VERO/FALSO $I(\text{TRUE}) = 1$)

DEFINITION: MAXIMUM PROBABILITY OF ERROR:

$$\lambda^{(n)} = \max_{i \in \{1, 2, \dots M\}} \lambda_i$$

DEFINITION: (ARITHMETIC AVERAGE) PROBABILITY OF ERROR $P_e^{(n)}$ FOR THE CODE IS DEFINED AS:

$$P_e^{(n)} = \frac{1}{M} \sum_{i=1}^{M} \lambda_i$$

IF THE INDEX IS CHOSEN ACCORDING TO UNIFORM DISTRIBUTION OVER SET $\{1, 2, \dots, M\}$ AND $X^n = x^n(W)$ THEN:

$$P_e^{(n)} = Pr(W \neq g(Y^n))$$

i.e. $P_e^{(n)}$ IS PROBABILITY OF ERROR. ALSO OBVIOUSLY

$$\boxed{P_e^{(n)} \leq \lambda^{(n)}}$$

DEFINITION: THE RATE R OF AN (M, n) CODE IS:

$$R = \frac{\log M}{n} \quad \text{BITS PER TRANSMISSION}$$

DEFINITION: A RATE R IS SAID TO BE ACHIEVABLE IF THERE EXISTS A SEQUENCE OF $(\lceil 2^{nR} \rceil, n)$ CODES SUCH THAT THE MAXIMAL PROBABILITY OF ERROR $\lambda^{(n)}$ TENDS TO 0 AS $n \to \infty$.
LATER, WE WRITE $(2^{nR}, n)$ CODES TO MEAN $(\lceil 2^{nR} \rceil, n)$ CODES. THIS WILL SIMPLIFY THE NOTATION.

DEFINITION: THE CAPACITY OF THE CHANNEL IS SUPREMUM OF ALL ACHIEVABLE RATES.

THUS RATES LESS THEN CAPACITY YIELD ARBITRARILY SMALL PROBABILITY OF ERROR FOR SUFFICIENTLY LARGE BLOCK LENGTHS.

$$\boxed{2^{-n(H+\varepsilon)} \leq p(x^n) \leq 2^{-n(H-\varepsilon)}} \text{ AEP}$$

JOINTLY TYPICAL SEQUENCES

WE DECODE THE CHANNEL Y^n AS THE i-TH INDEX IF THE CODEWORD $X^n(i)$ IS JOINTLY TYPICAL WITH RECEIVED SIGNAL Y^n.

DEFINITION: THE SET $A_\varepsilon^{(n)}$ OF JOINTLY TYPICAL SEQUENCES $\{(x^n, y^n)\}$ WITH RESPECT TO THE DISTRIBUTION $p(x,y)$ IS THE SET OF n-SEQUENCES WITH EMPIRICAL ENTROPIES ε-CLOSE TO THE TRUE ENTROPIES:

$$A_\varepsilon^{(n)} = \{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \left| -\frac{1}{n} \log p(x^n) - H(X) \right| < \varepsilon,$$

$$\left| -\frac{1}{n} \log p(y^n) - H(Y) \right| < \varepsilon, \quad \left| -\frac{1}{n} \log p(x^n, y^n) - H(X,Y) \right| < \varepsilon \}$$

$$-n(H+\varepsilon) \leq \log p(x^n) \leq -n(H-\varepsilon)$$

$$-H-\varepsilon \leq \frac{1}{n} p(x^n) \leq -H+\varepsilon \qquad \boxed{-\varepsilon \leq -\frac{1}{n} p(x^n) - H \leq \varepsilon}$$

$$H+\varepsilon \geq -\frac{1}{n} p(x^n) \geq H-\varepsilon \qquad H-\varepsilon \leq -\frac{1}{n} p(x^n) \leq H+\varepsilon$$

WHERE: $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$

THEOREM 7-6.1 (JOINT AEP) LET (X^n, Y^n) BE SEQUENCES OF LENGTH n DRAWN I.I.D ACCORDING TO $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$

THEN:

1.) $Pr\left[(X^n, Y^n) \in A_\varepsilon^{(n)}\right] \to 1$ AS $n \to \infty$

2.) $|A_\varepsilon^{(n)}| \leq 2^{n(H(X,Y)+\varepsilon)}$

3.) IF $\tilde{X}^n, \tilde{Y}^n \sim p(x^n) \cdot p(y^n)$ THAT IS $\tilde{X}^n \perp \tilde{Y}^n$ ARE INDEPENDENT WITH THE SAME MARGINALS $p(x^n, y^n)$, THEN:

$$Pr\left((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^{(n)}\right) \leq 2^{-n(I(X;Y) - 3\varepsilon)}$$

ALSO FOR SUFFICIENTLY LARGE n,

$$Pr\left((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^{(n)}\right) \geq (1-\varepsilon)\, 2^{-n(I(X;Y)+3\varepsilon)}$$

PROOF: $-\frac{1}{n} \operatorname{ld} p(X^n) \to -E[\operatorname{ld} p(X)] = H(X)$ IN PROBABILITY

HENCE GIVEN $\varepsilon > 0$, there exist n_1 SUCH THAT ALL $n > n_1$

$$Pr\left(\left|-\frac{1}{n} \operatorname{ld} p(X^n) - H(X)\right| \geq \varepsilon\right) < \frac{\varepsilon}{3}$$

- SIMILARLY BY THE WEAK LAW

$$-\frac{1}{n} \operatorname{ld} p(Y^n) \to E[p(Y)] = H(Y)$$ IN PROBABILITY

AND:

$$-\frac{1}{n} \operatorname{ld} p(X^n, Y^n) \to E[p(X,Y)] = H(X,Y)$$

AND THERE EXIST n_2 & n_3 SUCH AS:

$$Pr\left(\left|-\frac{1}{n} \operatorname{ld} p(Y^n) - H(Y)\right| \geq \varepsilon\right) < \frac{\varepsilon}{3} \quad n \geq n_2$$

$$Pr\left(\left|-\frac{1}{n} \operatorname{ld} p(X^n, Y^n) - H(X,Y)\right| \geq \varepsilon\right) < \frac{\varepsilon}{3} \quad n \geq n_3$$

- CHOOSING $n > \max\{n_1, n_2, n_3\}$

$\Rightarrow$ FOR SUFFICIENTLY LARGE $n \Rightarrow$

$Pr\{A_\varepsilon^{(n)}\} \Rightarrow 1$ i.e. $Pr\{A_\varepsilon^{(n)}\} = 1 - \varepsilon$

2) PROOF OF SECOND PART:

$$1 = \sum p(x^n, y^n) \geq \sum_{A_\varepsilon^{(n)}} p(x^n, y^n) \geq |A_\varepsilon^{(n)}| \cdot 2^{-n(H(X,Y)+\varepsilon)}$$

$$2^{-n(H+\varepsilon)} \leq p(x^n, y^n) \leq 2^{-n(H-\varepsilon)}$$

PETAR KOSTURSKI

0462855131

$$|A_\varepsilon^{(n)}| \leq 2^{n(H(X,Y)+\varepsilon)}$$

3) $$P_r\left((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^n\right) = \sum_{(x^n, y^n) \in A_\varepsilon^{(n)}} p(x^n) p(y^n) \leq$$

$$\leq 2^{n(H(X,Y)+\varepsilon)} \cdot 2^{-n(H(X)-\varepsilon)} \cdot 2^{-n(H(Y)-\varepsilon)} =$$

$$= 2^{nH(X) + nH(Y|X) + n\varepsilon - nH(X) + n\varepsilon - nH(Y) + n\varepsilon}$$

$$= 2^{-n[H(Y) - H(Y|X)] + 3n\varepsilon} = 2^{-n[I(X;Y) - 3\varepsilon]}$$

- FOR SUFFICIENTLY LARGE n $P_r(A_\varepsilon^{(n)}) \geq 1 - \varepsilon$ AND THEREFORE:

$$1 - \varepsilon \leq \sum_{(x^n, y^n) \in A_\varepsilon^{(n)}} p(x^n, y^n) \leq |A_\varepsilon^{(n)}| \, 2^{-n(H(X,Y)-\varepsilon)}$$

$$|A_\varepsilon^{(n)}| \geq (1-\varepsilon) \cdot 2^{n(H(X,Y)-\varepsilon)}$$

ODGOVARA NA TEOREMA 3.12

$$P_r\left((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^n\right) = \sum_{A_\varepsilon^{(n)}} p(x^n) \cdot p(y^n) \geq$$

$$\geq (1-\varepsilon) \cdot 2^{n(H(X,Y)-\varepsilon)} \cdot 2^{-n(H(X)+\varepsilon)} \cdot 2^{-n(H(Y)+\varepsilon)}$$

$$= (1-\varepsilon) \, 2^{nH(X,Y) - n\varepsilon - nH(X) - n\varepsilon - nH(Y) - n\varepsilon}$$

$$= (1-\varepsilon) \, 2^{n[H(X) + H(Y|X)] - 3n\varepsilon - nH(X) - nH(Y)}$$

$$= (1-\varepsilon) \, 2^{-n[H(Y) - H(Y|X)] - 3n\varepsilon} = (1-\varepsilon) \, 2^{-n[I(X;Y) + 3\varepsilon]}$$

THEOREM 7.6.1 E ANALOGNA NA 3.12

$$H(X) - \varepsilon \leq -\frac{1}{n} \log p(x^n) \leq H(X) + \varepsilon \qquad 2^{-n(H+\varepsilon)} \leq p(x^n) \leq 2^{-n(H-\varepsilon)}$$

1) $P_r\{A_\varepsilon^{(n)}\} = 1 - \varepsilon$

2) $|A_\varepsilon^{(n)}| \leq 2^{n(H(X)+\varepsilon)}$

3) $|A_\varepsilon^{(n)}| \geq (1-\varepsilon) \, 2^{n(H(X)-\varepsilon)}$ } OSVEN TEOREMA DOKAZ ... ZAD.

X

THERE ARE ABOUT $2^{nH(X)}$ TYPICAL X SEQUENCES AND ABOUT $2^{nH(Y)}$ TYPICAL Y SEQUENCES. HOWEVER SINCE THERE ARE ONLY $2^{n(H(X,Y))}$ JOINTLY TYPICAL SEQUENCES, NOT ALL PAIRS OF TYPICAL X^n AND TYPICAL Y^n ARE ALSO JOINTLY TYPICAL. ($H(X,Y) \leq H(X)+H(Y)$)

PROBABILITY THAT ANY RANDOMLY CHOSEN PAIR IS JOINTLY TYPICAL IS ABOUT $2^{-nI(X;Y)}$.

THIS SUGGESTS THAT THERE ARE ABOUT $2^{nI(X;Y)}$ DISTINGUISHABLE SIGNALS X^n.

- FOR THE FIXED OUTPUT SEQUENCE Y^n THERE ARE ABOUT $2^{nH(X|Y)}$ CONDITIONALLY TYPICAL INPUT SIGNALS. THE PROBABILITY THAT SOME RANDOMLY CHOSEN (OTHER) INPUT SIGNAL IS JOINTLY TYPICAL WITH Y^n IS ABOUT:

$$\frac{2^{nH(X|Y)}}{2^{nH(X)}} = 2^{-n(H(X)-H(X|Y))} = 2^{-nI(X,Y)}$$

HENCE WE CAN CHOSE $2^{nI(X,Y)}$ CODEWORDS $X^n(W)$ BEFORE ONE OF THIS CODEWORDS WILL GET CONFUSED WITH THE CODEWORD THAT CAUSED THE OUTPUT Y^n.

RECALL

$$2^{-n(H(X,Y)+\varepsilon)} \leq p(x^n, y^n) \leq 2^{-n(H(X,Y)-\varepsilon)}$$

1.) $\Pr\left[(X^n, Y^n) \in A_\varepsilon^{(n)}\right] \to 1$ AS $n \to \infty$

2) $|A_\varepsilon^n| \leq 2^{n(H(X,Y)+\varepsilon)}$

3.) $\exists n$ s.t. $\Pr\left[(\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^{(n)}\right] \leq 2^{-n(I(X;Y)-3\varepsilon)}$

$\exists n$ s.t. $\Pr\left[(\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^n\right] \geq (1-\varepsilon)\, 2^{-n(I(X,Y)+3\varepsilon)}$

(2) $1 = \sum_{A_\varepsilon^{(n)}} p(x^n y^n) \geq |A_\varepsilon^n| \cdot 2^{-n(H(X,Y)+\varepsilon)} \Rightarrow$

$$|A_\varepsilon^{(n)}| \leq 2^{n(H(X,Y)+\varepsilon)}$$

(3) $\Pr(\tilde{X}, \tilde{Y} \in A_\varepsilon^{(n)}) = \sum_{A_\varepsilon^{(n)}} p(x^n) p(y^n) \leq |A_\varepsilon^{(n)}| \cdot 2^{-n(H(X)-\varepsilon)}\, 2^{-n(H(Y)-\varepsilon)}$

$= 2^{n(H(X,Y)+\varepsilon) - nH(X) + n\varepsilon - nH(Y) + n\varepsilon} =$

$= 2^{nH(X) + nH(Y|X) + n\varepsilon - nH(X) + n\varepsilon - nH(Y) + n\varepsilon}$

$= 2^{-n[H(Y)-H(Y|X)] + 3\varepsilon n} = 2^{-n[I(X;Y)-3\varepsilon]}$

$\exists n \quad 1-\varepsilon \leq \sum_{A_\varepsilon^{(n)}} p(x^n, y^n) \leq |A_\varepsilon^{(n)}| \cdot 2^{-n[H(X,Y)-\varepsilon]}$

$$\boxed{|A_\varepsilon^{(n)}| \geq (1-\varepsilon)\, 2^{+n[H(X,Y)-\varepsilon]}}$$

$$\sum_{A_\varepsilon^{(n)}} p(x^n)p(y^n) \geq |A_\varepsilon^{(n)}| \cdot 2^{-n[H(X)+\varepsilon]} \cdot 2^{-n[H(Y)+\varepsilon]} \geq$$

$$\geq (1-\varepsilon)\, 2^{+n[H(X)+H(Y|X)-\varepsilon] - nH(X) - n\varepsilon - nH(Y) - \varepsilon n}$$

$$= (1-\varepsilon)\, 2^{nH(X)+nH(Y|X)-n\varepsilon - nH(X) - n\varepsilon - nH(Y) - n\varepsilon}$$

$$= (1-\varepsilon)\, 2^{-n[H(Y)-H(Y|X)] - 3n\varepsilon}$$

$$\Pr[(\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^{(n)}] \geq (1-\varepsilon)\, 2^{-n[I(X;Y)-3\varepsilon]}$$

7.7. Channel Coding Theorem

Theorem 7.7.1 (Channel Coding Theorem)

For a discrete memoryless channel all rates below capacity C are achievable. Specifically, for every rate $R<C$, there exist sequence of $(2^{nR}, n)$ codes with maximum probability of error:

$$\lambda^{(n)} \to 0$$

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

Achievability: Fix $p(x)$. Generate a $(2^{nR}, n)$ code at random according to distribution $p(x)$. Specifically we generate 2^{nR} codewords independently according to the distribution

$$p(x^n) = \prod_{i=1}^{n} p(x_i)$$

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ \vdots & & & \\ x_1(2^{nR}) & x_2(2^{nR}) & \dots & x_n(2^{nR}) \end{bmatrix}$$

The probability that we generate a particular code $\mathcal{C}$ is:

$$(*) \quad \boxed{P_r(\mathcal{C}) = \prod_{w=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(w))}$$

CONSIDER FOLLOWING SEQUENCE OF EVENTS:

1.) A RANDOM CODE IS GENERATED AS DESCRIBED IN 9.3.3. (*) ACCORDING TO $p(x)$.

2.) THE CODE IS REVEALED TO BOTH SENDER AND RECEIVER. BOTH THE SENDER AND THE RECEIVER ARE ALSO ASSUMED TO KNOW THE CHANNEL TRANSITION MATRIX $p(y|x)$ FOR CHANNEL.

3.) A MESSAGE W IS CHOSEN ACCORDING TO A UNIFORM DISTRIBUTION.

$$P_r(W=w) = 2^{-nR} \qquad w = 1, 2, \ldots, 2^{nR}$$

4.) THE w-TH CODEWORD $X^n(w)$ CORRESPONDS TO THE w-TH ROW OF C, IS SENT OVER THE CHANNEL.

5.) THE RECEIVER RECEIVES A SEQUENCE Y^n ACCORDING TO THE DISTRIBUTION

$$P(y^n|x^n(w)) = \prod_{i=1}^{n} p(y_i|x_i(w))$$

6.) THE RECEIVER GUESSES WHICH MESSAGE WAS SENT. THE OPTIMUM PROCEDURE TO MINIMIZE THE PROBABILITY OF ERROR IS MAXIMUM LIKELIHOOD DECODING (I.E. THE RECEIVER SHOULD CHOOSE THE **A POSTERIORI** MOST LIKELY MESSAGE)

INSTEAD (DUE TO COMPLEXITY) WE USE **JOINTLY TYPICAL DECODING**. THE RECEIVER DECLARES THAT THE INDEX $\hat{W}$ WAS SENT IF THE FOLLOWING CONDITIONS ARE SATISFIED:

- $(X^n(\hat{W}), Y^n)$ IS JOINTLY TYPICAL
- THERE IS NO OTHER INDEX $W' \neq \hat{W}$ SUCH THAT $(X^n(W'), Y^n) \in A_\epsilon^{(n)}$

IF NO SUCH $\hat{W}$ EXIST OR THERE IS MORE THAN ONE SUCH, AN ERROR IS DECLARED. (WE ASSUME THAT THE RECEIVER OUTPUTS DUMMY SYMBOL IN THIS CASE E.G. $\emptyset$)

7.) THERE IS DECODING ERROR IF $\hat{W} \neq W$. LET E BE THE EVENT $\{\hat{W} \neq W\}$.

- ANALYSIS OF THE PROBABILITY OF ERROR:

FOR TYPICAL CODEWORD, THERE ARE TWO DIFFERENT SOURCES OF ERROR WHEN WE USE JOINTLY TYPICAL DECODING: EITHER THE OUTPUT Y^n IS NOT JOINTLY TYPICAL WITH TRANSMITTED CODEWORD OR

THERE IS SOME OTHER Y THAT IS JOINTLY TYPICAL WITH Y^n. THE PROBABILITY THAT THE TRANSMITTED CODEWORD AND THE RECEIVED SEQUENCE ARE JOINTLY TYPICAL GOES TO 1, AS SHOWN BY THE JOINT AEP. FOR ANY RIVAL CODEWORD, THE PROBABILITY THAT IT IS JOINTLY TYPICAL WITH THE RECEIVED SEQUENCE IS APPROXIMATELY 2^{-nI} AND HENCE WE CAN USE ABOUT 2^{nI} CODEWORDS AND STILL HAVE LOW PROBABILITY OF ERROR.

DETAILED CALCULATION OF PROBABILITY OF ERROR

WE LET W BE DRAWN ACCORDING TO UNIFORM DISTRIBUTION OVER $\{1,2,\dots,2^{nR}\}$ AND USE JOINTLY TYPICAL DECODING $\hat{W}(Y^n)$.

LET: $\mathcal{E} = \{\hat{W}(Y^n) \neq W\}$ DENOTE THE ERROR EVENT

WE WILL CALCULATE THE AVERAGE PROBABILITY OF ERROR AVERAGED OVER ALL CODEWORDS IN THE CODEBOOK, AND OVER ALL CODEBOOKS:

$$\Pr(\mathcal{E}) = \sum_{C} \Pr(C) P_e^{(n)}(C) = \sum_{C} \Pr(C) \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \lambda_w(C) =$$

$$= \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{C} \Pr(C) \lambda_w(C)$$

BY THE SYMMETRY OF THE CODE CONSTRUCTION, THE AVERAGE PROBABILITY OF ERROR AVERAGED OVER ALL CODES DOES NOT DEPEND ON THE PARTICULAR INDEX THAT WAS SENT [i.e. $\sum_{C} \Pr(C) \cdot \lambda_w(C)$ DOES NOT DEPEND ON w]. THUS WE CAN ASSUME THAT $W=1$ MESSAGE WAS SENT

$$\Pr(\mathcal{E}) = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{C} \Pr(C) \cdot \lambda_w(C) = \sum_{C} \Pr(C) \cdot \lambda_1(C) = \Pr(\mathcal{E} \mid W=1)$$

DEFINE THE FOLLOWING EVENTS:

$$E_i = \{(X^n(i), Y^n) \text{ is in } A_\epsilon^{(n)}\}, \quad i \in \{1,2,\dots,2^{nR}\}$$

WHERE E_i IS THE EVENT THAT THE iTH CODEWORD AND Y^n ARE JOINTLY TYPICAL.

THEN ERROR OCCURS IF EITHER E_1^c OCCURS (THE TRANSMITTED CODEWORD AND THE RECEIVED SEQUENCE ARE NOT JOINTLY TYPICAL) OR $E_2 \cup E_3 \cup \dots \cup E_{2^{nR}}$ OCCURS (WHEN THE WRONG CODE IS JOINTLY TYPICAL WITH RECEIVED SEQUENCE.)

$$\Pr(\mathcal{E} \mid W=1) = P(E_1^c \cup E_2 \cup E_3 \dots \cup E_{2^{nR}} \mid W=1) \leq P(E_1^c \mid W=1) + \sum_{i=2}^{2^{nR}} P(E_i \mid W=1)$$

④

$P(A \cup B) = P(A+B) = P(A) + P(B) - P(A \cap B)$ ⇒ ① #35

BY THE JOINT AEP, $P(E_1^c | W=1) \to 0$

$P(E_1^c | W=1) \leq \varepsilon$ FOR SUFFICIENTLY LARGE „n"

SINCE BY CODE GENERATION PROCESS $X^n(1)$ AND $X^n(i)$ ARE INDEPENDENT FOR $i \neq 1$ SO ARE Y^n AND $X^n(i)$. HENCE THE PROBABILITY THAT $X^n(i)$ AND Y^n ARE JOINTLY TYPICAL IS $\leq$ $2^{-n(I(X;Y)-3\varepsilon)}$ BY THE JOINT AEP.

$$\frac{1}{n} \operatorname{ld} p(X^n) = \frac{1}{n} \sum_{i=1}^{n} \operatorname{ld} p(X_i) \xrightarrow[n \to \infty]{} -E[\operatorname{ld}(X)] = -H(X)$$

$$p(x^n) \to 2^{-n \cdot H(X)}$$

$$Pr(E) = P_1(E|W=1) \leq P(E_1^c | W=1) + \sum_{i=2}^{2^{nR}} P(E_i | W=1) =$$

$$= \varepsilon + (2^{nR} - 1)\, 2^{-n[I(X;Y) - 3\varepsilon]} \leq$$

$$\leq \varepsilon + 2^{nR} \cdot 2^{-n[I(X;Y)-3\varepsilon]} = \varepsilon + 2^{3n\varepsilon} \cdot 2^{-n[I(X;Y)-R]} \leq 2\varepsilon$$

IF „n" IS SUFFICIENTLY LARGE AND:

$$R < I(X;Y) - 3\varepsilon$$

$$2^{-n[-R + I(X;Y) - 3\varepsilon]} \qquad -R + I(X;Y) - 3\varepsilon \geq 0$$

$$R \leq I(X;Y) - 3\varepsilon$$

~~IF „n" IS SUFFICIENTLY LARGE AND ...~~

HENCE IF $R < I(X;Y)$, WE CAN CHOSE ε AND „n" SO THAT THE AVERAGE PROBABILITY OF ERROR, AVERAGED OVER CODEBOOKS AND CODEWORDS IS LESS THEN 2ε.

1.) CHOSE $p(x)$ IN THE PROOF TO BE $p^*(x)$, THE DISTRIBUTION OF X THAT ACHIEVES CAPACITY ($\max_{p(x)} [I(X;Y)]$). THEN THE CONDITION $R < I(X;Y)$ CAN BE REPLACED BY THE ACHIEVABILITY CONDITION: $R < C$

2.) SINCE AVERAGE PROBABILITY OF ERROR OVER CODEBOOKS IS SMALL ($\leq 2\varepsilon$) THERE EXISTS AT LEAST ONE CODEBOOK C^* WITH A SMALL AVERAGE PROBABILITY OF ERROR. THUS:

$$Pr(E | C^*) \leq 2\varepsilon$$

$$Pr(E | C^*) = \frac{1}{2^{nR}} \sum_{i=1}^{2^{nR}} \lambda_i(C^*)$$

NOW THAT THE WORST HALF OF THE CODEWORDS [OF] THE BEST CODEBOOK C^*. SINCE ARITHMETIC [AVERAGE] PROBABILITY OF ERROR $P_e^{(n)}(C^*)$ FOR THIS CODE IS LESS THAN 2ϵ, WE HAVE:

$$Pr(\mathcal{E}|C^*) \le \frac{1}{2^{nR}} \sum \lambda_i(C^*) \le 2\epsilon$$

WHICH IMPLIES THAT AT LEAST HALF THE INDICES i AND THEIR ASSOCIATED CODEWORDS $X^n(i)$ MUST HAVE CONDITIONAL PROBABILITIES λ_i LESS THAN 4ϵ (OTHERWISE, THE CODEWORDS THEMSELVES WOULD CONTRIBUTE MORE THAN 2ϵ TO THE SUM).

$$\frac{1}{2^{nR-1}} \cdot \frac{1}{2} \sum \lambda_i(C^*) \le 2\epsilon$$

$$\boxed{\frac{1}{2^{nR-1}} \sum \lambda_i(C^*) \le 4\epsilon}$$

HENCE THE BEST HALF OF THE CODEWORDS HAVE MAXIMAL PROBABILITY OF ERROR LESS THAN 4ϵ. THROWING OUT THE HALF [OF THE] CODEWORDS REDUCES THE RATE:

$$R' = \frac{\log(2^{nR-1})}{n} = \frac{nR-1}{n} = R - \frac{1}{n}$$

We HAVE CONSTRUCTED CODE WITH RATE: $R' = R - \frac{1}{n}$ WITH MAXIMAL PROBABILITY OF ERROR $\lambda^{(n)} \le 4\epsilon$. THIS PROVES THE ACHIEVABILITY OF ANY RATE BELOW CAPACITY.

$$\boxed{\frac{1}{2^{nR}} \sum \lambda_i(C^*) \le \frac{1}{2^{nR}} 2^{nR-1} \cdot 4\epsilon = \frac{4\epsilon}{2} \le 2\epsilon}$$

[illegible] $\lambda_i \ge 4\epsilon$

7.8 ZERO-ERROR CODES

$X \to Y \to Z \qquad I(X;Y) \ge I(X;Z)$

$$H(Z|X,Y) = H(Z|Y)$$

$$I(X;Y,Z) = I(X;Y) + I(X;Z|Y) = I(X;Y)$$

$$I(X;Z|Y) = H(Z|Y) - H(Z|X,Y) = H(Z|Y) - H(Z|Y)$$

$$I(X;Y,Z) = I(X;Z) + I(X;Y|Z) = I(X;Y)$$

$\ne 0 \qquad \boxed{I(X;Y) \ge I(X;Z)}$

• We WILL NOW PROVE THAT $P_e^{(n)} = 0$ IMPLIES THAT $R \le C$. ASSUME THAT WE HAVE $(2^{nR}, n)$ CODE WITH ZERO PROBABILITY OF ERROR [i.e. THE DECODER OUTPUT $g(Y^n)$ IS

EQUAL TO THE INPUT INDEX W WITH PROBABILITY 1]. THEN THE INPUT INDEX W IS DETERMINED BY THE OUTPUT SEQUENCE [i.e. $H(W|Y^n)=0$].

NOW TO OBTAIN STRONG BOUND, WE ARBITRARY ASSUME THAT W IS UNIFORMLY DISTRIBUTED OVER $\{1,2,\dots,2^{nR}\}$. THUS $H(W)=nR$.

$$nR = H(W) = H(W|Y^n) + H(W) - H(W|Y^n) =$$
$$= \underbrace{H(W|Y^n)}_{0} + I(W;Y^n) = I(W;Y^n)$$
$$\overset{(a)}{\leq} I(X^n;Y^n) \overset{(b)}{\leq} \sum_{i=1}^{n} I(X_i;Y_i) \overset{(c)}{\leq} nC$$

(a) DATA PROCESSING INEQUALITY $W \to X^n(W) \to Y^n$ FOR MARKOV CHAIN

$$I(W;X^n) \geq I(W;Y^n) \qquad I(X^n;Y^n) \geq I(W;Y^n)$$

$$X \to Y \to Z$$

$$I(X;YZ) = I(X;Y) + \underbrace{I(X;Z|Y)}_{0} = I(X;Z) + I(X;Y|Z)$$

$$I(Y;X,Z) = I(X;Y) + \underbrace{I(Y;Z|X)}_{=I(Y;Z)} = I(Y;Z) + I(Y;X|Z)$$

$$I(X;Y) = I(Y;X|Z)$$

$$I(Z;X,Y) = I(Z;X) + I(Z;Y|X) =$$
$$= I(X;Z) + I(Y;Z|X) = I(Z;Y) + I(Z;X|Y) =$$
$$= I(Y;Z) + \underbrace{I(X;Z|Y)}_{0} \Rightarrow \boxed{I(Y;Z) \geq I(X;Z)}$$

$$I(X;Z|Y) = H(X|Y) - H(X|Y,Z) = H(Z|Y) - H(Z|X,Y) =$$
$$= H(Z|Y) - H(Z|Y)$$

$$P(XZ|Y) = \frac{P(XYZ)}{P(Y)} = \frac{P(XY)\cdot P(Z|XY)}{P(Y)} =$$
$$= \frac{P(XY)\cdot P(Z|Y)}{P(Y)} = \frac{P(Y)\cdot P(X|Y)P(Z|Y)}{P(Y)} =$$
$$= P(X|Y)\cdot P(Z|Y); \qquad P(YZ|X) = \frac{P(XYZ)}{P(X)} =$$
$$= \frac{P(XY)\cdot P(Z|XY)}{P(X)} = \frac{P(X)\cdot P(Y|X)P(Z|Y)}{P(X)} = P(Y|X)P(Z|Y)$$

(b) WILL BE PROVED IN THE THEOREM 7.9.2 USING DISCRETE MEMORYLESS ASSUMPTION.

(c) FOLLOWS FROM THE DEFINITION OF INFORMATION CAPACITY.

HENCE FOR ANY ZERO-ERROR $(2^{nR}, n)$ CODE FOR ALL $n \geq 1$

$$\boxed{R \leq C}$$

$H(X,Y) = H(X) + H(Y|X)$

7.9 FANO'S INEQUALITY AND THE CONVERSE TO THE CODING THEOREM

WE NOW EXTEND THE PROOF THAT WAS DERIVED FOR ZERO-ERROR CODES TO THE CASE OF CODES WITH VERY SMALL PROBABILITY OF ERROR.

RECALL

$X \to Y \to \hat{X}$

$$E = \begin{cases} 1 & \hat{X} \neq X \\ 0 & \hat{X} = X \end{cases} \qquad P(E) = \begin{cases} P_e \\ 1-P_e \end{cases}$$

$$H(X|\hat{X},E) = P(E=0)\cdot H(X|\hat{X},E=0) + P(E=1)\, H(X|\hat{X},E=1)$$

$$= (1-P_e)\cdot 0 + P_e \cdot H(X|\hat{X},E=1) \leq P_e \cdot H(X|\hat{X}) \leq P_e H(X)$$

$$\leq P_e \operatorname{ld}|X|$$

$H(X|\hat{X}) = ?$ $\quad H(E,X|\hat{X}) = H(X|\hat{X}) + H(E|X,\hat{X})$

$$= H(E|\hat{X}) + H(X|\hat{X},E) \leq H(P_e) + P_e \operatorname{ld}|X|$$

~~$H(E|\hat{X}) = P(\hat{X}=x) H(E|\hat{X}=x) + P(\hat{X}\neq x) H(E|\hat{X}\neq x)$~~

$$H(E|\hat{X}) \leq H(E) = H(P_e) = P_e \operatorname{ld}\frac{1}{P_e} + (1-P_e)\operatorname{ld}\frac{1}{1-P_e}$$

$$H(X|\hat{X}) \leq H(P_e) + P_e \operatorname{ld}|X| \leq 1 + P_e \operatorname{ld}|X|$$

$$P_e \geq \frac{H(X|\hat{X})-1}{\operatorname{ld}|X|} \quad / \cdot(-1) \Rightarrow P_e \leq \frac{1-H(X|\hat{X})}{\operatorname{ld}|X|}$$

- STRONGER INEQUALITY: $H(X|\hat{X},E) \leq P_e \operatorname{ld}(|X|-1)$

$$P_e \leq \frac{1-H(X|\hat{X})}{\operatorname{ld}(|X|-1)} \qquad \boxed{P_e \geq \frac{H(X|Y)-1}{\operatorname{ld}(|X|-1)}}$$

- MARKOV CHAIN + DATA PROCESSING INEQ:

$I(X;\hat{X}) \leq I(X;Y)$ $\qquad H(X) - H(X|\hat{X}) \leq$

$H(X) - H(X|Y)$ $\qquad H(X|\hat{X}) \geq H(X|Y)$

$H(X|Y) \leq H(X|\hat{X}) \leq 1 + P_e \operatorname{ld}|X|$ $\qquad H(X|Y) \leq 1 + P_e \operatorname{ld}|X|$

$$\boxed{P_e \leq \frac{1-H(X|Y)}{\operatorname{ld}|X|}}$$

- THE INDEX IS UNIFORMLY DISTRIBUTED ON THE SET $W = \{1, 2, \dots 2^{nR}\}$ AND SEQUENCE Y^n IS RELATED PROBABILISTICALLY TO W.

FROM Y^n WE ESTIMATE THE INDEX W THAT WAS SENT. LET THE ESTIMATE BE $\hat{W} = g(Y^n)$.

THUS: $W \to X^n(W) \to Y^n \to \hat{W}$ FORMS MARKOV CHAIN. PROBABILITY OF ERROR IS:

$$P_r(\hat{W} \neq W) = \frac{1}{2^{nR}} \sum_i \lambda_i = P_e^{(n)}$$

LEMMA 7.9.1 (FANO'S INEQUALITY) FOR A DISCRETE MEMORYLESS CHANNEL WITH CODEBOOK C AND THE INPUT MESSAGE W UNIFORMLY DISTRIBUTED OVER 2^{nR}, WE HAVE:

$$H(W|\hat{W}) \leq 1 + P_e^{(n)} \cdot nR.$$

PROOF: $W \sim$ uniform distro $P_e^{(n)} = P_r(W \neq \hat{W})$ + APPLY FANO INEQUALITY

- CAPACITY PER TRANSMISSION IS NOT INCREASED IF WE USE A DISCRETE MEMORYLESS CHANNEL MANY TIMES.

LEMMA 7.9.2 LET $Y^{(n)}$ BE THE RESULT OF PASSING X^n THROUGH A DISCRETE MEMORYLESS CHANNEL OF CAPACITY C. THEN:

$$I(X^n; Y^n) \leq nC \quad \text{for all } p(x^n)$$

PROOF: $I(X^n; Y^n) = H(Y^n) - H(Y^n|X^n) =$

$$= H(Y^n) - \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}, X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i)$$

SINCE BY DEFINITION OF THE DISCRETE MEMORYLESS CHANNEL DMC, Y_i DEPENDS ONLY ON X_i AND IS CONDITIONALLY INDEPENDENT OF EVERYTHING ELSE.

$$I(X^n; Y^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i) \leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i)$$

$$= \sum_{i=1}^{n} I(X_i; Y_i) \leq \underline{nC}$$

ENTROPY OF COLLECTION OF RANDOM VARIABLES IS LESS THAN THE SUM OF INDIVIDUAL ENTROPIES:

$$H(Y_1^n) = \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}) \leq \sum_{i=1}^{n} H(Y_i)$$

CONDITIONING REDUCES ENTROPY

... 7.7.1 Channel ... theorem ...

$$W \to X^n \to Y^n \to \hat{W} \qquad I(W; X^n) \geq I(W; Y^n)$$

$$I(W; X^n, Y^n) \geq I(W; \hat{W})$$

$$I(W; X^n, Y^n) = I(X^n, Y^n; W) = I(X^n; W) + I(Y^n; W | X^n)$$

$$I(Y^n; W) \geq I(W; \hat{W}) \qquad I(W; Y^n) \geq I(W; \hat{W})$$

$$I(Y^n; W) + 0 = I(X^n; W) + I(Y^n; W | X^n)$$

$$I(W, X^n; Y^n) = I(W; Y^n) + I(X^n; Y^n | W) =$$

$$= I(X^n; Y^n) + I(W; Y^n | X^n) \Rightarrow I(X^n; Y^n) \geq I(W; Y^n)$$

$$I(X^n; Y^n) \geq I(W; Y^n) \geq I(W; \hat{W})$$

$$\boxed{I(X^n; Y^n) \geq I(W; \hat{W})} \quad (*)$$

MMV

Let W be drawn according uniform distro $\{1, 2, \ldots, 2^{nR}\}$ $\quad Pr(\hat{W} \neq W) = P_e^{(n)} = \frac{1}{2^{nR}} \sum \lambda_i$

$$= H(W) - H(W|\hat{W})$$

$$nR = H(W) = H(W|\hat{W}) + I(W; \hat{W}) \leq 1 + P_e \log 2^{nR} + I(W; \hat{W})$$

$$= 1 + P_e \cdot nR + I(W; \hat{W}) \leq 1 + P_e nR + I(X^n; Y^n) \leq \text{(lemma 7.7.2)}$$

$$\leq 1 + P_e^{(n)} nR + nC \quad / n$$

$$\boxed{R \leq \frac{1}{n} + P_e^{(n)} \cdot R + C} \quad \text{IF } n \to \infty \quad \boxed{R \leq C}$$

$$\boxed{P_e^{(n)} \geq 1 - \frac{1}{nR} - \frac{C}{R}} \quad \text{(MMV)} \qquad \text{IF } R > C$$

PROBABILITY OF ERROR IS BOUNDED AWAY FROM 0 FOR SUFFICIENTLY LARGE n. HENCE WE CANNOT ACHIEVE AN ARBITRARILY LOW PROBABILITY OF ERROR AT RATES ABOVE CAPACITY. THIS CONVERSE IS CALLED WEAK CONVERSE TO THE CHANNEL CODING THEOREM. IT IS ALSO POSSIBLE TO PROVE A STRONG CONVERSE WHICH STATES THAT FOR RATES ABOVE CAPACITY, PROBABILITY OF ERROR GOES EXPONENTIALLY TO 1.

HENCE CAPACITY IS A VERY CLEAR DIVIDING POINT – AT RATES BELOW CAPACITY $P_e^{(n)} \to 0$ EXPONENTIALLY AND AT RATES ABOVE CAPACITY $P_e^{(n)} \to 1$ EXPONENTIALLY.

7.10 EQUALITY IN THE CONVERSE TO THE CHANNEL CODING THEOREM

– WE REPEAT THE STEPS OF THE PROOF OF THE CONVERSE $P_e = 0$ WITH MORE IT WILL GIVE SOME IDEAS FOR THE CODES THAT ACHIEVE CAPACITY.

$$nR = H(W) = \underbrace{H(W|\hat{W})}_{\text{NO ERROR} = 0} + I(W;\hat{W}) = I(W;\hat{W})$$

$$\overset{(a)}{\leq} I(X^n;Y^n) = H(Y^n) - H(Y^n|X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1^{i-1},X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i) \leq$$

$$\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) = \sum_{i=1}^{n} I(X_i;Y_i)$$

$$\leq nC$$

(a) DATA PROCESSING

EQUALITY IF: (★) $I(X^n,Y^n) = I(W;Y^n) + \underbrace{I(X^n;Y^n|W)}_{\text{THEOR. DA} = 0}$

$$I(X_n,Y_n;\hat{W}) = I(X_n;\hat{W}) + \underbrace{I(X_n;\hat{W}|Y^n)}_{0} =$$

$$= I(X_n;\hat{W}) + I(Y_n;\hat{W}|X^n)$$

$$I(X_n;Y_n,\hat{W}) = I(X_n;Y_n) + \underbrace{I(X_n;\hat{W}|Y_n)}_{0} =$$

$$= I(X_n;\hat{W}) + I(X_n;Y_n|\hat{W})$$

$$I(X_n;Y_n) = I(X_n;\hat{W}) + \underbrace{I(X^n;Y^n|\hat{W})}_{\text{THEOR. DA} = 0}$$

$$I(W;X_nY_n;\hat{W}) = I(X_n,Y_n;\hat{W}) + \underbrace{I(W;\hat{W}|X_n,Y_n)}_{0} =$$

$$= I(W;\hat{W}) + I(X_n,Y_n;\hat{W}|W) = I(Y_n;\hat{W}|W)$$

(★) $\boxed{I(X_n,Y_n;\hat{W}) = I(Y_n;\hat{W})}$

$$I(Y_n;\hat{W}) = I(W;\hat{W}) + I(X_n,Y_n;\hat{W}|W) \quad (\triangle)$$

$$I(W, X^n; Y^n, \hat{W}) = I(X^n; Y^n, \hat{W}) + I(W; Y^n, \hat{W} | X^n)$$
$$= I(X^n; Y^n) + I(W; Y^n, \hat{W} | X^n)$$

$$I(Y_n; \hat{W}) = I(W; \hat{W}) + I(Y_n; \hat{W} | W)$$
$$I(Y_n; \hat{W}) - I(Y_n; \hat{W} | W) = I(W; \hat{W})$$

$$I(X_n; Y_n, W, \hat{W}) = I(X_n; Y_n) + I(X_n; W | Y_n) + I(X_n; \hat{W} | Y_n W)$$

SOLUTION

$$I(W, X^n; Y^n, \hat{W}) = I(X^n; Y^n \hat{W}) + I(W; Y^n \hat{W} | X^n)$$ 0

$$= I(W, X^n; Y^n) + I(W, X^n; \hat{W} | Y^n) = (*)$$ 0

$$I(X^n; Y^n, \hat{W}) = I(W, X^n; Y^n) = I(X^n; Y^n)$$ PROOFED!

$$(*) = I(W; Y^n \hat{W}) + I(X^n; Y^n \hat{W} | W) =$$
$$= I(W, X_n; \hat{W}) + I(W, X_n; Y_n | \hat{W})$$ $\oplus$

$$I(X^n; Y^n) = I(W, X^n; Y^n, \hat{W}) = I(W; Y^n \hat{W}) + I(X^n; Y^n | W)$$

$$I(W; Y^n \hat{W}) = I(W; \hat{W}) + I(W; Y^n | \hat{W}) =$$
$$= I(W; Y^n) + I(W; \hat{W} | Y_n)$$

$$I(X^n; Y^n) = I(W; \hat{W}) + I(W, Y^n | \hat{W}) + I(X^n; Y^n | W)$$ (4)

$$(4) = H(W | \hat{W}) - H(W | \hat{W}, Y^n) + H(X^n) - H(X^n | W Y^n)$$ 0 NO ERROR!! 0 $I(X^n; Y^n | W)$

$$I(X^n; Y^n) = I(W; \hat{W}) + I(X^n; Y^n | W)$$

$$I(X^n; Y^n) = I(W; \hat{W}) \quad IF \quad I(X^n; Y^n | W) = 0$$

$$I(X^n; Y^n) = I(W, X^n; Y^n, \hat{W}) = I(W X^n; \hat{W}) + I(X_n; Y_n | \hat{W})$$
$$= I(W; \hat{W}) + I(X^n; \hat{W} | W) + I(X_n; Y_n | \hat{W}) = I(W; \hat{W}) + I(X_n; Y_n | \hat{W})$$

$$I(X^n; \hat{W} | W) = H(\hat{W} | W) + H(\hat{W} | W, X^n)$$ 0 0 NO ERROR!!

$$I(X^n;Y^n) = I(W;\hat{W}) \quad \text{IF} \quad I(X^n;Y^n|\hat{W}) = 0$$

→ PREVIOUS CONDITION WAS $I(X^n;Y^n|W) = 0$

(A) WE HAVE INEQUALITY IF ONLY

$$I(X^n;Y^n|W) = I(X^n;Y^n|\hat{W}) = 0$$

WHICH IS TRUE IF ALL THE CODEWORDS ARE DISTINCT AND IF $\hat{W}$ IS SUFFICIENT STATISTIC FOR DECODING.

(B) WE HAVE EQUALITY IF Y_i'S ARE INDEPENDENT

(C) WE HAVE EQUALITY ONLY IF THE DISTRIBUTION OF X_i IS $p^*(x)$, THE DISTRIBUTION OF X THAT ACHIEVES THAT CAPACITY.

THIS INDICATES THAT A CAPACITY-ACHIEVING ZERO-ERROR CODE HAS DISTINCT CODEWORDS AND THE DISTRIBUTION OF THE Y_i'S MUST BE I.I.D WITH

$$p^*(y) = \sum_x p^*(x)\, p(y|x)$$

THE DISTRIBUTION OF Y INDUCED BY THE OPTIMUM DISTRIBUTION OF X^*.

$$p(x_i, y_i) = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} I(x_i(w) = x_i)\, p(y_i|x_i)$$

$$R \leq Pe \cdot R + \frac{1}{n} + C \qquad Pe \geq 1 - \frac{1}{Rn} - \frac{C}{R}$$

7.11 HAMMING CODES

THE CHANNEL CODING THEOREM PROMISES THE EXISTENCE OF BLOCK CODES THAT WILL ALLOW US TO TRANSMIT INFORMATION AT RATES BELLOW CAPACITY WITH AN ARBITRARY SMALL PROBABILITY OF ERROR IF THE BLOCK LENGTH IS LARGE ENOUGH.

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \qquad Z^T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

(i-TH COLUMN marked in H; first row of Z^T marked C_1)

→ ONE OF NULL MATRIX.

$$H \cdot Z = 0$$

SINCE THE SET OF CODEWORDS IS THE NULL SPACE OF A MATRIX IT IS LINEAR IN THE SENSE THAT THE SUM OF ANY TWO CODEWORDS IS ALSO A CODEWORD.

THE MINIMUM NUMBER OF 1'S IN ANY CODEWORD IS 3. THIS IS CALLED MINIMUM WEIGHT

OF THE CODE MINIMUM DISTANCE (Hamming) IS EQUAL TO THE MINIMUM WEIGHT OF THE LINEAR CODE.

SINCE THE CODE IS LINEAR THE DIFFERENCE BETWEEN TWO CODEWORDS IS ALSO A CODEWORD, AND HENCE ANY TWO CODEWORDS DIFFER IN AT LEAST THREE PLACES.

FOR THE CODE ABOVE THE MINIMUM DISTANCE IS 3. HENCE, IF THE CODEWORD "C" IS CORRUPTED IN ONLY ONE PLACE IT WILL DIFFER FROM ANY OTHER WORD IN AT LEAST TWO PLACES, AND HENCE IT WILL BE CLOSER TO "C" THAN TO ANY OTHER CODEWORD.

THE MATRIX H [illegible] CHECK MATRIX HAS THE PROPERTY [illegible] EVERY CODEWORD "C"

$$H \cdot C = 0$$

LET e_i BE A [illegible] POSITION AND [illegible] CORRUPTED AT [illegible]:

$$r = c + e_i$$

$$H \cdot r = H(c + e_i) = H \cdot c + H \cdot e_i = H \cdot e_i$$

HENCE LOOKING AT $H \cdot r$ WE CAN FIND WHICH POSITION OF THE VECTOR WAS CORRUPTED. REVERSING THIS BIT WILL GIVE US THE CODEWORD.

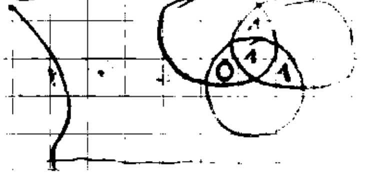
WE HAVE CONSTRUCTED 16 CODEWORDS OF LENGTH 7, WHICH CAN CORRECT UP TO ONE ERROR. THIS CODE IS CALLED HAMMING CODE.

0000	000	0100	101	1000	011	1100	110
0001	111	0101	010	1001	100	1101	001
0010	110	0110	011	1010	101	1110	000
0011	001	0111	100	1011	010	1111	111

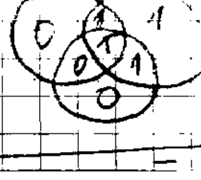
k - BITS REPRESENT THE MESSAGE

n-k - BITS OF THE CODEWORD REPRESENT THE PARITY CHECK BITS. SUCH CODE IS CALLED A SYSTEMATIC CODE.

- THE ABOVE CODE IS CALLED (7,4,3) HAMMING CODE (i.e. n=7, k=4, d=3). d ← MINIMUM DISTANCE OF THE CODE.

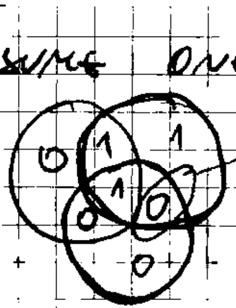


1101 INFORMATION (MESSAGE) BITS



PARITY CHECK BITS (EVEN PARITY FOR EACH CIRCLE)

- ASSUME ONE BIT IS CHANGED



— GREŠEN BIT (PRESEK NA KRUGOVITE SO POGREŠEN PARITY CHECK) — BITOT ŠTO PRIPAĞA NA "KRUGOVI" NA SO PRAVILEN PARITY CHECK.

• GENERALIZATION OF HAMMING CODE:

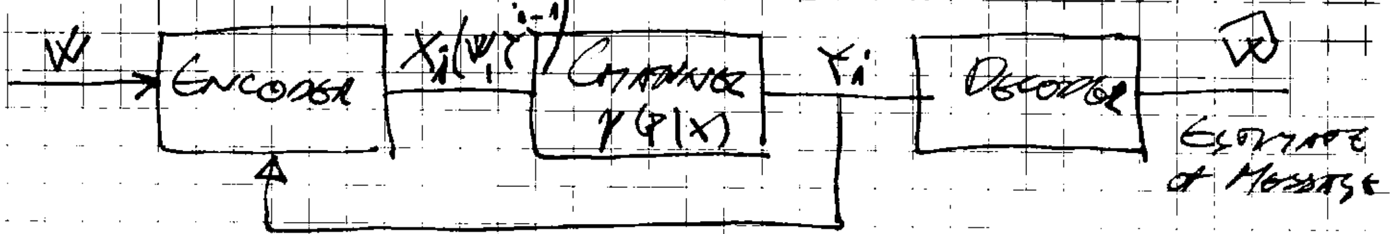
$n = 2^l - 1$ $\quad k = 2^l - l - 1$ AND MINIMUM DISTANCE (3)

l - NUMBER OF ROWS IN H MATRIX

e.g. $l = 3$ $\quad n = 2^3 - 1 = 7$ $\quad k = 2^3 - 3 - 1 = 8 - 4 = 4$

1 1 1 1 1

7.12 FEEDBACK CAPACITY



WE DEFINE A $(2^{nR}, n)$ feedback code AS A SEQUENCE OF MAPPINGS $X_i(W, Y^{i-1})$, WHERE EACH X_i IS A FUNCTION ONLY OF THE MESSAGE $W \in 2^{nR}$ AND PREVIOUS RECEIVED VALUES: $Y_1, Y_2, \ldots, Y_{i-1}$ AND A SEQUENCE OF DECODING FUNCTIONS $g: Y^n \to \{1, 2, \ldots, 2^{nR}\}$

$$P_e^{(n)} = P_r\{g(Y^n) = W\}$$

WHEN W IS UNIFORMLY DISTRIBUTED OVER $\{1, 2, \ldots, 2^{nR}\}$

DEFINITION: THE CAPACITY WITH FEEDBACK, C_{FB}, OF THE DISCRETE MEMORYLESS CHANNEL IS THE SUPREMUM OF ALL RATES ACHIEVABLE BY FEEDBACK CODES.

THEOREM 7.12.1 (FEEDBACK CAPACITY)

$$C_{FB} = C = \max_{p(x)} I(X;Y)$$

PROOF: LET W BE UNIFORMLY DISTRIBUTED OVER $\{1, 2, 3, \ldots, 2^{nR}\}$. THEN $P_r(W \neq \hat{W}) = P_e^{(n)}$

$$nR = H(W) = H(W|\hat{W}) + I(W;\hat{W}) \underset{(a)}{\leq} \to H(W) - H(W|\hat{W})$$

$$1 + P_e^{(n)} nR + I(W;\hat{W}) \underset{(b)}{\leq} 1 + P_e^{(n)} n \cdot R + I(W;Y^n)$$

(a) FANO INEQUALITY

(b) DATA-PROCESSING INEQUALITY

$$I(W;Y^n) = H(Y^n) - H(Y^n|W) =$$

$$= H(Y^n) - \sum_{i=1}^{n} H(Y_i | Y_1, Y_2 \dots Y_{i-1}, W) =$$

$$= \left| X_i = f(W, Y_1, Y_2, \dots Y_{i-1}) \right| = H(Y^n) - \sum_{i=1}^{n} H(Y_i | Y_1^{i-1}, W, X_i)$$

$$H(Z|X,Y) \qquad X = f(Y)$$

$$H(X,Z|Y) = \underbrace{H(X|Y)}_{0} + H(Z|X,Y) =$$

$$= H(Z|X,Y) = H(Z|Y) + \underbrace{H(X|Y,Z)}_{0} \qquad \boxed{H(Z|X,Y) = H(Z|Y)}$$

$$= H(Y^n) - \sum_{i=1}^{n} H(Y_i | X_i)$$

$W \to X^{(n)} \to Y^{(n)}$ } MARKOV CHAIN

$$H(Y|X,W) = H(Y|X)$$

$$I(W;Y^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i) \leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) = \sum_{i=1}^{n} I(Y_i;X_i) \leq nC$$

$$nR \leq P_e^{(n)} n \cdot R + 1 + nC \quad / n$$

$$R \leq P_e^{(n)} \cdot R + \frac{1}{n} + C \underset{n \to \infty}{\to} 0 + 0 + C$$

$\boxed{R \leq C}$ THUS WE CANNOT ACHIEVE ANY HIGHER RATES WITH FEEDBACK THAN WE CAN WITHOUT FEEDBACK, AND

$\boxed{C_{FB} = C}$

- AS WE HAVE SEEN IN THE EXAMPLE OF THE BINARY ERASURE CHANNEL, FEEDBACK CAN HELP ENORMOUSLY IN SIMPLIFYING ENCODING AND DECODING. HOWEVER IT CANNOT INCREASE THE CAPACITY OF THE CHANNEL

7.13 SOURCE - CHANNEL SEPARATION THEOREM

$$\frac{H(X^n)}{n} \leq \frac{E[L(X^n)]}{n} \leq \frac{H(X^n)+1}{n}$$

$$R = \frac{\log M}{n} = \frac{\log 2^{nR}}{n} \leq C$$

IT IS NOW TIME TO COMBINE THE TWO MAIN RESULTS THAT WE HAVE PROVED SO FAR:

- DATA COMPRESSION $R > H$ (THEOREM 5.4.2)
- DATA TRANSMISSION $R < C$ (THEOREM 7.7.1)

IS THE CONDITION $H < C$ NECESSARY AND SUFFICIENT FOR SENDING THE SOURCE OVER THE CHANNEL.

WE CAN CONSIDER THE DESIGN OF ~~THE~~ A COMMUNICATION SYSTEM AS A COMBINATION OF TWO PARTS, SOURCE CODING AND CHANNEL CODING.
WE CAN DESIGN SOURCE CODES FOR THE MOST EFFICIENT REPRESENTATION OF THE DATA. WE CAN, SEPARATELY AND INDEPENDENTLY DESIGN CHANNEL CODES APPROPRIATE FOR THE CHANNEL. THE COMBINATION WILL BE AS EFFICIENT AS ANYTHING WE COULD DESIGN BY CONSIDERING BOTH PROBLEMS TOGETHER.

LET US DEFINE SETUP UNDER CONSIDERATION. WE HAVE A SOURCE V THAT GENERATES SYMBOLS FROM AN ALPHABET $\mathcal{V}$. THE STOCHASTIC PROCESS GENERATED BY V IS FROM FINITE ALPHABET AND SATISFIES AEP. (E.G. SEQUENCE OF I.I.D RANDOM VARIABLES AND THE SEQUENCE OF STATES OF A STATIONARY IRREDUCIBLE MARKOV CHAIN.

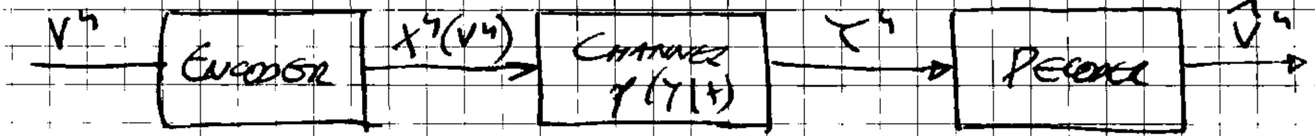
ANY STATIONARY ERGODIC SOURCE SATISFIES AEP

- WE WANT TO SEND SEQUENCE OF SYMBOLS:

$$V^n = V_1, V_2, \ldots, V_n$$

OVER THE CHANNEL SO THAT THE RECEIVER CAN RECONSTRUCT THE SEQUENCE. WE MAP THE SEQUENCE ONTO A CODEWORD $X^n(V^n)$ AND SEND THE CODEWORD OVER THE CHANNEL.
THE RECEIVER LOOKS AT HIS RECEIVED SEQUENCE Y^n AND MAKES ESTIMATE $\hat{V}^n$. ERROR IS DECLARED IF $\hat{V}^{(n)} \neq V^{(n)}$. WE DEFINE:

$$P_r(V^n \neq \hat{V}^n) = \sum_{y^n} \sum_{v^n} p(v^n)\, p(y^n \mid x^n(v^n))\, I(g(y^n) \neq v^n)$$



THEOREM 7.13.1 (JOINT SOURCE-CHANNEL CODING THEOREM)

IF $V_1, V_2, \dots V^n$ IS A FINITE ALPHABET STOCHASTIC PROCESS THAT SATISFIES THE AEP AND $H(\mathcal{V}) < C$ THERE EXISTS A SOURCE-CHANNEL CODE WITH PROBABILITY OF ERROR $P_r(\hat{V}^n \neq V^n) \to 0$. CONVERSELY, FOR ANY STATIONARY STOCHASTIC PROCESS, IF $H(\mathcal{V}) > C$, THE PROBABILITY OF ERROR IS BOUNDED AWAY FROM ZERO, AND IT IS NOT POSSIBLE TO SEND THE PROCESS OVER THE CHANNEL WITH ARBITRARY LOW PROBABILITY OF ERROR.

PROOF ACHIEVABILITY SINCE WE HAVE ASSUMED THAT THE STOCHASTIC PROCESS SATISFIES THE AEP, IT IMPLIES THAT THERE EXISTS A TYPICAL SET:

$$A_\varepsilon^{(n)} \text{ OF SIZE } \leq 2^{n(H(\mathcal{V})+\varepsilon)}$$

WE WILL ENCODE ONLY THE SOURCE SEQUENCES BELONGING TO THE TYPICAL SET, ALL OTHER SEQUENCES WILL RESULT IN ERROR.

SINCE THERE ARE AT MOST $2^{n(H+\varepsilon)}$ SUCH SEQUENCES $n(H+\varepsilon)$ BITS SUFFICE TO INDEX THEM.

WE CAN TRANSMIT THE DESIRED INDEX TO THE RECEIVER WITH PROBABILITY OF ERROR LESS THAN ε IF:

$$H(\mathcal{V}) + \varepsilon = R \leq C$$

$$P(V^n \neq \hat{V}^n) \leq P(V^n \notin A_\varepsilon^{(n)}) + P(g(Y^n) \neq V^n \mid V^n \in A_\varepsilon^{(n)})$$
$$\leq \varepsilon + \varepsilon = 2\varepsilon$$

FOR SUFFICIENTLY LARGE "n". HENCE WE CAN RECONSTRUCT THE SEQUENCE WITH LOW P_E FOR SUFFICIENTLY LARGE "n" IF:

$$H(\mathcal{V}) < C$$

CONVERSE WE WISH TO SHOW THAT $P_r(\hat{V}^n \neq V^n) \to 0$ IMPLIES THAT $H(\mathcal{V}) \leq C$ FOR ANY SEQUENCE OF SOURCE-CHANNEL CODES:

$$X^n(V^n): \mathcal{V}^n \to \mathcal{X}^n \qquad g_n(Y^n): \mathcal{Y}^n \to \hat{\mathcal{V}}^n$$

$$H(V^n|\hat{V}^n) \leq 1 + P_r(\hat{V}^n \neq V_n)\, ld|\mathcal{V}^n| = 1 + P_r(\hat{V}^n \neq V^n) \cdot n\, ld|\mathcal{V}|$$

$$H(V) \overset{(a)}{\leq} \frac{H(V_1, V_2 \dots V_n)}{n} = \frac{H(V^n)}{n} = \frac{1}{n} H(V^n|\hat{V}^n) +$$

$$+ \frac{1}{n} I(V^n;\hat{V}^n) \overset{(b)}{\leq} \frac{1}{n}\left(1 + P_r(\hat{V}_n \neq V_n)\cdot n \log|V|\right) + \frac{1}{n} I(V^n;\hat{V}^n)$$

$$\overset{(c)}{\leq} \frac{1}{n}\left(1+P_r(\hat{V}^n\neq V^n)\cdot n \log|V|\right) + \frac{1}{n} \underbrace{I(X^n;Y^n)}_{\leq n\cdot C} \overset{(d)}{\leq} \frac{1}{n} +$$

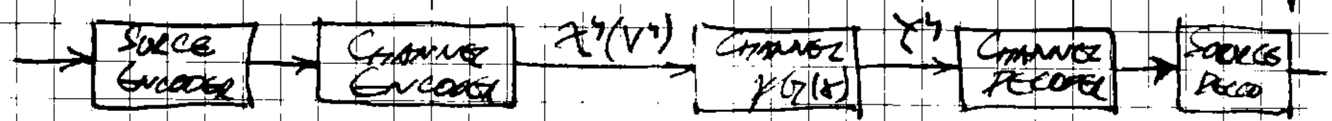
$$+ P_r(\hat{V}^n \neq V^n) \log|V| + C$$

- NOW LETTING $n \to \infty$ WE HAVE:

$$\boxed{H(V) \leq C}$$

HENCE WE CAN TRANSMIT A STATIONARY ERGODIC SOURCE OVER A CHANNEL IF AND ONLY IF ITS ENTROPY RATE IS LESS THEN THE CAPACITY OF THE CHANNEL.

THE SOURCE CODER TRIES TO FIND THE MOST EFFICIENT REPRESENTATION OF THE SOURCE, AND THE CHANNEL CODER ENCODES THE MESSAGE TO COMBAT THE NOISE AND ERRORS INTRODUCED BY THE CHANNEL.



- THE DATA COMPRESSION THEOREM IS CONSEQUENCE TO AEP WHICH SHOWS THAT THERE EXIST TYPICAL SET $\left(|A_\epsilon^n| \leq 2^{nH}\right)$ THAT WE CAN PRESENT THE SOURCE WITH SMALL PROBABILITY OF ERROR USING H BITS PER SYMBOL. THE DATA TRANSMISSION THEOREM IS BASED ON JOINT AEP; IT USES THE FACT THAT FOR LONG BLOCK LENGTHS, THE OUTPUT SEQUENCE OF THE CHANNEL IS VERY LIKELY TO BE JOINTLY TYPICAL WITH INPUT CODEWORD, WHILE ANY OTHER CODEWORD IS JOINTLY TYPICAL WITH PROBABILITY $= 2^{-nI}$. HENCE WE CAN USE 2^{nI} CODEWORDS AND STILL HAVE NEGLIGIBLE PROBABILITY OF ERROR.

SUMMARY

■ **CHANNEL CAPACITY**. THE LOGARITHM OF THE NUMBER OF DISTINGUISHABLE INPUTS IS GIVEN BY

$$C = \max_{p(x)} I(X;Y)$$

EXAMPLES:

- BINARY SIMETRIC CHANNEL: $C = 1 - H(p)$
- BINARY ERASURE CHANNEL: $C = 1 - \alpha$
- SIMETRIC CHANNEL $C = \log|Y| - H$(ROW OF TRANSITION MATRIX)

PROPERTIES OF C

1.) $0 \leq C \leq \min\{\log|\mathcal{X}|, \log|\mathcal{Y}|\}$

2.) $I(X;Y)$ IS A CONTINUOUS CONCAVE FUNCTION OF $p(x)$

JOINT TYPICALITY: THE SET $A_\epsilon^{(n)}$ OF JOINTLY TYPICAL SEQUENCES $\{(x^n, y^n)\}$ WITH RESPECT TO THE DISTRIBUTION $p(x,y)$ IS GIVEN BY:

$$A_\epsilon^{(n)} = \{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \left|-\tfrac{1}{n}\log p(x^n) - H(X)\right| < \epsilon,$$
$$\left|-\tfrac{1}{n}\log p(y^n) - H(Y)\right| < \epsilon, \left|-\tfrac{1}{n}\log p(x^n, y^n) - H(X,Y)\right| < \epsilon\}$$

WHERE: $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$

JOINT AEP LET (X^n, Y^n) BE SEQUENCES OF LENGTH n DRAWN I.I.D ACCORDING TO: $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$. THEN:

1. $\Pr((X^n, Y^n) \in A_\epsilon^{(n)}) \to 1 \quad n \to \infty$
2. $|A_\epsilon^{(n)}| \leq 2^{n(H(X,Y)+\epsilon)}$
3. IF $(\tilde{X}^n, \tilde{Y}^n) \sim p(x^n)p(y^n)$ THEN $\Pr((\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}) \leq 2^{-n(I(X;Y) - 3\epsilon)}$

CHANNEL CODING THEOREM. ALL RATES BELOW C ARE ACHIEVABLE, AND ALL RATES ABOVE CAPACITY ARE NOT; THAT IS, FOR ALL RATES $R<C$, THERE EXIST SEQUENCE $(2^{nR}, n)$ CODES WITH PROBABILITY OF ERROR $\lambda^{(n)} \to 0$. CONVERSELY, FOR RATES $R > C$, $\lambda^{(n)}$ IS BOUNDED AWAY FROM 0.

FEEDBACK CAPACITY. FEEDBACK DOES NOT INCREASE CAPACITY FOR DISCRETE MEMORYLESS CHANNEL. (i.e. $C_{FB} = C$)

SOURCE-CHANNEL THEOREM: A STOCHASTIC PROCESS WITH ENTROPY RATE H CANNOT BE SENT RELIABLY OVER A DISCRETE MEMORYLESS CHANNEL IF $H > C$. CONVERSELY, IF THE PROCESS SATISFIES AEP, THE SOURCE CAN BE TRANSMITTED RELIABLY IF $H < C$.

RECALL

$$1 \geq \sum_{x^n y^n \in A_\epsilon} p(x^n, y^n) \geq |A_\epsilon^{(n)}| \cdot 2^{-n(H(X,Y)+\epsilon)}$$

$$\boxed{|A_\epsilon^n| \leq 2^{+n(H(X,Y)+\epsilon)}} \quad 2^{-n(H(X,Y)+\epsilon)} \leq p(x^n, y^n) \leq 2^{-n(H(X,Y)-\epsilon)}$$

$$1 = \sum_{x^n y^n} p(x^n y^n) \geq \sum_{(x^n y^n)\in A_\varepsilon^{(n)}} p(x^n)\cdot p(y^n) \geq 2^{+n(H(X,Y)+\varepsilon)} \cdot 2^{-n(H(X)+\varepsilon)} \cdot 2^{-n(H(Y)+\varepsilon)} = 2^{nH(X,Y)+n\varepsilon - nH(X) - nH(Y) - 2n\varepsilon}$$

$$\Pr\left((\tilde{X}_n, \tilde{Y}_n) \in A_\varepsilon^{(n)}\right) = \sum_{\tilde{x}_n \tilde{y}_n \in A_\varepsilon^n} p(x_n)\cdot p(y_n) \leq |A_\varepsilon^{(n)}| \cdot 2^{-n(H(X)-\varepsilon)} \cdot 2^{-n(H(Y)-\varepsilon)} \leq 2^{\underbrace{n(H(X,Y)+\varepsilon) - nH(X) + n\varepsilon - nH(Y) + n\varepsilon}_{(*)}}$$

$$(*) = n\left(H(Y) + H(X|Y)\right) + 3n\varepsilon - nH(X) - nH(Y) =$$

$$= -n\left(H(X) - H(X|Y)\right) + 3n\varepsilon$$

$$\Pr\left((\tilde{X}_n, \tilde{Y}_n) \in A_\varepsilon^{(n)}\right) \leq 2^{-n[I(X;Y) - 3\varepsilon]}$$

- $$(1-\varepsilon) \leq \sum_{x^n y^n \in A_\varepsilon^{(n)}} p(x^n, y^n) \leq |A_\varepsilon^{(n)}|\, 2^{-n(H(X,Y)-\varepsilon)}$$

$$\boxed{|A_\varepsilon^{(n)}| \geq (1-\varepsilon)\, 2^{+n(H(X,Y)-\varepsilon)}}$$

$$\Pr\left((\tilde{X}_n, \tilde{Y}_n) \in A_\varepsilon^{(n)}\right) = \sum_{\tilde{x}_n \tilde{y}_n \in A_\varepsilon^n} p(x_n)\cdot p(y_n) \geq (1-\varepsilon)\, 2^{n(H(X,Y)-\varepsilon)} \cdot 2^{-n(H(X)+\varepsilon)} \cdot 2^{-n(H(Y)+\varepsilon)} = (1-\varepsilon)\, 2^{n(H(Y)+H(X|Y)) - n\varepsilon - n(H(X)) - n\varepsilon - nH(Y) - n\varepsilon} = (1-\varepsilon)\, 2^{-n[H(X) - H(X|Y) + 3\varepsilon]}$$

$$\Pr\left[(\tilde{X}_n, \tilde{Y}_n) \in A_\varepsilon^{(n)}\right] \geq (1-\varepsilon)\, 2^{-n[I(X;Y) + 3\varepsilon]}$$

THEOREM 7.7.1 (REVISITED)

$$C = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ \vdots & & & \vdots \\ x_1(2^{nR}) & x_2(2^{nR}) & \dots & x_n(2^{nR}) \end{bmatrix} \qquad W = \{1, 2, \dots, 2^{nR}\}$$

$$\Pr(C) = \prod_{w=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(w)) \qquad \Pr(W = w) = 2^{-nR}, \quad w = 1, 2, \dots, 2^{nR}$$

$$P(y^n \mid \underbrace{x^n(w)}_{w\text{-th codeword}}) = \prod_{i=1}^{n} P(y_i \mid x_i(w))$$

$$Pr(\mathcal{E}) = \sum_{C} Pr(C) P_e^{(n)}(C) = \sum_{C} Pr(C) \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \lambda_w(C) =$$

$$= \frac{1}{2^{nR}} \sum_{C} Pr(C) \sum_{w=1}^{2^{nR}} \lambda_w(C) = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{C} Pr(C) \lambda_w(C) =$$

$$= \frac{1}{2^{nR}} \cdot 2^{nR} \sum_{C} Pr(C) \lambda_1(C) = \sum_{C} Pr(C) \cdot \lambda_1(C) = Pr(\mathcal{E} \mid W=1)$$

$$\boxed{E_i = \{ (X^n(i), Y^n) \text{ is in } A_\varepsilon^{(n)} \} \quad i \in \{1, 2, \ldots, 2^{nR}\}}$$

$$Pr(\mathcal{E} \mid W=1) = P(E_1^c \cup E_2 \cup E_3 \cup \ldots \cup E_{2^{nR}} \mid W=1) \leq$$

$$\leq P(E_1^c \mid W=1) + \sum_{i=2}^{2^{nR}} P(E_i \mid W=1)$$

$$Pr(\mathcal{E}) = Pr(\mathcal{E} \mid W=1) \leq P(E_1^c \mid W=1) + \sum_{i=2}^{2^{nR}} \underline{P(E_i \mid W=1)} \leq$$

$$\varepsilon + \sum_{i=2}^{2^{nR}} 2^{-n(I(X;Y)-3\varepsilon)} = \varepsilon + (2^{nR}-1)\, 2^{-n(I(X;Y)-3\varepsilon)}$$

$$= \varepsilon + 2^{nR} \cdot 2^{-n(I(X;Y)-3\varepsilon)} - 2^{-n(I(X;Y)-3\varepsilon)} \leq$$

$$\leq \varepsilon + 2^{3n\varepsilon} \cdot 2^{-n(I(X;Y)-R)} \leq 2\varepsilon \qquad (\to 1)$$

IF n IS SUFFICIENTLY LARGE AND $\boxed{R < I(X;Y)}$

$\therefore\ P(\mathcal{E} \mid C^*) \leq 2\varepsilon$

$$P_r(\mathcal{E} \mid C^*) = \frac{1}{2^{nR}} \sum_{i=1} \lambda_i(C^*)$$

$$\lambda_i = \underbrace{P_r(\mathcal{E} \mid x_n(W=i))}_{P_r(g(Y_n) \neq i \mid X^n = x^n(i))} = \sum_{y^n} p(y^n \mid x^n(i))\, I(g(y^n) \neq i)$$

$$H(W) = \sum \frac{1}{2^{nR}} \cdot \log 2^{nR} = 2^{nR} \cdot \frac{1}{2^{nR}} \, ld\, 2^{nR} = \underline{\underline{nR}}$$

$$H(W \mid \hat{W}) \leq 1 + P_e^{(n)}\, ld\, 2^{nR} \leq 1 + P_e^{n} \cdot n \cdot R$$

$$I(X^{(n)}, Y^{(n)}) \leq n \cdot C$$

$$I(X^{(n)}, Y^{(n)}) = H(Y^n) - H(Y^n \mid X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i \mid Y_1^{i-1}, X^n)$$

$$= H(Y^n) - \sum_{i=1}^{n} H(Y_i \mid X_i) \leq \sum H(Y_i) - \sum_{i=1}^{n} H(Y_i \mid X_i) =$$

$$= \sum_{i=1}^{n} H(Y_i) - H(Y_i \mid X_i) = \sum_{i=1}^{n} I(X_i; Y_i) \leq \underline{n \cdot C}$$

$$n \cdot R = H(W) = H(W|\hat{W}) + I(W;\hat{W}) \le 1 + P_e^{(n)} \cdot nR + I(W;\hat{W})$$

$$\le 1 + P_e^{(n)} \cdot nR + I(X^n;Y^n) \le 1 + P_e^{(n)} \cdot nR + nC$$

$$nR \le 1 + P_e^{(n)} \cdot nR + nC \qquad R \le \frac{1}{n} + P_e^{(n)} R + C$$

$n \to \infty$ $\boxed{R \le C}$

$$\boxed{P_e^{(n)} \ge 1 - \frac{1}{nR} - \frac{C}{R}}$$

PROBLEMS

7.1 PREPROCESSING THE OUTPUT. ONE IS GIVEN A COMMUNICATION CHANNEL WITH TRANSITION PROBABILITIES $p(y|x)$ AND CHANNEL CAPACITY $C = \max_{p(x)} I(X;Y)$. A HELPFUL STATISTICIAN PREPROCESSES THE OUTPUT BY FORMING $\tilde{Y} = g(Y)$. HE CLAIMS THIS WILL STRICTLY IMPROVE THE CAPACITY

(a) SHOW THAT HE IS WRONG
(b) UNDER WHAT CONDITIONS DOES HE NOT STRICTLY DECREASE THE CAPACITY.

(a)

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \le H(Y) - H(g(Y)|X)$$

$$I(X;g(Y)) = H(X) - H(X|g(Y)) = H(g(Y)) - H(g(Y)|X)$$

$$H(Y;g(Y)) = H(Y) + \underbrace{H(g(Y)|Y)}_{0} = H(g(Y)) + \underbrace{H(Y|g(Y))}_{\ge 0}$$

$$\boxed{H(Y) \ge H(g(Y))} \quad (\$)$$

$$I(X;g(Y)) = \underline{\underline{H(g(Y))}} - H(g(Y)|X) \le H(Y) - H(g(Y)|X)$$

$$H(Y,g(Y)|X) = H(Y|X) + \underbrace{H(g(Y)|X,Y)}_{0} =$$

$$= H(g(Y)|X) + H(Y|X,g(Y))$$

$$H(Y|X) \ge H(g(Y)|X)$$

?

$$I(X;g(Y)) \le H(Y) - H(g(Y)|X) \ge H(Y) - H(Y|X)$$

$$\boxed{X \to Y \to \tilde{Y}} \qquad I(X;\tilde{Y}) \ge I(X;Y)$$

$I(X;Y) \leq H(Y) - H(g(Y)|X)$

$I(X;g(Y)) \leq H(Y) - H(g(Y)|X)$ DATA PROCESSING INEQUALITY

• $X \to Y \to \tilde{Y}$, $\tilde{Y} = g(Y)$; $I(X;Y) \geq I(X;\tilde{Y})$

$X \to Y \to g(Y)$

ZA BILO KAKOV IZBOR NA $g(Y)$: $\max_{p(x)} I(X;Y) \geq \max_{p(x)} I(X;\tilde{Y})$

(b) "=" KOGA IMAME "=>"

$I(X;\tilde{Y}) = H(\tilde{Y}) - H(\tilde{Y}|X) = H(X) - H(X|\tilde{Y})$

($) $I(X;\tilde{Y}) \leq H(Y) - H(\tilde{Y}|X)$

$H(Y,\tilde{Y}|X) = H(Y|X) + H(\tilde{Y}|X,Y) = H(\tilde{Y}|X) + H(Y|X,\tilde{Y})$

IF $\tilde{Y} = g(Y)$ ONE-TO-ONE MAPPING

$\Rightarrow H(Y|\tilde{Y},X) = 0 \Rightarrow H(Y|X) = H(\tilde{Y}|X)$

ALSO: ($) $H(Y) = H(\tilde{Y}) + H(Y|\tilde{Y}) = H(\tilde{Y})$

$H(Y|\tilde{Y}) = 0$ (ONE-TO-ONE MAPPING) $\Rightarrow I(X;Y) = I(X;\tilde{Y})$

Solutions Ed 1

$X \to \tilde{Y} \to Y$ $\quad I(X;\tilde{Y}) \geq I(X;Y)$

$X \to Y \to \tilde{Y}$ $\quad I(X;Y) \geq I(X;\tilde{Y})$

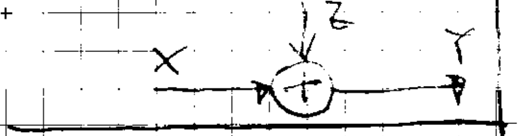
ZA DA VAZAT I DVETE

$I(X;Y) = I(X;\tilde{Y})$

OVA VAZI AKO: $H(Y) = H(\tilde{Y})$ & $H(Y|X) = H(\tilde{Y}|X)$

A TOA VAZI AKO $\tilde{Y} = g(Y)$ e ONE-TO-ONE

7.2 ADDITIVE NOISE CHANNEL. FIND THE CHANNEL CAPACITY OF THE FOLLOWING DISCRETE MEMORYLESS CHANNEL:



WHERE $\Pr\{Z=0\} = \Pr\{Z=a\} = \frac{1}{2}$

THE ALPHABET FOR X IS $\mathcal{X} = \{0,1\}$. ASSUME THAT Z IS INDEPENDENT OF X. OBSERVE THAT THE CHANNEL CAPACITY DEPENDS ON THE VALUE OF a.

$Y = X + Z$

$\boxed{I(X;Y) = ?}$ $\qquad I(X; X+Z) = ?$

$\max_{p(x)} I(X;Y)$ $\qquad \mathcal{X} = \{0, 1\} \quad p(x) = \{p, 1-p\}$

$\mathcal{Z} = \{0, a\}$

$p(Y=0) = P(X=0) \cdot P(Z=0) = p/2$

$p(Y=1) = P(X=1) \cdot P(Z=0) = (1-p) \cdot 1/2$

$P(Y=a) = P(X=0) \cdot P(Z=a) = p \cdot 1/2$

$p(Y=1+a) = P(X=1) \cdot P(Z=a) = (1-p) \cdot 1/2$

$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$

$H(Y) = 2 \frac{p}{2} \operatorname{ld} \frac{2}{p} + 2 \frac{(1-p)p}{2} \cdot \operatorname{ld} \frac{(1-p)p}{2}$

$H(Y|X) = P(X=0) \cdot H(Y|X=0) + H(Y|X=1) \cdot P(X=1)$

$H(Y|X=0) = -\sum_{Y} p(Y|X=0) \operatorname{ld} p(Y|X=0) = \frac{2}{2} \operatorname{ld} 2 = 1$

$Y \in \{0, a\} \qquad p(Y|X=0) = \{\frac{1}{2}, \frac{1}{2}\}$

$H(Y|X=1) = \sum_{Y} p(Y|X=1) \operatorname{ld} p(Y|X=1) = 2 \frac{1}{2} \operatorname{ld} 2 = 1$

$Y \in \{1, 1+a\} \qquad p(Y|X=1) = \{\frac{1}{2}, \frac{1}{2}\}$

$H(Y|X) = p \cdot 1 + (1-p) \cdot 1 = 1$

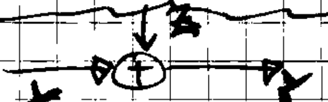
$I(X;Y) = -\underbrace{\left[p \operatorname{ld} \frac{p}{2} + (1-p) p \operatorname{ld} \frac{(1-p)p}{2}\right] - 1}_{*}$

$* = p \operatorname{ld} \frac{p}{2} + p \operatorname{ld} \frac{(1-p)p}{2} - p^2 \operatorname{ld} \frac{(1-p)p}{2} - 1$

$\frac{d}{dp}\left(p \operatorname{ld} \frac{(1-p)p^2}{4} - p^2 \operatorname{ld} \frac{(1-p)p}{2}\right) = 0 \qquad p = 0.59506$

$*(0{,}59506) = -1.77639$

$\max_{p(x)} I(X;Y) = +1.77639$

RECALL $Y = \bmod(X+Z)$  $X \in \{0,1\}$, $Z \in \{0,1\}$

$I(X;Y) = ?$ $\quad I(X;Y) = H(Y) - H(Y|X)$

$I(X,Z;Y) = I(X;Y) + I(Z;Y|X) = I(Z;Y) + I(X;Y|Z)$

$I(Z;Y|X) = H(Y|X) - \underbrace{H(Y|X,Z)}_{0}$ — ako gi znaes x,z go znaes y

$I(Z;Y|X) = H(Y|X)$ $\qquad$ $I(X,Z;Y) = I(X;Y) + H(Y|X)$

$I(X;Y) = H(Y) - H(Y|X)$

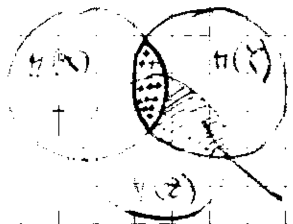
$I(X,Z;Y) = H(Y) - H(Y|X) + H(Y|X) = H(Y)$

$I(X;Y|Z) = H(Y|Z) - \underbrace{H(Y|X,Z)}_{0}$

$I(X,Z;Y) = I(Z;Y) + I(X;Y|Z) = I(Z;Y) + H(Y|Z) = H(Y) - H(Y|Z)$
$+ H(Y|Z) = H(Y)$

$I(Z;Y) + H(Y|Z) = I(X;Y) + H(Y|X)$

$H(X,Z)$ $\qquad$ $H(X,Y) = H(X) + H(Y|X)$



$I(X,Z;Y) = I(X;Y) + I(Z;Y|X)$

$I(Z;Y|X)$

logicno e: $I(X,Z;Y) = H(Y)$ — ako gi znaes x,z togas go znaes y

$I(X;Z,Y) = I(X;Y) + I(X;Z|Y) = I(X;Z) + I(X;Y|Z)$

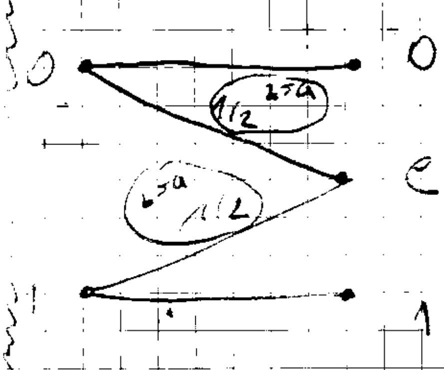
$I(X;Y|Z) = H(Y|Z) - \underbrace{H(Y|X,Z)}_{0} = H(X|Z) - \underbrace{H(X|Y,Z)}_{0}$

$Y = \bmod(X+Z)$ $\quad$ (go naogas x ako znaes y, z)

Y	Z	X
0	0	0
0	1	1
1	0	1
1	1	0

$H(Y|Z) = H(X|Z)$

x, z statisticki nezavisni ⇒ $H(X|Z) = H(X)$

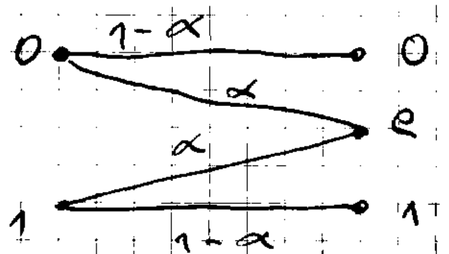


$I(Z;Y|X) = H(Y|X) - \underbrace{H(Y|X,Z)}_{0} = H(Y|X)$

$I(Z;Y|X) = H(Z|X) - \underbrace{H(Z|X,Y)}_{0}$ $= H(Z)$

$H(Y|X) = H(Z|X)$

$I(X;Y) = H(Y) - H(Y|X) = H(Y) - H(Z|X) =$
$= H(Y) - H(Z)$ $\quad$ **MAV**



$$I(X;Y) = H(X) - H(X|Y)$$
$$= H(Y) - H(Y|X)$$

$Y(X)$ $\begin{cases}\pi & x=0\\ 1-\pi & x=1\end{cases}$

$$H(Y|X) = P(X=0)\cdot H(Y|X=0) + P(X=1)\cdot H(Y|X=1) = \pi\, H(Y|X=0) + (1-\pi)\, H(Y|X=1) =$$
$$= \pi\left[\underbrace{(1-\alpha)\,\mathrm{ld}(1-\alpha)^{-1} + \alpha\,\mathrm{ld}\,\alpha^{-1}}_{H(\alpha)}\right] + (1-\pi)\left[\underbrace{(1-\alpha)\,\mathrm{ld}(1-\alpha)^{-1} + \alpha\,\mathrm{ld}\,\alpha^{-1}}_{H(\alpha)}\right]$$
$$= [\pi + (1-\pi)]\,H(\alpha) = \underline{H(\alpha)}$$

$$\boxed{I(X;Y) = H(Y) - H(\alpha)}$$

$$\max_{P(X)} I(X;Y) = (1-\alpha) + H(\alpha) - H(\alpha) = \underline{1-\alpha}$$

$$\boxed{\alpha = \frac{1}{2}} \qquad \max I(X;Y) = 1 - \frac{1}{2} = 1/2$$

$H(Y) \geq H(X) \qquad H(Y) \geq H(X)$

$$I(X,Z;Y) = H(Y) = I(Z;Y) + I(X;Y|Z) = I(Z;Y) + H(X)$$

$$I(X;Y|Z) = \underbrace{H(X|Z)}_{=H(X)} - \underbrace{H(X|Y,Z)}_{0}$$

$$H(Y) = \underbrace{I(Z;Y)}_{\geq 0} + H(X) \quad \Rightarrow \quad \boxed{H(Y) \geq H(X)}$$

$Z = X + Y \qquad I(X;Z) = H(Z) - H(Z|X) =$

$$= H(Z) - H(X+Y|X) = H(Z) - \underbrace{H(Y|X)}_{\text{INDEPENDENT}} = H(Z) - H(Y)$$

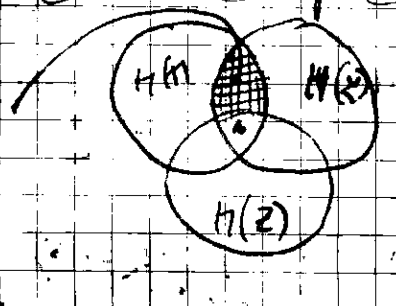
$$\boxed{H(Z|X) = H(Y|X)}$$

ENTROPIJA NA IZLEZ − ENTROPIJA NA VLEZ (?)

$$H(f(X),X) = H(X) + \underbrace{H(f(X)|X)}_{0} = H(f(X)) + H(X|f(X))$$

$$H(X) \geq H(f(X))$$

$H(X) = H(f(X))$ IF THE MAPPING IS ONE-TO-ONE



$Z = (X+Y) \bmod (2) \qquad Y = (Z-X) \bmod 2$

$$I(X;Y|Z) = 1$$

$$I(Y,Z;X) = I(X;Z) + \underbrace{I(X;Y|Z)}_{1}$$

$$I(X;Y) = 0 \Rightarrow \boxed{A = 1}$$

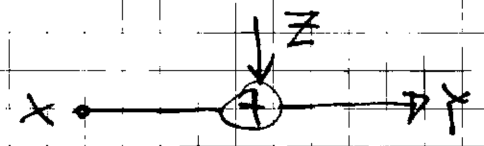
$I(X;Y|Z) = H(X|Z) - \underbrace{H(X|Y,Z)}_{0} = \underline{H(X|Z)}$ $\quad$ $Y=(Z-X) \bmod 2$, $X=(Z-Y) \bmod 2$

$H(X|Z) = \sum_{z \in \mathcal{Z}} p(z) \cdot H(X|Z=z) = \underbrace{p(z=0)}_{1/2} \cdot \Big[\underbrace{p(x=0|0)}_{1/2}\, ld \frac{1}{p(0|0)}$

$+ \underbrace{p(x=1|0)}_{1/2}\, ld \frac{1}{p(1|0)} \Big] + \underbrace{p(z=1)}_{1/2} \Big[p(x=0|1)\, ld \frac{1}{p(0|1)} +$

$+ p(x=1|1)\, ld \frac{1}{p(1|1)} \Big] = \frac{1}{2}\Big[2 \cdot \frac{1}{2} \cdot ld\, 2\Big] + \frac{1}{2}\Big[2 \cdot \frac{1}{2} \cdot ld\, 2\Big]$

$= \underline{\frac{1}{2} + \frac{1}{2} = 1}$ bit



$Z = \{0, a\}$ $\quad$ $p(Z) = \{\frac{1}{2}, \frac{1}{2}\}$

$X = \{0, 1\}$

$Y = X + Z$ $\quad$ $\boxed{X = Y - Z}$

$I(X;Z|Y) = H(X|Y) - \underbrace{H(X|Y,Z)}_{0}$

• IF $\boxed{a=0}$

$Z = \{0, 0\}$ $\quad$ $p(Z) = \{\frac{1}{2}, \frac{1}{2}\}$

$Y = X$

$I(X;Y) = H(X) - H(X|Y) = H(X) \geq$

• IF $\boxed{a=1}$ $\quad$ $I(X;Y) = H(Y) - H(Z) = H(Y) - 1 = H(X) - H(Z)$

$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \underbrace{H(Z|X)}_{INDEP.} = H(Y) - H(Z)$

$H(Y|X) = H(Z|X)$ $\quad$ $\boxed{Z = Y - X}$

$\underline{H(Y|X)} = \sum_x p(x) H(Y|X=x) = \sum p(x) \cdot H(Z + x | X) =$

$= \sum_x p(x) \sum_z p(z+x|x) \cdot ld \frac{1}{p(z+x,x)} = \sum_x p(x) H(Z|X=x)$

$= \underline{H(Z|X)}$

ALTERNAT.:

$I(X,Z;Y) = I(X;Y) + I(Z;Y|X)$

$I(Z;Y|X) = H(Z|X) - \underbrace{H(Z|X,Y)}_{0} =$

$= H(Y|X) - \underbrace{H(Y|X,Z)}_{0}$ $\quad \Rightarrow \quad \boxed{H(Z|X) = H(Y|X)}$

$I(X;Y) = H(X) - H(X|Y) = H(X) - H(Z|X) = H(X) - H(Z)$

$I(Y;Z;X) = I(Y;X) - I(X;Z|Y) = I(X;Y) - H(Z|X) +$

$+ \underbrace{H(Z|X,Y)}_{0}$ $\quad$ $I(Z;X|Y) = H(Z|X) + 0 - H(X|Y) -$

$- \underbrace{H(X|Y,Z)}_{0} = H(X|Y)$

$$I(X;Y) = \begin{cases} H(X) & a=0 \\ H(X) - \underbrace{H(Z)}_{1} & a=1 \end{cases}$$

$$C = \max_{p(x)} I(X;Y) = \begin{cases} 1 & a=0 \\ 1-1=0 & a=1 \end{cases}$$

$$C = \begin{cases} 1 & a=0 \\ 0 & a=1 \end{cases}$$ ⇒ CHANNEL CAPACITY DEPENDS ON THE VALUE OF a.

EDITION 1 SOLUTION!

$a=0$ $\quad Y = X \Rightarrow I(X;Y) = H(X)$

$$C = \max_{p(x)} I(X;Y) = \log m = \log 2 = 1$$

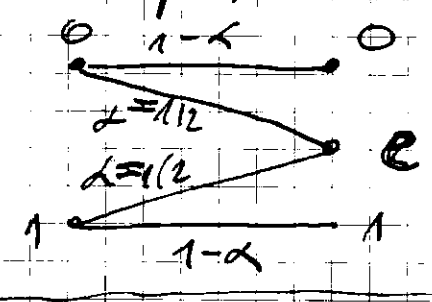
$a \neq 0, \pm 1$ $\quad Y \in \{0, 1, a, 1+a\}$

KNOWING Y WE KNOW X:

$$I(X;Y) = H(X) - \underbrace{H(X|Y)}_{0} = H(X)$$

$$C = \max_{p(x)} H(X) = 1$$

$a=1$



$$C = 1 - \alpha = 1 - \frac{1}{2} = \frac{1}{2}$$

$$I(X;Y) = H(Y) - H(Y|X)$$

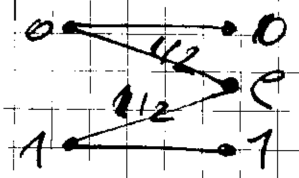
$$H(Y|X) = \sum_{x \in X} p(x) H(Y|X=x) = p(0) \cdot H(Y|X=0) +$$

$$+ p(1) \cdot H(Y|X=1) = p(0) \cdot H(\alpha) + p(1) \cdot H(\alpha) = H(\alpha)$$

$$H(Y) = H(\alpha) + 1 - \alpha \qquad I(X;Y) = H(\alpha) + 1 - \alpha - H(\alpha) = 1 - \alpha$$

$$I(X;Y) = 1 - \frac{1}{2} = \frac{1}{2}$$ (BITS PER TRANSMISSION)

$a=-1$ $\quad Y \in \{-1, 0, 1\}$ ⇒



$$I(X;Y) = 1 - \alpha = \frac{1}{2}$$ bits/transmission

$$H(Y) = \frac{(1-\alpha)}{2} ld \left[\frac{(1-\alpha)}{2}\right]^{-1} + \alpha\, ld\, \alpha^{-1} + \frac{1-\alpha}{2} ld \left(\frac{2}{1-\alpha}\right)$$

$$= (1-\alpha) ld \frac{2}{(1-\alpha)} + \alpha\, ld \frac{1}{\alpha} = (1-\alpha) ld 2 + (1-\alpha) ld \frac{1}{(1-\alpha)}$$

$$+ \alpha \cdot ld \frac{1}{\alpha} = (1-\alpha) + H(\alpha)$$

$$H(Y) = H(Y, E) = H(Y) + \underbrace{H(E|Y)}_{0} = H(E) + H(Y|E)$$

$$H(Y,E) = H((1-\pi)\cdot(1-\alpha), \alpha, \pi(1-\alpha)) = \Big[(1-\pi)(1-\alpha) ld \frac{1}{(1-\pi)(1-\alpha)}$$

$$+ \alpha\, ld \frac{1}{\alpha} + \pi(1-\alpha) ld \frac{1}{\pi(1-\alpha)}\Big] =$$

$$= (1-\pi)(1-\alpha) ld \frac{1}{(1-\pi)} + (1-\pi)(1-\alpha) ld \frac{1}{(1-\alpha)} + \alpha\, ld \frac{1}{\alpha} +$$

$$+ \pi(1-\alpha) ld \frac{1}{\pi} + \pi(1-\alpha) ld \frac{1}{(1-\alpha)} =$$

$$= (1-\alpha)\left[(1-\pi) ld \frac{1}{1-\pi} + \pi\, ld \frac{1}{\pi}\right] + H(\alpha) = (1-\alpha) H(\pi) + H(\alpha)$$

$\boxed{\pi = \frac{1}{2}}$ $\quad H(Y) = (1-\alpha) + H(\alpha)$

$$I(X;Y) = H(Y) - H(Y|X) = (1-\alpha) + H(\alpha) - H(\alpha)$$

$\boxed{I(X;Y) = 1-\alpha} \Rightarrow \boxed{C = 1-\alpha}$ CAPACITY OF BINARY CHANNEL WITH ERASURE.

PROBLEM 7.3 CHANNELS WITH MEMORY HAVE HIGHER CAPACITY. CONSIDER BINARY SYMMETRIC CHANNEL WITH $Y_i = X_i \oplus Z_i$ $X_i, Y_i \in \{0,1\}$. SUPPOSE THAT $\{Z_i\}$ HAS CONSTANT MARGINAL PROBABILITIES:

$$Pr\{Z_i = 1\} = p = 1 - Pr\{Z_i = 0\}$$ BUT THAT

$Z_1, Z_2, \ldots, Z_n$ ARE NOT NECESSARY INDEPENDENT. ASSUME THAT Z^n IS INDEPENDENT FROM THE INPUT X^n!
LET $\boxed{C = 1 - H(p, 1-p)}$. SHOW THAT

$$\max_{p(x_1, x_2 \ldots x_n)} I(X_1, X_2, \ldots X_n; Y_1, Y_2, \ldots Y_n) \geq nC$$

$$R \leq P_e^{(n)} \cdot R + \frac{1}{n} + C$$

$$P_e^{(n)} = Pr\{g(Y^n) \neq W\} \qquad C_{FB} = C = \max_{p(x)} I(X;Y)$$

$nR = H(W) = H(W|\hat{W}) + I(W;\hat{W}) \le 1 + P_e nR + I(W;\hat{W})$

$\le 1 + P_e nR + I(W;Y^n)$

$I(W;Y^n) = H(Y^n) - H(Y^n|W) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}, W)$

$/X_i = f(Y_1^{i-1}, W)/ = H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}, W, X_i) =$

$= H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i) \le \sum_{i=1}^{n} H(Y_i) + \sum_{i=1}^{n} H(Y_i|X_i)$

$= \sum_{i=1}^{n} I(X_i;Y_i) \le n \cdot C$

$nR \le 1 + P_e nR + n \cdot C$ $\quad$ $\boxed{R \le \frac{1}{n} + P_e R + C}$

$n \to \infty$ $\quad$ $\boxed{R \le C}$ $\quad$ $\boxed{C_{FB} = C}$

$P_r(V^n \ne \hat{V}^n) = \sum_{y^n}\sum_{v^n} p(v^n)\, p(y^n|x^n(v^n))\, I(g(y^n) \ne v^n)$

$\boxed{H(V) + \epsilon = R < C}$ $\quad$ $V^n \to X^n \to Y^n \to \hat{V}^n$

$H(V) \le \frac{H(V_1^n)}{n} = H(V|\hat{V}) + I(V;\hat{V}) \le \frac{1}{n}(1 + P(V \ne \hat{V})\, n \log|V|$

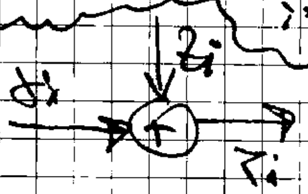
$+ I(V;\hat{V})) \le \frac{1}{n} + P(V \ne \hat{V}) \log|V| + \frac{1}{n} \underbrace{I(X^n;Y^n)}_{\le nC} \le$

$\le \frac{1}{n} + P(V \ne \hat{V}) \log|V| + C$

$n \to \infty$ $\quad$ $\boxed{H(V) \le C}$

$I(X^n;Y^n) = H(Y^n) - H(Y^n|X^n) =$

$= H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}, X_1^n) = H(Y^n) - \sum_{i=1}^{n} H(Z_i + X_i|Y_1^{i-1}, X_1^n)$



$I(X_i;Y_i) = H(Y_i) - H(Y_i|X_i) =$

$/H(Y_i|X_i) = H(Z_i|X_i)/ = H(Y_i) - H(Z_i|X_i)$

$= H(Y_i) - H(Z_i)$ $\quad$ $\boxed{I(X_i;Y_i) = H(Y_i) - H(Z_i)}$

[illegible]

$$H(Y_i|X_i) = \sum_{x_i} p(x_i)\cdot H(Y_i|X_i=x_i)$$

$$H(Z_i|X_i) = \sum_{x_i} p(x_i) H(Z_i|X_i=x_i) = \sum_{x_i} p(x_i) H(\underbrace{Z_i+x_i}_{Y_i}|X_i=x_i)$$

$$= \sum_{x_i} p(x_i) H(Y_i|X_i=x_i) = H(Y_i|X_i)$$

$$I(X^n;Y^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1^{i-1},X_1^n) = (*)$$

$$H(Y^n|X^n) = H(Z^n|X^n)$$

$$(*) = H(Y^n) - H(Z^n|X^n) \geq H(Y^n) - H(Z^n)$$

$$H(Z^n) = \sum_{i=1}^{n} H(Z_i|Z_1^{i}) \leq H(Z)\cdot n \quad \geq$$

$$\geq H(Y^n) - n\cdot H(Z) = n\cdot H(Y) - n H(Z) = n\underbrace{(H(Y)-H(Z))}_{I(X;Y)}$$

$$H(Y^n) = \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}) = n\cdot H(Y)$$

INDEPENDENT

$$= n\cdot I(X;Y) = n\cdot C$$

OVA E DEFINITIVO $\geq$ ZBOG ŠTO

$$H(Z^n) = \sum_{i=1}^{n} H(Z_i|Z_1^{i-1}) \leq \sum_{i=1}^{n} H(Z_i) = n\cdot H(Z)$$

$$H(Z) = 1 - H(p) \Rightarrow$$

$$\boxed{I(X^n;Y^n) \geq n\cdot C} = n(1-H(p))$$

EDITORIAL SOLUTION (WE RELEASE NO PRIOR X^n):

$$I(X^n;Y^n) = H(X^n) - H(X^n|Y^n) = H(X^n) - H(Z^n|Y^n) = (*)$$

$$\boxed{H(X^n|Y^n) = H(Z^n|Y^n)} \quad Z_i = Y_i - X_i \text{ CONST}$$

$$H(X^n|Y^n) = \sum_{y_i^n} p(y_i^n)\, H(X^n|Y^n = y_i^n) =$$

$$= \sum_{y_i^n} p(y_i^n) \cdot H(y_i^n - x_i^n \mid y_i^n) = \underbrace{\sum_{y_i^n} p(y_i^n) H(Z_i^n \mid y_i^n)}_{H(Z^n \mid Y^n)}$$

[p63 ①] $\geq H(X^n) - H(Z^n) = H(X^n) - \sum_{i=1}^{n} H(Z_i \mid Z_1^{i-1}) \geq$

$\geq H(X^n) - \sum_{i=1}^{n} H(Z_i) = H(X^n) - n \cdot \underbrace{H(Z)}_{H(p)} =$

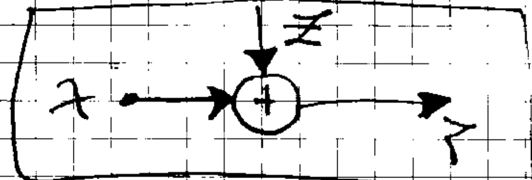
$= H(X^n) - n \cdot H(p) = n \cdot H(X) - n\, H(p)$

$\Big/\; H(X^n) = \sum H(X_i \mid X_1^{i-1}) = n H(X)$; $H(X) \leq \mathrm{ld}\, m = \mathrm{ld}\, 2 = 1 \;\Big/$ $= n \underbrace{(1 - H(p))}_{= C}$

$= n \cdot C \Rightarrow \boxed{I(X^n; Y^n) \geq n \cdot C}$ PROVED!!!

PROBLEM 7.4 CHANNEL CAPACITY. CONSIDER THE DISCRETE MEMORYLESS CHANNEL $Y = X + Z \mod 11$ WHERE:

$$Z = \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$



AND $X \in \{0, 1, \ldots, 10\}$. ASSUME THAT Z IS INDEPENDENT OF X.

(a) FIND THE CAPACITY.

(b) WHAT IS MAXIMIZING $p^*(x)$?

(a) $I(X;Y) = H(Y) - H(Y \mid X) = H(Y) - H(Z \mid X) =$
$= H(Y) - H(Z) = \mathrm{ld}\, m - H\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) =$
$= \mathrm{ld}\, 11 - 3 \cdot \frac{1}{3} \cdot \mathrm{ld}\, 3 = \mathrm{ld}\, 11 - \mathrm{ld}\, 3 = \mathrm{ld}\, \frac{11}{3} = \underline{\underline{1.88}}$

(b) $I(X;Y) = H(X) - H(X \mid Y) = H(X) - H(Z \mid Y) =$
$H(X) - \underbrace{H(Z)}_{?} = \mathrm{ld}\, m - \mathrm{ld}\, 3 = \mathrm{ld}\, 10 - \mathrm{ld}\, 3$

? NE ZNAM ALI Z NE ZAVISI OD Y

64 $\max_{p(x)} I(X;Y) = C = \mathrm{ld}\, \frac{11}{3}$ FOR $p(x) = \{\frac{1}{11} \ldots \frac{1}{11}\}$

$$p(X=i)=\frac{1}{11} \qquad i=0,1,\dots,10$$

PROBLEM 7.5 USING TWO CHANNELS AT ONCE. CONSIDER TWO DISCRETE MEMORYLESS CHANNELS $(\mathcal{X}_1, p(y_1|x_1), \mathcal{Y}_1)$ AND $(\mathcal{X}_2, p(y_2|x_2), \mathcal{Y}_2)$ WITH CAPACITIES C_1 AND C_2, RESPECTIVELY. A NEW CHANNEL $(\mathcal{X}_1\times\mathcal{X}_2, p(y_1|x_1)\times p(y_2|x_2), \mathcal{Y}_1\times\mathcal{Y}_2)$ IS FORMED IN WHICH $x_1\in\mathcal{X}_1$ AND $x_2\in\mathcal{X}_2$ ARE SENT SIMULTANEOUSLY, RESULTING IN y_1, y_2. FIND THE CAPACITY OF THIS CHANNEL.

$$=\max_{p(x)} I(X;Y) = H(X)-H(X|Y) = H(Y)-H(Y|X)$$

$$C_1 = \max_{p(x_1)} I(X_1;Y_1) = H(Y_1)-H(Y_1|X_1)$$

$$C_2 = \max_{p(x_2)} I(X_2;Y_2) = H(Y_2)-H(Y_2|X_2)$$

$$X = X_1\times X_2 \qquad Y = Y_1\times Y_2$$

$$p(x) = p(x_1)\cdot p(x_2) \qquad X_1, X_2 \text{ INDEPENDENT}$$

$$Z = X+Y \qquad p(z) = p(x)+p(y) = p(x)\,p(y)$$

$$\boxed{H(X) = H(X_1)+H(X_2)}$$

$$H(Y) = H(Y_1)+H(Y_2)$$

$$H(Y|X) = H(Y_1\times Y_2 | X_1\times X_2)$$

$$H(Y_1\times Y_2|X_1\times X_2) = \sum_{x_1 x_2} p(x_1\times x_2)\, H(Y_1\times Y_2 | X_1 = x_1 \times X_2 = x_2)$$

$$= \sum_{x_1 x_2} p(x_1\times x_2) \sum_{y_1 y_2} p(y_1\times y_2 | X_1 = x_1 \times X_2 = x_2)\cdot ld\frac{1}{p(y_1\times y_2|x_1 x_2)}$$

$$= \sum_{\substack{x_1 x_2\\ y_1 y_2}} p(x_1\times x_2, y_1\times y_2)\, ld\frac{1}{p(y_1\times y_2|x_1\times x_2)} =$$

$$= \sum_{\substack{x_1 y_1\\ x_2 y_2}} p(x_1\times x_2, y_1\times y_2)\, ld\, p(y_1|x_1\times x_2)\cdot p(y_2|x_1 x_2) =$$

$$= \sum p(x_1\times x_2, y_1\times y_2)\, ld\, p(y_1|x_1) + \sum p(x_1 x_2, y_1\times y_2)\, ld\,(y_2|x_2)$$

$$= H(Y_1|X_1) + H(Y_2|X_2)$$

$$C = \max_{p(x_1)p(x_2)} \left[H(Y_1) + H(Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \right]$$

$$= \max_{p(x_1)p(x_2)} \left[H(Y_1) - H(Y_1|X_1) \right] + \max_{p(x_2)} \left[H(Y_2) - H(Y_2|X_2) \right]$$

$$= \max_{p(x_1)} I(X_1;Y_1) + \max_{p(x_2)} I(X_2;Y_2) = C_1 + C_2$$

$$\boxed{C = C_1 + C_2}$$

ΕΡΙΠΟΙ ΛΥΟΣΣΙΠΟΝ

- MARKOVITY!

$$\boxed{X \rightarrow Y \rightarrow Z}$$

$$p(x,z|y) = \frac{p(x,y,z)}{p(y)} = \frac{p(x,y)\cdot p(z|x,y)}{p(y)} =$$

$$= \frac{p(x,y)\cdot p(z|y)}{p(y)} = \frac{p(y)\cdot p(x|y)\cdot p(z|y)}{p(y)}$$

$$p(x,z|y) = p(x|y)\cdot p(z|y)$$

$$p(y,z|x) = \frac{p(x,y,z)}{p(x)} = \frac{p(x,y)\, p(z|x,y)}{p(x)} =$$

$$= \frac{p(x)\cdot p(y|x)\cdot p(z|y)}{p(x)} = p(y|x)\cdot p(z|y)$$

$$\boxed{Y_1 \rightarrow X_1 \rightarrow X_2 \rightarrow Y_2}$$

$$p(y_1,y_2|x_1,x_2) = \frac{p(x_1,x_2,y_1,y_2)}{p(x_1 x_2)} =$$

$$= \frac{p(x_1 y_1)\cdot p(y_2 x_2|x_1,y_1)}{p(x_1 x_2)} = \frac{p(x_1 y_1)\cdot p(y_2 x_2|x_1)}{p(x_1 x_2)}$$

$$= p(y_2 x_2|x_1) = p(x_2)\cdot p(y_2|x_1 x_2) = p(x_2)\cdot p(y_2|x_2)$$

$$= \frac{p(x_1)\cdot p(y_1|x_1)\cdot p(x_2)\cdot p(y_2|x_2)}{p(x_1)\cdot p(x_2)} = p(y_1|x_1)\cdot p(y_2|x_2)$$

$$\frac{p(x_1,x_2,y_1 y_2)}{p(x_1 x_2)} = p(y_1|x_1)\cdot p(y_2|x_2) = p(y_1,y_2|x_1 x_2)$$

$$p(x_1 x_2 y_1 y_2) = p(x_1 x_2) \cdot p(y_1|x_1) \cdot p(y_2|x_2)$$

$$p(y_1 y_2|x_1 x_2) = p(y_1|x_1) \cdot p(y_2|x_2)$$

$$I(x_1 x_2; y_1 y_2) = H(y_1 y_2) - H(y_1 y_2|x_1 x_2) = \circledast$$

$$H(y_1 y_2|x_1 x_2) = H(y_1|x_1 x_2) + H(y_2|x_1 x_2 y_1) =$$

$$= H(y_1|x_1 x_2) + H(y_2|x_2) = H(y_1|x_1) + H(y_2|x_2)$$

(*) ■ NE MI E DOKAZANO

$$H(y_1|x_1 x_2) = ? \quad H(x_1 y_1|x_2) = H(x_1|x_2) + H(x_1|x_2 y_1)$$

$$= H(y_1|x_2) + H(x_1|y_1)$$

$$y_1 \to x_1 \to x_2$$

$$p(y_1|x_1 x_2) = \frac{p(x_1 x_2 y_1)}{p(x_1 x_2)} = \frac{p(x_1 y_1) \cdot p(x_2|x_1 y_1)}{p(x_1) \cdot p(x_2)}$$

$$= \frac{p(x_1) \cdot p(y_1|x_1) \cdot p(x_2|x_1)}{p(x_1) \cdot p(x_2)}$$

$$p(x_2 y_1|x_1) = p(y_1|x_1) \cdot p(x_2|x_1)$$

$$H(y_1|x_1 x_2) = \sum_{x_1 x_2 y_1} \frac{p(y_1|x_1)\, p(x_2|x_1)}{p(x_2)} \operatorname{ld} \frac{p(x_2)}{p(y_1|x_1)\, p(x_2|x_1)}$$

$$\circledast = H(y_1 y_2) - H(y_1|x_1) - H(y_2|x_2) \leq H(y_1) + H(y_2)$$

$$- H(y_1|x_1) - H(y_2|x_2) = I(x_1; y_1) + I(x_2; y_2)$$

$$C = \max_{p(x_1 x_2)} I(x_1 x_2; y_1 y_2) \leq \max_{p(x_1 x_2)} I(x_1; y_1) +$$

$$+ \max_{p(x_1 x_2)} I(x_2; y_2) = \max_{p(x_1)} I(x_1; y_1) + \max_{p(x_2)} I(x_2; y_2)$$

$$= C_1 + C_2 \qquad \boxed{C = C_1 + C_2}$$

■ $X \to Y \to Z$ $\quad H(Z|XY) = H(Z|Y) \quad / \cdot H(XY)$

$$H(XY) \cdot H(Z|XY) = H(XY) \cdot H(Z|Y)$$

$$H(X, Y, Z) = H(Y) \cdot H(X|Y) \cdot H(Z|Y) = H(X|Y) \cdot H(Y, Z)$$

$$\boxed{H(X|YZ) = H(X|Y)}$$

DOKAZANO MMV !!!

$(1-p-q)^2 = (1-q)^2 - 2(1-q)p + p^2$

$(1-p-q)^3 = (1-p)^3 - 3(1-p)\cdot p^2 + 3(1-p)^2\cdot q - q^3$

$\frac{1}{2}\left(1-(1-2q)^4\right)$

- VO MATLAB PODIV:

$[V\ D] = \mathrm{eig}(A)$: $V = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ $D = \begin{bmatrix} 1 & 0 \\ 0 & 1-2q \end{bmatrix}$

$A = V\cdot D\cdot V^{-1}$ $A^4 = V\cdot D^4\cdot V^{-1} =$

$D^4 = \begin{bmatrix} 1 & 0 \\ 0 & (1-2q)^4 \end{bmatrix}$ $A^4 = \begin{bmatrix} \frac{1}{2}(1+(1-2q)^4); & \frac{1}{2}(1-(1-2q)^4) \\ \frac{1}{2}(1-(1-2q)^4); & \frac{1}{2}(1+(1-2q)^4) \end{bmatrix}$

$$P_e = \frac{1}{2}\left(1-(1-2q)^4\right)$$

- EDITION 1 SOLUTION: PROBABILITY OF ERROR FOR CASCADE CHANNEL IS SUM OF ODD TERMS OF BINOMIAL EXPANSION: $(x+y)^4$

WITH $x = q$ $y = 1-q$

$\frac{1}{2}(x+y)^4 - \frac{1}{2}(y-x)^4 = \frac{1}{2}(q+1-q) - \frac{1}{2}(1-q-q)^4 =$

$= \frac{1}{2}\left(1-(1-2q)^4\right)$

MNOGU BABA FRTA | MMV

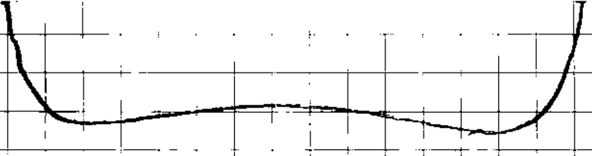
- **PROBLEM 7.8** Z-CHANNEL. THE Z CHANNEL HAS BINARY INPUT AND OUTPUT ALPHABETS AND TRANSITION PROBABILITIES $p(y|x)$ GIVEN BY THE FOLLOWING MATRIX:

$$Q = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} \qquad x,y \in [0,1]$$

FIND THE CAPACITY OF THE Z CHANNEL AND MAXIMIZING INPUT PROBABILITY DISTRIBUTION.

$$C = \max_{p(x)} I(X;Y) = H(Y) - H(Y|X)$$

$$H(Y|X) = \sum_x p(x)\cdot H(Y|X=x)$$



$$\begin{bmatrix} p(y_1) \\ p(y_2) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}^T \begin{bmatrix} p(x_1) \\ p(x_2) \end{bmatrix}$$

$x \in \{x_1, x_2\}$, $y \in \{y_1, y_2\}$

$p(y_1) = p_{11} p(x_1) + p_{21} p(x_2)$ $\quad p(y_2) = p_{12} p(x_1) + p_{22} p(x_2)$

$H(Y|X=x_1) = p_{11} \, ld \frac{1}{p_{11}} + p_{12} \, ld \frac{1}{p_{12}}$

$H(Y|X=x_2) = p_{21} \, ld \frac{1}{p_{21}} + p_{22} \, ld \frac{1}{p_{22}}$

$H(Y|X=x_1) = 1 \, ld \frac{1}{1} = 0$

$H(Y|X=x_2) = \frac{1}{2} ld 2 + \frac{1}{2} ld 2 = 1$

$H(Y|X) = \sum p(x) H(Y|X=x) = p(x_1) \cdot H(Y|X=x_1) +$

$p(x_2) \cdot H(Y|X=x_2) = 0 + p(x_2)$

$C = H(Y) - p(x_2)$ $\qquad \max H(Y) = ld\, m = ld\, 2 = 1$

$C = 1 - p(x_2)$ $\qquad p(x_2) = \frac{1}{2}$ } UNIFORM DISTRO

$C = 1 - 1/2 = 1/2$

• EDITION 2 SOLUTION

$H(Y|X) = p(x_2) = Pr(X=x_2) = Pr(X=1) = x$

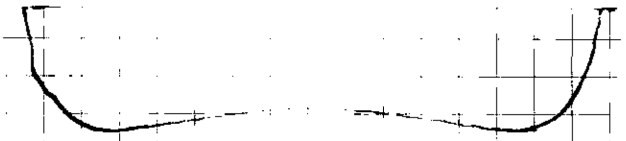
$H(Y) = Pr(Y=y_1) \, ld \frac{1}{Pr(Y=y_1)} + Pr(Y=y_2) \, ld \frac{1}{Pr(Y=y_2)} =$

$H(Y) = H(Pr(Y=1)) = H(Pr(X=x_2) \, Pr(Y=y_2|X=x_2)) =$

[illegible] $= H\left(\frac{Pr(X=x_2)}{2}\right) = H\left(\frac{Pr(X=1)}{2}\right) = H\left(\frac{x}{2}\right)$

$I(X,Y) = H(Y) - x = \frac{x}{2} ld \frac{2}{x} - x$

$\frac{dI(X,Y)}{dx} = \frac{1}{2} ld \frac{2}{x} + \frac{x}{2}$ [illegible] $- 1 = \frac{1}{2} ld \frac{2}{x} - \frac{1}{2}$



(70)

$$\begin{bmatrix} Y(y_1) \\ Y(y_2) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} Y(x_1) \\ Y(x_2) \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} Y(x_1) \\ Y(x_2) \end{bmatrix}$$

(p_{11}, p_{12}, p_{21}, p_{22})

$$H(Y|X) = P_r(X = x_2) = P_r(X = 1) = x$$

$$H(Y) = -P_r(Y=y_1)\, ld\, P_r(Y=y_1) - P_r(Y=y_2)\, ld\, P_r(Y=y_2)$$

$$P_r(Y=y_1) = \underbrace{P_r(X=x_1)}_{x} \cdot \underbrace{P_r(Y=y_1|X=x_1)}_{p_{11}=1} + \underbrace{P_r(X=x_2)}_{1-x} \cdot \underbrace{P(y_1|x_2)}_{p_{21}=1/2}$$

$$P_r(Y=y_2) = P_r(X=x_2) \cdot \underbrace{P_r(Y=y_2|X=x_2)}_{p_{22}} + P(X=x_1)\underbrace{P_r(y_2|x_1)}_{p_{12}}$$

$$= (1-x)\cdot p_{22} + x\, p_{12} = \frac{1}{2}(1-x) + 0$$

$$P_r(Y=y_1) = x + (1-x)\frac{1}{2} = x + \frac{1}{2} - \frac{x}{2} = \frac{x}{2} + \frac{1}{2} = \frac{1}{2}(x+1)$$

$$P_r(Y=y_1) + P_r(Y=y_2) = \frac{1}{2}(1-x) + \frac{1}{2}(x+1) =$$
$$= \frac{1}{2} - \frac{x}{2} + \frac{x}{2} + \frac{1}{2} = 1$$

$$H(Y) = -\frac{x+1}{2}\, ld\, \frac{x+1}{2} - \frac{1-x}{2}\, ld\, \frac{1-x}{2} =$$

$$= -\left[\frac{x}{2}\, ld\, \frac{x+1}{2} - \frac{x}{2}\, ld\, \frac{1-x}{2} + \frac{1}{2}\, ld\, \frac{x+1}{2} + \frac{1}{2}\, ld\, \frac{1-x}{2}\right]$$ (minus)

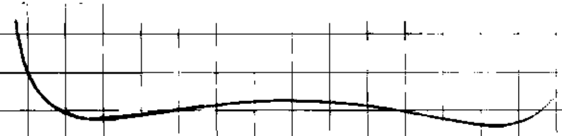
$$= -\left[\frac{x}{2}\, ld\, \frac{x+1}{1-x} + \frac{1}{2}\, ld\, \frac{x+1}{1-x}\right] = \frac{x+1}{2}\, ld\, \frac{x+1}{1-x}$$

$$I(X;Y) = -\frac{x+1}{2}\, ld\, \frac{x+1}{1-x} - x = \frac{x+1}{2}\, ld\, \frac{1-x}{x+1} - x$$

$$\frac{dI(X;Y)}{dx} = \frac{1}{2}\, ld\, \frac{1-x}{x+1} + \frac{x+1}{2} \cdot \frac{x+1}{1+x} \cdot \frac{-1\cdot(x+1) - (1-x)\cdot 1}{\ln 2\,(x+1)^2} - 1$$

$$= \frac{1}{2}\, ld\, \frac{1-x}{x+1} + \frac{1}{2}\, \frac{-x-1-1+x}{\ln 2} - 1 = \frac{1}{2}\, ld\, \frac{1-x}{1+x} - \frac{1}{\ln 2} - 1$$

$$= \frac{1}{2}\, ld\, \frac{1+x}{1+x} - \frac{\ln e}{\ln 2} - 1 = \frac{1}{2}\, ld\, \frac{1+x}{1+x} - ld\, e - 1$$



$P(Y=y_1) = (1-x) + x\cdot\frac{1}{2} = 1-\frac{x}{2}$

$P(Y=y_2) = x\cdot p_{22} + (1-x)\,p_{12} = \frac{x}{2}$

$$\boxed{H(Y) = H(y) = H\left(\frac{x}{2}\right)}$$

NA SKETMORANJU STANJA IMAM GREŠKU !!! IMAM ZAMJENU $P(X=x_1)=x$, A TREBA $P(X=x_2)=x$

$$I(X;Y) = \frac{x}{2}\,ld\frac{2}{x} + \left(1-\frac{x}{2}\right)ld\frac{2}{2-x} - x =$$

$$= \frac{x}{2}\,ld\frac{2}{x} - \frac{x}{2}\,ld\frac{2}{2-x} + ld\frac{2}{2-x} - x =$$

$$= \frac{x}{2}\,ld\,\frac{2}{x}\,\frac{2-x}{2} + ld\frac{2}{2-x} - x = \frac{x}{2}\,ld\frac{2-x}{x} + ld\frac{2}{2-x} - x$$

$$\frac{dI(x;Y)}{dx} = \frac{1}{2}\,ld\frac{2-x}{x} + \frac{x}{2}\,\frac{x}{2-x}\cdot\frac{-1\cdot x-(2-x)}{x^2(\ln 2)} + \frac{2(2-x)(2-x)^{-2}}{2(\ln 2)} - 1$$

$$= \frac{1}{2}\,ld\frac{2-x}{x} + \frac{-x-2+x}{2\ln 2\,(2-x)} + \frac{(2-x)}{\ln 2\,(2-x)^2} - 1 =$$

$$= \frac{1}{2}\,ld\frac{2-x}{x} - \frac{1}{2(\ln 2)(2-x)} + \frac{1}{2(\ln 2)(2-x)} - 1 =$$

$$= \frac{1}{2}\,ld\frac{2-x}{x} - 1 = \frac{1}{2}\,ld\left(\frac{1-\frac{x}{2}}{\frac{x}{2}}\right) - 1 = 0$$

$$\frac{1}{2}\,ld\frac{1-\frac{x}{2}}{\frac{x}{2}} = 1 \qquad ld\frac{1-\frac{x}{2}}{\frac{x}{2}} = 2$$

$$\frac{1-\frac{x}{2}}{\frac{x}{2}} = 4 \qquad 1-\frac{x}{2} = 2x \qquad 2x+\frac{x}{2} = 1$$

$$\frac{5x}{2} = 1 \qquad \boxed{x = \frac{2}{5}} \qquad C = \max_{p(x)} I(X;Y) = H\left(\frac{2}{5}\cdot\frac{1}{2}\right) - \frac{2}{5}$$

$$H\left(\frac{1}{5}\right) = \frac{1}{5}\,ld\,5 + \frac{4}{5}\,ld\frac{5}{4} = 0.722$$

$$\boxed{C = H\left(\frac{1}{5}\right) - \frac{2}{5} = 0.322}$$

IT IS REASONABLE THAT $P_r(X=1) < \frac{1}{2}$ BECAUSE $X=1$ IS THE MOST INPUT TO THE CHANNEL.

7.9 **Suboptimal codes** For the Z-channel of problem 7.8, assume that we chose $(2^{nR}, n)$ code at random where each codeword is a sequence of fair coin tosses. This will not achieve capacity. Find the maximum rate R such that the probability of error $P_e^{(n)}$ averaged over the randomly generated codes tends to zero as the block length n tends to infinity.

$$nR = H(W) = H(W|\hat{W}) + I(W;\hat{W}) \le \underbrace{1 + P_e^{(n)} nR}_{\text{Fano}} + nC.$$

$$P_e^{(n)} = \frac{1}{2^{nR}} \sum_{i=1}^{n} \lambda(i)$$

$n = 2$ 00 10 / 01 11 $\qquad \frac{1}{2}\frac{1}{2} = \frac{1}{2^n} = \frac{1}{2^2} = \frac{1}{4}$

$n = 3$ 000 100 / 001 101 / 010 110 / 011 111 $\qquad \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2^3}$

$$Q = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \qquad \begin{bmatrix} \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$13.5 = 10 \log \frac{P}{1\,mW} \qquad 10^{1.35}\, 1\,mW = P \qquad \boxed{P = 22.4\,mW}$$

$$I(W;\hat{W}) \le I(X^n;Y^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i)$$
$$\le \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) = \sum_{i=1}^{n} I(X_i;Y_i) \le nC$$

$$\lambda_i = P_r\left(g(Y^n) \ne i \,|\, X^n = X^n(i)\right) = \sum_{Y^n} P(Y^n|X^n(i)) \cdot I(g(Y^n) \ne i)$$

$$P_e^{(n)} = \sum_{i=1}^{n} p(X^n(i))\,\lambda_i = \frac{1}{2^{nR}}\sum_{i=1}^{n}\lambda_i$$

(uniform)

$$P(Y^n|X^n(i)) = P(Y_1|X^n(i))\,P(Y_2|Y_1X^n(i)) \cdots P(Y_n|Y_1^{n-1}X^n(i))$$
$$= P(Y_1|X_1)\cdot P(Y_2|X_2) \cdots P(Y_n|X_n)$$

(72)

$$\underbrace{\begin{bmatrix} P(0)\\ P(1)\end{bmatrix}}_{Y} = \begin{bmatrix}1 & 1/2\\ 0 & 1/2\end{bmatrix}\underbrace{\begin{bmatrix}P(0)\\ P(1)\end{bmatrix}}_{X}$$

$$P_Y(Y=0) = P(X=0)\cdot \overbrace{P(Y=0|X=0)}^{1} + P(X=1)\cdot \overbrace{P(Y=0|X=1)}^{1/2}$$

$$P_Y(Y=1) = P(X=0)\cdot \underbrace{P(Y=1|X=0)}_{0} + P(X=1)\cdot \underbrace{P(Y=1|X=1)}_{1/2}$$

- АКО ПРАТАМ „0" НА ИЗЛЕЗ СЕКОГАШ ДОБИВАМ 0 Т.Е НЕМА ДА СЕ ЈАВИ ГРЕШКА Т.Е ВЕРОЈАТНОСТА НА ГРЕШКА Е 0.
- АКО ПРАТАМ „1" НА ИЗЛЕЗ МОЖЕ ДА ПРИМАМ „0" СО ВЕРОЈАТНОСТ 1/2 ИЛИ „1" СО ВЕРОЈАТНОСТ 1/2.

• ВКУПНИОТ БРОЈ НА ЛАЗНИЧНИ КОДНИ ЗБОРОВИ Е 2^n

• ВКУПНИОТ БРОЈ НА КОДНИ ЗБОРОВИ СО „i" ЕДИНИЦИ Е: $\binom{n}{i}$

$i=1$: $\binom{n}{1} = n$ $\quad P(Y^n|X^n(i)) = \frac{1}{2}$

$$\lambda_1 = \binom{n}{1}\cdot P(Y^n|X^n(1)) = \binom{n}{1}\cdot\frac{1}{2^1}$$

$i=2$: $$\lambda_2 = \binom{n}{2}\cdot P(Y^n|X^n(i=2)) = \binom{n}{2}\frac{1}{2^2}$$

$i=n$: $$\lambda_n = \binom{n}{n}\cdot\frac{1}{2^n}$$

$i=k$: $$\lambda_k = \binom{n}{k}\frac{1}{2^k}$$

МАКЕ!!

$$P_e^{(n)} = \frac{1}{2^{nR}}\sum_{i=1}^{n}\binom{n}{i}\frac{1}{2^i} \qquad P_e = \frac{1}{2^{nR}}\left(\left(\frac{3}{2}\right)^n - 1\right)$$

$$P_e = 0 \qquad \frac{1}{2^{nR}}\left(\frac{3^n}{2^n} - 1\right) = \frac{3^n}{2^{(R+1)n}} - 1$$

$$P_e^n = \frac{3^n}{2^{(n+1)R}} - \frac{1}{2^{nR}} \qquad \frac{3^n}{2^{(n+1)R}} = \frac{1}{2^{nR}}$$

$$2^R = 3 \;/\; \log \qquad \boxed{R = \log 3}$$

CONTINUE FROM 86

7.10 ZERO-ERROR CAPACITY. A CHANNEL WITH ALPHABET $\{0,1,2,3,4\}$ HAS TRANSITION PROBABILITIES OF FORM

$$p(y|x) = \begin{cases} 1/2 & y = (x \pm 1) \bmod 5 \\ 0 & \text{OTHERWISE} \end{cases}$$

ANALOGY TO NOISY TYPEWRITER.

(a) COMPUTE THE CAPACITY OF THIS CHANNEL IN BITS
(b) THE ZERO-ERROR CAPACITY OF A CHANNEL IS THE NUMBER OF BITS PER CHANNEL USE THAT CAN BE TRANSMITTED WITH ZERO PROBABILITY OF ERROR. CLEARLY, THE ZERO-ERROR CAPACITY OF THIS PENTAGONAL CHANNEL IS AT LEAST 1 BIT (TRANSMIT 0 OR 1 WITH PROBABILITY 1/2). FIND A BLOCK CODE THAT SHOWS THAT ZERO ERROR CAPACITY IS GREATER THAN 1 BIT. CAN YOU ESTIMATE THE EXACT VALUE OF THE ZERO-ERROR CAPACITY? (HINT: CONSIDER CODES OF LENGTH 2 FOR THIS CHANNEL) THE ZERO-ERROR CAPACITY OF THIS CHANNEL WAS FINALLY FOUND BY LOVASZ [365].

$$nR = H(W) = H(W|\hat{W}) + I(W;\hat{W}) \le 1 + P_e \cdot nR + I(X^n;Y^n)$$

$$\le \underbrace{1 + P_e \cdot nR}_{0} + nC \qquad n \cdot R \le n\, C$$

(ZERO ERROR: $P_e = 0$)

$$\boxed{R \le C}$$

$y = (x \pm 1) \bmod 5$

$$\begin{bmatrix} p(y_1) \\ p(y_2) \\ p(y_3) \\ p(y_4) \\ p(y_5) \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 & 0 \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \\ p(x_4) \\ p(x_5) \end{bmatrix}$$

(rows/columns indexed 0, 1, 2, 3, 4; the matrix is Π^T)

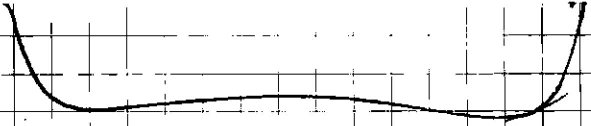
$$\Pi = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 & 0 \end{bmatrix}$$

$p(y_1) = p_{21} p_2 + p_{51} p_5$

$p(y_2) = p_{12} p_1 + p_{32} p_3$

$\dots$

$p(y_5) = p_{15}\, p(x_1) + p_{45}\, p(x_4)$



$$C = \max_{p(x)} I(X,Z) = \max(H(Z)) - H(Z|X)$$

$$\begin{bmatrix} p(z_1) \\ p(z_2) \\ \vdots \\ p(z_5) \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ & & \cdots & & \\ 1/2 & 0 & 0 & 1/2 & 0 \end{bmatrix} \begin{bmatrix} 1/5 \\ 1/5 \\ \vdots \\ 1/5 \end{bmatrix} = \begin{bmatrix} 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \\ 1/5 \end{bmatrix}$$

$$\max(H(Z)) = 5 \cdot \frac{1}{5}\, ld\, 5 = ld\, 5$$

$$H(Z|X) = \sum_x p(x)\, H(Z|X=x) = \sum_x p(x)\left[\frac{1}{2} ld\, 2 + \frac{1}{2} ld\, 2\right] = 1$$

$$C = ld\, 5 - 1 = 2.3219 - 1 = \underline{1.3219}$$

NA PR. 78. SE DOBIVA ISTI REZULTAT NO ZA PROIZVOLJNO $p(x)$

(6) $nR \leq 1 + P_e nR + nC$

$$R \leq \frac{1}{n} + P_e \cdot R + C$$

$W = \{0, 1, 2, 3, 4\}$

$x^n = \{0, 1, 00, 01, 10\}$ $\qquad$ $x^n = \{0, 1, 01, 10, 11\}$

$p(x)$				$c(x)$	$l(x)$
0.2	0.4	0.4	0.6]0	01	2
0.2	0.2	0.4]0	0.4]1	000	3
0.2	0.2]0	0.2]1		001	3
0.2]0	0.2]1			10	2
0.2]1				11	2

$$E[l] = \frac{1}{5}(6+6) = \frac{12}{5} = 2.4$$

$x^n = \{14; 02; 13; 24; 03\}$
$W = \{0; 1; 2; 3; 4\}$

$x^n = \{00, 11, 22, 33, 44\}$

$$p(y_1) = p_{21}\, p(x_2) + p_{51}\, p(x_5) = (p_{21} + p_{51}) \cdot \frac{1}{5}$$

$$p(y_2) = p_{12}\, p(x_1) + p_{32}\, p(x_3)$$

$$p(y_3) = p_{23}\, p(x_2) + p_{43}\, p(x_4)$$

$$p(y_4) = p_{34}\, p(x_3) + p_{54}\, p(x_5)$$

$$p(y_5) = p_{15}\, p(x_1) + p_{45}\, p(x_4)$$

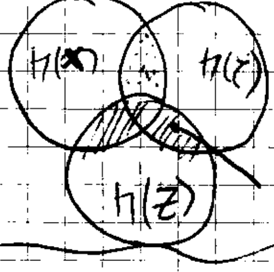
$Y^n = \{14;\ 02;\ 13;\ 24;\ 03\}$ $\quad W = \{0, 1, 2, 3, 4\}$

$nR = \underbrace{H(W|Y^n)}_{\emptyset \text{ zero error}} + I(W;Y^n) \le I(X^n;Y^n) \le nC$

$I(X^n;Y^n) = I(X_1X_2;Y_1Y_2) = I(X_1;Y_1Y_2) + I(X_2;Y_1Y_2|X_1)$

$I(X_1X_2;Y) = I(X_1;Y) + I(X_2;Y|X_1)$

$I(XY;Z) = I(X;Z) + I(Y;Z|X)$



$I(X_2;Y_1Y_2|X_1) = H(X_2|X_1) - \underbrace{H(X_2|X_1Y_1Y_2)}_{\emptyset}$

$H(X_2|X_1) = \sum_{x_1} p(x_1)\, H(X_2|X_1=x_1) =$

$= p(X_1=1)\, H(X_2|X_1=1) + p(X_1=0)\, H(X_2|X_1=0) + p(X_1=2)\cdot H(X_2|X_1=2) = \frac{2}{5}\left(\frac{1}{2}\,ld\,2 + \frac{1}{2}\,ld\,2\right) + \frac{2}{5}\,1 + \frac{1}{5}\,ld\,\frac{1}{0} =$
$= 4/5$

$I(X_1;Y_1Y_2) = I(Y_1;X_1) + I(Y_2;X_1|Y_1)$

$I(X_1;Y_1) = H(Y_1) - H(Y_1|X_1) = H(X_1) - H(X_1|Y_1)$

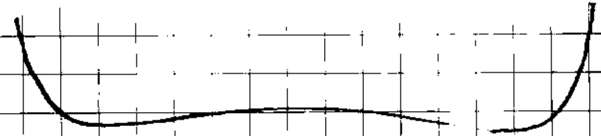
$I(X_1;Y_n) = I(X_1;Y_1Y_2) = H(X_1) - \underbrace{H(X_1|Y^{(n)})}_{0 \text{ zero error}}$

$I(X^n;Y^n) = H(X_1) + H(X_2|X_1) = \underline{H(X_1,X_2)}$

$H(X_1) = p(X_1=1)\, ld\frac{1}{p(X_1=1)} + p(X_1=2)\, ld\frac{1}{p(X_1=2)} + p(X_1=0)\, ld\frac{1}{p(X_1=0)} = \frac{2}{5}\, ld\frac{5}{2} + \frac{1}{5}\, ld\, 5 + \frac{2}{5}\, ld\frac{5}{2} = \frac{4}{5}\, ld\frac{5}{2} + \frac{1}{5}\, ld\, 5$

$= \frac{4}{5}\, ld\, 5 - \frac{4}{5}\, ld\, 2 + \frac{1}{5}\, ld\, 5 = ld\, 5 - \frac{4}{5}$

$I(X^n;Y^n) = ld\, 5 - \frac{4}{5} + \frac{4}{5} = ld\, 5$



(77)

$$\begin{bmatrix} p(Y_1)\\ p(Y_2)\\ p(Y_3)\\ p(Y_4)\\ p(Y_5)\end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 1/2\\ 1/2 & 0 & 1/2 & 0 & 0\\ 0 & 1/2 & 0 & 1/2 & 0\\ 0 & 0 & 1/2 & 0 & 1/2\\ 1/2 & 0 & 0 & 1/2 & 0\end{bmatrix} \begin{bmatrix} p(X_1)\\ p(X_2)\\ p(X_3)\\ p(X_4)\\ p(X_5)\end{bmatrix} \quad \begin{matrix}0\\1\\2\\3\\4\end{matrix}$$

(columns labelled 0 1 2 3 4)

$$p(g(Y^n)=0) = \left(\frac{1}{2}p(x_2)+\frac{1}{2}p(x_5)\right)\cdot\left(\frac{1}{2}p(x_2)+\frac{1}{2}p(x_5)\right)$$

$$p(g(Y^n)=0) = p(Y_1Y_2 = 14) = \frac{1}{2}p(x_2)\cdot\frac{1}{2}p(x_5)$$

$$I(X^n, Y^n) = H(Y^n) - \underbrace{H(Y^n|X^{(n)})}_{0}$$ (ZERO ERROR)

$\boxed{H(Y^n) = ?}$

$$H(Y^n) = p(Y^n=14)\,ld\,\frac{1}{p(Y^n=14)} \quad p(Y^n=03)\,ld\,\frac{1}{p(Y^n=03)}$$

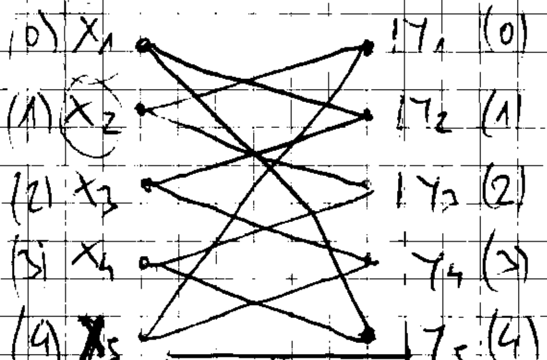
$$p(Y^n=14) = 1 - P(Y^n=11) - P(Y^n=44) - P(Y^n=41) =$$

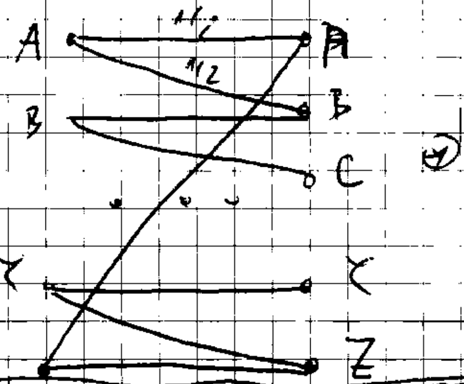
$Y^1 = 1 \quad Y^2 = 4$

$$p(g(Y^n)=0) = p(\ \ ,\ \ )$$

$Y^1 = Y_1$

$$p(\ \ ) = p(x_2)\,p_{21} + p(x_5)\,p_{51} = \frac{1}{5}\cdot 1$$





$\boxed{C = ld\,2 = 1}$ → CAPACITY IS LOGARITHM OF THE NUMBER OF RECOGNIZABLE INPUTS !!!

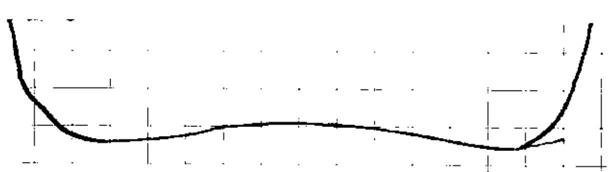
(*) $\max I(X,Y) = \max H(Y) - H(Y|X) = ld\,26 - H(Y|X)$

$$H(Y|X) = \sum_x p(x)\,H(Y|X=x) = p(X=A)\,H(Y|X=A) +$$

$$+ p(X=B)\,H(Y|X=B) + \dots + P(X=Z)\,H(Y|X=Z) =$$

$$= \frac{1}{26}\left(\frac{1}{2}ld\,2 + \frac{1}{2}ld\,2\right) + \dots + \frac{1}{26}(1) = 26\cdot\frac{1}{26} = 1$$

$$C = ld\,26 - 1 = ld\,26 - ld\,2 = \underline{ld\,13}$$



(78)

$$\begin{bmatrix} p(y_1) \\ p(y_2) \\ p(y_3) \\ p(y_4) \\ p(y_5) \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 & 0 \end{bmatrix} \cdot \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \\ p(x_4) \\ p(x_5) \end{bmatrix}$$

$C = ld(5) - H(Y|X)$

$H(Y|X) = \sum_x p(x)\, H(Y|X=x) =$

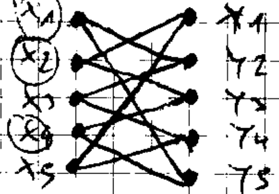
$= p(x_1)\left[p(Y=y_2|x_1)\, ld \frac{1}{p(Y=y_2|x_1)} + p(Y=y_5|x_1)\, ld \frac{1}{p(Y=y_5|x_1)} \right] + p(x_2)[\ldots] + p(x_5)[\ldots] =$

$= p(x_1)\left[\frac{1}{2} ld 2 + \frac{1}{2} ld 2\right] + p(x_2) + \ldots + p(x_5) = 1$

$\boxed{H(Y|X) = 1}$ $\quad \boxed{C = ld(5) - 1 = 1{,}3219}$

ISTOTO OD $p(x)$ NO ZA IZVORNO $p(x)$!!!

$W = \{0, 1, 2, 3, 4\}$ $\quad X = \{\overset{0}{\underset{x_1x_1}{00}}, \overset{1}{\underset{x_2x_2}{11}}, \overset{2}{\underset{x_4x_4}{}} \ldots\}$



$\boxed{C = ld 3 = 1.6}$

DA NO DA POKAŽEM DEKA E ETHODECSS

$I(X^1X^2; Y^1Y^2) =$

$H(Y|X_1=0) = \frac{1}{2} ld 2 + \frac{1}{2} ld 2 = 1 = p(Y=y_2|X_1=0)\, ld \frac{1}{p(Y=y_2|x_1=0)}$

$+ p(Y=y_4|X_1=0) \cdot ld \frac{1}{p(Y=y_4|X_1=0)} = 1$

$H(Y^{(n)}|X^{(n)}=00) = ? = p(Y^{(n)} = y_2y_2|X^n=00) \cdot ld \frac{1}{p(y_2y_2|00)}$

$+ \left[ld(p(y_2y_4|00))^{-1}\right] p(y_2y_4|00) + p(y_4y_2|00)\, ld\, p(y_4y_2|00)$

$+ p(y_4y_4|00)\, ld\, p(y_4y_4|00) = \left(\frac{1}{4} \cdot ld 4\right) 4 = 2$ (?)

$I(X^1X^2; Y^1Y^2) = H(Y^1Y^2) - \underbrace{H(Y^1Y^2|X^1X^2)}_{2}$

$H(Y^1Y^2|X^1X^2) = \sum_{x^1x^2} P(x^1x^2)\, H(Y^1Y^2|X^1X^2 = x_1x_2)$

$H(Y^2Y^4|X^1X^2=00) = -\sum_{y_2y_2} P(y_2y_4|00)\, ld\, \underbrace{P(y_2y_4|00)}_{\text{PODJEDNAKO}} = 4 \cdot \frac{1}{4} ld 4 = 2$ RAMO

78 ČETIRI MOŽNOSTI $(y_2y_4, y_2y_2, y_4y_2, y_4y_4)$

$H(Y^2Y^4 \mid X^1X^2 = 00) = \left(\frac{1}{4}\, ld\, 4\right) 4 = 2 \qquad H(Y^1Y^3 \mid$ [redacted] $) = 4 \cdot \frac{1}{4}\, ld\, 4 = 2$

$H(Y^4 \mid X_1 X_2) = ?$ $\qquad Y^4 = (Y_2Y_4,\ Y_1Y_2,\ Y_3Y_5)$

$H(Y^4 \mid \underbrace{X^1X^2}_{X^{(4)}}) = \sum_{x_1^1 x_2^2} p(x^1x^2)\, H(Y^4 \mid x^1x^2 = x_1^1 x_2^2)$

$H($ [redacted] $\mid$ [redacted] $) = ?$ (boxed; Y^4 written above the first blot, X_4X_4 under the second)

$X^4 = \{x_1x_1,\ x_2x_2,\ x_4x_4\}$

$H(Y^4 \mid 11) = p(Y^4 = Y_1Y_3 \mid 11)\ ld\ p(Y^4 = Y_1Y_3 \mid 11) + p(Y^4 = Y_3Y_4 \mid 11)$
$ld\ p(Y^4 = Y_3Y_1) + p(Y^4 = Y_4Y_1 \mid 11)\ ld\ p(Y^4 = Y_4Y_1 \mid 11) +$
$p(Y^4 = Y_3Y_3 \mid 11)\ ld\ p(Y^4 = Y_3Y_3 \mid 11)$ → OVA MOŽE DA GO GENERIRA $X_1^{(4)} - X_4X_4$

$H(Y^4 \mid X^{4} = x_1x_1) = 4 \cdot \frac{1}{4}\, ld\, 4 = 2 \qquad H(Y^4 \mid X^4 = x_2x_2) = 2$

$H(Y^4 \mid x_4x_4) = p(Y^4 = Y_3Y_5 \mid x_4x_4)\ ld\ (Y^4 = Y_3Y_5 \mid x_4x_4) +$
$+ p(Y^{(4)} = Y_5Y_3 \mid x_4x_4)\ ld\ (Y^4 = Y_5Y_3 \mid x_4x_4) + p(Y^4 = Y_5Y_5 \mid x_4x_4) \cdot$
$\cdot\ ld\ (Y^{(4)} = Y_5Y_5 \mid x_4x_4) + p(Y^4 = Y_3Y_3 \mid x_4x_4)\ ld\ p(Y^4 = Y_3Y_3 \mid x_4x_4)$

↳ OVA VODI VON GREŠKA !!!

① ‖ ② ‖ ③ ‖ ④ ‖ ⑤ ‖

$X^4 = \{14, 41;\ 02;\ 13;\ 24;\ 03, 30\} = \{x_2x_5;\ x_1x_3,\ x_2x_4,\ x_3x_5,$
$x_1x_4\}$

E.G. $X^4 = (x_2x_5) \to$ [redacted] $Y^4 \in (Y_1 \vee Y_3,\ Y_1 \vee Y_4) =$
$= (Y_1Y_1,\ Y_1Y_4,\ Y_3Y_1,\ Y_3Y_4)$

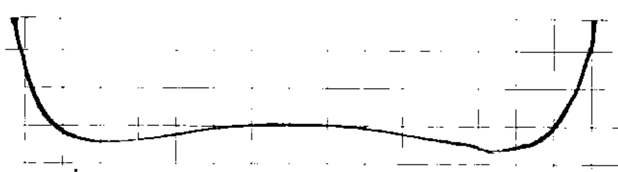
$X^4 = (x_1x_3) \to Y^4 \in (Y_2 \vee Y_5,\ Y_2 \vee Y_4) = (Y_2Y_2,\ Y_2Y_4,\ Y_5Y_2,\ Y_5Y_4)$

$X^4 = (X_2X_4) \to Y^4 \in (Y_1 \vee Y_3,\ Y_3 \vee Y_5) = (Y_1Y_3,\ Y_1Y_5,\ Y_3Y_3,\ Y_3Y_5)$

$X^4 = (x_5x_3) \to Y^4 \in (Y_4 \vee Y_4,\ Y_2 \vee Y_4) = (Y_4Y_2,\ Y_4Y_4,\ Y_4Y_2,\ Y_4Y_4)$

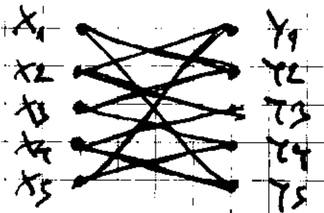
$X^4 = (X_1X_4) \to Y^4 \in (Y_2 \vee Y_5,\ Y_3 \vee Y_5) = (Y_2Y_3,\ Y_2Y_5,\ Y_5Y_3,\ Y_5Y_5)$

$X^4 = (X_5X_2) \to Y^4 \in (Y_4 \vee Y_4,\ Y_1 \vee Y_3) = (Y_4Y_1,\ Y_4Y_3,\ Y_4Y_1,\ Y_4Y_3)$



(204)

$X^4 = \{14, 02, 13, 24, 03\}$ $\quad X^4 = \{14, 20, 31, 42, 03\}$
$\quad\quad (x_2x_5,\ x_3x_1,\ x_4x_2,\ x_5x_3,\ x_1x_4)$



W	
0	$(x_2x_5) \to Y^4 = (y_1 \vee y_3,\ y_1 \vee y_4) = y_1y_1,\ y_1y_4,\ y_3y_1,\ y_3y_4$
1	$(x_3x_1) \to Y^4 = (y_2 \vee y_4,\ y_2 \vee y_5) = y_2y_2,\ y_2y_5,\ y_4y_2,\ y_4y_5$
2	$(x_4x_2) \to Y^4 = (y_3 \vee y_5,\ y_1 \vee y_3) = y_3y_1,\ y_3y_3,\ y_5y_1,\ y_5y_3$
3	$(x_5x_3) \to Y^4 = (y_1 \vee y_5,\ y_2 \vee y_4) = y_1y_2,\ y_1y_4,\ y_5y_2,\ y_5y_4$
4	$(x_1x_4) \to Y^4 = (y_2 \vee y_5,\ y_3 \vee y_5) = y_2y_3,\ y_2y_5,\ y_5y_3,\ y_5y_5$

$(25),\ (13),\ (24),\ (35),\ (14)$
$\ \ y_1 \quad\ \ y_2 \quad\ \ y_3 \quad\ \ y_4 \quad\ \ y_5$

$$\begin{bmatrix} p(y_1) \\ p(y_2) \\ p(y_3) \\ p(y_4) \\ p(y_5) \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 & 0 \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \\ p(x_4) \\ p(x_5) \end{bmatrix}$$

RETURN FROM N16.10b

I/O IZVEDUVANJETO ZA $H(Y^4|X^4)$

SUM PISAL DEKA:

$H(Y^4|x_2x_5) = H(Y^4|x_4x_2) = H(Y^4|x_5x_3)$ $= 0$, A DEKA $H(Y^4|x_3x_1) = H(Y^4|x_1x_4) = \frac{0.32}{2} = 0.16$

SEPAK IMAJKI GO VO PREDVID RAZVOJOT GORE KAKO I NA N19.9b IZLEGUVA DEKA I (x_2x_5) I (x_5x_3) ISTO IMAAT MALA GRESKA T.E

$H(Y^4|x_2x_5) = H(Y^4|x_5x_3) = \frac{0.32}{2} = 0.16$

SAMO $H(Y^4|x_4x_2)$ NEMA GRESKA T.E

$H(Y^4|x_4x_2) = 0$

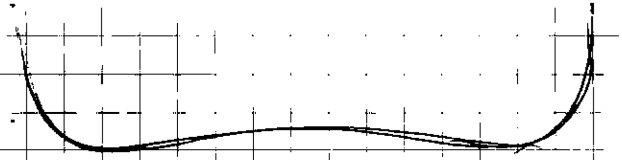
$C_2 = H(Y^4) - 4 \cdot 0.16 = H(Y^4) - 0.64$

$C_2 = 1.5 - 0.64 = 1.673 = 1.7$

$C = \frac{1.7}{2} = 0.85 \doteq 1$

VREDNO E DA SE ZABELEZI DEKA „2" SO TOA SO DOBIVAS REALNIOT KAPACITET NA KANALOT. SE PRAVI TAKA ZOSTO ZA EDEN IZLEZEN SIMBOL DA SE DEKODIRA ILANIZNO SE I IOPEZNITI DVA VLEZNI SIMBOLI T.E DVE USUITEDI NA KANALOT.

• EDITION 1 SOLUTIONS WE WILL PICK 5 CODEWORDS WITH DISTINCT FIRST SYMBOLS: $\{0a, 1b, 2c, 3d, 4e\}$



(81a)

$C = \{u, v\}$ $\quad v = (u+3) \bmod 5$

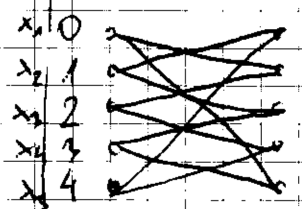
$u = 0 \quad v = 3 \bmod 5 = 3$ | $u = 2 \quad v = 5 \bmod 5 = 0$

$u = 1 \quad v = 4 \bmod 5 = 4$ | $u = 3 \quad v = 6 \bmod 5 = 1$

$u = 4 \quad v = (4+3) \bmod 5 = 2$

$\rightarrow \{03, 14, 20, 31, 42\}$

$p(y|x) = \begin{cases} 1/2 & y = (x \pm 1) \bmod 5 \\ 0 & \end{cases}$



$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \\ y(4) \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 0 & 1/2 & 0 \end{bmatrix} \begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \\ y(4) \end{bmatrix}$$

$(x_2x_5;\ x_1x_3;\ x_2x_4;\ x_3x_5;\ x_1x_4)$

$(14,\ 02;\ 13;\ 24;\ 03)$

$(14;\ 20;\ 31;\ 42;\ 03)$

$x_2x_5 \quad x_3x_1 \quad x_4x_2 \quad x_5x_3 \quad x_1x_4$

$(x_2x_5) \rightarrow (y_1 \vee y_3,\ y_1 \vee y_4) \rightarrow y_1y_1, \underline{y_1y_4}, y_3y_1, y_3y_4$

$(x_3x_1) \rightarrow (y_2 \vee y_4,\ y_2 \vee y_5) \rightarrow y_2y_2, \underline{y_2y_5}, y_4y_2, y_4y_5$

$(x_4x_2) \rightarrow (y_3 \vee y_5,\ y_1 \vee y_3) \rightarrow y_3y_1, y_3y_3, y_5y_1, y_5y_3$

$(x_5x_3) \rightarrow (y_1 \vee y_4,\ y_2 \vee y_4) \rightarrow y_1y_2, \underline{y_1y_4}, y_4y_2, y_4y_4$

$(x_1x_4) \rightarrow (y_2 \vee y_5,\ y_3 \vee y_5) \rightarrow y_2y_3, \underline{y_2y_5}, y_5y_3, y_5y_5$

$= 1/? \quad = 15$ TVRDAT DEKA $H(Y^n|X^n) = 0$

EATUA

$C = \frac{1}{2} \log 5$ — ZERO ERROR CAPACITY

ZATOA OVOJ KOD ... NA KANALOT

MMV

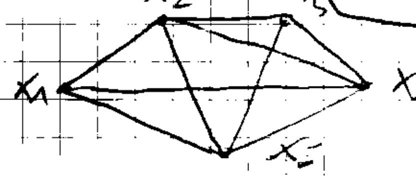
$x_1x_1 \rightarrow (y_2 \vee y_5;\ y_2 \vee y_5) \quad y_2y_2, y_2y_5, y_5y_2, y_5y_5$

$x_2x_3 \rightarrow (y_1 \vee y_3;\ y_2 \vee y_4) \quad y_1y_2, y_1y_4, y_3y_2, y_3y_4$

$x_3x_5 \rightarrow (y_2 \vee y_4;\ y_1 \vee y_4) \quad y_2y_1, y_2y_4, y_4y_1, y_4y_4$

$x_4x_2 \rightarrow (y_3 \vee y_5;\ y_1 \vee y_3) \quad y_3y_1, y_3y_3, y_5y_1, y_5y_3$

$x_5x_4 \rightarrow (y_1 \vee y_4;\ y_3 \vee y_5) \quad y_1y_3, y_1y_5, y_4y_3, y_4y_5$



OVA E PRAVILNOT KOD !!!

OVOJ KOD NEMA DA GREŠI !

GO NAJDOV OD ČLANAKOT NA LOVASZ !

7.11 TIME-VARYING CHANNELS. CONSIDER TIME-VARYING DISCRETE MEMORY-LESS CHANNEL

LET $Y_1, Y_2, \ldots Y_n$ BE CONDITIONALLY INDEPENDENT GIVEN $X_1, X_2, \ldots X_n$ WITH CONDITIONAL DISTRIBUTION GIVEN BY

$$p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i)$$

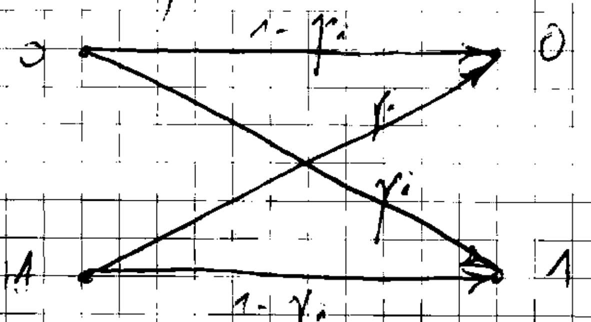
LET $\mathbf{X} = (X_1, X_2 \ldots X_n)$ $\mathbf{Y} = (Y_1, Y_2, \ldots Y_n)$

FIND $\max_{p(x)} I(X;Y)$

$$I(X;Y) = H(Y_1^n) - H(Y_1^n|X_1^n)$$

$$H(Y_1^n|X_1^n) = \sum_{i=1}^{n} H(Y_i|X_1^n Y_1^{i-1}) = \sum_{i=1}^{n} H(Y_i|X_1^n) = \sum_{i=1}^{n} H(Y_i|X_i)$$

$$H(X_1 X_2 \ldots X_n) = \sum_{i=1}^{n} H(X_i|X_1^{i-1}) = H(X_1) + H(X_2|X_1) + H(X_3|X_1 X_2) + \ldots H(X_n|X_1^{n-1})$$



$$I(X_i;Y_i) = H(Y_i) - H(Y_i|X_i) = 1 - H(Y_i|X_i)$$

$$H(Y_i|X_i) = p(X_i=0)\underbrace{H(Y_i|X_i=0)}_{H(p_i)} + p(X_i=1)\underbrace{H(Y_i|X_i=1)}_{H(p_i)}$$

$$\boxed{H(Y_i|X_i) = H(p_i)} \qquad I(X_i;Y_i) = 1 - H(p_i)$$

- MARKOV CHAIN $X \to Y \to Z$

$$p(X,Z|Y) = \frac{p(X,Y,Z)}{p(Y)} = \frac{p(XY)\cdot p(Z|XY)}{p(Y)} = p(X|Y)\cdot p(Z|Y)$$

$$p(Y,Z|X) = \frac{p(XYZ)}{p(X)} = \frac{p(XY)p(Z|Y)}{p(X)} = p(Y|X)\,p(Z|Y)$$

(83)

$$H(X,Y,Z) = H(X) + H(Y|X) + H(Z|X,Y) = H(X) + H(Y|X) + H(X)$$

$$H(\mathcal{X}) = \lim_{n\to\infty} \frac{H(X_1^n)}{n} = \lim_{n\to\infty} \frac{n\,H(X_2|X_1)}{n} = H(X_2|X_1)$$

$$H(\mathcal{X}) = \lim_{n\to\infty} H(X_n|X_1^{n-1}) = H(X_n|X_{n-1}) = H(X_2|X_1)$$ TIME INVARIANT

$$I(X;Y) = H(Y_1^n) - \sum_{i=1}^{n} \underbrace{H(Y_i|X_i)}_{H(p_i)}$$

$$I(X;Y) = H(Y_1^n) - \sum_{i=1}^{n} H(p_i) \leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(p_i)$$

$$\leq \sum_{i=1}^{n} I(X_i;Y_i) = \sum_{i=1}^{n} (1 - H(p_i))$$

$$I(X;Y) \leq \sum_{i=1}^{n} I(X_i;Y_i) \qquad I(X;Y) \leq n - \underbrace{\sum_{i=1}^{n} H(p_i)}_{0} \leq n$$

$H(p_i) = 0$ FOR $p_i = 0$ OR $p_i = 1$

$$C = n$$

$$C = \max_{p(x)} I(X;Y)$$

EDITION 1 SOLUTION

$$C = \sum_{i=1}^{n} (1 - H(p_i))$$

PROBLEM 7.12 UNUSED SYMBOLS. SHOW THAT CAPACITY OF THE CHANNEL WITH PROBABILITY TRANSITION MATRIX

$$P_{Y|X} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

IS ACHIEVED BY A DISTRIBUTION THAT PLACES ZERO PROBABILITY ON ONE OF INPUT SYMBOLS. WHAT IS THE CAPACITY OF THIS CHANNEL? GIVE AN INTUITIVE REASON WHY THAT LETTER IS NOT USED.

$$\begin{bmatrix} p(y_1) \\ p(y_2) \\ p(y_3) \end{bmatrix} = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \end{bmatrix}$$

$$\begin{bmatrix} p(y_1) \\ p(y_2) \\ p(y_3) \end{bmatrix} = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \end{bmatrix}$$

$$I(X;Y) = H(Y) - H(Y|X)$$

$$H(Y|X) = \sum_x p(x)\, H(Y|x)$$

$$= p(x_1)\cdot H(Y|x_1) + p(x_2)\cdot H(Y|x_2) + p(x_3)\cdot H(Y|x_3)$$

$$= p(x_1)\left[\tfrac{1}{3} ld 3 + \tfrac{2}{3} ld \tfrac{3}{2}\right] + p(x_2)\left[3\cdot\tfrac{1}{3} ld 3\right] + p(x_3)\left[\tfrac{1}{3} ld 3 + \tfrac{2}{3} ld \tfrac{3}{2}\right] =$$

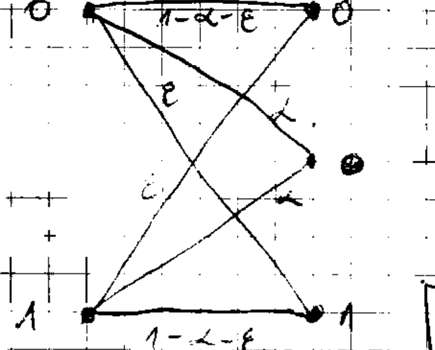
$$= p(x_1)\left[\tfrac{1}{3} ld 3 + \tfrac{2}{3} ld 3 - \tfrac{2}{3}\right] + p(x_2)\left[ld 3\right] + p(x_3)\left[ld 3 - \tfrac{2}{3}\right]$$

$$\min\left[H(Y|X)\right] = \left[p(x_1) + p(x_3)\right]\cdot\left(ld 3 - \tfrac{2}{3}\right) \quad \text{for } p(x_2) = 0$$

$$C = \max_{p(x)} I(X;Y) = H(Y) - ld 3 + \tfrac{2}{3} = ld 3 - ld 3 + \tfrac{2}{3}$$

$C = 2/3$ $\qquad p(x_1) = p(x_3) = \frac{1}{2}$ (UNIFORM) $\qquad p(Y) = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right]$ (UNIFORM)

PROBLEM 7.13 ERASURES AND ERRORS IN A BINARY CHANNEL. CONSIDER A CHANNEL WITH BINARY INPUTS THAT HAS BOTH ERRORS AND ERASURES. LET THE PROBABILITY OF ERROR BE ε AND THE PROBABILITY OF ERASURE BE α, SO THE CHANNEL IS AS FOLLOWS:



(a) FIND THE CAPACITY OF THIS CHANNEL

(b) SPECIALIZE TO THE CASE OF THE BINARY SYMMETRIC CHANNEL ($\alpha = 0$)

(c) SPECIALIZE TO THE CASE OF THE BINARY ERASURE CHANNEL ($\varepsilon = 0$)

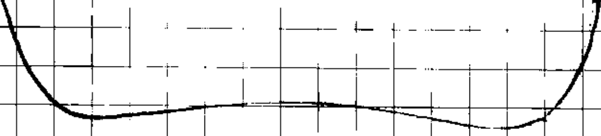
(a) $I(X;Y) = H(Y) - H(Y|X)$

$$H(Y) = p(y=0)\, ld\, p(y=0) + p(y=e)\, ld\, p(y=e) + p(y=1)\, ld\, p(y=1)$$

$$p(y=0) = p(x=0)\cdot(1-\alpha-\varepsilon) + p(x=1)\cdot\varepsilon = 1-\alpha-\varepsilon+\varepsilon = \frac{1-\alpha}{2}$$

$$p(y=1) = p(x=1)\cdot(1-\alpha-\varepsilon) + p(x=0)\cdot\varepsilon = (1-\alpha)/2$$

$$p(y=e) = p(x=0)\,\alpha + p(x=1)\cdot\alpha = \alpha \qquad p(x=0) = p(x=1) = \tfrac{1}{2}$$



$$\gamma(\tau) = \left\{ \frac{1-\alpha}{2}, \frac{1-\alpha}{2} \right\} \qquad \frac{1-\alpha}{2} + \frac{1-\alpha}{2} + \frac{2\alpha}{2} = \frac{2}{2} = 1$$

IF $\gamma(\tau=0) = \sigma \qquad \gamma(\tau=1) = 1-\sigma$

$$\gamma(\tau=0) = \sigma(1-\alpha-\varepsilon) + (1-\sigma)\varepsilon = \sigma(1-\alpha) - \varepsilon\sigma + (1-\sigma)\varepsilon$$

$$= \sigma - \alpha\sigma - \varepsilon\sigma - \varepsilon - \sigma\varepsilon = \sigma(1-\alpha-2\varepsilon) + \varepsilon$$

$$\gamma(\tau=\theta) = \sigma\cdot\alpha + (1-\sigma)\cdot\alpha = \alpha$$

$$\gamma(\tau=1) = (1-\sigma)(1-\alpha-\varepsilon) + \sigma\cdot\varepsilon = 1-\alpha-\varepsilon-\sigma+\sigma\alpha+\sigma\varepsilon+\sigma\varepsilon$$

$$\gamma(\tau=1) = 1-\alpha-\varepsilon+\sigma(2\varepsilon+\alpha-1) = 1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)$$

$$\gamma(\tau=0) + \gamma(\tau=1) + \gamma(\tau=\theta) = \varepsilon + \sigma(1-\alpha-2\varepsilon) + 1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon) + \alpha = 1$$

$$H(\tau) = \alpha \cdot \mathrm{ld}\frac{1}{\alpha} + \left(\varepsilon + \sigma(1-\alpha-2\varepsilon)\right)\mathrm{ld}\left[\sigma(1-\alpha-2\varepsilon)+\varepsilon\right]^{-1} +$$

$$+ \left[1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)\right]\cdot \mathrm{ld}\left[1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)\right]^{-1}$$

$$\frac{dH(\tau)}{d\sigma} = 0 + (1-\alpha-2\varepsilon)\cdot \mathrm{ld}\left[\sigma(1-\alpha-2\varepsilon)+\varepsilon\right]^{-1} +$$

$$+ \frac{\left[\varepsilon+\sigma(1-\alpha-2\varepsilon)\right]}{\ln 2}\cdot\frac{(1-\alpha-2\varepsilon)}{\sigma(1-\alpha-2\varepsilon)+\varepsilon} - (1-\alpha-2\varepsilon)\,\mathrm{ld}\left[1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)\right]^{-1} + \frac{\left[1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)\right]}{\ln 2}\cdot\frac{-(1-\alpha-2\varepsilon)}{\left[1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)\right]}$$

$$= (1-\alpha-2\varepsilon)\cdot \mathrm{ld}\left[\sigma(1-\alpha-2\varepsilon)+\varepsilon\right]^{-1} - (1-\alpha-2\varepsilon)\,\mathrm{ld}\left[1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)\right]^{-1} = 0$$

$$(1-\alpha-2\varepsilon)\,\mathrm{ld}\,\frac{1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)}{\sigma(1-\alpha-2\varepsilon)+\varepsilon} = 0$$

$$\frac{1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon)}{\sigma(1-\alpha-2\varepsilon)+\varepsilon} = 1 \qquad 1-\alpha-\varepsilon-\sigma(1-\alpha-2\varepsilon) = \varepsilon+\sigma(1-\alpha-2\varepsilon)$$

$$2\sigma(1-\alpha-2\varepsilon) = (1-\alpha-2\varepsilon) \qquad \boxed{\sigma = \frac{1}{2}}$$

$$H(\tau) = \alpha\,\mathrm{ld}\frac{1}{\alpha} + \left(\varepsilon + \frac{1-\alpha-2\varepsilon}{2}\right)\mathrm{ld}\left[\frac{1-\alpha-2\varepsilon}{2}+\varepsilon\right]^{-1} +$$

$$+ \left(1-\alpha-\varepsilon - \frac{(1-\alpha-2\varepsilon)}{2}\right)\mathrm{ld}\left[1-\alpha-\varepsilon-\frac{1-\alpha-2\varepsilon}{2}\right]^{-1}$$

$$H(Y) = \alpha \, ld \frac{1}{\alpha} + \frac{1-\alpha}{2} \, ld \left(\frac{1-\alpha}{2}\right)^{-1} + \left(\frac{2(1-\alpha-\varepsilon) - (1-\alpha-2\varepsilon)}{2}\right.$$

$$\cdot \, ld \left[\frac{2(1-\alpha) - 2\varepsilon - (1-\alpha) + 2\varepsilon}{2}\right]^{-1} = \alpha \, ld \frac{1}{\alpha} + \frac{1-\alpha}{2} \, ld \left[\frac{1-\alpha}{2}\right]^{-1} +$$

$$+ \frac{1-\alpha}{2} \, ld \left[\frac{1-\alpha}{2}\right]^{-1} = \underline{\alpha \, ld \frac{1}{\alpha} + (1-\alpha) \, ld \frac{2}{(1-\alpha)}}$$

$$H(Y) = \alpha \, ld \frac{1}{\alpha} + (1-\alpha) \, ld \frac{2}{(1-\alpha)}$$

- ISTO SE DOBIVA SO DIREKTNA PRETPOSTAVKA DEKA

$$p(x=0) = p(x=1) = 1/2$$

$$\boxed{H(Y) = \alpha \, ld \frac{1}{\alpha} + \frac{1-\alpha}{2} \, ld \frac{2}{1-\alpha} + \frac{1-\alpha}{2} \, ld \frac{2}{1-\alpha}}$$

$$H(Y|X) = \sum_x p(x) \cdot H(Y|X=x) = p(x=0) \cdot H(Y|X=0) +$$

$$p(x=1) \cdot H(Y|X=1) = \frac{1}{2}\left[(1-\alpha-\varepsilon) \, ld (1-\alpha-\varepsilon)^{-1} + \alpha \cdot ld \, \alpha^{-1} +\right.$$

$$\left. + \varepsilon \, ld \, \varepsilon^{-1}\right] \cdot 2 = (1-\alpha-\varepsilon) \, ld \frac{1}{1-\alpha-\varepsilon} + \alpha \, ld \frac{1}{\alpha} + \varepsilon \, ld \frac{1}{\varepsilon}$$

$$= ld \frac{1}{1-\alpha-\varepsilon} - \alpha \, ld \frac{1}{1-\alpha-\varepsilon} - \varepsilon \, ld \frac{1}{1-\alpha-\varepsilon} + \alpha \, ld \frac{1}{\alpha} + \varepsilon \, ld \frac{1}{\varepsilon} =$$

$$= ld \frac{1}{1-\alpha-\varepsilon} + \alpha \, ld \left(\frac{1-\alpha-\varepsilon}{\alpha}\right) + \varepsilon \, ld \left(\frac{1-\alpha-\varepsilon}{\varepsilon}\right)$$

$$\boxed{I(X;Y) = \alpha \, ld \frac{1}{\alpha} + (1-\alpha) \, ld \frac{2}{1-\alpha} - ld \frac{1}{1-\alpha-\varepsilon} - \alpha \, ld \left(\frac{1-\alpha-\varepsilon}{\alpha}\right) - \varepsilon \, ld \left(\frac{1-\alpha-\varepsilon}{\varepsilon}\right)}$$

$$C = I(X;Y)$$

(b) $\boxed{\alpha = 0}$ $\qquad C = ?$

$$C = \alpha \, ld (1-\alpha-\varepsilon)^{-1} + (1-\alpha) \, ld \frac{2}{1-\alpha} - \varepsilon \, ld \left(\frac{1-\alpha-\varepsilon}{\varepsilon}\right) - ld \frac{1}{1-\alpha-\varepsilon} = /\alpha = 0/$$

$$= 0 + ld 2 - \varepsilon \, ld \frac{1-\varepsilon}{\varepsilon} - ld \frac{1}{1-\varepsilon} =$$

$$= 1 - \varepsilon \, ld \frac{1}{\varepsilon} - \varepsilon \, ld (1-\varepsilon) - ld \frac{1}{1-\varepsilon} = 1 - \varepsilon \, ld \frac{1}{\varepsilon} + \varepsilon \, ld \frac{1}{1-\varepsilon} - ld \frac{1}{1-\varepsilon}$$

$$= 1 - \varepsilon \, ld \, \varepsilon^{-1} - (1-\varepsilon) \, ld \frac{1}{1-\varepsilon} = 1 - H(\varepsilon)$$

(87a)

(c) $\boxed{\varepsilon=0}$ $C = \alpha\, ld\frac{1}{\alpha} + (1-\alpha)\, ld\frac{2}{1-\alpha} - ld\frac{1}{1-\alpha} - \alpha\, ld\left(\frac{1-\alpha}{\alpha}\right)$

$- \varepsilon\, ld\frac{1-\alpha-\varepsilon}{\varepsilon}$ $A = 0$

$A = \varepsilon\, ld\,\varepsilon - \varepsilon\, ld(1-\alpha)$

$\lim_{\varepsilon\to 0} (\varepsilon\, ld\,\varepsilon) = \lim_{\varepsilon\to 0} \frac{\varepsilon \ln \varepsilon}{\ln 2}$

$\lim_{x\to 0} x \ln x = -\lim_{x\to 0} \frac{\ln x}{-\frac{1}{x}} \left(\frac{-\infty}{-\infty}\right) = \lim_{x\to 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = -\lim_{x\to 0} \frac{x^2}{x}$

$= -\lim_{y\to\infty} \frac{\frac{1}{y^2}}{\frac{1}{y}} = -\lim_{y\to\infty} \frac{y}{y^2} \to 0$

$C = \alpha\, ld\frac{1}{\alpha} + (1-\alpha)\, ld\frac{1}{1-\alpha} + (1-\alpha)\, ld\, 2 - ld\frac{1}{1-\alpha} - \alpha\, ld\left(\frac{1-\alpha}{\alpha}\right)$

$= \alpha\, ld\frac{1}{1-\alpha} + (1-\alpha)\, ld\frac{1}{1-\alpha} + (1-\alpha) - ld\frac{1}{1-\alpha} =$

$= -(1-\alpha)\, ld\frac{1}{1-\alpha} + (1-\alpha)\, ld\frac{1}{1-\alpha} + (1-\alpha) = 1-\alpha$

1.14 CHANNEL WITH DEPENDENCE BETWEEN LETTERS.
CONSIDER THE FOLLOWING CHANNEL OVER A BINARY ALPHABET THAT TAKES IN 2-BIT SYMBOLS AND PRODUCES A 2-BIT OUTPUT AS DETERMINED BY THE FOLLOWING MAPPING: $00 \to 01$, $01 \to 10$, $10 \to 11$, AND $11 \to 00$. THUS IF THE TWO BIT SEQUENCE 01 IS THE INPUT TO THE CHANNEL, THE OUTPUT IS 10 WITH PROBABILITY 1. LET X_1X_2 DENOTE THE TWO INPUT SYMBOLS AND Y_1Y_2 DENOTE THE CORRESPONDING OUTPUT SYMBOLS.

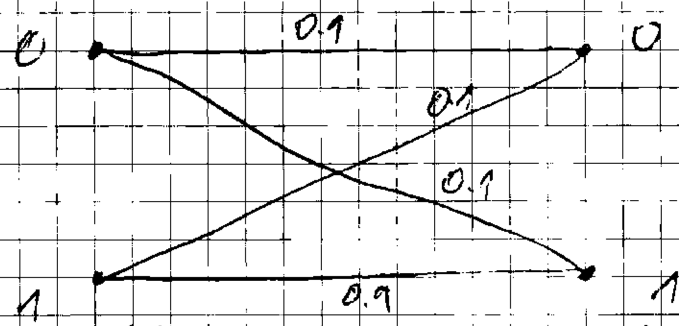
(a) CALCULATE THE MUTUAL INFORMATION $I(X_1X_2; Y_1Y_2)$ AS A FUNCTION OF THE INPUT DISTRIBUTION ON THE FOUR POSSIBLE PAIRS OF INPUTS.

(b) SHOW THAT THE CAPACITY OF PAIR OF TRANSMISSION ON THIS CHANNEL IS 2 BITS

(c) SHOW THAT UNDER THE MAXIMIZING INPUT DISTRIBUTION, $I(X_1; Y_1) = 0$. THUS, THE DISTRIBUTION ON THE INPUT SEQUENCE THAT ACHIEVES CAPACITY DOES NOT NECESSARY MAXIMIZE MUTUAL INFORMATION BETWEEN INDIVIDUAL SYMBOLS AND THEIR CORRESPONDING OUTPUTS.

CONTINUE FROM 128!

PROBLEM 7.15 [MMV] JOINTLY TYPICAL SEQUENCES. AS WE DID IN PROBLEM 3.13 FOR THE TYPICAL SET FOR A SINGLE RANDOM VARIABLE, WE WILL CALCULATE THE JOINTLY TYPICAL SET FOR A PAIR OF RANDOM VARIABLES CONNECTED BY A BINARY SYMMETRIC CHANNEL, AND THE PROBABILITY OF ERROR FOR JOINTLY TYPICAL DECODING FOR SUCH A CHANNEL.



WE CONSIDER A BINARY SYMMETRIC CHANNEL WITH CROSSOVER PROBABILITY 0.1. THE INPUT DISTRIBUTION THAT ACHIEVES CAPACITY IS THE UNIFORM DISTRIBUTION $[p(x) = \{\frac{1}{2}, \frac{1}{2}\}]$ WHICH YIELDS THE JOINT DISTRIBUTION $p(x,y)$ FOR THIS CHANNEL:

x \ y	0	1	p(x)
0	0.45	0.05	0.5
1	0.05	0.45	0.5
p(y)	0.5	0.5	

THE MARGINAL DISTRIBUTION OF Y IS: $p(y) = \{\frac{1}{2}, \frac{1}{2}\}$

(a) CALCULATE $H(X)$, $H(Y)$, $H(X,Y)$ AND $I(X;Y)$ FOR JOINT DISTRIBUTION ABOVE

(b) LET X^n BE DRAWN i.i.d. ACCORDING BERNOULLI(1/2) DISTRIBUTION. OF THE 2^n POSSIBLE INPUT SEQUENCES OF LENGTH n WHICH OF THEM ARE TYPICAL [i.e. MEMBER OF $A_\epsilon^{(n)}(X)$ FOR $\epsilon = 0.2$]? WHICH ARE THE TYPICAL SEQUENCES IN $A_\epsilon^{(n)}(Y)$?

(c) THE JOINTLY TYPICAL SET $A_\epsilon^{(n)}(X,Y)$ IS DEFINED AS THE SET OF SEQUENCES THAT SATISFY EQUATIONS (7.25 – 7.37). THE FIRST TWO EQUATIONS CORRESPOND TO THE CONDITIONS THAT x^n AND y^n ARE IN $A_\epsilon^n(X)$ AND $A_\epsilon^n(Y)$, RESPECTIVELY. CONSIDER THE LAST CONDITION, WHICH CAN BE REWRITTEN TO STATE THAT

$$-\frac{1}{n}\log p(x^n, y^n) \in \left(H(X,Y)-\epsilon,\ H(X,Y)+\epsilon\right).$$

LET k BE THE NUMBER OF PLACES IN WHICH THE SEQUENCE x^n DIFFERS FROM y^n (k IS A FUNCTION OF THE TWO SEQUENCES). THEN WE CAN WRITE

$$p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i) = (0.45)^{n-k}(0.05)^{k} = \left(\frac{1}{2}\right)^{n}(1-p)^{n-k}p^{k}$$

AN ALTERNATIVE WAY OF LOOKING AT THIS PROBABILITY IS

(87a)

TO LOOK AT THE BSC AS AN ADDITIVE CHANNEL $Y = X \oplus Z$, WHERE Z IS BINARY RANDOM VARIABLE THAT IS EQUAL TO 1 WITH PROBABILITY p, AND IS INDEPENDENT OF X. IN THIS CASE,

$$p(x^n, y^n) = p(x^n)\, p(y^n \mid x^n) = p(x^n) \cdot p(z^n \mid x^n) = p(x^n)\, p(z^n) = \left(\frac{1}{2}\right)^n \cdot (1-p)^{n-k}\, p^k$$

SHOW THAT THE CONDITION THAT (x^n, y^n) BEING JOINTLY TYPICAL IS EQUIVALENT TO THE CONDITION THAT x^n IS TYPICAL AND $z^n = y^n - x^n$ IS TYPICAL.

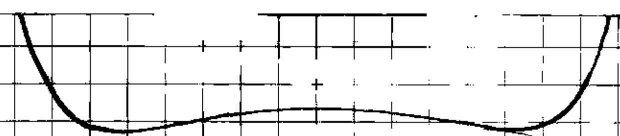
(d) WE NOW CALCULATE THE SIZE OF $A_\varepsilon^{(n)}(Z)$ FOR $n=25$ AND $\varepsilon = 0.2$ AS IN THEOREM 3.13. HERE IS TABLE OF THE PROBABILITIES AND NUMBER OF SEQUENCES WITH k ONES:

k	$\binom{n}{k}$	$\binom{n}{k} p^k (1-p)^{n-k}$	$-\frac{1}{n} \log p(x^n)$
0	1	0.071790	0.152003
1	25	0.199416	0.278800
2	300	0.265888	0.405597
3	2300	0.226497	0.532394
...		...	...

[SEQUENCES WITH MORE THAN 12 ONES ARE OMITTED SINCE THEIR TOTAL PROBABILITY IS NEGLIGIBLE (AND THEY ARE NOT IN THE TYPICAL SET).]

WHAT IS THE SIZE OF THE SET $A_\varepsilon^{(n)}(Z)$?

(e) NOW CONSIDER RANDOM CODING FOR THE CHANNEL AS IN THE PROOF OF THE CHANNEL CODING THEOREM. ASSUME THAT 2^{nR} CODEWORDS $X^n(1), X^n(2), \ldots, X^n(2^{nR})$ ARE CHOSEN UNIFORMLY OVER THE 2^n POSSIBLE BINARY SEQUENCES OF LENGTH n. ONE OF THE CODEWORDS IS CHOSEN AND SENT OVER THE CHANNEL. THE RECEIVER LOOKS AT THE RECEIVED SEQUENCE AND TRIES TO FIND A CODEWORD IN THE CODE THAT IS JOINTLY TYPICAL WITH RECEIVED SEQUENCE. AS ARGUED ABOVE, THIS CORRESPONDS TO FINDING A CODEWORD $X^n(i)$ SUCH THAT $Y^n - X^n(i) \in A_\varepsilon^{(n)}(Z)$. FOR A FIXED CODEWORD $x^n(i)$, WHAT IS THE PROBABILITY THAT THE RECEIVED SEQUENCE Y^n IS SUCH THAT $(x^n(i), Y^n)$ IS JOINTLY TYPICAL.



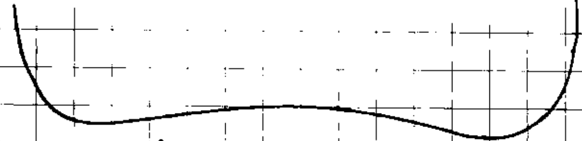
(f) NOW CONSIDER A PARTICULAR RECEIVED SEQUENCE $y^n = 000000\ldots0$ SAY. ASSUME THAT WE CHOSE A SEQUENCE x^n AT RANDOM, UNIFORMLY DISTRIBUTED AMONG ALL THE 2^n POSSIBLE BINARY n-SEQUENCES. WHAT IS THE PROBABILITY THAT THE CHOSEN SEQUENCE IS JOINTLY TYPICAL WITH y^n? [HINT: THIS IS THE PROBABILITY OF ALL SEQUENCES x^n SUCH THAT $y^n - x^n \in A_\epsilon^{(n)}(Z)$.]

(g) NOW CONSIDER A CODE WITH $2^9 = 512$ CODEWORDS OF LENGTH 12 CHOSEN AT RANDOM, UNIFORMLY DISTRIBUTED AMONG ALL THE 2^n SEQUENCES OF LENGTH n = 25. ONE OF THESE CODEWORDS, SAY THE ONE CORRESPONDING TO i=1, IS CHOSEN AND SENT OVER THE CHANNEL. AS CALCULATED IN PART (e), THE RECEIVED SEQUENCE WITH HIGH PROBABILITY IS JOINTLY TYPICAL WITH THE CODEWORD THAT WAS SENT. WHAT IS THE PROBABILITY THAT **ONE OR MORE** OF THE OTHER CODEWORDS (WHICH WERE CHOSEN AT RANDOM INDEPENDENT OF THE SENT CODEWORD) IS JOINTLY TYPICAL WITH THE RECEIVED SEQUENCE? [HINT: YOU COULD USE THE UNION BOUND, BUT YOU COULD ALSO CALCULATE THE PROBABILITY EXACTLY USING THE RESULT OF PART (f) AND INDEPENDENCE OF THE CODEWORDS.]

(h) GIVEN THAT A PARTICULAR CODEWORD WAS SENT, THE PROBABILITY OF ERROR (AVERAGED OVER THE PROBABILITY DISTRIBUTION OF THE CHANNEL AND OVER THE RANDOM CHOICE OF OTHER CODEWORDS) CAN BE WRITTEN AS

$$P_r(E \mid x^n(1) \text{ sent}) = \sum_{x^n, y^n \text{ cause error}} p(y^n \mid x^n(1)).$$

THERE ARE TWO TYPES OF ERROR: (1) RECEIVED SEQUENCE y^n IS NOT JOINTLY TYPICAL WITH THE TRANSMITTED CODEWORD, (2) THERE IS ANOTHER CODEWORD JOINTLY TYPICAL WITH THE RECEIVED SEQUENCE. USING THE RESULTS OF THE PRECEDING PARTS, CALCULATE THE PROBABILITY OF ERROR. BY THE SYMMETRY OF THE RANDOM CODING ARGUMENT, THIS DOESN'T DEPEND ON WHICH CODEWORD WAS SENT.



$$2^{-n(H+\varepsilon)} \le p(x_1^n) \le 2^{-n(H-\varepsilon)}$$ **AEP**

$$P_r(x \in A_\varepsilon^n) \ge 1-\varepsilon \qquad |A_\varepsilon^n| \le 2^{n(H+\varepsilon)} \qquad |A_\varepsilon^{(n)}| \ge (1-\varepsilon)2^{n(H-\varepsilon)}$$

$$-n(H+\varepsilon) \le \frac{1}{n} ld\, p(x_1^n) \le -n(H-\varepsilon) \qquad (H-\varepsilon) \le -\frac{1}{n} ld\, p(x_1^n) \le (H+\varepsilon)$$

$$-\varepsilon \le -H - \frac{1}{n} ld\, p(x_1^n) \le \varepsilon \qquad \left|-H - \frac{1}{n} ld\, p(x_1^n)\right| \le \varepsilon$$

KONTINUE FROM 1bc |

(f) $Y^n = 000000 \ldots 0$

MISLAM DEKA AKO E KEPSOT (3) OD JOINT AEP ... KEPSOT KADE X^n, Y^n SE NEZAVISNI

$$2^{-n(I(X,Y)+3\varepsilon)} \cdot (1-\varepsilon) \le P_r(\tilde{X}_n, \tilde{Y}^n \in A_\varepsilon^n) \le 2^{-n(I(X,Y)-3\varepsilon)}$$

$$I(X,Y) = 0.531$$

$n = 25$ $\qquad 2.5 \cdot 10^{-4} \le P_r(\tilde{X}_n, \tilde{Y}^n \in A_\varepsilon^n) \le 0.00027$

0.00025

$n = 100$ $\qquad 1.6 \cdot 10^{-18} \le P_r(\tilde{X}_n, \tilde{Y}^n \in A_\varepsilon^n) \le 6.6 \cdot 10^{-15}$

$n = 200$ $\qquad 2.6 \cdot 10^{-36} \le P_r(\tilde{X}_n, \tilde{Y}^n \in A_\varepsilon^n) \le 4 \cdot 10^{-29} \sim 2^{-nI(X,Y)}$

(g)

$$P_r(\tilde{X}_n, \tilde{Y}_n \in A_\varepsilon^n) = \sum_{x^n, y^n \in A_\varepsilon^n} p(x^n)\, p(y^n)$$

p ZEMI NA

$Y^n = 000 \ldots 0$ TI TREBA SUMA NA SITE X^n KOI SE RAZLIKUVAAT OD Y^n VO 1, 2, 3, 4 MESTA

$$P_r = \sum_{A_\varepsilon^n} \frac{1}{2^n} = |A_\varepsilon^n| \cdot \frac{1}{2^n}$$

$n = 25$ $\qquad P_r = \frac{15275}{2^{25}} = 0.00046$ $\qquad |A_\varepsilon^{(n)}| = 15275$

$$\sum_{k=1}^{4} \binom{n}{k} \frac{1}{2^n} \qquad \frac{\sum_{k=1}^{4} \binom{n}{k}}{2^n}$$

(g) $P_r(\tilde{X}_n, \tilde{Y}_n \in A_\varepsilon^n) = \frac{A_\varepsilon^9}{2^{25}} = \frac{255}{2^{25}} = 0.00001$

$$\sum_{k=1}^{4} \binom{9}{k} = 255$$

92 a

$$\left| -\frac{1}{n} p^k (1-p)^{12-k} - 0.469 \right| \le \varepsilon = 0.2 \qquad k = 0, 1, \dots, 12$$

$$P_r\left[(\tilde{X}^n, \tilde{Z}^n) \in A_\varepsilon^{n}\right] = \frac{781}{2^{12}} = 0.19$$

→ TRUE $k = 2, 3, 4, 5$

$$\sum_{q=2}^{4} \binom{n}{q} = \sum_{2=2}^{4} \binom{12}{2} = 781$$

$$\sum_{k=2}^{4} \binom{n}{k} p^2 (1-p)^{12-2} = 0.631$$

~~VIC SOLUTION of (e)~~

$$2^{-n \cdot 1.2} \le 2^{-n} \le 2^{-n \cdot 0.8} \qquad // \, ld$$

$$/ \cdot -\frac{1}{n}$$

$$-n \cdot 1.2 \le -n \le -n \cdot 0.8$$

$$1.2 \ge 1 \ge 0.8 \qquad 0.8 \le 1 \le 1.2 \Big\}$$ MORE RIGHT

solution of (e)

$$\left| -\frac{1}{n} ld \, p(x^n, z^n) - H(X, Z) \right| < \varepsilon$$

$$p(x^n, z^n) = p(x^n) \cdot p(z^n | x^n) = p(x^n) \cdot p(z^n)$$

$$H(X, Z) = H(X) + H(Z|X) = H(X) + H(Z)$$

$$\left| -\frac{1}{n} ld \, p(x^n) \cdot p(z^n) - H(X) - H(Z) \right| \le$$

$$\le \left| -\frac{1}{n} ld \, 2^{-n} - \frac{1}{n} ld \, p(z^n) - 1 - H(Z) \right| =$$

$$= \left| 1 - \frac{1}{n} ld \, p(z^n) - 1 - H(Z) \right| = \left| -\frac{1}{n} ld \, p(z^n) - H(Z) \right| \le \varepsilon$$

solution of (g)

$$P_r\left(\bigcup_i A_i\right) \le \sum_i P_r(A_i)$$

THE PROBABILITY THAT AT LEAST ONE OF THIS 512 CODEWORDS FALLS IN THE $A_\varepsilon^{(n)}$ SET IS NO GREATER THAN THE SUM OF PROBABILITIES THAT EACH CODEWORD FALLS IN SET OF $A^{(n)}$. THEREFORE

ALTERNATIVE SOLUTION ON NR. 130

$$P_r\left(\bigcup_i A_i\right) \le 512 \times 0.00046 = 512 \cdot \frac{15275}{2^{25}} = 0.23308$$

(h) $$P_r(\text{Error} | x^n(1)) = \sum_{y^n: y^n \text{ causes error}} p(y^n | x^n(1)) = 1 - P_r(x^n, z^n \in A_\varepsilon^{n}) +$$

$$+ P_r\left[(\tilde{X}^n, \tilde{Z}^n) \in A_\varepsilon^{(n)}\right] =$$

ANOTHER CODEWORD IS JOINTLY TYPICAL WITH REC. SEQ.

(x^n, z^n) ARE NOT JOINTLY TYPICAL

$$= 1 - \sum_{k=2}^{4} \binom{n}{k} p^k (1-p)^{n-k} + 512 \frac{15275}{2^{25}} = 1 - 0.83 + 0.23308$$

(93.1)

$$\Pr(\text{Error} \mid X^n(1)) = 0.17 + 0.23 = 0.4$$

We have error if received sequence is not jointly typical with the transmitted codeword or if another codeword is jointly typical with received sequence.

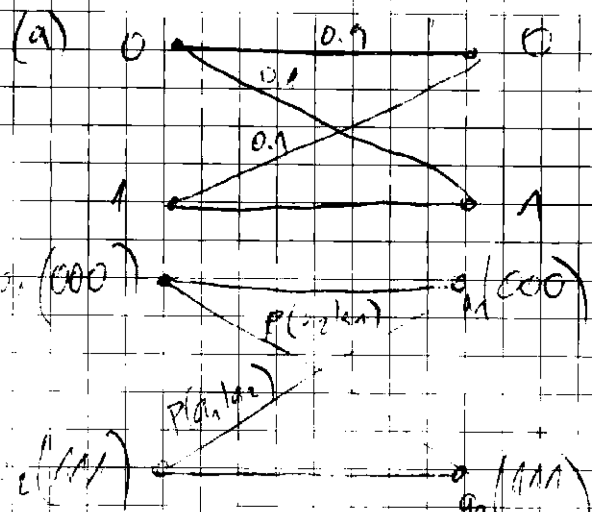
(→ OVA E SUŠTINATA NA [illegible] NA [illegible] NA GREŠKA.)

Problem 7.16 Encoder and decoder as part of the channel

Consider a binary symmetric channel with crossover probability 0.1. A possible coding scheme for this channel with two codewords of length 3 is to encode message a_1 as 000 and a_2 as 111. With this coding scheme we can consider the combination of encoder, channel, and decoder as forming a new BSC with two inputs a_1 and a_2 and two outputs a_1 and a_2.

(a) Calculate the crossover probability of this channel.
(b) What is the capacity of this channel?
(c) What is the capacity of original BSC with crossover probability 0.1?
(d) Prove a general result that for any channel, considering the encoder, channel and decoder together as a new channel from messages to estimated messages will not increase the capacity in bits per transmission of the original channel.

(a)



(a) $P(1|0) = 0.1$

$P(a_1) = 0.1$

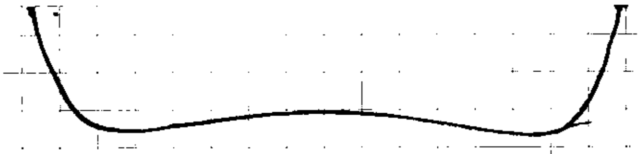
$P(a_2|a_1) = P(111|000) +$

$P(110|000) = \frac{1}{10^3} + \frac{1}{10^2} = \frac{11}{10^3}$

$P(a_1|a_2) = P(000|111) + P(00\ldots|111)$

$= \frac{1}{10^3} + \frac{1}{10^2} = \frac{11}{10^3}$

(b) $C = \max_{p(x)} I(X;Y)$, $I(X;Y) = H(Y) - H(Y|X)$



(9.1a)

$P(a_2|a_1) = \frac{11}{10^3} = P(a_1|a_2) = 0.01100$

$P(Y|X)$

X \ Y	a_1	a_2
a_1	0.989	0.011
a_2	0.011	0.989

$P(X,Y) = P(X)\cdot P(Y|X)$

X \ Y	a_1	a_2	P(X)
a_1	0.4945	0.0055	0.5
a_2	0.0055	0.4945	0.5
P(Y)	0.5	0.5	

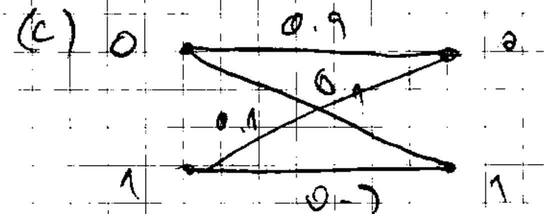
$H(X) = \left(\frac{1}{2} ld\, 2\right)\cdot 2 = 1$ $\qquad H(X|Y) = ? \qquad H(Y|X) = ?$

$H(Y) = \left(\frac{1}{2} ld\, 2\right) = 1$

$H(Y|X) = -\sum P(x,y)\, ld\, P(Y|X) = -\Big[0.4945\,(ld\, 0.989)\cdot 2 +$
$+\ 0.0055\,[ld\,(0.011)]\cdot 2\Big] = \left[0.989\, ld\, 0.989 + 0.011\, ld\, 0.011\right]$
$= 0.08735$

$I(X,Y) = H(X) - H(Y|X) = 1 - 0.08735 = 0.91265$

(c)



$P(Y|X)$

X \ Y	0	1
0	0.9	0.1
1	0.1	0.9

X \ Y	0	1	P(X)
0	0.45	0.05	0.5
1	0.05	0.45	0.5
P(Y)	0.5	0.5	

$C = 1 - H(Y|X)$

$H(Y|X) = -\left[0.45\,(ld\, 0.9)\cdot 2 + 0.05\,(ld\, 0.1)\cdot 2\right] =$
$= -\left[0.9\, ld\, 0.9 + 0.1\, ld\,(0.1)\right] = 0.4690$

$C = 1 - 0.4690 = \underline{0.531}$

(d) $\boxed{\frac{0.91265}{3} = 0.30422 \text{ bit/tranzm}} \leq 0.531$

- NE E DOVA PRESMETKATA NA CROSSOVER PROBABILITY. GI IMAM PROVERENO DVETE KOMBINACII OD SEKVENCI.

(b) REVISITED

$P(a_2|a_1) = P(\text{"111"}|\text{"000"}) = P(111|000) + P(110|000) + P(101|000) + P(011|000) = \left(\frac{1}{10}\right)^3 + \binom{3}{2}\left(\frac{1}{10}\right)^2 0.9 = 0.031$

$\binom{3}{2} = \frac{3!}{(3-2)! \cdot 2!} = \frac{6}{2} = 3$

X \ Y	a_1	a_2
a_1	0.969	0.031
a_2	0.031	0.969

2 · 0.4845

$H(Y|X) = 0.969 \cdot ld(0.969) + 0.031\ ld\ 0.031 = 0.19938$

$C = 1 - 0.19938 = 0.80062$

$\frac{C}{3} = 0.26687$ bits/tran

– VO GENERALEN SLUČAJ: $P(a_2|a_1) = P(\text{"}\overbrace{11\ldots1}^{n}\text{"}|\text{"}00\ldots0\text{"}) =$

$= \left(\frac{1}{10}\right)^n + \binom{n}{n-1}\left(\frac{1}{10}\right)^{n-1} + \ldots \binom{n}{\left[\frac{n}{2}\right]+1}\left(\frac{1}{10}\right)^{n-\left[\frac{n}{2}\right]} = \sum_{i=0}^{\left[\frac{n}{2}\right]} \binom{n}{n-i}\frac{1}{10^{n-i}}$

$n = 2k+1 \quad k = 1, 2, \ldots$

$n = 3, 5, 7, \ldots$

$P(a_2|a_1) = \sum_{i=0}^{\left[\frac{n}{2}\right]} \binom{n}{n-i}\frac{1}{10^{n-i}} \xrightarrow[n\to\infty]{} 0$ (*)

→ KAKO GO ZGOLEMUVAŠ „n" TAKA CROSSOVER PROBABILITY TEŽI KON „0"

$H(Y|X) = -\left[P(a_2|a_1)\ ld\ P(a_2|a_1) + (1-P(a_2|a_1))\ ld(1-P(a_2|a_1))\right]$

$C = 1 - H(Y|X) \qquad C_n = \frac{1-H(Y|X)}{n} \xrightarrow[n\to\infty]{} 0$

VIDI MARKE!!!

• NAMESTO (*) MOŽEŠ DA PIŠ SO:

$P(a_2|a_1) = \sum_{i=0}^{\frac{n-1}{2}} \binom{n}{n-i}\frac{1}{10^{n-i}}$

John Hopkins Solution

(a) ISTIOT REZON SUM GO IMAL SAMO SUM PROPUSTIL DA MNOZAM SO 0.9.

$$P(a_2|a_1) = (0.1)^3 + 3(0.1)^2\, 0.9 = 0.02800$$

→ OVA GO IMAV ZABORAVENO.

- GENERALIZED

$$P(a_2|a_1) = \sum_{i=0}^{\frac{n-1}{2}} \binom{n}{n-i} \frac{1}{10^{n-i}} \left(\frac{9}{10}\right)^i$$

(b) $C_3 = \frac{1}{3}\left(1 - H(P(a_2|a_1))\right) = |n=3| = \frac{0.81574}{3} = 0.272$

(c) $C = 1 - H(0.1) = 0.531$

(d) $C = \max_{p(x^n)} I(X^n;Y^n)$

$$I(X^n;Y^n) = H(Y^n) - H(Y^n|X^n) = H(Y^n) - \sum_{x^n} p(x^n) H(Y^n|x^n)$$

$$\le H(Y^n) - \sum_{k=1}^{n} H(Y_k|X_k) \le \sum_{k=1}^{n}\left(H(Y_k) - H(Y_k|X_k)\right)$$

$$= \sum_{k=1}^{n} I(X_k;Y_k) = n\cdot I(X_k;Y_k) \le n\cdot C$$

$$C(X^n;Y^n) = \frac{I(X^n;Y^n)}{n} \le C(X_k;Y_k)$$

→ MOZE DIREKTNO DA SE ZAKLUCI OD LEMA 7.9.2.

MATLAB: $\frac{C(3)}{3} = 0.27$; $\frac{C(5)}{5} = 0.18$; $\frac{C(7)}{7} = 0.14$;

7.17 Codes of length 3 for BSC and BEC. In Problem 7.16 the probability of error was calculated for a code with two codewords of length 3 (000 and 111) sent over a binary symmetric channel with crossover probability ϵ. For this problem take $\epsilon = 0.1$.

(a) Find the best code of length 3 with four codewords for this channel. What is the probability of error for this channel. Note that

ALL POSSIBLE RECEIVED SEQUENCES SHOULD BE MAPPED ONTO POSSIBLE CODEWORDS)

(b) WHAT IS THE PROBABILITY OF ERROR IF WE USED ALL EIGHT POSSIBLE SEQUENCES OF LENGTH 3 AS CODEWORDS?

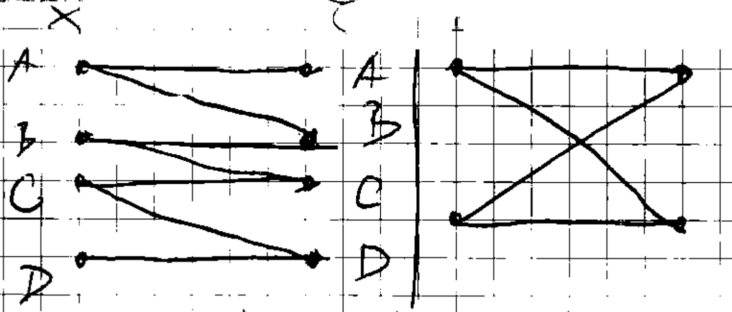
(c) NOW CONSIDER A BINARY ERASURE CHANNEL WITH ERASURE PROBABILITY 0.1. AGAIN, IF WE USED THE TWO-CODEWORD CODE 000 AND 111, RECEIVED SEQUENCES 00E, 0E0, E00, 0EE, E0E, EE0 WILL ALL BE DECODED AS 0, AND SIMILARLY, WE WOULD DECODE 11E, 1E1, E11, 1EE, E1E, EE1 AS 1.
IF WE RECEIVED THE SEQUENCE EEE, WE WOULD NOT KNOW IF IT WAS 000 OR A 111 THAT WAS SENT SO WE CHOOSE ONE OF THESE TWO AT RANDOM, AND ARE WRONG HALF THE TIME. WHAT IS THE PROBABILITY OF ERROR FOR THIS CODE OVER THE ERASURE CHANNEL

(d) WHAT IS THE PROBABILITY OF ERROR FOR CODES OF PART (a) AND (b) WHEN USED OVER THE BINARY ERASURE CHANNEL.

(a)

NOISY TYPEW.

W	x^n	Y^n	$\hat{W}$
1	000	000	1
		001	
2	010	010	
		011	2
3	100	100	
		101	3
4	110	110	4
		111	



$C = I(X;Y) = H(Y) - H(Y|X)$;

$H(Y) = ld\,26$

$H(Y|X) = \sum_x p(x)\, H(Y|X=x) =$

$= p(x=A)\cdot H(Y|X=A) + p(x=B)\cdot H(Y|X=B) + \dots$

$H(Y|X=A) = -\big[P(Y=A|X=A)\cdot ld\,P(Y=A|X=A) + P(Y=B|X=A)\; ld\,P(Y=B|X=A)\big] =$

$= +\frac{1}{2}\, ld\,2 + \frac{1}{2}\, ld\,2 = 1$

$C = I(X;Y)\big|_{max} = ld\,26 - 1 = ld\,\frac{26}{2} = ld\,13$

$P(Err|W=000) = P(010) + P(011) + P(100) + P(101) + P(110) + P(111)$

$(0.1)^3 + (0.1)(0.1)^2 + (0.1)^2\, 0.9 + (0.1)\, 0.9^2 + (0.1)^2 + 0.1\cdot 0.9$

(a) $P(Err \mid 010) = P(000) + P(001) + P(100) + P(101) + P(110) + P(111)$

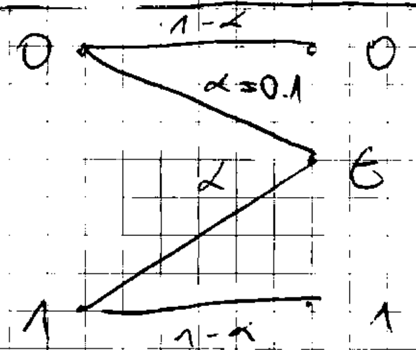
$= (0.1)(0.9)^2 + (0.1)^2 \cdot 0.9 + (0.1)^2\, 0.9 + (0.1)^3 + (0.1)(0.9)^2 + (0.1)^2 0.9$

$= (0.1)^3 + 3 \cdot (0.1)^2 \cdot 0.9 + 2\,(0.1)(0.9)^2 = 0.19$

$P(\varepsilon) = \sum_x P(x)\, P(Err \mid z = x) = \left(\frac{1}{4} \cdot 0.19\right) \cdot 4 = \mathbf{0.19}$

(b) $P(ERR \mid 000) = \sum_{i=1}^{3} P(i) = 1 - (0.9)^3 = 0.271$

$P(\varepsilon) = \sum_x P(x)\, P(Err \mid z = x) = 8 \cdot \frac{1}{8}\, 0.271 = \mathbf{0.271}$



$C = I(X, Y)\big|_{max} = H(Y) - H(Y \mid X)$

$H(Y) = P(Y=0)\, ld\, P(Y=0) + P(Y=1)\, ld\, P(Y=1) + P(Y=\epsilon)\, ld\, P(Y=\epsilon)$

$P(Y=0) = P(X=0)\, P(0 \mid 0) = \frac{1}{2}(1-\alpha)$

$P(Y=1) = \frac{1}{2}(1-\alpha)$ $\quad P(Y=\epsilon) = P(X=0)\, P(\epsilon \mid 0) + P(X=1) \cdot P(\epsilon \mid 1) = \frac{1}{2}\alpha + \frac{1}{2}\alpha = \alpha$

$\sum P(y) = \frac{1}{2}(1-\alpha) \cdot 2 + \alpha = 1 - \alpha + \alpha = 1$

$H(Y) = 2\left(\frac{1-\alpha}{2}\, ld \frac{2}{1-\alpha}\right) + \alpha\, ld \frac{1}{\alpha} = (1-\alpha)\, ld \frac{2}{1-\alpha} + \alpha\, ld \frac{1}{\alpha}$

$H(Y \mid X) = \sum_x p(x)\, H(Y \mid x)$ $\quad H(Y \mid X=0) = H(\alpha)$

$H(Y \mid X=1) = H(\alpha)$ $\quad H(Y \mid X) = 2\, \frac{1}{2}\, H(\alpha) = H(\alpha)$

$C = (1-\alpha)\, ld \frac{1}{1-\alpha} + (1-\alpha) + \alpha\, ld \frac{1}{\alpha} - H(\alpha) = \underline{1-\alpha}$

$P(\text{"1"} \mid \text{"0"}) = 1 - P(\text{"0"} \mid \text{"0"}) = 1 - [P(00\epsilon \mid 000) + P(0\epsilon 0 \mid 000) + P(\epsilon 00 \mid 000) + P(0\epsilon\epsilon \mid 000) + P(\epsilon 0 \epsilon \mid 000) + P(\epsilon\epsilon 0 \mid 000)] - \frac{P(\epsilon\epsilon\epsilon)}{2} - P(000)$

$= 1 - \left[(1-\alpha)^2 \alpha \cdot 3 + (1-\alpha)\alpha^2 \cdot 3 + \frac{\alpha^3}{2} + (1-\alpha)^3\right] = \ldots$

99a

$P(\text{"1"}|\text{"0"}) = 1 - P(\text{"0"}|\text{"0"}) = $ [illegible] $+(1-\alpha)^3$

[illegible] $= 0.2$ [illegible]

$= 1 - \left(3\alpha - 3\alpha^2 + \frac{1}{2}\alpha^3\right) =$ [illegible]

000	εεε	[illegible]	
001	εε1	εε0	01ε
010	ε1ε	ε0ε	0ε1
011	ε11	ε00	10ε
100	1εε	0εε	1ε0
101	1ε1	0ε0	ε01
110	11ε	00ε	ε10
111			
8	7	6	6

$12 + 8 + 7 = 27 = 3^3$

$P(\text{"1"}|\text{"0"}) = 1 - \left[3(1-\alpha)^2\alpha + 3(1-\alpha)\alpha^2 + \frac{\alpha^3}{2} + (1-\alpha)^3\right] = = 1\left(1 - \frac{\alpha^3}{2}\right) = \frac{\alpha^3}{2}$

$P(\alpha) = \frac{\alpha^3}{2}$ $\qquad P(0.1) = (0.1)^3/2 = 0.00050$

$P(\varepsilon) = \underbrace{P(x=\text{"0"})}_{} \cdot \underbrace{P(\text{"1"}|\text{"0"})}_{\frac{\alpha^3}{2}} + \underbrace{P(x=\text{"1"})}_{1/2} \underbrace{P(\text{"0"}|\text{"1"})}_{\frac{\alpha^3}{2}} = \frac{\alpha^3}{2}$

$P(\varepsilon)\Big|_{\alpha=0.1}^{1/2} = \frac{(0.1)^3}{2} = 0.0005$

MMV

(d) (b)' $P(\text{err}|000) = 1 - (0.9)^3 - ?$

W	x^4	y^4	$\tilde{x}$
1	000	000 00ε	
2	010	010 01ε	
3	100	100 10ε	
4	110	110 11ε	

000 → 000, 00ε

010 → 010, 01ε

$P(\text{err}|000) = P(010) + P(01\varepsilon) + P(100) + P(10\varepsilon) + P(110) + P(11\varepsilon) =$

$= 0.9^3 + 0.9^2 \cdot 0.1 +$

$P(\text{err}|000) =$ [illegible] $P(\varepsilon\varepsilon\varepsilon) + P(0\varepsilon 0) + P(0\varepsilon\varepsilon) + P(\varepsilon 0 \varepsilon) + P(\varepsilon\varepsilon 0) = 0.1^3 + 0.1 \cdot 0.9^2 + 0.1^2 \cdot 0.9 + 0.1^2 \cdot 0.9 + 0.1^2 \cdot 0.9 =$

$= 0.1^3 + 2 \cdot 0.1 \cdot 0.9^2 + 3 \cdot 0.1^2 \cdot 0.9 = 0.19$

$P(\varepsilon) = \sum_x P(x) P(\text{err}|x=x) = \frac{1}{4} \cdot 0.19 \cdot 4 = 0.19$

(d) (b)' ⇒ $P(\varepsilon) = 1 - 0.9^3 = 0.271$

JOHN HOPKINS'S UNIVERSITY SOLUTIONS

(d) (a)

	$\underline{y}$			$\hat{W}$
X_1	100	100 1E0	E00 10E	1
X_2	010	010 0E0	E10 01E	2
X_3	001	001 0E1	E01 00E	3
X_4	111	111 1E1	E11 11E	4

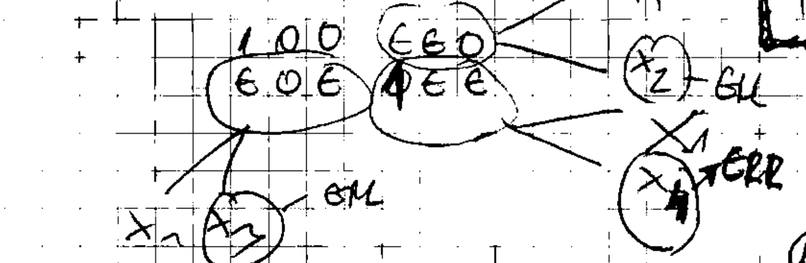
• WITH ONE ERASURE BITS RECEIVED

$$P_1(\mathcal{E}|100) = 1 - [P(E00) + P(100) + P(1E0) + P(10E)] =$$

$$= 1 - \alpha(1-\alpha)^2 - (1-\alpha)^3 - \alpha(1-\alpha)^2 - \alpha(1-\alpha)^2 = 1 - 3\alpha(1-\alpha)^2 - (1-\alpha)^3$$

NO ERROR !!!

• WITH TWO ERASURE BITS



$$P_2(\mathcal{E}|100) = \frac{(1-\alpha)\alpha^2 \cdot 3}{2}$$

$$P(\mathcal{E}|100) = p \cdot P(\mathcal{E}|EE0) + p \cdot P(\mathcal{E}|E0E) + p \cdot P(\mathcal{E}|0EE) = (1-\alpha)\alpha^2 \cdot \frac{1}{2} + \frac{1}{2}(1-\alpha)\alpha^2 + \frac{1}{2}(1-\alpha)\alpha^2 = \frac{3}{2}(1-\alpha)\alpha^2$$

$p = P(EE0) = P(E0E) = P(0EE)$

MAP

• WITH THREE ERASURE BITS RECEIVED: $P_3(\mathcal{E}|100) = \frac{3 \cdot P(EEE)}{4} =$

$$= [P_3(010|100) + P(001|100) + P(111|100)] = \frac{3}{4} \cdot P(EEE) \Rightarrow$$

$$P_3(\mathcal{E}) = \sum_x P(x) P_3(\mathcal{E}|x) = \frac{1}{4} \cdot P_3(\mathcal{E}|\cdot) \cdot 4 = \frac{3}{4} P(EEE) = \frac{3}{4}\alpha^3$$

$$\boxed{P_3(\mathcal{E}) = \frac{3\alpha^2}{4}}$$

$$P(\mathcal{E}) = P_1(\mathcal{E}) + P_2(\mathcal{E}) + P_3(\mathcal{E}) = \frac{3(1-\alpha)\alpha^2}{2} + \frac{3}{4}\alpha^3$$

(d) (b)

000	100
001	101
010	110
011	111

• ONE ERASED BIT

$$P_1(\mathcal{E}|000) = P(00E) + P(0E0) + P(E00)$$

"0" V "1" WITH $\frac{1}{2}$ PROBABILITY

"0" V "2"

"0" V "4" } WITH $\frac{1}{2}$ PROBABILITY

$$P(\mathcal{E}|000) = \frac{\alpha(1-\alpha)^2 \cdot 3}{2}$$

$$P(\mathcal{E}) = \frac{1}{8} 8 P(\mathcal{E}|000) \quad 2$$

• TWO ERASED BITS

$$P_2(\mathcal{E}|000) = P(0EE) + P(E0E) + P(EE0)$$

$\frac{3}{4}(1-\alpha)\alpha^2 \quad \frac{3}{4}(1-\alpha)\alpha^2 \quad \frac{3}{4}(1-\alpha)\alpha^2$

$$P_2(\mathcal{E}|000) = \frac{3}{4} 3(1-\alpha)^2 \cdot \alpha$$

• THREE ERASED BITS

$$P_3(\mathcal{E}|000) = \frac{7}{8} \alpha^3$$

$$P(\mathcal{E}) = \frac{3\alpha(1-\alpha)^2}{2} + \frac{3}{4} \cdot 3(1-\alpha)\alpha^2 + \frac{7}{8}\alpha^3 = 0.1426$$

7.18 CHANNEL CAPACITY ... CAPACITY OF FOLLOWING CHANNELS

(a) $\mathcal{X} = \mathcal{Y} = \{0, ...$

$$p(y|x) = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

(b) $\mathcal{X} = \mathcal{Y} = \{0, 1, 2\}$

$$p(y|x) = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

(c) $\mathcal{X} = \mathcal{Y} = \{0, 1, 2, 3\}$

$$p(y|x) = \begin{bmatrix} p & 1-p & 0 & 0 \\ 1-p & p & 0 & 0 \\ 0 & 0 & q & 1-q \\ 0 & 0 & 1-q & q \end{bmatrix}$$

(a) $C = \max I(x;y) = H(Y) - H(Y|X) = \text{ld}\,3 - H(Y|X)$

$H(Y|X) = \sum p(x) H(Y|X=x) = \frac{1}{3} P(0) H(Y|X=0) + \frac{1}{3} P(1) H(Y|X=1) + \frac{1}{3} P(2) H(Y|X=2) = \frac{1}{3}[\,3 \cdot \text{ld}\,3\,] = \text{ld}\,3$

$C = \text{ld}\,3 - \text{ld}\,3 = 0$

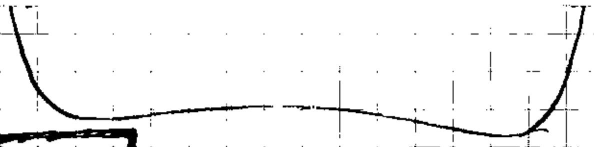
(b) $H(Y|X) = \frac{1}{3}\left[\left(\frac{1}{2}\,\text{ld}\,2\right) \cdot 2\right] \cdot 3 = \text{ld}\,2$

$C = \text{ld}\,3 - \text{ld}\,2 = 0.58496$ bits

(c) $H(Y|X) = \frac{1}{4}\left[p\,\text{ld}\frac{1}{p} + (1-p)\,\text{ld}\frac{1}{1-p}\right] \cdot 2 + \frac{1}{4}\left[q\,\text{ld}\frac{1}{q} + (1-q)\,\text{ld}\frac{1}{1-q}\right] \cdot 2$

$H(Y|X) = \frac{1}{2} H(p) + \frac{1}{2} H(q)$ $\qquad C = \text{ld}\,4 - \frac{1}{2}[H(p) + H(q)]$

$C = 2 - \frac{1}{2}[H(p) + H(q)]$



JOHN HOPKINS SOLUTION (102)

$$H(\tau|X) = \alpha H(\gamma) + (1-\alpha) H(q)$$

$$H(\tau) = \gamma(\tau_1)\, ld \frac{1}{\gamma(\tau_1)} + \gamma(\tau_2)\, ld \frac{1}{\gamma(\tau_2)} + \gamma(\tau_3)\, ld \frac{1}{\gamma(\tau_3)} +$$

$$+ \gamma(\tau_4)\, ld \frac{1}{\gamma(\tau_4)} = ?$$

$$\gamma(\tau_1) + \gamma(\tau_2) = \alpha \qquad \gamma(\tau_1) = \rho \qquad \gamma(\tau_2) = \alpha - \rho$$

$$\rho\, ld \frac{1}{\rho} + (\alpha-\rho)\, ld \frac{1}{\alpha-\rho} = \rho\, ld \frac{1}{\rho} + \alpha\, ld \frac{1}{\alpha-\rho}$$

$$- \rho\, ld \frac{1}{\alpha-\rho} = \rho\, ld \frac{\alpha-\rho}{\rho} + \alpha\, ld \frac{1}{\alpha-\rho} =$$

$$= \rho\, ld \frac{\alpha-\rho}{\rho} + \alpha\, ld \frac{\rho}{\alpha-\rho} - \alpha\, ld\, \rho =$$

$$= \rho\, ld \frac{\alpha-\rho}{\rho} - \alpha\, ld \frac{\alpha-\rho}{\rho} - \alpha\, ld\, \rho = (\rho-\alpha)\, ld \frac{\alpha-\rho}{\rho} + \alpha\, ld \frac{1}{\rho}$$

$$C = \max_{?} I(X;\tau)$$

$$H(\tau) = \frac{\alpha}{2}\, ld \frac{2}{\alpha} + \frac{\alpha}{2}\, ld \frac{2}{\alpha} + 2\frac{1-\alpha}{2}\, ld \frac{2}{1-\alpha} =$$

$$= \alpha\, ld \frac{2}{\alpha} + (1-\alpha)\, ld \frac{2}{1-\alpha}$$

$$C = \max I(X;\tau) = \alpha\, ld \frac{2}{\alpha} + (1-\alpha)\, ld \frac{2}{1-\alpha} - \alpha H(\gamma)$$

$$- (1-\alpha) H(q) = \alpha\, ld\, 2 + \alpha\, ld \frac{1}{\alpha} + (1-\alpha)\, ld \frac{1}{1-\alpha} + (1-\alpha)\, ld\, 2$$

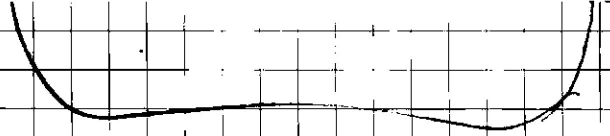
$$- \alpha H(\gamma) - (1-\alpha) H(q) = 1 + H(\alpha) - \alpha H(\gamma) - (1-\alpha) H(q)$$

$$\frac{dC}{d\alpha} = \frac{d}{d\alpha}\left[\alpha\, ld \frac{1}{\alpha} + (1-\alpha)\, ld \frac{1}{1-\alpha}\right] - H(\gamma) + H(q) =$$

$$= ld \frac{1}{\alpha} + \alpha \frac{1/\ln 2}{\alpha} + - ld \frac{1}{1-\alpha} + (1-\alpha) \frac{1/\ln 2}{1-\alpha} - H(\gamma) + H(q)$$

$$= ld \frac{1}{\alpha} - \frac{2}{\ln 2} - ld \frac{1}{1-\alpha} - H(\gamma) + H(q)$$

$$(*) = \frac{d}{d\alpha}\left[\alpha\, ld\, \alpha + (1-\alpha)\, ld (1-\alpha)\right] = -\left[ld\, \alpha + \alpha \frac{1}{\alpha \cdot \ln 2} - 1 \cdot ld(1-\alpha) + (1-\alpha) \frac{1}{(1-\alpha)\ln 2}\right]$$



$$(*) = -\left[ld\,\alpha - ld(1-\alpha) + \frac{2}{\ln 2} \right] = -\left[ld\left(\frac{\alpha}{1-\alpha}\right) + \frac{2}{\ln 2} \right] =$$

$$= -\frac{\ln\left(\frac{\alpha}{1-\alpha}\right) + 2}{\ln 2}$$

$$\frac{d}{d\alpha}\left[\alpha\, ld\,\alpha - \alpha\, ld(1-\alpha) + ld(1-\alpha) \right] =$$

$$= ld\,\alpha + \alpha \frac{1}{\alpha \ln 2} - ld(1-\alpha) - \alpha \frac{(-1)}{(1-\alpha)\ln 2} + \frac{-1}{(1-\alpha)\ln 2}$$

$$= ld\,\alpha + \frac{1}{\ln 2} - ld(1-\alpha) + \frac{\alpha}{(1-\alpha)\ln 2} - \frac{1}{(1-\alpha)\ln 2} =$$

$$= ld\,\alpha + \frac{1}{\ln 2} - ld(1-\alpha) - \frac{(1-\alpha)}{(1-\alpha)\ln 2} = ld \frac{\alpha}{1-\alpha}$$

$$(*) = -ld\frac{\alpha}{(1-\alpha)} = ld\frac{(1-\alpha)}{\alpha}$$

$$\frac{dC}{d\alpha} = ld\frac{(1-\alpha)}{\alpha} - H(\gamma) + H(2) = 0$$

$$ld\frac{1-\alpha}{\alpha} = H(\gamma) - H(2) \qquad \frac{1-\alpha}{\alpha} = 2^{H(\gamma)-H(2)}$$

$$1 - \alpha = \alpha\, 2^{H(\gamma)-H(2)} \qquad 1 = \alpha\left(1 + 2^{H(\gamma)-H(2)}\right)$$

$$\alpha = \frac{1}{1 + 2^{H(\gamma)-H(2)}} = \frac{1}{1 + \frac{2^{H(\gamma)}}{2^{H(2)}}} = \frac{2^{H(2)}}{2^{H(\gamma)} + 2^{H(2)}}$$

$$C = 1 - \alpha\, ld\,\alpha - (1-\alpha)\, ld(1-\alpha) - \frac{2^{H_2}}{2^{H_\gamma} + 2^{H_2}} H(\gamma) -$$

$$- \frac{2^{H_\gamma}}{2^{H_\gamma} + 2^{H_2}} H(2) = 1 - \frac{2^{H_2}}{2^{H_\gamma} + 2^{H_2}}\, ld \frac{2^{H_2}}{2^{H_\gamma} + 2^{H_2}} - \frac{2^{H_\gamma}}{2^{H_\gamma} + 2^{H_2}}\, ld \frac{2^{H_\gamma}}{2^{H_\gamma} + 2^{H_2}}$$

$$- \frac{2^{H_2} H(\gamma) + 2^{H_\gamma} H(2)}{2^{H_\gamma} + 2^{H_2}} = 1 - \frac{2^{H_2}}{2^{H_\gamma} + 2^{H_2}}\, ld\, 2^{H_2} - \frac{2^{H_\gamma}}{2^{H_\gamma} + 2^{H_2}}\, ld\, 2^{H_\gamma}$$

$$+ \frac{2^{H_2} + 2^{H_\gamma}}{2^{H_\gamma} + 2^{H_2}}\, ld\left(2^{H_\gamma} + 2^{H_2}\right) - \frac{2^{H_2} H_\gamma + 2^{H_\gamma} H_2}{2^{H_\gamma} + 2^{H_2}} =$$

$$C = 1 - \frac{H_q 2^{H_2} + H_\gamma 2^{H_p}}{2^{H_p} + 2^{H_2}} + ld\left(2^{H_1} + 2^{H_2}\right) -$$

$$- \frac{H_p 2^{H_2} + H_q 2^{H_p}}{2^{H_p} + 2^{H_q}} = 1 - H_p - H_q + ld\left(2^{H_p} + 2^{H_2}\right)$$

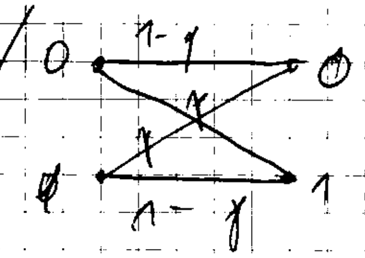
$$C = ld\, 2 + ld\, 2^{-H_p}\, 2^{-H_2} + ld\left(2^{H_1} + 2^{H_2}\right) =$$

$$= ld\left[2^{1-H_p-H_2} \cdot \left(2^{H_p} + 2^{H_2}\right)\right] = ld\left[2^{1-H_2} + 2^{1-H_p}\right]$$

$$2^C = 2^{1-H_2} + 2^{1-H_\gamma}$$

$$\boxed{2^C = 2^{C_1} + 2^{C_2}}$$

$$C_1 = 1 - H(\gamma) \qquad C_2 = 1 - H(q)$$



$$C = 1 - H(Y|X) = 1 - P(0)\left[H(q)\right] -$$
$$- P(1)\left[H(q)\right] = 1 - H(\gamma)$$

KAPACITETOT NA DVA NEZAVISNI KANALI E:

$$C = ld\left[2^{C_1} + 2^{C_2}\right]$$

IF $C_1 = C_2 = 1$
$C = 2$

IF: $C_1 = 3 \quad C_2 = 3 \qquad C = ld\left[8+8\right] = ld\,16 = 4$

PROBLEM 7.20 CHANNEL WITH TWO INDEPENDENT LOOKS AT X. LET Z_1 AND Z_2 BE CONDITIONALLY INDEPENDENT AND CONDITIONALLY IDENTICALLY DISTRIBUTED GIVEN X.

(a) SHOW THAT $I(X; Z_1, Z_2) = 2I(X; Z_1) - I(Z_1, Z_2)$

(b) CONCLUDE THAT THE CAPACITY OF THE CHANNEL

X → [ ] → $(Z_1 Z_2)$

IS LESS THAN TWICE THE CAPACITY OF THE CHANNEL

X → [ ] → Z_1

(a) $I(X;Y_1,Y_2) = I(X;Y_1) + I(X;Y_2|Y_1)$

$I(X;Y_2|Y_1) = H(X|Y_1) - H(X|Y_1Y_2) =$

$= H(Y_2|Y_1) - \underbrace{H(Y_2|Y_1X)}_{= H(Y_2|X)} = \underline{H(Y_2|Y_1) - H(Y_2|X)}$ INDEPENDENT GIVEN x

$= \underbrace{-H(Y_2) + H(Y_2|Y_1)}_{-I(Y_1;Y_2)} + \underbrace{H(Y_2) - H(Y_2|X)}_{I(X;Y_2)}$

$I(X;Y_1,Y_2) = \underbrace{I(X;Y_1)} + \underbrace{I(X;Y_2)} - I(Y_1;Y_2)$

$Y_1 Y_2$ ARE IDENTICALLY DISTRIBUTED

$I(X;Y_1) = H(Y_1) - H(Y_1|X)$ $\quad H(Y_1|X) = H(Y_2|X)$

$I(X;Y_2) = H(Y_2) - H(Y_2|X)$ $\quad \underline{H(Y_1) = H(Y_2)}$ IDENTICALLY DISTRIB.

$I(X;Y_1) = I(X;Y_2)$

$\boxed{I(X;Y_1,Y_2) = 2I(X;Y_1) - I(Y_1;Y_2)}$

(b) $C_2 = \max_{p(x)} I(X;Y_1Y_2) = \underbrace{2\max\left(I(X;Y_1)\right)}_{2\cdot C_1} - \underbrace{I(Y_1;Y_2)}_{\text{CONSTANT OF } x}$

$\Rightarrow C_2 = 2\cdot C_1 - I(Y_1;Y_2)$ $\quad \boxed{C_2 \leq 2C_1}$

$-I(Y_1;Y_2)$

CONFIRM OF ELLIN'S SOLUTION

$I(X;Y_1Y_2) = H(Y_1Y_2) - H(Y_1Y_2|X) = H(Y_1) + H(Y_2) -$
$- H(Y_1|X) - H(Y_2|X) = I(X;Y_1) + I(X;Y_1) - I(Y_1;Y_2)$

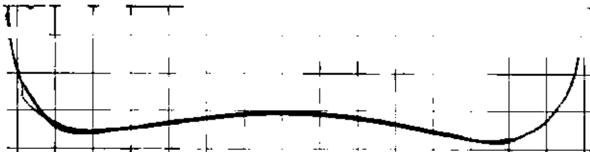
$H(Y_1Y_2) = H(Y_1) + H(Y_2|Y_1) = H(Y_1) - H(Y_2) + H(Y_2|Y_1) + H(Y_2)$
$= H(Y_1) + H(Y_2) - I(Y_1;Y_2)$

PROBLEM 7.21 TALL, FAT PEOPLE. SUPPOSE THAT THE AVERAGE HEIGHT OF PEOPLE IN A ROOM IS 5 FEET. SUPPOSE THAT THE AVERAGE WEIGHT IS 100 lb.

(a) ARGUE THAT NO MORE THAN ONE-THIRD OF THE POPULATION IS 15 FEET TALL.

(b) FIND UPPER BOUND ON THE FRACTION OF 300 lb 10-footers IN THE ROOM.

$\bar{h} = 5$ feet $= 152$ cm $\qquad \bar{w} = 100$ lb $= 45$ kg



1054

(a) $\overline{h} = 5$ $\quad h\left(\frac{n}{3}\right) = 15$ feet

$n_3 = \frac{n}{3}$ $\quad \frac{1}{n}\left(h_1 + h_2 + \dots + h_n\right) = 5$

$\sum_{i=1}^{n} h_i = 5n$

$\frac{1}{n_3} \sum_{i=1}^{n/3} h_i \leq 15$ $\quad 15 \cdot \frac{n}{3} = 5n$

IF $n_3 = \left(\frac{n}{3} + 1\right) \Rightarrow \left(\frac{n}{3} + 1\right) \cdot 15 = 5n + 15$

OVA E PODOLGO OD VKUPNATA VISINA OD SITE LUĐE VO SOBATA PA ZATOA VO SOBATA NE MOŽE DA IMA POVEĆE OD $\frac{n}{3}$ LUĐE SO VISINA POGOLEMA OD 15 feet-1.

(b) $\overline{w} = 100$ lb $\quad \frac{1}{n}\left(w_1 + w_2 + \dots + w_n\right) = 100$

$\sum_{i=1}^{n} w_i = 100n$

$n_2 \cdot 300 \leq 100n$ $\quad n_2 \leq \frac{n}{3}$

NE POVEĆE OD $n/3$ LUĐE IMAAT 300 lb.

$\lim_{n\to\infty} H(h_1, h_2, \dots h_n) = H(h) = \lim H(h_n | h_1^{n-1})$

$H(X) = \lim_{n\to\infty} \frac{H(X_1, X_2, \dots X_n)}{n} = \lim_{n\to\infty} H(X_n | X_1 \dots X_{n-1})$

$|A_\varepsilon^{(n)}| \leq 2^{n(H(h,w)+\varepsilon)}$ $\quad H(h,w) = H(H,W)$

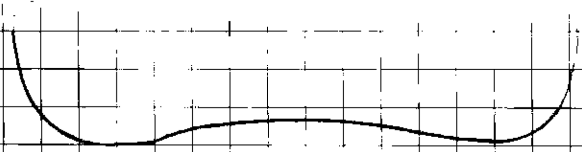
$2^{-n(I(X;Z)+\varepsilon)} \leq Pr\left[(\tilde{x}_n, \tilde{z}_n) \in A_\varepsilon^{(n)}\right] \leq 2^{-n(I(X;Z)-\varepsilon)}$

$Pr\left[(\tilde{x}_n, \tilde{z}_n) \in A_\varepsilon^{(n)}\right] \approx 2^{-nI(X;Z)}$

$H(h,w) = H(h) + H(w|h) = H(h) + H(w) \leq \log n + \log n = \log 2n$

$H(h,w) = 1 + \log n$

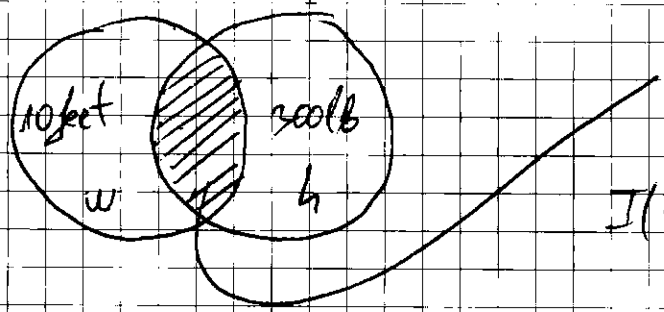
BIDEJKI NA RAZLIČNI MOŽE VISINI NA LUĐE KAKO I TEŽINI NA LUĐE.



(106a)

$$u_1 \cdot 10(\text{foot}) \leq 5 \cdot u \qquad u_1 \leq \frac{5u}{10} = \frac{u}{2}$$

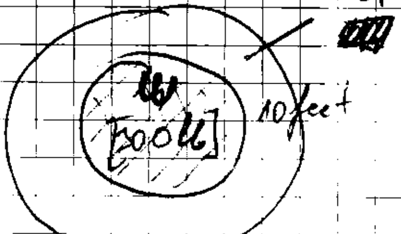
$$\boxed{u_1 \leq \frac{u}{2}}$$ → VO SOBATA NE MOŽE DA IMA POVEĆE OD TOČA LUĐE ŠTO SE VISOKI OD ~~10 feet~~ = 3m



$$\boxed{I(w;h)}$$

$$I(w;h) = H(w) - H(w|h)$$

UPPER BOUND E SITE ŠTO SE TEŠKI 300 lb DA BIDAT VISOKI 10 feet - i.e.



VO OVOJ SLUČAJ $H(w|h) = 0$

$$I(w;h) = H(w)$$

- UPPER BOUND OF 300 lb - 10 footers is equal to $max(u_2)$ i.e.

$$\boxed{U.b = [\text{redacted}] = \frac{u}{3}}$$

(106a)

STACKOVERFLOW SOLUTIONS

MARKOV INEQUALITY $\quad \boxed{P(X \geq c) \leq \frac{u_x}{c}}$

$$P(X > 15) \leq \frac{5}{15} \qquad P(X \geq c) = \sum_{x:\, x \geq c} P(x) \leq \sum_x P(x)$$

- PROOF OF MARKOV'S INEQUALITY

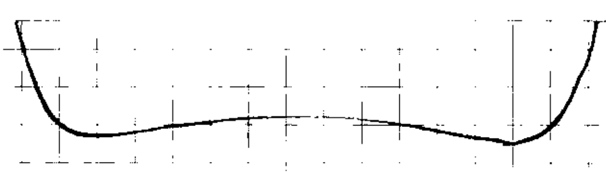
$I_E = 1$ IF E OCCURS $\quad I_E = 0$ OTHERWISE

$$I(x \geq a) = \begin{cases} 1 & x \geq a \\ 0 & x < a \end{cases}$$

$$a I(x \geq a) \leq X$$

IF $x < a$: $a \cdot 0 \leq X$

IF $x \geq a$: $\boxed{a \cdot 1 \leq X}$



(106a)

• $a \cdot I(X \geq a) \leq X \quad / E[\,] \qquad E[a I(X\geq a)] \leq E[X]$

$E[a\, I(X \geq a)] = a[1 \cdot P(X \geq a) + 0 \cdot P(X < a)] = aP(X \geq a)$

$\boxed{a P(X \geq a) \leq E[X]} \qquad \boxed{P(X \geq a) \leq \frac{E[X]}{a}}$

• CHEBISHEV INEQUALITY

$P_r[(Y-\mu)^2 \geq \varepsilon^2] \leq P_r((Y-\mu)^2 \geq \varepsilon^2) \leq \frac{E[(Y-\mu)^2]}{\varepsilon^2}$

$= \frac{\sigma^2}{\varepsilon^2}$ ↑ USE MARKOV'S INEQUALITY !!!

(б) $P(X > 300lb) = \frac{100lb}{300} = \frac{1}{3}$

(в) $P(Y \geq 10f) = \frac{5}{10} = \frac{1}{2}$

ТАКА ГИ ДОБИВАМ ИСТИТЕ ГРАНИЦИ ШТО ГИ ДОБИВ САМОСТОЈНО

– СОГЛАСНО ЗАКОНОТ ЗА ГОЛЕМИ БРОЕВИ: СРЕДНАТА ВРЕДНОСТ НА ПРИМЕРОКОТ ТЕЖНЕЕ КОН СТВАРНАТА ОЧЕКУВАНА ВРЕДНОСТ

ОД (в) СЛЕДУВА ДЕКА (ЗА ГОЛЕМО n)

• $\frac{n}{3}$ ЛУЃЕ ЌЕ БИДАТ СО ТЕЖИНА ПОГОЛЕМА ОД 300 lb

• $\frac{n}{2}$ ЛУЃЕ ЌЕ БИДАТ СО ВИСИНА > ОД 10f.

$P(X > 300lb, Y \geq 10f) = $ / X, Y СЕ СТАТИСТИЧКИ НЕЗАВИСНИ / $=$

$= P(X \geq 300)\, P(Y \geq 10) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$

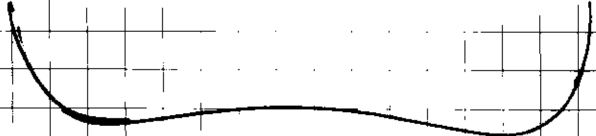
ОТТУКА: $\frac{n}{6}$ ЛУЃЕ СЕ ПОТЕШКИ ОД 300lb И ПОВИСОКИ ОД 10f.

ТУКА ЈА МОДЕЛИРАМ СО МОНТЕ-КАРЛО СИМУЛАЦИЈА

PROBLEM 7.22 CAN SIGNAL ALTERNATIVES LOWER CAPACITY? SHOW THAT ADDING A ROW TO A CHANNEL TRANSMISSION MATRIX DOESN'T DECREASE CAPACITY. Π^T

$\begin{bmatrix} p(y_1) \\ p(y_2) \end{bmatrix}_{2\times 1} = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \end{bmatrix}_{2\times 1} \qquad C = \max_{p(x)} I(X;Y)$

$I(X;Y) = H(Y) - H(Y|X)$



10/9

$p(y_1) = p_{11}\, p(x_1) + p_{21}\, p(x_2)$
$p(y_2) = p_{12}\, p(x_1) + p_{22}\, p(x_2)$

$-H(Y) = p(y_1)\,\mathrm{ld}\, p(y_1) + p(y_2)\,\mathrm{ld}\, p(y_2)$

$$\Pi = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \qquad \Pi_1 = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \\ p_{31} & p_{32} \end{bmatrix}$$

$$\begin{bmatrix} p(y_1) \\ p(y_2) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \end{bmatrix} \qquad \sum_{i=1}^{N} p(y_i \mid x_1) = 1$$

$\emptyset$

$p(y_1) = p_{11}\, p(x_1) + p_{21}\, p(x_2) + p_{31}\, p(x_3)$

$C = \mathrm{ld}\, m - H(Y \mid X)$ → OVA E VAŽNO!!!

$H(Y \mid X) = \sum_{x} p(x) \cdot H(Y \mid X = x)$

$-H(Y \mid X = x_1) = p_{11}\,\mathrm{ld}\, p_{11} + p_{12}\,\mathrm{ld}\, p_{12}$

$= \sum_{i=1}^{2} p_{1i}\,\mathrm{ld}\, p_{1i}$

$H(Y \mid X) = \sum_{i=1}^{2} p_{1i}\,\mathrm{ld}\, p_{1i} + \sum_{i=1}^{2} p_{2i}\,\mathrm{ld}\, p_{2i}$

$H(Y \mid X) = \sum_{i=1}^{2} \sum_{j=1}^{2} p_{ji}\,\mathrm{ld}\, p_{ji}$; 2×2 TRANSITION MATRIX.

$H(Y \mid X = x_2) = p_{21}\,(\mathrm{ld}\, p_{21}) + p_{22}\,(\mathrm{ld}\, p_{22})$
$H(Y \mid X = x_3) = p_{31}\,\mathrm{ld}\, p_{31} + p_{32}\,\mathrm{ld}\, p_{32}$

- TRANSITION PROBABILITY MATRIX FOR GRADUAL CHANGE

$$\Pi = \begin{bmatrix} 1-f & f & 0 \\ 0 & f & 1-f \end{bmatrix} \qquad Y = \Pi^T X \qquad \begin{bmatrix} p(y_1) \\ p(y_2) \\ p(y_3) \end{bmatrix} = \begin{bmatrix} 1-f & 0 \\ f & f \\ 0 & 1-f \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \end{bmatrix}$$

(1079)

$$z = \Pi^T x$$

$$\begin{bmatrix} p(z_1) \\ p(z_2) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \\ p_{31} & p_{32} \end{bmatrix}^T \begin{bmatrix} p(x_1) \\ p(x_2) \\ 0 \end{bmatrix} =$$

$0 \rightarrow p(x_3)$

$$= \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ 0 \end{bmatrix}$$

$$p(z_1) = p_{11}\, p(x_1) + p_{21}\, p(x_2)$$
$$p(z_2) = p_{12}\, p(x_1) + p_{22}\, p(x_2)$$

$$H_3(z) = H_2(z) = -\sum_{i=1}^{2} p(z_i) \operatorname{ld} p(z_i)$$

$$C = \max I(x; z) = H(z) - H(z|x)$$

$$H(z|x) = \sum_{i=1}^{3} p(x_i)\, H(z|x_i) = p(x_1) \cdot H(z|x_1) + p(x_2) \cdot$$
$$H(z|x_2) + \underbrace{p(x_3)}_{0} H(z|x_3) = \sum_{i=1}^{2} p(x_i)\, H(z|x_i)$$

$$H(z|x_1) = p_{11} \operatorname{ld}(p_{11}) + p_{12} \operatorname{ld} p_{12}$$

$$H(z|x_2) = p_{21} \operatorname{ld} p_{21} + p_{22} \operatorname{ld} p_{22}$$

VO GENERALEN SLUČAJ:

$$H(z) = \sum_{i=1} p(z_i) \operatorname{ld} p(z_i)$$

$$\Pi = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ p_{m1} & p_{m2} & \cdots & p_{mr} \\ p_{m+1,1} & p_{m+1,2} & \cdots & p_{m+1,r} \end{bmatrix}$$

NOVIOT RED !!!

$$p(z_1) = \sum_{i=1}^{m+1} p_{i1} \cdot p(x_i) = \sum_{i=1}^{m} p_{i1}\, p(x_i) + p_{m+1}\, \underbrace{p(x_{m+1})}_{0}$$

$$p(z_1) = \sum_{i=1}^{m} p_{i1} \cdot p(x_i) \Big\}$$ ISTA VAKO ZA KODOT NA NEMODULIRANATA

$$H(z|x) = \sum_{i=1}^{m+1} p(x_i)\, H(z|x_i) = \sum_{i=1}^{m} p(x_i)\, H(z|x_i) +$$
$$+ \underbrace{p(x_{m+1})}_{0} \cdot H(z|x_{m+1}) = \sum_{i=1}^{m} p(x_i)\, H(z|x_i) \Big\}$$ ISTO VAKO ZA NEMODIFICIRANATA MATRICA

$$H(z|x_1) = p_{11} \operatorname{ld} p_{11} + p_{12} \operatorname{ld} p_{12} + \cdots + p_{1r} \operatorname{ld} p_{1r} \Big\}$$ TRANZICIONITE ENTROPII NE SE MENUVAAT.

PROBLEM 7.23 BINARY MULTIPLIER CHANNEL.

(a) CONSIDER THE CHANNEL $Y = XZ$, WHERE X AND Z ARE INDEPENDENT BINARY RANDOM VARIABLES THAT TAKE ON VALUES 0 AND 1. Z IS BERNOULLI(α) [$P(Z=1)=\alpha$]

(b) NOW SUPPOSE THAT THE RECEIVER CAN OBSERVE Z AS WELL AS Y. WHAT IS THE CAPACITY

→ FIND THE CAPACITY OF THIS CHANNEL AND THE MAXIMIZING DISTRIBUTION ON X.

(a)

X \ Z	0	1
0	0	0
1	0	1

$$Y = \begin{cases} 1 & x=1, z=1 \\ 0 & \text{OTHERWISE} \end{cases}$$

$$p(Y) = \{ p(X=1)P(Z=1),\ P(X=1)P(Z=0) + P(X=0)\cdot P(Z=1) + P(X=0)\cdot P(Z=0) \}$$

X \ Y	0	1
0	1	0
1	1-α	α

P(Y\|X)

$$P(Y=0|X=0) = 1$$
$$P(Y=1|X=1) = P(Z=1) = \alpha$$
$$P(Y=0|X=1) = P(Z=0) = 1-\alpha$$

$$C = \left[\underbrace{H(Y)}_{1} - H(Y|X)\right]\Big|_{max} = 1 - H(Y|X) = 1 - p(X=0)H(Y|X=0) - p(X=1)H(Y|X=1) = 1 - \frac{1}{2}0 - H(\alpha)$$

$$\boxed{C = 1 - H(\alpha)}$$

(b)

X \ Y	0	1
0	1	0
1	0	1

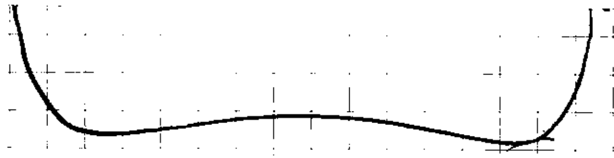
$$\boxed{C = 1 - \frac{1}{2}0 - \frac{1}{2}\cdot 0 = 1}$$

Berkley Solution

$$I(X;Y) = H(Y) - \overbrace{p(X=0)}^{1-\beta} H(Y|X=0) - p(X=1)H(Y|X=1) = H(Y) - (1-\beta)\cdot 0 - \beta H(\alpha) = H(Y) - \beta H(\alpha)$$

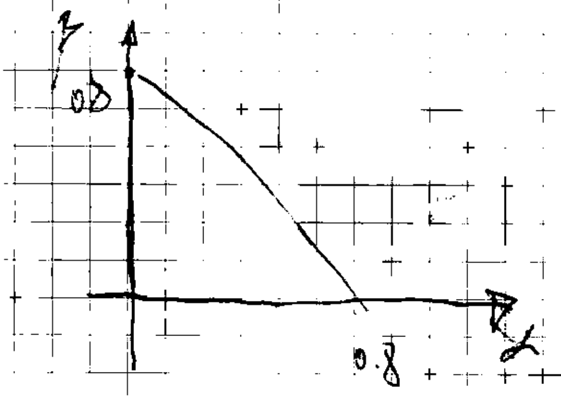
$$p(Y) = \{ \beta\alpha,\ \beta(1-\alpha) + (1-\beta)\alpha + (1-\beta)(1-\alpha) \} =$$
$$= \{ \beta\alpha,\ \beta - \beta\alpha + \alpha - \beta\alpha + 1 - \alpha - \beta + \alpha\beta \} = \{ \alpha\beta,\ 1-\alpha\beta \}$$

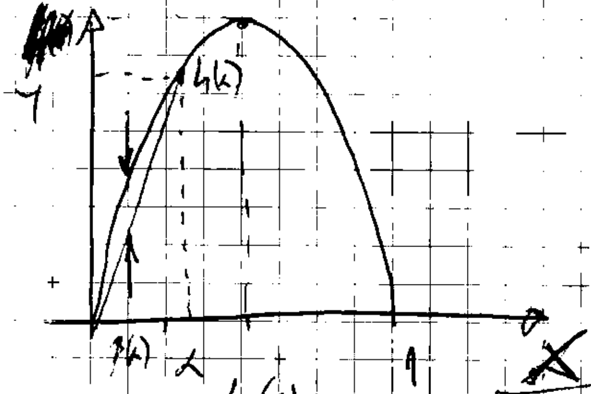
$$H(Y) = \alpha\beta \log \alpha\beta + (1-\alpha\beta) \log (1-\alpha\beta)$$



$I(X;Z) = H(Z) - \beta \cdot H(\alpha)$ $\quad H(Z) = H(\alpha\beta)$

$I(X;Z) = H(\alpha\beta) - \beta \cdot H(\alpha)$





$y = \frac{h(\alpha)}{\alpha} \cdot x$ $\quad h'(\beta(\alpha)\cdot\alpha) = \frac{h(\alpha)}{\alpha}$

$$C = h(\beta(\alpha)\cdot\alpha) - \beta(\alpha)\cdot h(\alpha)$$

(6)

X \ YZ	00	01	10	11
0	1-α	α	0	0
1	1-α	0	0	α

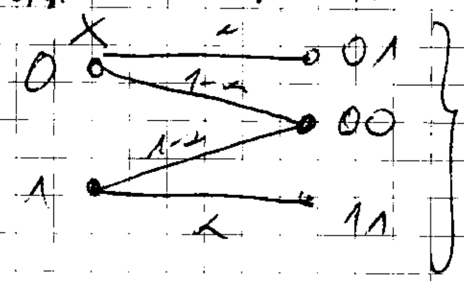
$Z = X \cdot \tilde{Z}$ $\quad p(\tilde{Z}) = \{\alpha$ "1", $1-\alpha$ "0"$\}$

$P(00|1) = P(Z=0) = 1-\alpha$

$P(11|1) = \alpha$

X \ YZ	00	01	11
0	1-α	α	0
1	1-α	0	1

NEMA VAŽAN IZLEZEN SYMBOL



ERASURE CHANNEL

$C = (1-\alpha')H(q) =$

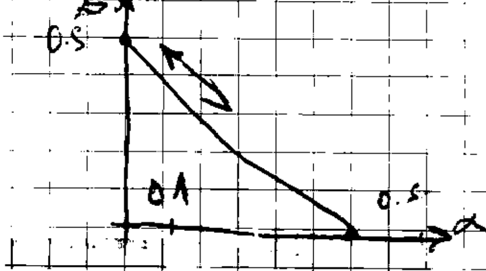
$= 1-\alpha' = / \alpha' = 1-\alpha /$

$= 1-1+\alpha = \alpha$

$C = \alpha$

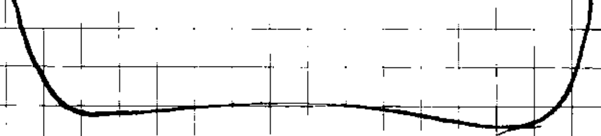
WE HAVE DISCOVERED THAT $\alpha \geq H(\beta(\alpha)\alpha) - \beta(\alpha)\cdot H(\alpha)$

- OD GRAFIKOT VO MATLAB ZA $\beta = \frac{1}{2}$ SE DOBIVA DEKA KAPACITETOT E MAKSIMALEN ZA $\alpha = 0.4$
- AKO GO KORISTAŠ: $f(x,y) = xy\,ld(xy) + (1-xy)\,ld(xy)$

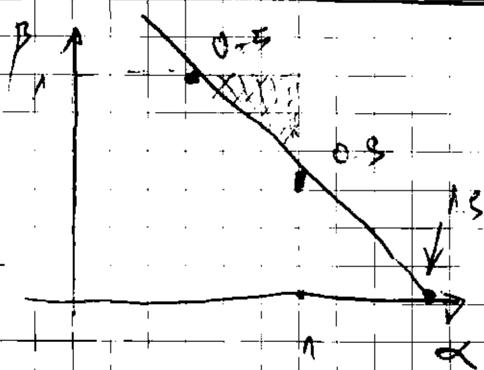


$y = -x + \frac{1}{2}$

$y = -\frac{1}{10} + \frac{1}{2} = \frac{-1+5}{10} = \frac{4}{10} = \frac{2}{5}$



122a



② $y = -x + 1.5$ → NA OVAA LINIJA LEŽAT α & β

$x = 1 \quad y = 0.5$

$x = 0.5 \quad y = 1$

ZA BILO KOI $y, x \in 0-1$ PRESMETANI SO ②

$H(\alpha\beta) = 1$ T.E MAKSIMALEN

$\beta = -\alpha + 1.5$

$C = 1 - p(\alpha)\, H(\alpha)$

IF $\alpha = \frac{1}{2} \Rightarrow p = 1 \quad C = 1 - \underbrace{H(\alpha)}_{1}$ } BINARY SYMETRIC CHANNEL

$C = 0$

IF $\beta = 1/2 \qquad \alpha = +1.5 - 1 = 0.5$

$\beta \in 0.5 - 1$
$\alpha \in 0.5 - 1$

OVIE VREDNOSTI VODAT KON MAKSIMALEN $H(\alpha, \beta)$

→ $C = 1 - \frac{1}{2} H(\frac{1}{2}) = 1 - \frac{1}{2} = \frac{1}{2}$

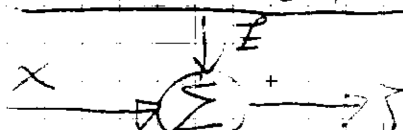
AKO $Z = 1$ CELO VREME T.E $\alpha = 1 \Rightarrow p = 0.5$

$C = 1 - \frac{1}{2} \cdot 0 = 1$ → VO OVOJ SLUČAJ SE POSTIGNUVA MAKSIMALEN KAPACITET

KAKO ŠTO SE ZGOLEMUVA α OD 0.5 → 1 TAKA CELO VREME SE ZGOLEMUVA KAPACITETOT ZA KANALOT A NEGOVIOT MAKSIMUM SLUČUVA ZA $p = 1/2$

T.E. ZA UNIFORMNA RASPREDELBA NA NEKOLKUTE SIMBOLI (VIDI SLIKA VO MALIOT)

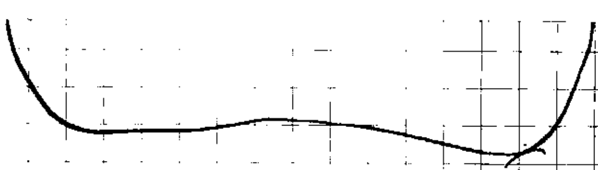
PROBLEM 7.24 UNUSE CHANNELS CONSIDER THE CHANNEL



$X \in \{0, 1, 2, 3\}$ WHICH $Y = X + Z$ AND

Z IS UNIFORMLY DISTRIBUTED OVER THREE DISTINCT INTEGER VALUES $Z = \{Z_1, Z_2, Z_3\}$

(a) WHAT IS THE MAXIMUM CAPACITY OVER ALL CHOICES OF THE Z ALPHABET? GIVE DISTINCT INTEGER VALUES Z_1, Z_2, Z_3 AND A DISTRIBUTION ON X ACHIEVING THIS.



(4.13a)

(b) WHAT IS THE MINIMUM CAPACITY OVER ALL CHOICES FOR THE Z ALPHABET? GIVE DISTINCT INTEGER VALUES z_1, z_2, z_3 AND A DISTRIBUTION ON X ACHIEVING THIS

(a)

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \sum_x p(x) H(Y|X=x)$$

$$= H(Y) - \sum_x p(x) H(x+Z/x) = H(Y) - H(Z)$$

$Z = \{z_1, z_2, z_3\}$ $\quad p(Z) = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$

$H(Z) = \log 3$ $\quad I(X,Y) = H(Y) - \log 3$

(4.13b)

X \ Y	0+Z			1+Z			2+Z			3+Z		
0	z_1	z_2	z_3	z_1	z_2	z_3	z_1	z_2	z_3	z_1	z_2	z_3
	z_1	z_2	z_3									
1												
2												
3												

$Z = \{0, 1, 2\} \Rightarrow p(Z) = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$

X \ Y	Z_1			Z_2			Z_3			Z_4			
	0	1	2	1	2	3	2	3	4	3	4	5	
0	1	3	5	2	4	6	3	5	7	4	6	8	{0, 1, 2}
1	3	5	9	4	6	10	5	7	11	6	8	12	{1, 3, 5}
2	1	2	3	2	3	4	3	4	5	4	5	6	{3, 5, 9}
3													{1, 2, 3}

1° $0 \div 5$ (N = 6) $\quad$ 2° $1 \div 8$ (N = 8) $\quad$ 3° $3 \div 12$ (N = 10)

4° $1 \div 6$ (N = 6)

$P(Y_4 = 1) = P(X=0) \cdot P(Z=1) = \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12}$

$P(Y_4 = 2) = P(X=0) \cdot P(Z=2) + P(X=1) \cdot P(Z=1) = \frac{1}{6}$

$P(Y_4 = 3) = P(X=0) \cdot P(Z=3) + P(X=1) \cdot P(Z=2) + P(X=2) \cdot P(Z=1) = \frac{1}{4}$

$P(Y_4 = 4) = (1,3) + (2,2) + (3,1) = \frac{1}{4}$

$P(Y_4 = 5) = (2,3)\ (3,2) = \frac{1}{6}$ $\quad$ $P(Y = 6) = \frac{1}{12}$

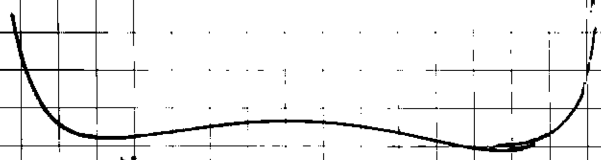
$$P = 2 \cdot \frac{1}{12} + 2 \cdot \frac{1}{6} + 2 \cdot \frac{1}{4} = \frac{1}{6} + \frac{1}{3} + \frac{1}{2} = \frac{1+2+3}{6} = 1$$

$$H(Y) = 2 \cdot \frac{1}{12} \log 12 + 2 \cdot \frac{1}{6} \log 6 + 2 \cdot \frac{1}{4} \log 4 = \frac{1}{6} \log 12 + \frac{1}{3} \log 6 + 1$$

$$= \frac{1}{6} \log 12 + \frac{2}{6} \log 6 + 1 = \frac{1}{6} \left(\log (12 \cdot 36) \right) + 1$$

$$= \frac{1}{6} \log \ldots = 1.5 + 1 = 2.5$$

$$C_{min} = 2.5 - \log 3 = 2.5 - 1.6 = 0.9$$



- MAKSIMALNI KAPACITET SE DOBIVA AKO NE POSTOJI PRESEK MEĐU X I Z TOKOM SE DOBIVAT IZ NEZAVISNI IZLAZI ZA Y.

- ZNAČI REŠENJE ZA (9) E AKO X I Z SE UNIFORMNO RASPOREĐENI AKO NEMA PRESEK MEĐU MV VC TOJ SLUČAJ

$H(Z) = \left(\frac{1}{12}\, ld\, 12\right)\cdot 12 = ld\, 12$

$I(X;Z)_{max} = C_{max} = ld\,\frac{12}{3} = ld\, 4 = 2$

$C \in [0.9, 2]$

SO OVA MAŽINA ŠTO MOŽDO REŠENJE SO NAMAME PRESEK, E POKE-MENATO !!!

Handout 19 Solution

- Max. TOT ... MODIV ... IST MATN

- Minimum: $Z=\{0,1,2\}$ $\quad p(z)=\{\frac{1}{2}, 0, 0, \frac{1}{2}\}$

$P(Z=0) = \frac{1}{2}\cdot P(Z=0) = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6}$

$P(Z=1) = \underbrace{\frac{1}{2}}_{P(X=0)}\cdot P(Z=1) = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6}$

$P(Z=2) = P(X=0)\cdot P(Z=2) = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6}$

$P(Z=3) = P(X=3)\cdot P(Z=0) = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6}$

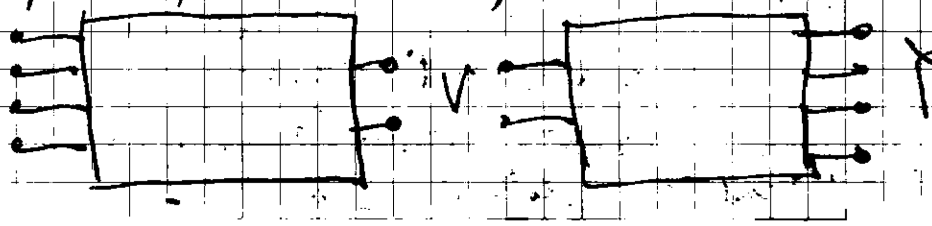
$P(Z=4) = P(X=3)\cdot P(Z=1) = \frac{1}{2}\cdot\frac{1}{3} = 1/6$

$P(Z=5) = P(X=3)\cdot P(Z=2) = \frac{1}{2}\cdot\frac{1}{3} = 1/6$

$H(Z) = \left(\frac{1}{6}\, ld\, 6\right)\cdot 6 = ld\, 6 \qquad I(X;Z) = ld\, 6 - \ldots$

$I(X;Z) = ld\,\frac{6}{3} = ld\, 2 = 1 = C_{max}$

PROBLEM 7.25 BOTTLENECK CHANNEL. SUPPOSE THAT A SIGNAL $X \in \mathcal{X} = \{1, 2, 3, \ldots, m\}$ GOES THROUGH INTERVENING TRANSITION $X \to V \to Y$



(115) WHERE $X = \{1, 2, \dots, n\}$ $Y = \{1, 2, \dots, m\}$ AND $U = \{1, 2, \dots, k\}$. THEN $p(u|x)$ AND $p(y|u)$ ARE ARBITRARY AND THE CHANNEL HAS TRANSITION PROBABILITY:

$$p(y|x) = \sum_{u} p(u|x)\, p(y|u)$$

SHOW THAT $C \leq \mathrm{ld}\, k$

PROBLEM 7.26 (CONTINUE FROM N16.6)

(C) $x \in \{0, 1, 2, 3, 4\}$ $Z \in \{A, B\}$ $C = ?$

$$P(z|x) = P(Z = z | X = x) = \sum P(Y = y_0 | X = x)$$

$g(y_0) = z$

$P(x, y)$

x \ y	A: 0	A: 1	B: 2	B: 3
0	p_{11}	p_{12}	p_{13}	p_{14}
1	p_{21}	p_{22}	p_{23}	p_{24}
2	p_{31}	p_{32}	p_{33}	p_{34}
3	p_{41}	p_{42}	p_{43}	p_{44}
$p(z)$	$P(A)$		$P(B)$	

x \ Z	A	B
0	$p_{11}+p_{12}$	$p_{13}+p_{14}$
1	$p_{21}+p_{22}$	$p_{23}+p_{24}$
2	$p_{31}+p_{32}$	$p_{33}+p_{34}$
3	$p_{41}+p_{42}$	$p_{43}+p_{44}$
$P(Z)$	$P(A)$	$P(B)$

$$I(X;Z) = H(Z) - H(Z|X)$$

$$C = \underbrace{1}_{\mathrm{ld}\,2} - H(Z|X)$$

$$P(Z = A) = \sum_{i=1}^{4} \sum_{j=1}^{2} p_{ij} = \sum_{i=1}^{4} (p_{i1} + p_{i2})$$

$$P(Z = B) = \sum_{i=1}^{4} (p_{i3} + p_{i4})$$

$$-H(Z) = \left[\sum_{i=1}^{4} (p_{i1} + p_{i2})\right] \mathrm{ld} \sum_{i=1}^{4} (p_{i1} + p_{i2}) + \left[\sum_{i=1}^{4} (p_{i3} + p_{i4})\right] \mathrm{ld} \left[\sum_{i=1}^{4} (p_{i3} + p_{i4})\right]$$

TRANSITION PROBAB.

$$p'_{ij} = \frac{p_{ij}}{p(i)}$$

$$H(Z|X) = \sum_{x} p(x) \cdot H(Z | X = x)$$

$$H(Z | X = 0) = -\left[(p'_{11} + p'_{12}) \cdot \mathrm{ld}(p'_{11} + p'_{12}) + (p'_{13} + p'_{14})\, \mathrm{ld}(p'_{13} + p'_{14})\right]$$

$$p(X=0) \cdot H(Z | X = 0) = -\left[(p_{11} + p_{12}) \cdot \mathrm{ld} \frac{p_{11} + p_{12}}{p(X=0)} + (p_{13} + p_{14})\, \mathrm{ld} \frac{p_{13} + p_{14}}{p(X=0)}\right]$$

X \ Y	0 (A)	1	2 (B)	3
0	1/2	1/2	0	0
1	0	1/2	1/2	0
2	0	0	1/2	1/2
3	1/2	0	0	1/2

(TBL.1)

$$H(Z | X = 0) = -\left[\underbrace{1 \cdot \mathrm{ld}\, 1}_{0} + \dots\right]$$

$$H(Z | X = 1) = -\left[P(Z = A | X = 1)\, \mathrm{ld}\, P(Z = A | X = 1) + P(Z = B | X = 1) \cdot \mathrm{ld}\, P(Z = B | X = 1)\right] = \frac{1}{2}\mathrm{ld}\,2 + \frac{1}{2}\mathrm{ld}\,2 = 1$$

$H(Z|X=2) = -P(Z=3|X=2)\cdot ld\, P(Z=3|X=2) = 1\, ld 1 = 0$

$H(Z|X=3) = -P(Z=1|X=3)\cdot ld\, P(Z=1|X=3) - P(Z=3|X=3)\, ld\, P(Z=3|X=3)$
$= \frac{1}{2}\, ld 2 + \frac{1}{2}\, ld 2 = 1$

$H(Z|X) = P(X=0)\cdot 0 + P(X=1)\cdot 1 + P(X=2)\cdot 0 + P(X=3)\cdot 1$

$C = \max_{p(x)} [I(X;Y)] = 1 - H(Z|X) = 1 - [P(X=1) + P(X=3)]$

$C = P(X=0) + P(X=2)$ ($)

(b) REVISITED:

$P(X,Z)$

TBL. 2

x \ z	0	1	2	3
0	0	0	0	0
1	0	1/4	1/4	0
2	0	0	0	0
3	1/4	0	0	1/4
P(Z)	1/2		1/2	

$P(X=1) = P(X=3) = 1/2$, $P(X=0) = P(X=2) = 0$

$H(Z|X=0) = 0$

$H(Z|X=2) = 0$

$H(Z|X=1) = P(Z=1|X=1)\cdot ld\, P(Z=1|X=1) + P(Z=2|X=1)\cdot ld\, P(Z=2|X=1)$

$H(Z|X=1) = \frac{1}{4}\cdot ld 4 + \frac{1}{4}\, ld 4 = \frac{1}{2} + \frac{1}{2} = 1$

$H(Z|X) = P(X=1)\cdot H(Z|X=1) + P(X=3)\cdot H(Z|X=3) = \frac{1}{2} + \frac{1}{2}$

$I(X;Y) = H(Z) - H(Z|X)$

$P(Z) = \left\{\frac{1}{4}+\frac{1}{4};\ \frac{1}{4}+\frac{1}{4}\right\}$ $\quad H(Z) = H\left(\frac{1}{2}, \frac{1}{2}\right)$

$H(Z) = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 1 = 1$ $\quad I(X;Y) = 1 - 1 = 0$

ii)

x \ y	0	1	2	3
0	1/4	1/4	0	0
1	0	0	0	0
2	0	0	1/4	1/4
3	0	0	0	0

$P(X=0) = P(X=2) = 1/2$

$H(Z|X=0) = P(Z=1|X=0)\, ld\, P(Z=1|X=0) = 1\cdot ld 1 = 0$

$H(Z|X) = \sum_x p(x)\, H(Z|X=x) = 0$

$I(X;Y) = H(Z) - 0 = \left(\frac{1}{2}\cdot ld 2\right)\cdot 2 - 0 = 1$ (*)

OBSTATA FORMULA ($) FUNKCIONIRA I ZA ⊕ I ZA (*)

d) Does $X \to Z \to Y$ FORM MARKOV CHAIN?

$P(Z,Y|X) = \frac{P(X,Y,Z)}{P(X)} = \frac{P(X,Z)\, P(Y|X,Z)}{P(X)} = P(Z|X)\, P(Y|Z)$

$I(XZ;Y) = I(Z;Y) + I(X;Y|Z)$ $\quad I(X;Y|Z) = H(X|Z) - H(X|YZ) = I(Y|Z) - I(X|Z) = 0$

(116) $I(X,Z;Y) = I(Z;Y) + \overbrace{I(X;Y|Z)}^{0}$ $= 0$

$I(X;Y|Z) = H(Y|Z) - H(Y|XZ) = H(Y|Z) - H(Y|Z)$

$H(Y|Z) = ?$ $H(Y|Z=A) = P(Y=0|Z=A)\cdot ld\, P(Y=0|Z=A) + P(Y=1|Z=A)\cdot ld\, P(Y=1|Z=A) = \left(\frac{1}{2} ld 2\right)\cdot 2 = 1$

$H(Y|Z) = P(Z=A)\cdot 1 + P(Z=B)\cdot 1 = \frac{1}{2} + \frac{1}{2} = 1$

$H(Y|XZ) = ?$ $H(Y|0A) = 0$ $H(Y|0B) = 0$

$H(Y|1A) = 0$; $H(Y|1B) = 0$; $H(Y|2A) = 0$;
$H(Y|2B) = 0$; $H(Y|3A) = 0$; $H(Y|3B) = 0$

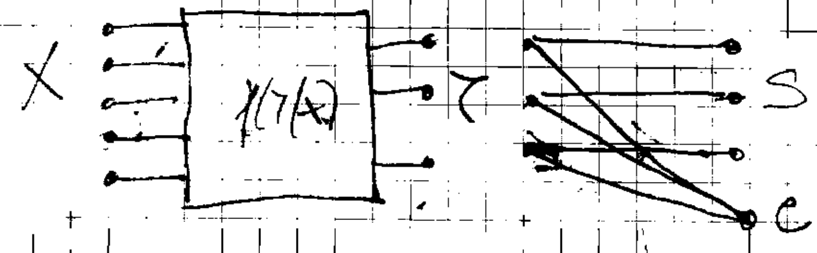
AKO GI ZNAES X, Z NEMA NEIZVESNOSTTA ZA ZNAES „Y". $H(Y|XZ) = 0 \neq H(Y|Z) = 1$

X-Z-Y NE E MARKOV LANEC !!!

(OD TBL.2 STR.116 t ZASTO DEKA AKO E POZNATO Z, Y ZAVISI OD VREDNOSTA NA „X". NA PRIMER AKO Z=A „Y" MOZE DA BIDE „1" SO VEROJATNOST 1/4 A MOZE DA BIDE 1, 3 SO VEROJATNOST 1/2 VO PRVIOT SLUCAJ X=1, A VO VTORIOT X=3 (TBL.1)

• ZA (b) (ii) $X \to Z \to Y$: FORMIRA MARKOV LANEC ZOSTO AKO SO ZNAES „Z" Y NE ZAVISI OD X. NA PRIMER, AKO Z=A Y MOZE DA BIDE 0 ILI 1 SO VEROJATNOST 1/2 (TBL.1) NO NE ZAVISI OD X ZOSTO X SEKOGAS E = 0.

PROBLEM 7.27 ERASURE CHANNEL. LET $\{X, P(Y|X), Y\}$ BE A DISCRETE MEMORYLESS CHANNEL WITH CAPACITY C. SUPPOSE THAT CHANNEL IS CASCADED IMMEDIATELY WITH ERASURE CHANNEL $\{Y, P(S|Y), S\}$ THAT ERASES α OF ITS SYMBOLS.

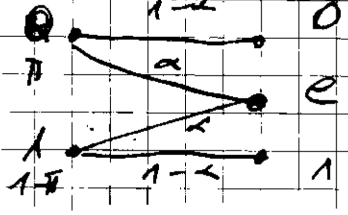


SPECIFICALLY, $S = \{\gamma_1, \gamma_2, \dots, \gamma_m, e\}$ AND

$Pr\{S = \gamma \mid \Upsilon = \gamma\} = \bar{\alpha}\, p(\gamma \mid x)$, $\gamma \in \Upsilon$

$Pr\{S = e \mid X = x\} = \alpha$

DETERMINE THE CAPACITY OF THIS CHANNEL.



$I(X;Y) = H(Y) - H(Y|X)$

$p(Y) = \{\pi \cdot (1-\alpha),\ \pi\alpha + (1-\pi)\cdot\alpha,\ (1-\pi)(1-\alpha)\}$

$= /\pi = \tfrac{1}{2}/ = \left\{\frac{1-\alpha}{2}, \alpha, \frac{1-\alpha}{2}\right\}$

$H(Y) = 2\left(\frac{1-\alpha}{2}\right)\cdot ld\frac{2}{1-\alpha} + \alpha\, ld\frac{1}{\alpha} = (1-\alpha)\, ld\frac{1}{1-\alpha} + \alpha\, ld\frac{1}{\alpha} + (1-\alpha)$

$H(Y) = (1-\alpha) + H(\alpha)$; $H(Y|X) = \pi\cdot H(\alpha) + (1-\pi)H(\alpha) = H(\alpha)$

$I(X;Y) = H(Y) - H(Y|X) = H(\alpha) + (1-\alpha) - H(\alpha) = 1-\alpha$

- $I(X;Y,Z)$; $I(X,Y;Z) = I(Y;Z) + I(X;Z|Y)$, $I(Y;Z) = I(X;Z) + I(Y;Z|X)$, $I(X;Z|Y) = 0$

$I(X;Z) = H(Z|X) - H(Z|XY) = H(Z|Y) - H(Z|Y) = 0$

$I(X;Z) \le I(Y;Z)$ $C_e = \max I(X;Z) = I(Y;Z)$

$I(Y;S) = H(S) - H(S|Y)$

$H(S|Y) = \sum_{\gamma_i} p(\gamma_i)\, H(S \mid Y=\gamma) =$

$= \sum_{\gamma} p(\gamma)\cdot H(\alpha) = H(\alpha)$

$H(S) = ?$ $p(S) = \{p(\gamma_1)\cdot(1-\alpha),\ p(\gamma_2)(1-\alpha), \dots p(\gamma_m)(1-\alpha), \alpha\}$

$H(S) = -\sum_{S} p(s)\, ld\, p(s) = -\{p(\gamma_1)(1-\alpha)\, ld[p(\gamma_1)(1-\alpha)] +$

$+ \dots + p(\gamma_m)(1-\alpha)\, ld\, p(\gamma_m)(1-\alpha)\} - \alpha\, ld\alpha =$

$= -[p(\gamma_1)(1-\alpha)\, ld(p(\gamma_1)) + \dots + p(\gamma_m)(1-\alpha)\, ld(p(\gamma_m))] -$

$- [p(\gamma_1)\cdot ld(1-\alpha) + \dots + p(\gamma_m)\, ld(1-\alpha)] + \alpha[p(\gamma_1)\cdot ld(1-\alpha) \dots p(\gamma_m)\dots]$

$- \alpha\, ld\alpha = +(1-\alpha)\cdot H(Y) - ld(1-\alpha) + \alpha\, ld(1-\alpha) - \alpha\cdot ld\alpha =$

$= (1-\alpha)H(Y) - (1-\alpha)\, ld(1-\alpha) - \alpha\, ld\alpha = (1-\alpha)H(Y) + H(\alpha)$

$C_e = (1-\alpha)\, H(Y) + H(\alpha) - H(\alpha) = (1-\alpha)\, H(Y)$ $\Rightarrow = (1-\alpha)\, ld\, m$

(15) $H(Y) = ?$

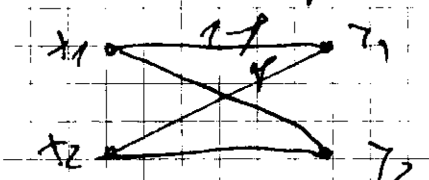
$$\begin{bmatrix} p(y_1) \\ p(y_2) \\ \vdots \\ p(y_n) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{m1} \\ p_{21} & p_{22} & & p_{m2} \\ p_{n1} & p_{n2} & & p_{nm} \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ \vdots \\ p(x_m) \end{bmatrix}$$

$p(y_1) = \sum_i p_{i1} \cdot p(x_i) \qquad p(y_j) = \sum_i p_{ij}\, p(x_i)$

$H(Y) = \sum_j p(y_j) \log p(y_j) = -\sum_j \sum_i p_{ij}\, p(x_i) \log \sum_i p_{ij} p(x_i)$

- KAPACITETOT NA DISCRETEN MEMORYLESS KANAL E:

$C = \max_{p(x)} I(X;Y) = \underbrace{H(Y)}_{\log m}\Big|_{max} - H(Y|X)$



$H(Y|X) = -\sum_{x,y} p(x,y) \log p(y|x)$

$= -\sum_{x,y} p(x)\, p(y|x) \log p(y|x)$

$H(Y) = C + H(Y|X)$

- UNIVERSITY OF TORONTO SOLUTIONS (sol 3,4)

$I(X;Y,S) = I(X;S) + \underbrace{I(X;Y|S)}_{\neq 0} = I(X;Y) + \underbrace{I(X;S|Y)}_{=0}$

$I(X;S) = I(X;Y) - I(X;Y|S) = H(Y) - H(Y|X) - I(X;Y|S)$

$I(X;Y|S) = H(X|S) - H(X|YS)$

$H(X|S) = \sum_s p(s) H(X|S=s) = (1-\alpha) H(X|S=\bar{e}) +$

$+ \alpha H(X|S=e)$

$H(X|YS) = \sum_s p(s) H(X|YS=s) = (1-\alpha) H(X|Y) +$

$\alpha\, H(X|YS=e)$

$I(X;Y|S) = H(Y|S) - H(Y|XS) \quad \alpha$

KOGA NEMA GRESKA $S=Y$ (MHV)

$H(Y|S) = \sum p(s) H(Y|S=s) = p(s=e) \cdot H(Y|S=e) +$

$+ \underbrace{p(s=\bar{e})}_{1-\alpha} H(Y|S=Y) = \alpha \cdot H(Y|S=e) + (1-\alpha) H(Y|S=Y)$

$\bar{e} = Y \qquad = \alpha \cdot H(Y|S=e) \quad (1)$

$H(Y|XS) = \sum_s p(s)\cdot H(Y|X, S=s) = p(e)\, H(Y|X, S=e) +$

$+ p(Y)\cdot H(Y|X, S=Y) = \alpha\, H(Y|X, S=e)$ (*)

~~$I(X;Y|S) = \alpha H(Y|S=e) + (1-\alpha) H(Y) - \alpha H(Y|X, S=e) -$~~
~~$- H(Y|X) + \alpha H(Y|X)$~~
~~$I(X;S) = H(Y) - H(Y|X) - \alpha H(Y|S=e) - (1-\alpha) H(Y) + \alpha H(Y|X, S=e)$~~
~~$+ H(Y|X) - \alpha H(Y|X) = H(Y) - H(Y|X) - \alpha H(Y|S=e) - H(Y) + \alpha H(Y)$~~
~~$+ \alpha H(Y|X, S=e) - H(Y|X) + \alpha H(Y|X)$~~
~~$I(X;S) = \alpha H(Y) + \alpha H(Y|X) - \alpha H(Y|S=e) + \alpha H(Y|X, S=e)$~~

(*), (1) $\Rightarrow I(X;Y|S) = \alpha \cdot H(Y|S=e) - \alpha \cdot H(Y|X, S=e)$

$I(X;S) = H(Y) - H(Y|X) - \alpha H(Y|S=e) + \alpha H(Y|X, S=e)$

$= H(Y) - H(Y|X) - \alpha H(Y) + \alpha H(Y|X) =$ $|Y|$

$= H(Y) - (1-\alpha) H(Y|X) - \alpha \log(\ldots)$

$C = \max I(X;S) = \max_{p(x)} H(Y) - (1-\alpha) H(Y|X) - \alpha \log \ldots$

$I(X;S) = (1-\alpha) H(Y) - (1-\alpha) H(Y|X) = (1-\alpha) I(X;Y)$

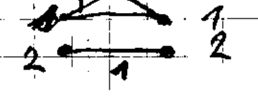
→ VALUATA FORMA VERSE DICI NA MOSTO ORIGINALNO REŠENIE NA [N16.118]

PROBLEM 7.28 CHOICE OF CHANNELS. FIND THE CAPACITY "C" OF THE UNION OF TWO CHANNELS $(X_1, p_1(y_1|x_1), Y_1)$ AND $(X_2, p_2(y_2|x_2), Y_2)$ WHERE AT EACH TIME ONE CAN SEND A SYMBOL OVER CHANNEL 1 OR CHANNEL 2 BUT NOT BOTH. ASSUME THAT THE OUTPUT ALPHABETS ARE DISTINCT AND DON'T INTERSECT.

(a) SHOW THAT $2^C = 2^{C_1} + 2^{C_2}$. THUS, 2^C IS EFFECTIVE ALPHABET SIZE OF A CHANNEL WITH CAPACITY C.

(b) COMPARE WITH PROBLEM 2.10 WHERE $2^H = 2^{H_1} + 2^{H_2}$, AND INTERPRET (a) IN TERMS OF NOISE FREE SYMBOLS.

(c) USE THE ABOVE RESULT TO CALCULATE THE CAPACITY OF THE FOLLOWING CHANNEL.



(121)

$n \cdot R = \underbrace{H(W)}_{ld\, 2^{nR}} = I(W; \hat{W}) + H(W|\hat{W})$

FANO: $H(W|\hat{W}) \leq 1 + P_e \cdot ld\, |W|$

$nR \leq 1 + P_e \cdot n \cdot R + I(X^n; Y^n) \leq 1 + P_e \cdot nR + nC$

$R \leq \frac{1}{n} + P_e \cdot R + C \qquad n \to \infty \quad \boxed{R \to C}$

$nR\, P_e^n \geq nR - 1 - nC \qquad \boxed{P_e^n \geq 1 - \frac{C}{R} - \frac{1}{nR}}$

$R > C \Rightarrow P_e^n$ IS BOUNDED AWAY FROM "0"

---

(9) $2^C = 2^{C_1} + 2^{C_2} \qquad C_1 = 1 - H(q)$

$C_2 = 0$

$2^C = 2^{1-H(q)} + 1 \qquad C = ld\left[2^{1-H(q)} + 1\right]$

$C = ld\left[\frac{2 + 2^{H(q)}}{2^{H(q)}}\right]$

(a)

$$\begin{bmatrix} p(y_{11}) \\ p(y_{12}) \\ \vdots \\ p(y_{1n}) \\ p(y_{21}) \\ p(y_{22}) \\ \vdots \\ p(y_{2n}) \end{bmatrix} = \begin{bmatrix} p_{11}^{(1)} & p_{21}^{(1)} & p_{31}^{(1)} & \dots & p_{n1}^{(1)} \\ p_{21}^{(1)} & p_{22}^{(1)} & \dots & & p_{n2}^{(1)} \\ p_{1n}^{(1)} & p_{2n}^{(1)} & \dots & & p_{nn}^{(1)} \\ p_{11}^{(2)} & p_{21}^{(2)} & & & p_{2n}^{(2)} \\ p_{1n}^{(2)} & p_{2n}^{(2)} & & & p_{nn}^{(2)} \end{bmatrix} \begin{bmatrix} p(x_1^1) \\ p(x_2^1) \\ p(x_n^1) \\ p(x_1^2) \\ p(x_2^2) \\ p(x_n^2) \end{bmatrix} \begin{matrix} \}(\alpha) \\ \\ \}(1-\alpha) \end{matrix}$$

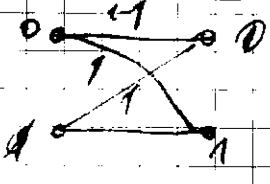
$I(X;Y) = H(Y) - H(Y|X)$

$-H(Y) = \left(\frac{\alpha}{n}\, ld\, \frac{\alpha}{n}\right) \cdot n + \left(\frac{1-\alpha}{n} \cdot ld\, \frac{1-\alpha}{n}\right) \cdot n$

UNIFORM DISTRIBUTION GIVES MAXIMUM "C"

$\boxed{-H(Y) = \alpha\, ld\, \frac{\alpha}{n} + (1-\alpha)\, ld\, \frac{1-\alpha}{n}}$

$\boxed{H(Y|X) = \alpha H(Y_1|X_1) + (1-\alpha) H(Y_2|X_2)}$ ← (4P)



$H(Y|X) = p(x=0) \cdot H(q) + p(x=1) \cdot H(q) = H(q)$

(1) $\begin{cases} C_1 = \max_{p(x_1)} H(Y_1) - H(Y_1|X_1) = ld\, n - H(Y_1|X_1) \\ C_2 = \max_{p(x_2)} H(Y_2) - H(Y_2|X_2) = ld\, n - H(Y_2|X_2) \end{cases}$

$I(X;Y) = -\alpha\, ld\frac{\alpha}{1} - (1-\alpha)\, ld\left(\frac{1-\alpha}{1}\right) - \alpha H(Y_1|X_1) - (1-\alpha) H(Y_2|X_2)$

$\frac{d I(X;Y)}{d\alpha} = 0 = ld\left(\frac{1-\alpha}{\alpha}\right) - H(Y_1|X_1) + H(Y_2|X_2)$

$ld\left(\frac{1-\alpha}{\alpha}\right) = \underbrace{H(Y_1|X_1)}_{H_1} - \underbrace{H(Y_2|X_2)}_{H_2}$ $\quad \frac{1-\alpha}{\alpha} = 2^{H(Y_1|X_1)} \cdot 2^{-H(Y_2|X_2)}$

$1-\alpha = \alpha \cdot 2^{H_1} \cdot 2^{-H_2}$ $\quad 1 = \alpha\left(1 + \frac{2^{H_1}}{2^{H_2}}\right)$ $\quad \alpha = \frac{2^{H_2}}{2^{H_1}+2^{H_2}}$

$C = ld\left[2^{1-H_2} + 2^{1-H_1}\right]$ $\quad 2^C = 2^{1-H_2} + 2^{1-H_1}$ ← (1)

$$\boxed{2^C = 2^{C_1} + 2^{C_2}}$$

~~0208~~ 0237163971

(b) COMPARE WITH (2.10)

$I(X;Y) = -\alpha\, ld\alpha - (1-\alpha)\, ld(1-\alpha) + \cancel{\alpha\, ld_1} + (1-\alpha)\, ld_1 - \alpha H(Y_1|X_1) - (1-\alpha) H(Y_2|X_2) = -\alpha\, ld\alpha - (1-\alpha)\, ld(1-\alpha) - \alpha H(Y_1|X_1) - (1-\alpha) H(Y_2|X_2)$

$2^{H(X)} \le 2^{H(X_1)} + 2^{H(X_2)}$

UNIVERSITY OF TORONTO SOLUTION

$p(z) = \begin{cases} 1 & \alpha \\ 2 & 1-\alpha \end{cases}$

$Z \in \{1,2\}$ $\quad Z=1$ CHANNEL 1, $Z=2$ CHANNEL 2 $\quad X \in \{X_1 \cup X_2\}$, $Y \in \{Y_1 \cup Y_2\}$

(9) $I(X,Z;Y) = H(Z,X) - H(X,Z|Y) = \underline{H(Z)} + H(X|Z) - H(X|Y) - H(Z|\underline{X},\underline{Y}) = H_2(\alpha) + \alpha H(X_1) + (1-\alpha) H(X_2) - H(X|Y)$ ← (1?)

$H(Z|X,Y) = 0$

$= H_2(\alpha) + \alpha H(X_1) + (1-\alpha) H(X_2) - \alpha H(X_1|Y_1) - (1-\alpha) H(X_2|Y_2) =$

$= H_2(\alpha) + \alpha I(X_1;Y_1) + (1-\alpha) I(X_2;Y_2)$

$$\boxed{C = \max_\alpha H_2(\alpha) + \alpha C_1 + (1-\alpha) C_2}$$

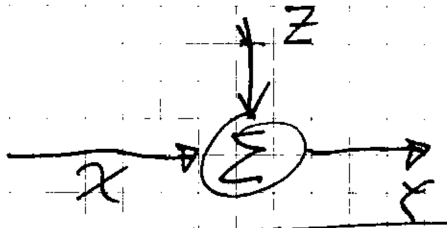
$-H_2(\alpha) = \alpha\, ld\alpha + (1-\alpha)\, ld(1-\alpha)$ $\quad \frac{d H_2(\alpha)}{d\alpha} = ld\left(\frac{1-\alpha}{\alpha}\right)$

$0 = ld\frac{1-\alpha}{\alpha} + C_1 - C_2$ $\quad ld\frac{1-\alpha}{\alpha} = C_2 - C_1$ $\quad \frac{1-\alpha}{\alpha} = 2^{C_2-C_1}$

$1-\alpha = \alpha\, 2^{C_2-C_1}$ $\quad 1 = \alpha\left(1 + 2^{C_2-C_1}\right)$ $\quad \alpha = \frac{2^{C_1}}{2^{C_1}+2^{C_2}}$

$$\boxed{C = 2^{C_1} + 2^{C_2}}$$

PROBLEM 7.29 NOISE ALPHABETS. CONSIDER CHANNEL

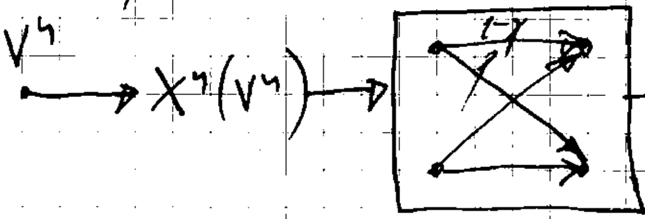


$\mathcal{X} = \{0,1,2,3\}$ $Y = X + Z$

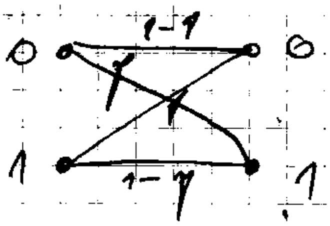
AND Z IS UNIFORMLY DISTRIBUTED.

ISTA KAKO 7.24

PROBLEM 7.31 WE WISH TO ENCODE A BERNOULLI(α) PROCESS $V_1 V_2 \ldots$ FOR TRANSMISSION OVER A BINARY SYMMETRIC CHANNEL, WITH CROSSOVER PROBABILITY p.



FIND CONDITIONS ON "α" AND "p" SO THAT THE PROBABILITY OF "ERROR" $P(\hat{V}^n \neq V^n)$ CAN BE MADE TO GO TO ZERO AS $n \to \infty$.



$C_2 = 1 - H(p)$

$I(X;Y) = H(Y) - H(Y|X)$

$\max_{p(x)} I(X;Y) = \max_{p(x)} H(Y) - H(Y|X)$

$$I(X^n;Y^n) \leq n \cdot I(X;Y)$$

$n \cdot R \leq 1 + P_e \cdot n \cdot R + nC$ $\quad R \leq \frac{1}{n} + P_e \cdot R + C$

$C = H(\alpha) - H(p)$ $\quad R(1 - P_e^{n}) \leq \frac{1}{n} + H(\alpha) - H(p)$

$1 - P_e^n \leq \frac{1}{nR} + \frac{H(\alpha) - H(p)}{R}$ $\quad P_e \geq 1 - \frac{1}{nR} + \frac{H(\alpha) - H(p)}{R}$

$n \to \infty$ $\quad P_e \geq 1 - \frac{H(\alpha) - H(p)}{R} \to 0 \to$

$\boxed{H(\alpha) - H(p) = R}$ e.g. $\alpha = \frac{1}{2}$ $\quad 1 - H(p) = R$

$p = 0.1 \Rightarrow H(p) = 0.469$ $\quad R = 1 - 0.469 = \underline{0.531}$

$H(\alpha) - H(p) \leq R$ (BSC $\max R = 1$)

$H(\alpha) - H(p) \leq 1$ $\quad \boxed{H(\alpha) \leq 1 - H(p)}$ — IS NOT REZULTAT

ŠTO GO DODILE VO REŠENIETO VO UIC.

SOLUTION UNIVERSITY OF ILLINOIS CHICAGO

WE CAN TRANSMIT A STATIONARY ERGODIC SOURCE OVER A CHANNEL IF AND ONLY IF ITS ENTROPY RATE IS LESS THEN THE CAPACITY OF THE CHANNEL.

$$H(V) < C$$

$$H(V) = \lim_{n\to\infty} \frac{1}{n} H(V_1, V_2, \dots V_n) = \frac{n H(\alpha)}{n} = H(\alpha)$$

$$H(\alpha) < \underbrace{1 - H(q)}_{C_{BSC}}$$

SHORD IS A HARD 5.45

PROBLEM 7.32 RANDOM 20 QUESTIONS. LET X BE UNIFORMELY DISTRIBUTED OVER $\{1, 2, \dots m\}$. ASSUME THAT $m = 2^n$. WE ASK RANDOM QUESTIONS IS $X \in S_1$? IS $X \in S_2$? ... UNTIL ONE INTEGER REMAINS. ALL 2^m SUBSETS S OF $\{1, 2, \dots m\}$ ARE EQUALLY LIKELY.

(a) HOW MANY DETERMINISTIC QUESTIONS ARE NEEDED TO DETERMINE X?

(b) WITHOUT LOSS OF GENERALITY, SUPPOSE THAT $X = 1$ IS THE RANDOM OBJECT. WHAT IS THE PROBABILITY THAT OBJECT 2 YIELDS THE SAME ANSWER AS OBJECT 1 FOR k QUESTIONS.

(c) WHAT IS THE EXPECTED NUMBER OF OBJECTS IN $\{2, 3, \dots m\}$ THAT HAVE THE SAME ANSWERS TO THE QUESTIONS AS THOSE OF THE CORRECT OBJECT 1?

(d) SUPPOSE THAT WE ASK $n + \sqrt{n}$ RANDOM QUESTIONS. WHAT IS THE EXPECTED NUMBER OF WRONG OBJECTS AGREEING WITH ANSWERS?

(e) USE MARKOV INEQUALITY $\Pr\{X \geq t\mu\} \leq 1/t$ TO SHOW THAT PROBABILITY OF ERROR (ONE OR MORE WRONG OBJECTS REMAINING) GOES TO ZERO AS $n \to \infty$.

PROBLEM 7.36 CHANNEL WITH MEMORY. CONSIDER THE DISCRETE MEMORYLESS CHANNEL $Y_i = Z_i X_i$ WITH INPUT ALPHABET $X_i \in \{-1, 1\}$

(a) WHAT IS THE CAPACITY OF THIS CHANNEL WHEN $\{Z_i\}$ IS I.I.D.

$$Z_i = \begin{cases} 1 & p = 0.5 \\ -1 & p = 0.5 \end{cases}$$?

(b) NOW CONSIDER THE CHANNEL WITH MEMORY. BEFORE TRANSMISSION BEGINS Z IS RANDOMLY CHOSEN AND FIXED FOR ALL TIME. THUS $Z_i = Z$

(b) WHAT IS THE CAPACITY IF:

$$Z = \begin{cases} 1 & p = 0.5 \\ -1 & q = 0.5 \end{cases} ?$$

(a) $Y_i = X_i \cdot Z_i$

$I(X;Y) = ?$

$P(X,Z,Y)$ / $P(X,Z)$

$X_i Z_i$ \ Y_i		-1	1	$P(X,Z)$
-1	-1	0	$1/4$	$1/4$
	1	$1/4$	0	$1/4$
1	-1	$1/4$	0	$1/4$
	1	0	$1/4$	$1/4$
$P(Y)$		$1/2$	$1/2$	

$$I(X;Y) = \underbrace{H(Y)}_{ld 2} - H(Y|X)$$

$$H(Y|X) = P(X=-1) \cdot H(Y|X=-1) + P(X=1) \cdot H(Y|X=1) =$$

$$= \frac{1}{2}\left(\frac{1}{4} ld 4\right) \cdot 2 + \frac{1}{2} \frac{1}{2} ld 4 =$$

$$= \frac{1}{2} + \frac{1}{2} = 1 \quad (\$)$$

NA OVOJ NAČIN VSUŠNOST DOBIVAŠ BSC SO CROSSOVER PROBAB. OF 0.5

$I(X;Y) = 0$ → OVA JE ZA $P(Y) = \left\{\frac{1}{2}, \frac{1}{2}\right\}$

AKO: $P(X) = \{p, 1-p\}$

$X_i Z_i$ \ Y		-1	1	$P(X,Z)$
-1	-1	0	$p/2$	$p/2$
	1	$p/2$	0	$p/2$
1	-1	$(1-p)/2$	0	$(1-p)/2$
	1	0	$(1-p)/2$	$(1-p)/2$
$P(Y)$		$1/2$	$1/2$	

$$H(Y) = \frac{1}{2} ld 2 \cdot 2 = 1$$

$$H(Y|X) = p \cdot \left(\frac{p}{2} \cdot ld \frac{2}{p} \cdot 2\right.$$

$$\left.(1-p)\frac{(1-p)}{2} ld \frac{2}{(1-p)}\right) \cdot 2 = p^2 ld \frac{2}{p} +$$

$$+ (1-p)^2 ld \frac{2}{1-p} = p^2 ld \frac{2}{p} +$$

$$+ \frac{1-2p+p^2}{} ld \frac{2}{1-p} =$$

$$= p^2 ld \frac{2}{p} + \frac{-2p(1-p) + 1 - p^2}{} ld \frac{2}{1-p}$$

GREŠKA!!! TREBA USLOVNA VEROJATNOST DA SE ZEMA, T.E. DA DELIŠ SO $P(X)$. VIDI PR. 128

$$I(X;Y) = 1 - \left[p^2 ld \frac{2}{p} + (1-p)^2 ld \frac{2}{(1-p)}\right] \quad (*)$$

$$\frac{d(I(X;Y))}{dp} = 0 \qquad p^2 ld \frac{2}{p} + (1-p)^2 ld 2 = p^2 + 1 - 2p + p^2 = 1-2p+2p^2$$

$$2p \ln\left(\frac{1-p}{p} \cdot 2\right) = \ln\left[\left(\frac{1-p}{2}\right)^2 \cdot 2\right] \Rightarrow$$

$$\left[\frac{2(1-p)}{p}\right]^{2p} = \left(\frac{1-p}{\sqrt{2}}\right)^2 \Rightarrow \boxed{p = 1}$$

- SO PLOT VO MATLE ZA (4) SE DOBIVA MAKSIMUM: (26)

$p = 0.128$ $f(p) = 1 - p^2 ld\frac{2}{p} + (1-p)^2 ld\frac{2}{1-p}$

$f(p) = 0{,}024388$

$-2p\ln\frac{2}{p} + 2(1-p)\ln\left(\frac{2}{1-p}\right) - 1 + 2p = 0$

$-2p\ln\frac{2}{p} + 2\ln\frac{2}{1-p} - 2p\ln\frac{2}{1-p} - 1 + 2p = 0$

$-2p\left(\ln\frac{2}{p} + \ln\frac{2}{1-p} - 1\right) = 1 - 2\ln\frac{2}{1-p}$

$-2p\ln\frac{2\cdot 2}{p(1-p)}\frac{1}{2} = \ln 2 - \ln\frac{4}{(1-p)^2} = \ln x\cdot\frac{(1-p)^2}{2}$

$\left[\frac{p(1-p)}{2}\right]^{2p} = \frac{(1-p)^2}{2}$ $\left(\frac{p(1-p)}{2}\right)^p = \pm\frac{1-p}{\sqrt{2}}$ $p_1 = 0.5$, $p_2 = 1$

(6) $I(X^n, Y^n) = I(X_1^n, Y_1^n) = H(Y_1^n) - H(Y_1^n|X_1^n)$

$H(Y_1^n) = H(Y_1) + H(Y_2|Y_1) + H(Y_3|Y_1Y_2) + \dots H(Y_n|Y_1^{n-1})$

$H(Y_1^n|X_1^n) = H(Y_1|X_1^n) + H(Y_2|X_1^nY_1) + \dots + H(Y_n|X_1^nY_1^{n-1})$

$= H(Y_1|X_1) + H(Y_2|X_2) + \dots + H(Y_n|X_n)$

IMA ŠUM!! 0, 0 NEMA ŠUM!!! ŠTO PRENESUVAŠ TOA PRIMAŠ

$H(Y_1|X_1) = p^2 ld\frac{2}{p} + (1-p)^2 ld\frac{2}{1-p}$

$I(X^n, Y^n) = n\cdot 1 - H(Y_1|X_1) = n - p^2 ld\frac{2}{p} + (1-p)^2 ld\frac{2}{1-p}$

$I(X^n, Y^n) \leq \sum_{i=1} I(X_i; Y_i) = 1 - p^2 ld\frac{2}{p} + (1-p)^2 ld\frac{2}{(1-p)} +$

$+ \sum_{i=2}^{n} I(X_i; Y_i) = 1 - p^2 ld\frac{2}{p} - (1-p)^2 ld\frac{2}{(1-p)} + \sum_{i=2}^{n-1} H(Y_i)$

$= n - p^2 ld\frac{2}{p} - (1-p)^2 ld\frac{2}{(1-p)}$ } CHANNEL CAPACITY FOR "n" TIMES USAGE OF THE CHANNEL!!

$C_n = \frac{1}{n}\left(n - p^2 ld\frac{2}{p} - (1-p)^2 ld\frac{2}{(1-p)}\right)$

$n \to \infty$ $C \to 1$

- AKO ZEMEŠ DEKA $p(x) = \left\{\frac{1}{2}, \frac{1}{2}\right\} \to (\#) \to$

$H(Y_1|X_1) = 1 \Rightarrow I(X^n, Y^n) \leq n - 1$

$C = \lim_{n\to\infty}\frac{n-1}{n} = 1$

• CONTINUE FROM N16C NOTEBOOK (midterm-1-2009-solutions) (127)

$Y_i = (X_i + Z_i) \bmod 2$

$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \sum_x p(x) H(Y|X=x)$

$= H(Y) - \sum_x p(x) H((X+Z) \bmod 2 \,|\, x) = H(Y) - H(Z)$

$I(X;Y) = H(Y) - H(Z)$ $\qquad C = 1 - H(Z) = 1 - H(p)$

$P(XZ, Y)$

XZ \ Y	0	1	P(XZ)
0 0	(1-p)/2	0	(1-p)/2
0 1	0	p/2	p/2
1 0	0	(1-p)/2	(1-p)/2
1 1	p/2	0	p/2
P(Y)	1/2	1/2	

$I(X;Y) = H(Y) - H(Y|X) =$

$= H(Y) - P(X=0)\, H(Y|X=0) -$

$- P(X=1) \cdot H(Y|X=1) =$

$= H(Y) - H(p)\,[P(X=0) + P(X=1)]$

$= H(Y) - H(p) \Rightarrow \qquad H(Y) = 1$

$C = 1 - H(p)$

$P(ZY|X) = \frac{P(XZY)}{P(X)}$

X	Z \ Y	0	1	
0	0	1-p	0	
0	1	0	p	
1	0	(1-p)	-	
1	1	0	p	

Ne mora da ovi što metiju tavka $p(x) = \{\frac{1}{2}, \frac{1}{2}\}$ sendno uz iresmetila se kraći.

$P(ZY|X)$

XZ \ Y	0	1	
0 0	(1-p)	0	
0 1		p	
1 0		1-p	
1 1	p		

Na ovaj način empirijski se dokazuje (*). MMV

$p(X,Y,Z)$

XZ \ Y	0	1	H(XZ)
0 0	$p_1(1-p)$	0	$p_1(1-p)$
0 1	0	$p_1 p$	$p_1 p$
1 0	0	$p_2(1-p)$	$p_2(1-p)$
1 1	$p_2 p$	0	$p_2 p$
P(Y)	$p_1 - p_1 p + p_2 p$	$p_1 p + p_2 - p_2 p$	

This is binary symmetric channel with crossover probability "p".

(e) $C(K) = \max_{p(X_1^K)} \frac{I(X_1^K; Y_1^K)}{K} = \frac{1}{K} \max_{p(X_1^K)} \left[H(X_1^K) - H(X_1^K | Y_1^K) \right]$

$= \frac{1}{K} \max_{p(X_1^K)} \left[H(X_1^K) - \sum_{Y_1^K} p(Y_1^K) \cdot H\left((Y_i - Z) \bmod 2 \,/\, Y_i\right) \right]$

$= H(Z)$

$= \frac{1}{K} \cdot \left[K - H(Z) \right] = 1 - \frac{H(Z)}{K} = 1 - \frac{H(p)}{K}$

The channel with memory length "K" has larger

- PROBLEM 7.36 REVISITED:

$p(YZ\mid X)$		
X \ Y	-1	1
-1 -1	0	1/2
-1 1	1/2	0
1 -1	1/2	0
1 1	0	1/2

~~$H(Y\mid X) = \ldots$~~ [illegible]

$$H(Y\mid X) = p\,H(\tfrac{1}{2}) + (1-p)\,H(\tfrac{1}{2}) = \mathbf{1}$$

ZNACI - PR. 125 (4) VAZI I OVDE ISTO

$I(X;Y) = 1 - 1 = 0 \qquad C = 0$

(b) REVISITED $\quad H(Y_1\mid X_1) = 1 \Rightarrow$

$I(X^n;Y^n) = n\cdot 1 - 1 \qquad I(X^n;Y^n) \leq n\cdot C$

$$C \geq \frac{I(X^n,Y^n)}{n} \xrightarrow[n\to\infty]{} \lim_{n\to\infty}\left(\frac{n-1}{n}\right) = \mathbf{1}$$

PROBLEM 7.37 JOINT TYPICALITY. LET (X_i, Y_i, Z_i) BE i.i.d ACCORDING TO $p(x,y,z)$. WE WILL SAY THAT (X^n,Y^n,Z^n) IS JOINTLY TYPICAL [WRITTEN $(X^n,Y^n,Z^n)\in A_\varepsilon^{(n)}$] IF:

- $p(x^n)\in 2^{-n(H(X)\pm\epsilon)}$; $p(y^n)\in 2^{-n(H(Y)\pm\epsilon)}$; $p(z^n)\in 2^{-n(H(Z)\pm\epsilon)}$
- $p(x^n,y^n)\in 2^{-n(H(X,Y)\pm\epsilon)}$; $p(x^n,z^n)\in 2^{-n[H(X,Z)\pm\epsilon]}$;

$p(y^n,z^n)\in 2^{-n(H(Y,Z)\pm\epsilon)}$; $p(x^n,y^n,z^n)\in 2^{-n[H(X,Y,Z)\pm\epsilon]}$

NOW SUPPOSE THAT $(\tilde{X}^n,\tilde{Y}^n,\tilde{Z}^n)$ IS DRAWN ACCORDING TO $p(x^n)\cdot p(y^n)\cdot p(z^n)$. THUS $\tilde{X}^n\,\tilde{Y}^n\,\tilde{Z}^n$ HAVE THE SAME MARGINALS AS $p(x^n,y^n,z^n)$ BUT ARE INDEPENDENT. FIND BOUNDS ON $\Pr\{(\tilde{X}^n,\tilde{Y}^n,\tilde{Z}^n)\in A_\varepsilon^{(n)}\}$ IN TERMS OF ENTROPIES $H(X), H(Y), H(Z)$, $H(X,Y)$, $H(X,Z)$, $H(Y,Z)$, AND $H(X,Y,Z)$.

$\Pr(\tilde{X}^n,\tilde{Y}^n\in A_\varepsilon^n) = ?\qquad p(x^n,y^n)\in 2^{-n(H(X,Y)\pm\epsilon)}$

$$1 = \sum_{(x^n,y^n)} p(x^n y^n) \geq \sum_{(x^n,y^n)\in A_\varepsilon^n} p(x^n,y^n) \geq \sum_{(x^n,y^n)\in A_\varepsilon^n} 2^{-n(H(X,Y)+\epsilon)}$$

$$\Big\|\; 2^{-n(H(X,Y)+\varepsilon)} \leq p(x^n,y^n) \leq 2^{-n[H(X,Y)-\varepsilon]} \;\Big\| \geq |A_\varepsilon^n|\cdot 2^{-n(H(X,Y)+\epsilon)}$$

(128)

$$|A_\varepsilon^n| \leq 2^{+n(H(X,Y)+\varepsilon)} \qquad Pr((X_n,Y_n)\in A_\varepsilon^n) \geq 1-\varepsilon$$

$$Pr\left((\tilde{X}_n,\tilde{Y}_n)\in A_\varepsilon^n\right) \leq 2^{-n(I(X,Y)+3\varepsilon)}$$

$$(1-\varepsilon) \leq \sum_{x^n,y^n\in A_\varepsilon^n} p(x^n,y^n) \leq |A_\varepsilon^n| \cdot 2^{-n(H(X,Y)-\varepsilon)}$$

$$|A_\varepsilon^{(n)}| \geq (1-\varepsilon)\, 2^{n(H(X,Y)-\varepsilon)}$$

■ $Pr\left[(\tilde{X}_n,\tilde{Y}_n)\in A_\varepsilon^n\right] = p(\tilde{X}_n,\tilde{Y}_n)$

$$= \sum_{\tilde{x}_n\tilde{y}_n\in A_\varepsilon^n} p(\tilde{x}_n)\cdot p(\tilde{y}^n) \leq \sum_{\tilde{x}_n\tilde{y}_n\in A_\varepsilon^n} 2^{-n(H(X)-\varepsilon)}\cdot 2^{-n(H(Y)-\varepsilon)}$$

$$\leq \sum_{\tilde{x}^n\tilde{y}^n\in A_\varepsilon^n} 2^{-n(H(X)-\varepsilon)}\cdot 2^{-n(H(Y)-\varepsilon)} \leq \underbrace{2^{n(H(X,Y)+\varepsilon)}}_{|A_\varepsilon^n|}\cdot 2^{-nH(X)+n\varepsilon-nH(Y)+n\varepsilon}$$

$$= 2^{nH(X,Y)+n\varepsilon-nH(X)+n\varepsilon-nH(Y)+n\varepsilon}$$

$$= 2^{n(H(X)+H(Y|X)-H(X)-H(Y))+3n\varepsilon} = 2^{-nI(X;Y)+3n\varepsilon}$$

$$Pr\left[(\tilde{X}_n,\tilde{Y}_n)\in A_\varepsilon^n\right] \leq 2^{-n[I(X;Y)-3\varepsilon]}$$

■ $$Pr\left[(\tilde{X}_n,\tilde{Y}_n)\in A_\varepsilon^n\right] = \sum_{\tilde{x}_n\tilde{y}_n\in A_\varepsilon^n} p(\tilde{x}_n)\cdot p(\tilde{y}_n) \geq (1-\varepsilon)\, 2^{n(H(X,Y)-\varepsilon)}$$

$$\cdot\, 2^{-n(H(X)+\varepsilon)}\cdot 2^{-n(H(Y)+\varepsilon)} = (1-\varepsilon)\, 2^{nH(X)+nH(Y|X)-n\varepsilon-nH(X)-n\varepsilon-nH(Y)-n\varepsilon}$$

$$= (1-\varepsilon)\, 2^{n(-I(X;Y)-3n\varepsilon)}$$

$$Pr\left[(\tilde{X}_n,\tilde{Y}_n)\in A_\varepsilon^n\right] \geq (1-\varepsilon)\, 2^{-n(I(X;Y)+3n\varepsilon)}$$

$$Pr[(\tilde{X}_n,\tilde{Y}_n,\tilde{Z}_n)\in A_\varepsilon^n] = \sum_{x^n,y^n,z^n\in A_\varepsilon^n} p(\tilde{x}^n,\tilde{y}^n,\tilde{z}^n) = \sum_{x^n,y^n,z^n\in A_\varepsilon^n} p(\tilde{x}^n)\,p(\tilde{y}^n)\,p(\tilde{z}^n) \leq$$

$$\leq |A_\varepsilon^n|\cdot 2^{-n(H(X)-\varepsilon)}\cdot 2^{-n(H(Y)-\varepsilon)}\cdot 2^{-n(H(Z)-\varepsilon)}$$

$$\leq 2^{+n(H(X,Y,Z)+\varepsilon)-nH(X)+n\varepsilon-nH(Y)+n\varepsilon-nH(Z)+n\varepsilon}$$

$$\leq 2^{-n(-H(X,Y,Z)+H(X)+H(Y)+H(Z)-4n\varepsilon)}$$

(130)

$$1 = \sum_{x^n y^n z^n} p(x^n, y^n, z^n) \geq \sum_{x^n y^n z^n \in A_\epsilon^n} p(x^n, y^n, z^n) \geq |A_\epsilon^n| \cdot 2^{-n(H(X,Y,Z)+\epsilon)}$$

$$|A_\epsilon^n| \leq 2^{+n(H(X,Y,Z)+\epsilon)}$$

$$\Pr\left((X^n, Y^n, Z^n) \in A_\epsilon^n\right) \geq 1-\epsilon \quad (\to 1 \text{ if } n \to \infty)$$

$$(1-\epsilon) \leq \sum_{x^n y^n z^n \in A_\epsilon^n} p(x^n, y^n, z^n) \leq |A_\epsilon^n| \cdot 2^{-n(H(X,Y,Z)-\epsilon)}$$

$$(1-\epsilon) \leq |A_\epsilon^n| \, 2^{-n(H(X,Y,Z)-\epsilon)}$$

$$|A_\epsilon^n| \geq (1-\epsilon) 2^{+n(H(X,Y,Z)-\epsilon)}$$

$$(1) \Rightarrow \Pr\left[(\tilde{X}_n, \tilde{Y}_n, \tilde{Z}_n) \in A_\epsilon^n\right] \leq 2^{-n[-H(X,Y,Z)+H(X)+H(Y)+H(Z)-4\epsilon]}$$

$$\Pr\left[(\tilde{X}_n, \tilde{Y}_n, \tilde{Z}_n) \in A_\epsilon^n\right] = \sum_{\tilde{x}_n \tilde{y}_n \tilde{z}_n \in A_\epsilon^n} p(\tilde{x}_n) p(\tilde{y}_n) p(\tilde{z}_n) \geq (1-\epsilon) 2^{n(H(X,Y,Z)-\epsilon)}$$

$$\cdot\, 2^{-n(H(X)+\epsilon) - n(H(Y)+\epsilon) - n(H(Z)+\epsilon)} = (1-\epsilon)\, 2^{nH(X,Y,Z) - n\epsilon -}$$

$$^{- nH(X) - n\epsilon - nH(Y) - n\epsilon - nH(Z) - n\epsilon} = (1-\epsilon)\, 2^{n[H(X,Y,Z) - H(X) - H(Y) - H(Z)] - 4n\epsilon}$$

$$\Pr\left[(\tilde{X}_n, \tilde{Y}_n, \tilde{Z}_n) \in A_\epsilon^n\right] \geq (1-\epsilon)\, 2^{n[H(X,Y,Z) - H(X) - H(Y) - H(Z)] - 4n\epsilon}$$

$$\Pr\left[(\tilde{X}_n, \tilde{Y}_n, \tilde{Z}_n) \in A_\epsilon^n\right] \geq (1-\epsilon)\, 2^{-n[-H(X,Y,Z) + H(X) + H(Y) + H(Z) + 4\epsilon]}$$

$$-H(X,Y,Z) + H(X) + H(Y) + H(Z)$$

$$\lambda_i = \Pr\left[g(Y^n) \neq i \mid X^n = X^n(i)\right] = \sum_{y^n} P(y^n | X^n_i)\, I\left[g(y_i) \neq i\right]$$

PROBLEM 7.15 REVISITED (g)

$\Pr = 4.6 \cdot 10^{-4}$ ANY SEQUENCE TO BE JOINTLY TYPICAL WITH Y^n. THE PROBABILITY THAT NONE OF 511 CODEWORDS ARE JOINTLY TYPICAL WITH RECEIVED SEQUENCE IS THEREFORE:

$$\left(1 - \underbrace{4.552 \cdot 10^{-4}}_{\Pr}\right)^{511} = 0.792$$

THE PROBABILITY THAT AT LEAST ONE IS JOINTLY TYPICAL WITH RECEIVED SEQUENCE IS:

$1 - 0.792 = \underline{0.208}$

→ PODUDNA T.E. POETNICTA KORIŠTENJA ČO UNION BOUND (0.233)

CHAPTER 8 (CONTINUE FROM N.16c)

$h(\phi_K) \geq h(g)$ ϕ_K – GAUSSIAN DENSITY $N(0,K)$

IN PARTICULAR GAUSSIAN DISTRIBUTION MAXIMIZES THE ENTROPY OVER ALL DISTRIBUTIONS WITH SAME VARIANCE. THIS LEADS TO THE ESTIMATION COUNTERPART TO FANO'S INEQUALITY. LET X BE A RANDOM VARIABLE WITH DIFFERENTIAL ENTROPY $h(X)$. LET $\hat{X}$ BE AN ESTIMATE OF X, AND LET $E(X-\hat{X})^2$ BE THE EXPECTED PREDICTION ERROR. LET $h(X)$ BE IN NATS.

RECALL FANO INEQUALITY: $H(X|\hat{X}) \leq 1 + P_e \, ld|X|$

$x \in \{0,1\}$ $\hat{x} \in \{0,1\}$

$H(X|\hat{X}) = P(\hat{X}=0) \cdot H(X|\hat{X}=0) + P(\hat{X}=1) H(X|\hat{X}=1)$

$= -P(\hat{X}=0)\left[P(X=0|\hat{X}=0) \cdot ld\, P(X=0|\hat{X}=0) + P(X=1|\hat{X}=0) \cdot ld\, P(X=1|\hat{X}=0)\right] - P(\hat{X}=1)\left[P(X=0|\hat{X}=1)\, ld\, P(X=0|\hat{X}=1) + P(X=1|\hat{X}=1) \cdot ld\, P(X=1|\hat{X}=1)\right]$

$P_e = P(X=0) \cdot P(\hat{X}=1|X=0) + P(X=1) \cdot P(\hat{X}=0|X=1)$

$H(X,\hat{X}) = H(X) + H(\hat{X}|X) = H(\hat{X}) + H(X|\hat{X})$

$E = \begin{cases} 1 & \hat{X} \neq X \\ 0 & \hat{X} = X \end{cases}$ $P_e = P_r(\hat{X} \neq X)$

$H(E,X|\hat{X}) = H(E|\hat{X}) + \underbrace{H(X|E\hat{X})} = P(\hat{X}=0) \cdot H(E|\hat{X}=0) + P(\hat{X}=1) \cdot H(E|\hat{X}=1) = -P(\hat{X}=0) \cdot \left[P(E=1|\hat{X}=0) \cdot ld\, P(E=1|\hat{X}=0) + P(E=0|\hat{X}=0) \cdot ld\, P(E=0|\hat{X}=0)\right]$

$H(E,X|\hat{X}) = H(X|\hat{X}) + \underbrace{H(E|X,\hat{X})}_{0} = H(X|\hat{X}) =$

$= H(E|\hat{X}) + H(X|E\hat{X}) \leq H(P_e) + P_e \, ld|X|$

$\begin{cases} H(E|\hat{X}) \leq H(E) = H(P_e) \\ H(X|E\hat{X}) = \underbrace{P(E=0)}_{1-P_e} \cdot \underbrace{H(X|E=0,\hat{X})}_{0} + P(E=1) \cdot H(X|E=1,\hat{X}) \end{cases}$

$= P_e \cdot H(X|E=1,\hat{X}) \leq P_e \cdot H(X) \leq P_e \, ld|X|$

$H(X|\hat{X}) \le \overset{H(e)\le 1}{1} + P_e \operatorname{ld}|X|$ $\quad e = \{1, 0\} \quad P(e) = \{P_e, 1-P_e\}$

$H(e, X|\hat{X}) = H(X|\hat{X}) + \underbrace{H(e|X,\hat{X})}_{0} = \underbrace{H(e|\hat{X})}_{\le H(e) = P(P_e)} + H(X|e\hat{X}) =$

$H(X|e\hat{X}) = P(e=0)\cdot \underbrace{H(X|e=0,\hat{X})}_{0} + \underbrace{P(e=1)}_{P_e} H(X|e=1,\hat{X})$

$H(X|e\hat{X}) \le P_e H(X) \le P_e \operatorname{ld}|X|$

$H(X|\hat{X}) \le H(P_e) + P_e \operatorname{ld}|X| \le 1 + P_e \operatorname{ld}|X|$

$X \to Y \to \hat{X} \qquad I(X;Y) \ge I(X;\hat{X})$

$I(X;Y) \ge I(Y;\hat{X})$

$I(X;Y) = H(X) - H(X|Y) \ge I(X;\hat{X}) = H(X) - H(X|\hat{X})$

$H(X|Y) \le H(X|\hat{X}) \le 1 + P_e \operatorname{ld}|X|$

$\boxed{H(X|Y) \le 1 + P_e \operatorname{ld}|X|} \le 1 + P_e \operatorname{ld}(|X|-1)$

THEOREM 8.6.6 (ESTIMATION ERROR AND DIFFERENTIAL ENTROPY)

FOR ANY RANDOM VARIABLE X AND ESTIMATOR $\hat{X}$,

$$\boxed{E(X-\hat{X})^2 \ge \frac{1}{2\pi e} e^{2h(X)}}$$

WITH EQUALITY IF AND ONLY IF X IS GAUSSIAN AND $\hat{X}$ IS THE MEAN OF X.

PROOF: $E(X-\hat{X})^2 \ge \underbrace{\min_{\hat{X}} E[X-\hat{X}]^2 = E[X-E(X)]^2}_{(*)}$

$= \operatorname{Var}(X) \ge \frac{1}{2\pi e} e^{2h(X)}$

$h(X) \le h(\phi) = \frac{1}{2}\ln(2\pi e \sigma^2)$ nats $\qquad h(X) = \frac{1}{2}\ln(2\pi e)^n|K|$ nats

$\boxed{\sigma^2 \ge \frac{1}{2\pi e} e^{2h(X)}}$

$(*) \Rightarrow$ FOLLOWS FROM THE FACT THAT THE MEAN OF X IS THE BEST ESTIMATOR FOR X.

COROLLARY. GIVEN SIDE INFORMATION Y AND ESTIMATOR $\hat{X}(Y)$ IT FOLLOWS THAT

$$\boxed{E(X-\hat{X}(Y))^2 \ge \frac{1}{2\pi e} e^{2h(X|Y)}} \qquad h(X) \ge$$

(133)

SUMMARY

$h(X) = h(f) = -\int_S f(x) \operatorname{ld} f(x)\, dx$

$f(X^n) \doteq 2^{-nh(X)}$

$Vol(A_\varepsilon^{(n)}) \doteq 2^{nh(X)}$

$H([X]_{2^{-n}}) \doteq h(X) + n$

$h(\mathcal{N}(0,\sigma^2)) = \frac{1}{2} \operatorname{ld} 2\pi e \sigma^2$

$h(\mathcal{N}(\mu, K)) = \frac{1}{2} \operatorname{ld}[(2\pi e)^n |K|]$

$D(f\|g) = \int f \operatorname{ld} \frac{f}{g} \geq 0$

$h(X_1, X_2, \dots, X_n) = \sum_{i=1}^{n} h(X_i | X_1^{i-1})$

$h(X|Y) \leq h(X)$

$h(aX) = h(X) + \operatorname{ld}|a|$

$I(X;Y) = \int f(x,y) \operatorname{ld} \frac{f(x,y)}{f(x) \cdot f(y)} \geq 0$

$\max_{EXX^T = K} h(\mathbf{X}) = \frac{1}{2} \operatorname{ld} (2\pi e)^n |K|$

$E[X - \hat{X}(Y)]^2 \geq \frac{1}{2\pi e} e^{2h(X|Y)}$

$2^{nH(X)}$ → IS THE EFFECTIVE ALPHABET SIZE OF DISCRETE RANDOM VARIABLE.

$2^{nh(X)}$ → IS THE EFFECTIVE SUPPORT SET SIZE FOR A CONTINUOUS RANDOM VARIABLE.

2^C → IS THE EFFECTIVE ALPHABET SIZE OF A CHANNEL OF CAPACITY „C".

CONTINUE FROM N°6a

(c) $f(x_1) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}}$ $\quad f(x_2) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}}$

$x = x_1 + x_2$

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

~~$\mu = \mu_1 + \mu_2$~~
~~$\sigma = \sigma_1 + \sigma_2$~~

$h(x) = \frac{1}{2} \operatorname{ld} 2\pi e \sigma^2$

~~$\mu = \mu_1 + \mu_2$~~
~~$\sigma^2 = \sigma_1^2 + \sigma_2^2$~~

ТЕОРИЯ НА ИЗПИТИ!!!

$$h(x) = \frac{1}{2} \operatorname{ld}\left[2\pi e(\sigma_1^2+\sigma_2^2)\right]$$

$$p(x) = \frac{dF(x)}{dx} \qquad F(x) = \int_{-\infty}^{x} p(t)\,dt$$

$$M(s) = \int p(x) e^{-sx} dx \qquad M(s) = \mathcal{L}[p(x)]$$

$$\mathcal{L}\left[\frac{dF(x)}{dx}\right] = M(s) \qquad \mathcal{L}\left[\frac{dF(x)}{dx}\right] = s\,\mathcal{L}[F(x)]$$

$$s\hat{F}(x) = M(s) \qquad \hat{F}(x) = \frac{M(s)}{s}$$

$$F(x) = \mathcal{L}^{-1}\left[\frac{M(s)}{s}\right] \qquad p(x) = \frac{1}{(\sqrt{2\pi})^n |K|^{1/2}} e^{-\frac{x^T K^{-1} x}{2}}$$

PROBLEM 8.2 CONCAVITY OF DETERMINANTS. LET K_1 AND K_2 BE TWO SYMMETRIC NON-NEGATIVE DEFINITE $n \times n$ MATRICES. PROVE THE RESULT OF KY FAN [199]:

$$|\lambda K_1 + \bar{\lambda} K_2| \geq |K_1|^{\lambda} |K_2|^{\bar{\lambda}} \quad \text{FOR } 0 \leq \lambda \leq 1;\ \bar{\lambda} = 1-\lambda;$$

WHERE $|K|$ DENOTES THE DETERMINANT OF K.
[HINT: LET $Z = X_\theta$, $X_1 \sim N(0, K_1)$, $X_2 \sim N(0, K_2)$ AND θ = BERNOULLI (λ). THEN USE $h(Z|\theta) \leq h(Z)$]

$$h(Z|\theta) = p(\theta=1) \cdot h(Z|\theta=1) + p(\theta=0)\, h(Z|\theta=2) =$$
$$= \lambda h(Z|\theta=1) + (1-\lambda) h(Z|\theta=2) = \lambda h(X_1) + (1-\lambda) h(X_2)$$

$$h(X_1) = \frac{1}{2} \operatorname{ld}(2\pi e)^n |K_1| \qquad h(X_2) = \frac{1}{2} \operatorname{ld}(2\pi e)^n |K_2|$$

CONCAVITY OF RELATIVE ENTROPY

$$D(\lambda p_1 + (1-\lambda) q_1 \| \lambda p_2 + (1-\lambda) q_2) \leq \lambda D(p_1 \| p_2) + (1-\lambda) D(q_1 \| q_2)$$ CONVEXITY OF RELATIVE ENTROPY

$$H(\lambda p_1 + (1-\lambda) p_2) \geq \lambda H(p_1) + (1-\lambda) H(p_2)$$ CONCAVITY OF ENTROPY!

$Z = X_\theta$ $\quad H(Z) \geq H(Z|\theta) = p(\theta=1) \cdot H(Z|\theta=1) + p(\theta=2) \cdot H(Z|\theta=2)$

$Z = \begin{cases} X_1 & p=\lambda \\ X_2 & q=1-\lambda \end{cases}$

$$H(Z) \geq \lambda H(X_1) + (1-\lambda) H(X_2)$$

$$\lambda h(X_1) + (1-\lambda) h(X_2) = \frac{\lambda}{2} \operatorname{ld}(2\pi e)^n |K_1| + \frac{(1-\lambda)}{2} \operatorname{ld}(2\pi e)^n |K_2|$$
$$= \frac{\lambda}{2} \operatorname{ld}(2\pi e)^n [|K_1| / |K_2|] + \frac{1}{2} \operatorname{ld}(2\pi e)^n |K_2| \leq h(Z)$$

(135) $h(Z) = ?$

$$Z = \begin{cases} X_1 & \lambda \\ X_2 & 1-\lambda \end{cases} \Rightarrow$$ DISJOINT MIXTURE

$$Z = \underbrace{\lambda \cdot X_1}_{\tau_1} + \underbrace{(1-\lambda) X_2}_{\tau_2} \qquad E[Z] = E[\lambda X_1] + E[(1-\lambda) X_2]$$

$$E[Z] = \lambda \mu_1 + (1-\lambda)\mu_2$$

$$\sigma_1 = \int \lambda x_1 \, p(\tau_1)\, d\tau_1 \qquad p(\tau_1) = ?$$

TREBA SO JACOBIAN pr. 138

$$\boxed{\tau_1 = \lambda \cdot X_1}$$

$$p(\tau_1) = \left. \frac{p(x_1)}{\frac{d\tau_1}{dx_1}} \right|_{x = \frac{\tau_1}{\lambda}}$$

$$p(\tau_1) = \frac{p\left(\frac{\tau_1}{\lambda}\right)}{\lambda} = \frac{1}{\lambda} \frac{1}{(2\pi)^n \sqrt{|K|}} e^{-\left(\frac{\tau_1}{\lambda} - \mu_1\right)^T K^{-1} \left(\frac{\tau_1}{\lambda} - \mu_1\right)}$$

$$h(Z) \geq h(Z|\theta) = \lambda h(X_1) + (1-\lambda) h(X_2) = \frac{\lambda}{2} ld(2\pi e)^n |K_1| + \frac{1}{2} ld(2\pi e)^n |K_2|$$

$$h(Z) \geq \frac{\lambda}{2} ld (2\pi e)^n \frac{|K_1|}{|K_2|} + \frac{1}{2} ld (2\pi e)^n |K_2|$$

$$p(\tau_1) = \frac{1}{\lambda} \frac{1}{(2\pi)^n \sqrt{|K|}} e^{-\frac{1}{2\lambda^2}(\tau_1 - \lambda\mu_1)^T K^{-1} (\tau_1 - \lambda \mu_1)}$$

$$p(\tau_1) = \frac{1}{(2\pi)^n \sqrt{|K| \cdot \lambda^2}} e^{-\frac{1}{2}(\tau_1 - \lambda\mu_1)^T (\lambda^2 K)^{-1} (\tau_1 - \lambda\mu_1)}$$

$$p(\tau_1) = \mathcal{N}(\lambda\mu_1, \lambda^2 |K|) \qquad p(\tau_2) = \mathcal{N}[(1-\lambda)\mu_1, (1-\lambda)^2 |K|]$$

$$\boxed{Z = \tau_1 + \tau_2} \qquad p(Z) = \mathcal{N}\left[E[\tau_1] + E[\tau_2];\ var(\tau_1) + var(\tau_2)\right]$$

$$\boxed{p(Z) = \mathcal{N}\left[\lambda\mu_1 + (1-\lambda)\mu_2;\ \lambda^2 |K_1| + (1-\lambda)^2 |K_2|\right]} \ (*)$$

$$\boxed{h(Z) = \frac{1}{2} ld \left[(2\pi e)^n \cdot (\lambda^2 |K_1| + (1-\lambda)^2 |K_2|)\right]}$$

$$\frac{1}{2} ld\left[(2\pi e)^n (\lambda^2 |K_1| + (1-\lambda)^2 |K_2|)\right] \geq \frac{\lambda}{2} ld (2\pi e)^n |K_1| + \frac{1-\lambda}{2} ld (2\pi e)^n |K_2|$$

$$\frac{1}{2} ld\left[(2\pi e)^n (\lambda^2 |K_1| + (1-\lambda)^2 |K_2|)\right] \geq \frac{1}{2} ld \left[(2\pi e)^n |K_1|\right]^\lambda \cdot \left[(2\pi e)^n |K_2|\right]^{\bar\lambda}$$

$$(2\pi e)^n (\lambda^2 |K_1| + (1-\lambda)^2 |K_2|) \geq (2\pi e)^{n\lambda} |K_1|^\lambda \cdot (2\pi e)^{n\bar\lambda} |K_2|^{\bar\lambda}$$

$\bar\lambda = 1-\lambda$

136

$$(2\pi e)^4 \left(\lambda^2 |K_1| + (1-\lambda)^2 |K_2|\right) \geq (2\pi e)^{\ldots} \cdot |K_1|^{\lambda} \cdot |K_2|^{\bar{\lambda}}$$

$$\left(\lambda^2 |K_1| + (1-\lambda)^2 |K_2|\right) \geq |K_1|^{\lambda} \cdot |K_2|^{\bar{\lambda}}$$

$$\left|\lambda K_1 + (1-\lambda) K_2\right| \qquad \begin{bmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{bmatrix} + \begin{bmatrix} (1-\lambda)e & (1-\lambda)f \\ (1-\lambda)g & (1-\lambda)h \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda a + (1-\lambda)e \ ; & \lambda b + (1-\lambda)f \\ \lambda c + (1-\lambda)g \ ; & \lambda d + (1-\lambda)h \end{bmatrix} = \begin{matrix} \lambda^2 ad + \lambda a(1-\lambda)h + \lambda d(1-\lambda)e + (1-\lambda)^2 eh \\ - \lambda^2 bc - \lambda b(1-\lambda)g - \lambda c(1-\lambda)f - (1-\lambda)^2 fg \end{matrix}$$

$$= \lambda^2 \underbrace{(ad - bc)}_{|K_1|} + (1-\lambda)^2 \underbrace{(eh - fg)}_{|K_2|} + \lambda(1-\lambda)\underbrace{[ah + de - bg - cf]}_{①}$$

POSITIVE DEFINITE

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a+c \ ; & b+d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = a+c+b+d > 0$$

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} a+2c \ ; & b+2d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = a+2c+b+2d >$$

$$\begin{bmatrix} h & e \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} h \\ e \end{bmatrix} = \begin{bmatrix} h & e \end{bmatrix}\begin{bmatrix} ah + be \\ ch + de \end{bmatrix} = ah^2 + beh + che + de^2 > 0$$

$$ah^2 + beh > -che - de^2$$

$$h(ah + be) > -e(ch + de)$$

IF: ① $ah + de - bg - cf > 0$ i.e. $ah + de > bg + cf$

$$\Rightarrow \underbrace{\lambda^2 |K_1| + (1-\lambda)^2 |K_2| + \lambda(1-\lambda) \cdot ①}_{*} \geq \underbrace{\lambda^2 |K_1| + (1-\lambda)^2 |K_2|}_{\geq |K_1|^{\lambda} |K_2|^{\bar{\lambda}} \ (**)}$$

$$(*), (**) \Rightarrow \left|\lambda K_1 + (1-\lambda) K_2\right| \geq |K_1|^{\lambda} |K_2|^{\bar{\lambda}}$$

(BERKELEY SOLUTIONS (sol_4_22?april07.pdf))

$$E[ZZ^T] = E\left[E[ZZ^T] \mid \theta\right] = E\left[|K_1| \mid \theta = 1 + |K_2| \mid \theta = 2\right] =$$

$$\lambda |K_1| + \bar{\lambda} |K_2| = E\left[E[X_1 X_1^T] \mid \theta = 1 + E[X_2 X_2^T] \mid \theta = 2\right]$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + bc & ab + bd \\ ca + dc & cb + d^2 \end{bmatrix} \qquad \begin{bmatrix} a & e \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2 + bb & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix}$$

(137)

$|\lambda K_1 + (1-\lambda)K_2| = \lambda^2(ad-bc) + (1-\lambda)^2(eh-fg) + \lambda(1-\lambda)[ah+de-bg-cf]$

K_1 & K_2 ARE SYMETRIC $\Rightarrow$ $b=c$ $f=g$

$|\lambda K_1 + (1-\lambda)K_2| = \lambda^2|K_1| + (1-\lambda)^2|K_2| + \lambda(1-\lambda)[ah+de-g-g]$
$= \lambda^2(K_1) + (1-\lambda)^2 + \lambda(1-\lambda)[ah+de-2cg]$

$MGF(N(\mu,\sigma^2)) = e^{\mu t + \sigma^2 t^2/2}$ [LECTURE 21]

- $(X_1, X_2, \dots X_n)$ ARE JOINTLY GAUSSIAN

$M(t_1, t_2, \dots t_n) = \exp(t_1\mu_1 + \dots + t_n\mu_n)\exp\left(\frac{1}{2}\sum_{i,j=1}^{n} t_i a_{ij} t_j\right)$

$M(t_1^n) = \exp(\underline{t}'\underline{\mu})\exp\left(\frac{1}{2}\underline{t}'A\cdot\underline{t}\right)$ $\underline{a}$ — COLUMN VECTOR

21.2 THEOREM JOINT GAUSSIAN RANDOM VARIABLES ARISE FROM NON-SINGULAR LINEAR TRANSFORMATIONS ON INDEPENDENT NORMAL RANDOM VARIABLES

PROOF: $X_1, \dots X_n$ BE INDEPENDENT WITH $X_i \sim N(0, \lambda_i)$
$\underline{X} = (X_1, \dots X_n)'$. LET $\underline{Y} = B\underline{X} + \mu$ WHERE B IS NONSINGULAR THEN $\underline{Y}$ IS GAUSSIAN.

$M_{\underline{Y}}(\underline{t}) = E[\exp(t'Y)] = E[\exp(t'BX)]\exp(\underline{t}'\mu)$.

$E[\exp(\mu'X)] = \prod_{i=1}^{n} E[\exp(\mu_i X_i)] = \exp\left(\sum_{i=1}^{n}\lambda_i \mu_i^2/2\right) = \exp\left(\frac{1}{2}\mu' D\mu\right)$

$[\mu_1\ \mu_2]\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} = [\lambda_1\mu_1\ \lambda_2\mu_2]\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} = \lambda_1\mu_1^2 + \lambda_2\mu_2^2$

$\mu = B't$ $\mu' = t'B$

$M_{\underline{Y}}(\underline{t}) = \exp(\underline{t}'\mu)\cdot\exp\left(\frac{1}{2}t'BDB't\right)$ BDB' IS SYMETRIC AND POSITIVE DEFINITE MATRIX $\Rightarrow$ Y IS GAUSSIAN.

$M_{\underline{Y}}(\underline{t}) = \exp(\underline{t}'\mu)\cdot\exp\left(\frac{1}{2}t'At\right)$

• DA SO ПОКАЖЕ ИЗВЕДОТ Н 135 ⊛ ЗА ДВУДИМЕНЗ.

$$f(x) = \frac{1}{\sqrt{(2\pi)^3}|K|^{1/2}} e^{-\frac{X^T \cdot K^{-1} \cdot X}{2}}$$

$$|K| = \sigma_1^2 \sigma_2^2 - \sigma_{12}^2$$

$$K = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \sigma_{12} = Cov(x_1 x_2)$$

$$f(x_1 x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{x_1^2}{\sigma_1^2} - 2\rho\frac{x_1}{\sigma_1}\frac{x_2}{\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right]\right)$$

$$|K| = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = \sigma_1^2\sigma_2^2(1-\rho^2)$$

$$Y = \lambda \cdot X \qquad f(Y) = ?$$

$$Y_1 = \lambda x_1 \quad Y_2 = \lambda x_2 \qquad J = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{vmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \lambda^2$$

$$f(Y) = \frac{f(x)}{J}\Big|_{x = \frac{y}{\lambda}}$$

⊛ $$f(Y_1 Y_2) = \frac{1}{\lambda^2}\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{Y_1^2}{\lambda^2\sigma_1^2} - 2\rho\frac{Y_1Y_2}{\lambda^2\sigma_1\sigma_2} + \frac{Y_2^2}{\lambda^2\sigma_2^2}\right]\right)$$

$$\sqrt{|K_Y|} = \lambda^2\sigma_1\sigma_2\sqrt{1-\rho^2} \qquad |K_Y| = \lambda^4\sigma_1^2\sigma_2^2(1-\rho^2)$$

$$K_Y = \lambda K_X \; ?$$

$$K_Y = \lambda\begin{vmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{vmatrix} = \begin{bmatrix} \lambda\sigma_1^2 & \lambda\sigma_{12} \\ \lambda\sigma_{12} & \lambda\sigma_2^2 \end{bmatrix} = \begin{bmatrix} \lambda\sigma_1^2 & \lambda\rho\sigma_1\sigma_2 \\ \lambda\rho\sigma_1\sigma_2 & \lambda\sigma_2^2 \end{bmatrix}$$

$$|K_Y| = \lambda^2(\sigma_1^2\sigma_2^2 - \sigma_{12}^2) = \lambda^2(\sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2) = \lambda^2\sigma_1^2\sigma_2^2(1-\rho^2)$$

$$K_Y^{-1} = \frac{1}{\lambda^2\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix} \lambda\sigma_2^2 & -\lambda\rho\sigma_1\sigma_2 \\ -\lambda\rho\sigma_1\sigma_2 & +\lambda\sigma_1^2 \end{bmatrix} \qquad \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

(W)

$$\frac{1}{(W)}\begin{bmatrix} Y_1 & Y_2 \end{bmatrix}\begin{bmatrix} \lambda\sigma_2^2 & -\lambda\rho\sigma_1\sigma_2 \\ -\lambda\rho\sigma_1\sigma_2 & \lambda\sigma_1^2 \end{bmatrix}\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \frac{1}{(W)}\begin{bmatrix} \lambda\sigma_2^2 Y_1 - \rho\lambda\sigma_1\sigma_2 Y_2 \\ -\lambda\rho\sigma_1\sigma_2 Y_1 + \lambda\sigma_1^2 Y_2 \end{bmatrix}$$

$$= \frac{1}{(W)}(\lambda\sigma_2^2 Y_1^2 - \rho\lambda\sigma_1\sigma_2 Y_1Y_2 - \lambda\rho\sigma_1\sigma_2 Y_1Y_2 + \lambda\sigma_1^2 Y_2^2) =$$

$$= \frac{\lambda\sigma_2^2 Y_1^2 - 2\rho\lambda\sigma_1\sigma_2 Y_1Y_2 + \lambda\sigma_1^2 Y_2^2}{2\lambda^2\sigma_1^2\sigma_2^2(1-\rho^2)} = \frac{1}{2(1-\rho^2)}\left[\frac{Y_1^2}{\lambda\sigma_1^2} + \frac{2Y_1Y_2}{\lambda\sigma_1\sigma_2} + \frac{Y_2^2}{\lambda\sigma_2^2}\right]$$

(139)

$$\begin{bmatrix} \frac{y_1}{\lambda} & \frac{y_2}{\lambda} \end{bmatrix} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} \begin{bmatrix} \frac{y_1}{\lambda} \\ \frac{y_2}{\lambda} \end{bmatrix} = \quad |K| = \sigma_1^2\sigma_2^2(1-\rho^2) =$$

$$= \frac{1}{2|K|}\begin{bmatrix} \frac{\sigma_2^2 y_1}{\lambda} - \frac{\rho\sigma_1\sigma_2 y_2}{\lambda} & -\frac{\rho\sigma_1\sigma_2 y_1}{\lambda} + \frac{\sigma_1^2 y_2}{\lambda} \end{bmatrix} \begin{bmatrix} \frac{y_1}{\lambda} \\ \frac{y_2}{\lambda} \end{bmatrix} =$$

$$= \frac{1}{2|K|}\left[\frac{\sigma_2^2 y_1^2}{\lambda^2} - \frac{\sigma_1\sigma_2 y_1 y_2}{\lambda^2} - \frac{\rho\sigma_1\sigma_2 y_1 y_2}{\lambda^2} + \frac{\sigma_1^2 y_2^2}{\lambda^2}\right]$$

$$= \frac{1}{2(1-\rho^2)}\left[\frac{y_1}{\lambda^2\sigma_1^2} - \frac{2y_1 y_2}{\sigma_1\sigma_2\lambda^2} + \frac{y_2^2}{\sigma_2^2\lambda^2}\right]$$

IST KAKO 139 (7)

$$= \quad \sigma_{y_1} = \lambda \sigma_{x_1} \quad \sigma_{y_2} = \lambda\sigma_{x_2} =$$

$\sigma_{x_1} = \sigma_1$, $\sigma_{x_2} = \sigma_2$

$$= \frac{1}{2(1-\rho^2)}\left[\frac{y_1}{\sigma_{y_1}^2} - \frac{2y_1 y_2}{\sigma_{y_1}\sigma_{y_2}} + \frac{y_2^2}{\sigma_{y_2}^2}\right]$$

$$K_y = \begin{bmatrix} \sigma_{y_1}^2 & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma_{y_2}^2 \end{bmatrix} = \lambda^2 \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$|K_y| = \lambda^4\sigma_1^2\sigma_2^2 - \lambda^4\sigma_1^2\sigma_2^2\rho^2 = \lambda^4(1-\rho^2)\sigma_1^2\sigma_2^2$$

$$f_Y = \frac{1}{(\sqrt{2\pi})^2\sqrt{|K_y|}} e^{-y^T K_y^{-1} y} = \frac{1}{\lambda^2 2\pi\sigma_{1x}\sigma_{2x}\sqrt{1-\rho^2}} e^{-\frac{1}{\lambda^2} y^T K_y^{-1} y}$$

$$|K_y| = \lambda^4 |K_x|$$

ZNAČI AKO: $Z = \lambda X_1 + (1-\lambda)X_2$, X_1, X_2
NE SE DISJOINT (FAZA BOGU NE E ZA DISJOINT. TUKA LEZI LEJBATOTO NA PRODLEMOT.

$$Z = \lambda X_1 + (1-\lambda)X_2 \rightarrow$$

$$f(Z) = N\left(\lambda\mu_1 + (1-\lambda)\mu_2,\ \lambda^4|K_1| + (1-\lambda)^4|K_2|\right)$$

~~SEKAK~~ ~~ZADAČATA~~ E ~~ZA~~ DISJOINT $Z = \begin{cases} X_1 & \lambda \\ X_2 & 1-\lambda \end{cases}$

PROBLEM 8.7 DIFFERENTIAL ENTROPY (CONTINUE FROM N16c) BOUND ON DISCRETE ENTROPY

$X \in \{a_1, a_2, \dots a_i \dots\}$ $p(X) = \{p_1\ p_2\ \dots\}$ $(H(X) = H(X'))$

$X' \in \{1, 2, \dots i, \dots\}$ $p(X') = \{p_1\ p_2\ \dots\}$

$Y = X' + U$ $U \in [0 \dots 1]$

$0 \dots 0.5$	$\to 0$
$0.5 \dots 1$	$\to 1$

$I(X';Y) = \underbrace{H(X')}_{=H(X)} - H(X'|Y) = H(Y) - H(Y|X') = H(Y) - H(U)$ (*)

$Y \in [1 \vee 2;\ 2 \vee 3;\ \dots,\ i \vee i+1,\ \dots]$ $Y = \{Y_1, Y_2, \dots Y_i \dots\}$

$H(Y) = \lim_{n\to\infty} \frac{1}{n} H(Y_1, Y_2 \dots Y_n) = \frac{n}{n} H(Y_i) = H(Y_i)$

$H(Y_i) = ?$ $Y_i = \begin{cases} X' & u = 0 \\ X'+1 & u = 1 \end{cases}$ $H(Y_i) = \frac{1}{2}\operatorname{ld} 2 + \frac{1}{2}\operatorname{ld} 2 = 1$

$\boxed{H(Y) = 1}$ (*) $\Rightarrow$ $H(X) - H(X'|Y) = H(Y) = 1$

$H(X) = 1 + H(X'|Y)$; $H(X'|Y) = P(Y_i = i) \cdot H(X'|Y_i = i) +$

$+ P(Y_i = i+1)\ H(X'|Y_i = i+1) = \frac{1}{2}1 + \frac{1}{2}\cdot 1 = 1$

$\boxed{H(X) = 1 + 1 = 2}$

$E[(X - \hat{X})^2] \geq \min (X - \hat{X})^2 = (X - \bar{X})^2 = \operatorname{var} X = \sigma^2$

$h(X) \leq \frac{1}{2}\log(2\pi e)\sigma^2$ $\boxed{\sigma^2 \geq \frac{e^{2h(X)}}{(2\pi e)}}$

$E[(X - \bar{X})^2] \geq \frac{e^{2h(X)}}{(2\pi e)}$ $\underline{h(X|Y) = h(X'|Y)}$

$E[(X - \hat{X}(Y))^2] \geq \frac{e^{2h(X|Y)}}{2\pi e} = \frac{e^2}{2\pi e} \geq \frac{e}{2\pi}$.

$E[(X - \bar{X})^2] = E[X^2] - \bar{X}^2 = E[X^2] - E[X]^2 = \sum_{i=1}^{\infty} i^2 p_i - \left(\sum_{i=1}^{\infty} i p_i\right)^2$

$E[X] = \sum_{i=1}^{\infty} a_i p_i = a_1 p_1 + a_2 p_2 + \dots$

$E[X'] = \sum_{i=1}^{\infty} i\, p_i = 1\cdot p_1 + 2\cdot p_2 + \dots$

$H(X) = H(X') \leq \frac{1}{2}\operatorname{ld} 2\pi e\left[\sum_{i=1}^{\infty} i^2 p_i - \left(\sum_{i=1}^{\infty} i p_i\right)^2\right]$

(141) $Y = \underline{X'} + U$ $\qquad E[(Y - \bar{Y})^2] = ?$

$Y = f(X', U) = f(t) \qquad f(X', U) = \int p(t-x) \cdot q(x)\, dx \quad , 1 = 0 \ldots 1$

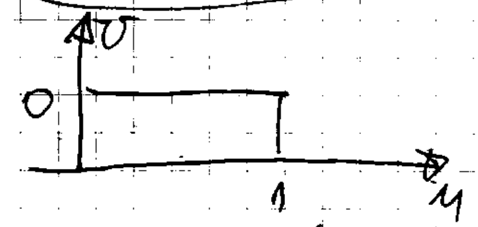
$f(t) = \int_0^1 p(t-x)\, dx = 1$

$p(X') = \{p_1, p_2, p_3, \ldots p_n \ldots \infty\}$

$H(U) + ld\Delta = h(U) \qquad H(U^\Delta) = -ld\Delta$

$\qquad\qquad\qquad\qquad\qquad H(U^\Delta) = -ld(2^n)^{-1}$

$\boxed{H(U^\Delta) = ld \frac{1}{2^n} = n}$



$var(U) = \int (u - \bar{u})^2 p(u)\, du$

$\bar{u} = \int_0^1 u\, p(u)\, du = \frac{u^2}{2}\Big|_0^1 = \frac{1}{2}$

$var(U) = \int_0^1 \left(u - \frac{1}{2}\right)^2 du = \int_0^1 \left(u^2 - u + \frac{1}{4}\right) du = \left(\frac{u^3}{3} - \frac{u^2}{2} + \frac{1}{4}u\right)\Big|_0^1 =$

$= \left(\frac{1}{3} - \frac{1}{2} + \frac{1}{4}\right) = \frac{4 - 6 + 3}{12} = \frac{7-6}{12} = \frac{1}{12}$

$E[(Y - \bar{Y})^2] = ?$ $\quad$ solus4.pdf

- BONUS SOLUTIONS: $Y = X' + U$ AND THEREFORE THE DISTRIBUTION OF Y HAS THE SHAPE OF A HISTOGRAM (i.e. $f_Y(y) = p_i$ for $i \le y \le i+1$)

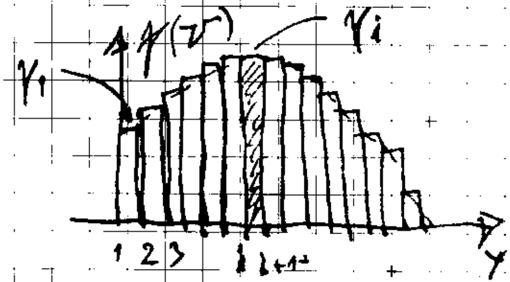
$H(X) = H(X')$ SINCE DISCRETE ENTROPIES DEPENDS ON THE PROBABILITIES AND NOT ON THE VALUES OF THE OUTCOMES.

$H(X') = -\sum_{i=1}^{\infty} p_i\, ld\, p_i = -\sum_{i=1}^{\infty} \left(\int_i^{i+1} f_Y(y)\, dy\right) ld \left(\int_i^{i+1} f_Y(y)\, dy\right)$ (CONSTANTS TO EVOLVE)

$= \sum_{i=1}^{\infty} c_i (i+1-i)\, ld\, c_i (i+1-i) = \sum_{i=1}^{\infty} c_i\, ld\, c_i =$

$= -\sum_{i=1}^{\infty} \int_i^{i+1} f_Y(y)\, ld\, f_Y(y)\, dy = -\int_1^{\infty} f_Y(y)\, ld\, f_Y(y)\, dy =$

$= h(Y)$ $\quad$ $H(X) = H(X') = h(Y)$



$h(Y) \le \frac{1}{2} \, ld(2\pi e)\sigma_Y^2 =$

$= \frac{1}{2} \, ld(2\pi e)\left(\sigma_{X'}^2 + \sigma_U^2\right) =$

$= \frac{1}{2} \, ld(2\pi e)\left[\sum_{i=1}^{\infty} i^2 p_i - \left(\sum_{i=1}^{\infty} i p_i\right)^2 + \frac{1}{12}\right]$

SINCE ENTROPY IS INVARIANT WITH RESPECT TO PERMUTATIONS $p_1, p_2, \dots$

$$h(Y) \le \frac{1}{2} \, ld(2\pi e)\left[\sum_{i=1}^{\infty} i^2 p_i - \left(\sum_{i=1}^{\infty} i p_i\right)^2 + \frac{1}{12}\right]$$

- PRESENTE E DATE $\sigma_Y^2 = \sigma_{X'}^2 + \sigma_U^2$?

SPOMED BIENAYME FORMULA FOR E TORNO

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$\iint_{-\infty}^{\infty} \left[(x+y) - (\overline{x+y})\right]^2 p(x,y)\,dx\,dy =$$

$$\iint_{-\infty}^{\infty} (x+y)^2 p(x,y)\,dx\,dy - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(x\bar{x} + x\bar{y} + y\bar{x} + y\bar{y}\right) p(x,y)\,dx\,dy$$

$$+ \iint (\overline{x+y})^2 p(x,y)\,dx\,dy$$

$$\left(Cov(X,Y) = E[(X-\mu_1)(Y-\mu_2)]\right)$$

$$= \iint x^2 p(x,y)\,dx\,dy + 2\iint xy\,p(x,y)\,dx\,dy - \dots (\overline{x+y}) \iint x\,p(x,y) + (\overline{x+y})$$

$$+ (\bar{x}+\bar{y})^2 = E[X^2] + E[Y^2] + Cov(X,Y) - \dots$$

$$- \bar{x}^2 - \bar{x}\bar{y} - \bar{x}\bar{y} - \bar{y}^2 + \bar{x}^2 + \bar{y}^2 + 2\bar{x}\bar{y} \quad 0 = E[X^2] + E[Y^2]$$

PROBLEM 8.9 (CONTINUE FROM NOTEBOOK) UIC SOLUTION

$X \to Y \to Z$ $\quad$ $(XY)\ \rho_1$ $\quad$ $(YZ)\ \rho_2$ $\quad$ $I(X;Z) = ?$

$I(X;Z) = h(X) - h(X|Z) = \big|\, h(X,Z) = h(Z) + h(X|Z) \,\big| =$

$= h(X) - h(X,Z) + h(Z)$

SINCE X, Y, Z ARE JOINTLY GAUSSIAN, HENCE

(14?) X AND Z ARE JOINTLY GAUSSIAN

$$K = \begin{bmatrix} \sigma_x^2 & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xz}\sigma_x\sigma_z & \sigma_z^2 \end{bmatrix} \qquad I(x;z) = h(x) + h(z) - h(x,z)$$

$$I(x;z) = \frac{1}{2}\,ld\,(2\pi e)\sigma_x^2 + \frac{1}{2}\,ld\,(2\pi e)\sigma_z^2 - \frac{1}{2}\,ld\,(2\pi e)^2 \cdot \sigma_x^2\sigma_z^2(1-\rho_{xz}^2) = \frac{1}{2}\,ld\,(2\pi e) + \frac{1}{2}\,ld\,(\sigma_x^2 \cdot \sigma_z^2) -$$

$$- \frac{1}{2}\,ld\,(2\pi e) - \frac{1}{2}\,ld\,\sigma_x^2\sigma_z^2 - \frac{1}{2}\,ld\,(1-\rho_{xz}^2) = -\frac{1}{2}\,ld\,(1-\rho_{xz}^2)$$

$\rho_{xz} = ?$ BY USING MARKOVITY:

$$p(x,z|y) = p(x|y)\,p(z|y)$$

$$p(yz|x) = \frac{p(x)p(y|x)\cdot p(z|xy)}{p(x)} = \frac{p(xy)\cdot p(z|y)}{p(x)}$$

$$p(yz|x) = \frac{p(xyz)}{p(x)} = \frac{p(xy)\cdot p(z|xy)}{p(x)} = \frac{p(x)\cdot p(y|x)\cdot p(z|y)}{p(x)}$$

$$p(yz|x) = p(y|x)\cdot p(z|y)$$

$$\rho_{xz} = \frac{E[x,z]}{\sigma_x\sigma_z} = \frac{E[(xz|y)]}{\sigma_x\sigma_z} = \frac{E[E[x|y]\,E[z|y]]}{\sigma_x\sigma_z}.$$

SINCE x, y, z ARE JOINTLY GAUSSIAN

$$E[x|y] = \frac{\sigma_x\rho_{xy}}{\sigma_y}y \qquad E[z|y] = \frac{\sigma_z\rho_{yz}}{\sigma_y}y$$

$$E[E[x|y]E[z|y]] = E\left[\frac{\sigma_x\rho_{xy}}{\sigma_y}E[y]\cdot\frac{\sigma_z\rho_{yz}}{\sigma_y}E[y]\right]$$

$$\rho_{xz} = \frac{1}{\sigma_x\sigma_z}\frac{\sigma_x\sigma_z\rho_{xy}\rho_{yz}}{\sigma_y^2}E[E[y]\cdot E[y]] = \frac{\sigma_x\sigma_z\rho_{xy}\rho_{yz}\,\sigma_y^2}{\sigma_y^2\,\sigma_x\sigma_z}$$

$$\rho_{xz} = \rho_{xy}\cdot\rho_{yz} \Rightarrow I(x;z) = -\frac{1}{2}\,ld\,(1-\rho_{xy}^2\cdot\rho_{yz}^2)$$
$$= -\frac{1}{2}\,ld\,(1-\rho_1^2\cdot\rho_2^2)$$

$$p(x,z|y) = \frac{p(x,y,z)}{p(y)} = \frac{p(xy)\cdot p(z|xy)}{p(y)} = \frac{p(y)p(x|y)p(z|y)}{p(y)}$$
$$= p(x|y)\,p(z|y)$$

$$E[(x,z)|z] = \iiint xz \, P(xz|z) \, dx\,dz$$

PROBLEM 8.10 SHAPE OF THE TYPICAL SET.

LET X_i BE i.i.d $\sim f(x)$ WHERE

$$f(x) = c\, e^{-x^4}$$

LET: $h = -\int f \ln f$. DESCRIBE THE SHAPE (OR FORM) OF THE TYPICAL SET: $A_\epsilon^{(n)} = \{x^n \in R^n : f(x^n) \in 2^{-n(h \pm \epsilon)}\}$

$$h(x) - \epsilon \le \frac{1}{n} \log^{-1} f(x) \le h(x) + \epsilon \qquad 2^{-n[h(x)+\epsilon]} \le f(x) \le 2^{-n[h(x)-\epsilon]}$$

$$2^{+n(h(x)-\epsilon)} \cdot (1-\epsilon) \le Vol\{A_\epsilon^{(n)}\} \le 2^{+n(h(x)+\epsilon)}$$

$$h(c) = \frac{1}{2} \frac{c \ln(c)\, \pi \sqrt{2}}{\Gamma(\frac{3}{4})} + \frac{1}{8} \frac{c\pi\sqrt{2}}{\Gamma(\frac{3}{4})} = \frac{1}{2} \frac{c\pi\sqrt{2}}{\Gamma(\frac{3}{4})} \left(\ln c + \frac{1}{4}\right)$$

$$h(c) = \frac{c\pi}{\sqrt{2}\,\Gamma(\frac{3}{4})} \ln\left(c\, e^{\frac{1}{4}}\right)$$

$$Vol\{A_\epsilon^{(n)}\} \doteq 2^{n h(x)}$$ — n - DIMENSIONAL VOLUME

$$\text{side length} = \left(2^{n h(x)}\right)^{\frac{1}{n}} = 2^{h(x)} = 2^{\frac{c\pi}{\sqrt{2}\,\Gamma(\frac{3}{4})} \ln(c e^{\frac{1}{4}})}$$

$$= \left(c\, \sqrt[4]{e}\right)^{\frac{c\pi}{\sqrt{2}\,\Gamma(\frac{3}{4})}} = \left(c^4 e\right)^{\frac{1}{4} \cdot \frac{c\pi}{\sqrt{2}\,\Gamma(\frac{3}{4})}}$$

- UNIVERSITY OF TORONTO SOLUTION:

$$f(x^n) = c^n e^{-(x_1^4 + x_2^4 + \cdots + x_n^4)}$$

$$2^{-n(h(x)+\epsilon)} \le c^n e^{-\sum_{i=1}^{n} x_i^4} \le 2^{-n(h(x)-\epsilon)} \quad / \ln$$

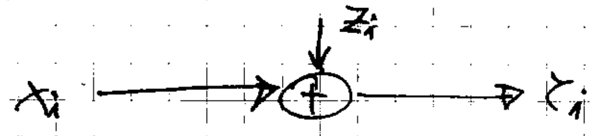
$$-n(h(x)+\epsilon)\ln 2 \le n \ln c - \sum_{i=1}^{n} x_i^4 \le -n(h(x)-\epsilon)\cdot \ln 2$$

$$A_\epsilon^{(n)} = \left\{x^n \in R^n \,\middle|\, n\left[(h(x)-\epsilon)\cdot \ln 2 + \ln c\right] \le \sum_{i=1}^{n} x_i^4 \le n[(h+\epsilon)\ln 2 + \ln c]\right\}$$

- THIS IS REMINISCENT OF THE FACT THAT TYPICAL SET OF A GAUSSIAN DISTRIBUTION IS A THIN SPHERICAL SHELL. IN THIS CASE, THE SPHERE HAS BEEN REPLACED BY A SHAPE OF FORM: $x_1^4 + x_2^4 + \cdots x_n^4 = r$, BUT THE "THIN SHELL" PROPERTY PERSISTS.

(144)

CHAPTER 9 - GAUSSIAN CHANNEL



$$Y_i = X_i + Z_i \qquad Z_i \sim \mathcal{N}(0,N)$$

WE ASSUME AN AVERAGE POWER CONSTRAINT. FOR ANY CODEWORD $(x_1, x_2, \ldots, x_n)$ TRANSMITTED OVER THE CHANNEL, WE REQUIRE THAT

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \le P$$

i.e. THE MOST COMMON LIMITATION ON THE INPUT IS ENERGY OR POWER CONST.

$$P_e = \frac{1}{2}P_r(Y<0 \mid X=\sqrt{P}) + \frac{1}{2}P_r(Y>0 \mid X=-\sqrt{P}) =$$

$$= \frac{1}{2}P_r(Z < -\sqrt{P} \mid X=\sqrt{P}) + \frac{1}{2}P_r(Z>\sqrt{P} \mid X=-\sqrt{P}) =$$

$$= P_r(Z>\sqrt{P}) = 1 - \int_{-\infty}^{\sqrt{P}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}} dz = \frac{1-\phi(\sqrt{P})}{2}$$

$$= 1 - \phi\left(\sqrt{\frac{P}{N}}\right) \qquad \phi(x) = \int_{-\infty}^{x} \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dx$$

PROBLEM 9.4 EXPONENTIAL NOISE CHANNEL. $Y_i = X_i + Z_i$ WHERE Z_i IS I.I.D. EXPONENTIALLY DISTRIBUTED NOISE WITH MEAN μ. ASSUME THAT WE HAVE MEAN CONSTRAINT ON THE SIGNAL ($EX_i \le \lambda$). SHOW THAT CAPACITY OF SUCH CHANNEL IS: $C = \operatorname{ld}\left(1+\frac{\lambda}{\mu}\right)$

$$I(X_1^n; Y_1^n) = H(Y_1^n) - H(Y_1^n \mid X_1^n) \le \sum_{i=1}^{n} h(Y_i) - \sum h(Z_i)$$

$$\le \left| Z_i = \lambda e^{-\lambda x} \right| \le \sum_{i=1}^{n} \frac{1}{2}\operatorname{ld}(2\pi e)\left(\sigma_{x_i}^2 + \sigma_{z_i}^2\right) - \sum_{i=1}^{n} h(Z_i)$$

($\mu = 1/\lambda$)

$$h(Z) = \frac{1-\ln\lambda}{\ln 2} = \operatorname{ld} e - \operatorname{ld}\lambda = \operatorname{ld} e + \operatorname{ld}\mu = \operatorname{ld} e\mu = \frac{1}{2}\operatorname{ld} e^2\mu^2$$

MATCH:

$$E[Z_i^2] = \frac{2}{\lambda^2} = 2\mu^2$$

$$I(X_1^n; Z_1^n) \le \sum_{i=1}^{n} \frac{1}{2}\operatorname{ld}(2\pi e)\cdot\frac{\left(\sigma_x^2 + 1/\lambda^2\right)}{e^2\mu^2}$$

$$\le \sum_{i=1}^{n} \frac{1}{2}\operatorname{ld}\left(\frac{2\pi}{e}\right)\left[\frac{\sigma_x^2 + \mu^2}{\mu^2}\right] = (*)$$

$$\sigma_z^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} = \mu^2$$

($\mu = \frac{1}{\lambda}$; $E[Z^2]$, $E[Z]^2$)

$$\sigma_y^2 = \sigma_x^2 + \sigma_z^2 \qquad E[Z_i] = \frac{1}{\lambda}$$

$$\sigma_x^2 = E[(X-\lambda)^2] = E[X^2] - 2\lambda E[X] + \lambda^2 = E[X^2] - \lambda^2$$

$\sigma_x^2 = E[x_i^2] - \lambda^2$ | $(*) = \sum_{i=1}^{n} \frac{1}{2} ld\left(\frac{2\pi}{e}\right) \cdot \frac{E[x_i^2] - \lambda^2 + \mu^2}{\mu^2} = (*)$

CONVEX

PROOF

JENSEN INEQUALITY:

$E[f(x)] \geq f(E[x])$ $\quad f(x) = x^2$

$E[x_i] \leq \lambda \quad E^2[x_i] \leq \lambda \quad E^2[x_i] \leq E[x_i^2]$

$(*) = \frac{n}{2} \cdot \frac{1}{n} \sum_{i=1}^{n} ld\left(E[x_i^2] - \lambda^2 + \mu^2\right) + \frac{1}{2}\sum_{i=1}^{n} ld \ldots \underset{\text{JENSEN}}{\leq}$

$\leq \frac{n}{2} \; ld \frac{1}{n}\sum_{i=1}^{n}\left(E[x_i^2] - \lambda^2 + \mu^2\right) + \frac{n}{2} \ldots$

MARKOV'S INEQUALITY: $\quad P_r(x > a) \leq \frac{E[x]}{a}$

J_1

$J(p_1^n) = -\sum_{i=1}^{n} p(x_i) \, ld \, p(x_i) - ld(e\mu) + \eta \sum_{i=1}^{n} x_i p(x_i)$

$C = \max_{p(x):\, \sum x p(x) \leq \lambda} I(x; z) \qquad J_{max} = \sum_{i=1}^{n} \frac{1}{2} ld(2\pi e)(\sigma_x^2 + \sigma_z^2)$

$\frac{dJ(p_1^n)}{dp_i} = -\sum_{i=1}^{n} \frac{1}{2} \frac{d}{dp_i}\left[ld(2\pi e)(\sigma^2 + \mu^2)\right] + \eta x_i = 0$

$I(x_1^n; z_1^n) \leq \sum_{i=1}^{n} h(z_i) - \sum_{i=1}^{n} h(z_i) = \sum_{i=1}^{n} h(y_i) - h(z_i)$

$\leq$ | GAUSSIAN DISTRO MAXIMIZES ENTROPY | $= \sum_{i=1}^{n}\left[\frac{1}{2} ld(2\pi e)(\sigma_x^2 + \mu^2) - \right.$

$\left. - \frac{1}{2} ld \, e^2 \mu^2\right] = \sum_{i=1}^{n}\left[\frac{1}{2} ld \, e^2(\lambda^2 + \mu^2) - \frac{1}{2} ld \, e^2 \cdot \mu^2\right] = \sum_{i=1}^{n} \frac{1}{2} ld \frac{\lambda^2 + \mu^2}{\mu^2}$

$y = x + z$ $\quad$ AKO PRETPOSTAVIMO DA X E EKSPONENCIJALNA TOGA:

$p(x) = \nu \cdot e^{-\nu x} \qquad E[x] = \frac{1}{\nu} = \lambda \qquad E[(x-\lambda)^2] = \frac{1}{\nu^2} = \lambda^2$

BUDUĆI X, Z SE EKSPONENCIJALNI DISTRIBUCIJI (TIP)

$\nu = \lambda$

$E[y^2] = E[x^2] + E[z^2] \qquad \sigma_y^2 = \sigma_x^2 + \sigma_z^2 = \lambda^2 + \mu^2$

$\leq \frac{1}{2}\sum_{i=1}^{n} ld \frac{(\lambda + \mu)^2}{\mu^2} = \sum_{i=1}^{n} ld\left(1 + \frac{\lambda}{\mu}\right) = n \cdot ld\left(1 + \frac{\lambda}{\mu}\right) \cdot \frac{\text{bits}}{n \text{ channel uses}}$

(147)

$$I(X_i;Y_i) \le \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Z_i) \le$$

$$= \sum_{i=1}^{n} \frac{1}{2}\log_{(2e)}[\sigma_x^2+\mu^2] - \sum_{i=1}^{n} \frac{1}{2}\log e^2\mu^2$$

$$J = \frac{1}{2}\log(\sigma_x^2+\mu^2) - \eta\sum x\,p(x)$$

$$\frac{dJ}{dp(x)} = 0 \quad \frac{d}{dp}\left[\frac{1}{2}\log\left[\int_{-\infty}^{\infty}(x^2-\lambda^2)\,p(x)\,dx + \mu^2\right] - \eta\int_{-\infty}^{\infty} x\,p(x)\,dx\right] = 0$$

- AKO SO SUMI NAMESTO $\int$

$$\frac{d}{dp}\left[\frac{1}{2}\log\left(\sum_{i=1}^{n}(x^2-\lambda^2)\,p(x)+\mu^2\right) - \eta\sum x\,p(x)\right] = 0$$

$$\frac{1}{2}\frac{\sum_{i=1}^{n}(x^2-\lambda^2)}{\sum_{i=1}^{n}(x^2-\lambda^2)\,p(x)+\mu^2}\cdot\frac{1}{\ln e} - \eta\sum_{i=1}^{n} x = 0$$

CHERNOFF BOUND: $\boxed{\Pr(Y \ge a) \le M(s)\cdot e^{-sa}}$

- PRAŠANJETO E SLEDNO: STOM $Y \sim N(E[Y];\sigma_Y^2+\mu^2)$ TOGAŠ KOJA DISTRIBUCIJA $p(x)$ KONVOLUIRANA SO EKSPONENCIJALNA DISTRIBUCIJA ĆE DADE GAUSOVA RASPREDELBA

$$p(Y) \sim N(E[Y];(\sigma_Y^2+\mu^2)) \quad \text{i} \quad p(x) \circledast p(Z)$$

$$M_Z(s) = E[e^{+sZ}] = \left(\frac{\lambda}{-s+\lambda}\right) = \left(\frac{-s+\lambda}{\lambda}\right)^{-1} = \left(\frac{\lambda-s}{\lambda}\right)^{-1} = \left(1-\frac{s}{\lambda}\right)^{-1}$$

$$M_Y = M_x\cdot M_Z \qquad M_x = \frac{M_Y}{M_Z} = \frac{M_Y}{\left(1-\frac{s}{\lambda}\right)^{-1}}$$

$$M(-s) = \int_{-\infty}^{\infty} p(x)\,e^{-sx}\,dx = \mathcal{L}[p(x)]$$

$$p(x) = \mathcal{L}[M(-s)] \qquad M_Y(s) = e^{\frac{1}{2}s(s\sigma^2+2\mu)}$$

$$\boxed{M_Z(s) = \frac{\lambda}{\lambda-s}}$$

$$M_x = \frac{\lambda-s}{\lambda}\cdot e^{\frac{1}{2}s(s\sigma^2+2\mu)} = e^{\frac{1}{2}s(s\sigma^2+2\mu)} - s\,e^{\frac{1}{2}s(s\sigma^2+2\mu)}\cdot\frac{1}{\lambda}$$

$$p(x) = p(Y) + \frac{1}{\lambda}\frac{d}{dY}[p(Y)]$$

(PROVERKA NA SLEDNATA STRANA)

$$\mathcal{L}[p(\tau)] = M_\tau(-s) = e^{s(s\sigma^2-2\mu)/2} \qquad \boxed{\mathcal{L}[p(z)] = M_z(-s) = \frac{\lambda}{s+\lambda}}$$

$$p(x) = \mathcal{L}^{-1}[M_z(-s)\cdot M_\tau(-s)] = \mathcal{L}^{-1}\left[\frac{s+\lambda}{\lambda} e^{s(s\sigma^2-2\mu)/2}\right] + \mathcal{L}^{-1}[p(0)] + p(0)/\lambda$$

$$= \frac{1}{\lambda}\mathcal{L}^{-1}[s\cdot M_\tau(-s)] + p(\tau) = \frac{1}{\lambda}\frac{d\,p(\tau)}{d\tau} + p(\tau)$$

$$\boxed{\mathcal{L}\left[\frac{d}{dt}f(t)\right] = s\cdot F(s) - f(0^+)} \qquad \frac{df(t)}{dt} = \mathcal{L}^{-1}[sF(s)] - \mathcal{L}^{-1}[f(0)]$$

$$p(x) = \frac{1}{\lambda}\frac{d\,p(x)}{dx} + \frac{\mathcal{L}^{-1}(p(0))}{\lambda} + p(x) = \frac{1}{\lambda}\frac{d\,p_\tau(x)}{dx} + \frac{\mathcal{L}^{-1}(p_\tau(0))}{\lambda} + p_\tau(x)$$

$$p(x) = -\frac{1}{\sqrt{2}}\frac{(-\mu-\lambda\sigma^2)}{\sqrt{\pi}\,\lambda\sigma^3}\cdot e^{\frac{2x\mu}{2\sigma^2} - \frac{x^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2}} + \underbrace{\frac{1}{\sqrt{2}}\frac{1}{\sigma\sqrt{\pi}}\cdot e^{-\frac{\mu^2}{\sigma^2}}}_{p_\tau(0)}$$

$$p(x) = \frac{\lambda\sigma^2+\mu-x}{\lambda\sigma^3\sqrt{2\pi}}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{\mu^2}{\sigma^2}}$$

$$p(x) = \frac{(\lambda\sigma^2+\mu)}{\lambda\sigma^3\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \mathcal{L}^{-1}\left(\frac{1}{\sigma\sqrt{2\pi}}e^{\dots}\right) - \frac{x}{\lambda\sigma^3\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\boxed{p(x) = \frac{(\lambda\sigma^2+\mu)}{\lambda\sigma^3\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \frac{1}{\lambda\sigma\sqrt{2\pi}}e^{-\frac{\mu^2}{\sigma^2}}\delta(x) - \frac{x}{\lambda\sigma^3\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

$$E[x^2] = \frac{(\mu+\lambda\sigma^2)(\mu^2+\sigma^2)}{\lambda\sigma^2} + 0 - \frac{\mu(\mu^2+3\sigma^2)}{\lambda\sigma^2} = -\frac{2\mu}{\lambda} + \mu^2 + \sigma^2$$

$$\boxed{E[x] = \frac{\mu_\tau\lambda_z - 1}{\lambda_z}} \qquad \sigma_x^2 = E[x^2] - E^2[x] = \sigma_\tau^2 - \frac{1}{\lambda_z^2} = \sigma_\tau^2 - \mu_z^2$$

$$E[x] \le \lambda_x \qquad \frac{\mu_\tau\lambda_z - 1}{\lambda_z} \le \lambda_x \qquad \mu_\tau\lambda_z - 1 \le \lambda_x\lambda_z$$

$$\lambda_z(\mu_\tau - \lambda_x) \le 1 \qquad \boxed{\lambda_z \le \frac{1}{\mu_\tau - \lambda_x}}$$

$$\mu_\tau - \mu_z \qquad \mu_z = \frac{1}{\lambda_z}$$

$$E[x] = \mu_\tau - \mu_z \le \lambda_x \qquad I(x;\tau) \le \sum_{i=1}^{n}\frac{1}{2}\operatorname{ld}(2\pi e)\sigma_\tau^2 - \sum_{i=1}^{n}\frac{1}{2}\operatorname{ld} e\mu$$

$$\le \frac{1}{2}\sum_{i=1}^{n}\operatorname{ld}(2\pi e)\frac{\sigma_\tau^2}{e^2\mu^2} = \frac{1}{2}\sum_{i=1}^{n}\operatorname{ld}(2\pi e)\frac{\sigma_x^2+\mu_z^2}{e^2\mu_z^2} = \frac{1}{2}\sum_{i=1}^{n}\operatorname{ld}\frac{2\pi}{e}\left(1+\frac{\sigma_x^2}{\mu_z^2}\right)$$

(14)

$E[X] = \mu_Y - \mu_Z \qquad \sigma_X^2 = \sigma_Y^2 - \mu_Z^2 \qquad \sigma_Z^2 = \frac{1}{\lambda_Z^2} = \mu_Z^2$

$I(X;Y) \leq \frac{1}{2}\sum_{i=1}^{n} \ln\left(\frac{2\pi}{e}\right)\cdot\left(1 + \frac{\sigma_X^2}{\mu_Z^2}\right)$

$\mu_Y - \mu_Z \leq \lambda \qquad E[X] = \mu_Y - \mu_Z = \lambda$

$E[X^2] = -\frac{2\mu_Y}{\lambda_Z} + \mu_Y^2 + \sigma^2 \; ; \; \mu_Y = \mu_Z \Rightarrow E[X] = 0$

$E[X^2] = \sigma^2$

$E[X^2] = -2\mu_Y \cdot \mu_Z + \mu_Y^2 + \sigma^2 \qquad \mu_Y = \lambda + \mu_Z$

$E[X^2] = -2(\lambda + \mu_Z)\mu_Z + (\lambda + \mu_Z)^2 + \sigma^2 =$

$= -2\lambda\mu_Z - 2\mu_Z^2 + \lambda^2 + 2\lambda\mu_Z + \mu_Z^2 + \sigma^2 = \lambda^2 + \sigma^2 - \mu_Z^2$

$\sigma_X^2 = E[X^2] - E[X]^2 = \lambda^2 + \sigma^2 - \mu_Z^2 - \lambda^2 = \sigma^2 - \mu_Z^2$

$I(X;Y) \leq \frac{1}{2}\sum_{i=1}^{n} \ln\left(\frac{2\pi}{e}\right)\left(1 + \frac{\sigma^2 - \mu_Z^2}{\mu_Z^2}\right) = \frac{1}{2}\sum \ln\frac{2\pi}{e}\left(\frac{\sigma^2}{\mu_Z^2}\right)$

$(\mu_Y - \mu_Z \leq \lambda)$

BETTER SOLUTION: $\qquad p(z) = \frac{1}{\alpha} e^{-z/\alpha}$

$h(z) = -\int_0^\infty p(z) \ln p(z) = 1 + \ln\alpha \qquad \boxed{\alpha \triangleq \mu}$ MEAN OF EXPONENTIAL.

• AMONG ALL NONNEGATIVE RANDOM VARIABLES HAVING DENSITY AND HAVING MEAN α, THE EXPONENTIAL DENSITY WITH MEAN α MAXIMIZES THE DIFFERENTIAL ENTROPY.

LET $f(x)$ BE DENSITY (WITH SUPPORT OF NONNEGATIVE VALUES) HAVING MEAN α:

$0 \leq D\left(f(x) \,\|\, \frac{1}{\alpha}e^{-\frac{x}{\alpha}}\right) = \int_0^\infty f(x) \ln \frac{f(x)}{\frac{1}{\alpha}e^{-x/\alpha}}\, dx$

$= +\int_0^\infty f(x)\ln f(x) - \int_0^\infty f(x) \ln\left(\frac{1}{\alpha}e^{-x/\alpha}\right) dx =$

$= -h(f(x)) - \int_0^\infty f(x)\left[-\frac{x}{\alpha}\ln e - \ln\alpha\right] dx =$

$= -h(f(x)) + \frac{1}{\alpha}\int_0^\infty x f(x)\, dx + \ln\alpha \int_0^\infty f(x)\, dx = -h(f(x)) + \frac{1}{\alpha}\cdot\alpha + \ln\alpha$

$= -h(f(x)) + 1 + \ln\alpha \Rightarrow \boxed{h(f(x)) \leq 1 + \ln\alpha}$

$$C = \max_{f(y)} I(X;Y); \quad I(X;Y) = h(Y) - h(Y|X) =$$

$$= h(Y) - h(Z) = h(Y) - (1 + \log \mu) \leq$$

EXPONENTIAL DISTRIB. MAXIMIZES $D(f\|g)$, $E[X] \leq \lambda$

$$\leq (1 + \log(\mu + \lambda)) - (1 + \log \mu) = 1 + \log(\mu+\lambda) - 1 - \log \mu$$

$$= \log\left(\frac{\mu+\lambda}{\mu}\right) = \log\left(1 + \frac{\lambda}{\mu}\right)$$

THE EXPONENTIAL DISTRIBUTION OF A GIVEN MEAN MAXIMIZES THE DIFFERENTIAL ENTROPY AMONG ALL NONNEGATIVE DENSITIES HAVING THAT MEAN

$$D(f\|g) = \sum f \operatorname{ld} \frac{f}{g} \qquad I(X;Y) = \sum_{xy} f(xy) \operatorname{ld} \frac{f(xy)}{f(x)f(y)}$$

$$I(X;Y) = D(f(x,y) \| f(x)f(y));$$

$$I(X;Y) = \sum_{xy} f(xy) \operatorname{ld} \frac{1}{f(x)} + \sum f(xy) \operatorname{ld} \frac{f(xy)}{f(y)}$$

$$= \sum_x \operatorname{ld} \frac{1}{f(x)} \sum_y f(xy) \;\; H(XY) \ldots$$

$$= H(X) + H(Y) - H(XY) = H(X) - H(X|Y).$$

$$- H(X|Y)$$

$$M_Z(-s) = \int_0^\infty e^{-sz} \frac{1}{\alpha} e^{-\frac{z}{\alpha}} dz = \frac{\frac{1}{\alpha}}{s + \frac{1}{\alpha}} = \frac{1}{1+\alpha s}$$

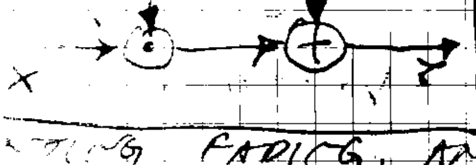
$$M_Y(-s) \cdot \frac{1}{1+\alpha s} = \frac{1}{1+(\alpha+\mu)s} \qquad \alpha = \lambda$$

$$M_Y(-s) = \frac{\alpha s}{1+(\alpha+\mu)s} + \frac{1}{1+(\alpha+\mu)s} = \frac{\lambda s}{1+(\lambda+\mu)s} + \frac{1}{1+(\lambda+\mu)s}$$

$$\text{convert}\left(\frac{1+\lambda s}{1+(\lambda+\mu)s}, \text{parfrac}, s\right) = \frac{\lambda}{\lambda+\mu} + \frac{\mu}{(\lambda+\mu)(1+(\lambda+\mu)s)}$$

$$f(y) = \mathcal{L}^{-1}[M_Y(-s)] = \frac{\lambda}{\lambda+\mu}\delta(y) + \frac{\mu}{\lambda+\mu} \cdot \frac{1}{\lambda+\mu} e^{-\frac{y}{\lambda+\mu}}$$

PROBLEM 9.5 FADING CHANNEL. CONSIDER AN ADDITIVE NOISE FADING CHANNEL



CONSIDER AN ADDITIVE NOISE FADING CHANNEL: $Y = X \cdot V + Z$ WHERE Z IS ADDITIVE NOISE; V IS RANDOM VARIABLE REPRESENTING FADING, AND Z AND V ARE INDEPENDENT

(158) OF EACH OTHER AND OF X. ARGUE THAT KNOWLEDGE OF THE FADING FACTOR V IMPROVES CAPACITY BY SHOWING THAT

$$I(X;Y|V) \geq I(X;Y)$$

$Y = X \cdot V + Z$ $\qquad I(X;Y|V) = H(Y) - H(Y|XV)$

$$H(Y|XV) = \sum_{xv} p(x,v)\, H(xv + Z \,/\, X=x, V=v) \;(\$)$$

$$= H(Z)$$

$$I(X;Y|V) = H(Y) - H(Z) \qquad \boxed{H(Y) = I(X;Y|V) + H(Z)}$$

$$I(X;Y) = H(Y) - \sum_x p(x)\, H(Y|X=x) =$$

$$= H(Y) - \sum_x p(x)\, H(x \cdot V + Z | x) = H(Y) - H(V+Z)$$

$\boxed{\Psi = V + Z}$ $\qquad H(V+Z) = \sum_{i=1} p(\Psi)\, ld \frac{1}{p(\Psi)}$

ADDITIVE NOISE CHANNEL

$$H(\Psi) \leq \frac{1}{2} \log (2\pi e)\, \sigma_\Psi^2 \qquad \sigma_\Psi^2 = \sigma_V^2 + \sigma_Z^2$$

$$I(X;Y) \geq H(Y) - \frac{1}{2} ld (2\pi e)[\sigma_V^2 + \sigma_Z^2]$$

$$I(X;Y|V) = H(Y) - \frac{1}{2} ld (2\pi e)[\sigma_Z^2]$$

$$\boxed{H(Y) = I(X;Y|V) + \frac{1}{2} ld (2\pi e)\sigma_Z^2}$$

$$I(X;Y) = I(X;Y|V) + \frac{1}{2} ld (2\pi e\, \sigma_Z^2) - H(\Psi)$$

$$H(Y) - H(\Psi) \geq H(Y) - \frac{1}{2} ld (2\pi e\, \sigma_\Psi^2)$$

$$I(X;Y|V) = I(X;Y) + H(\Psi) - \frac{1}{2} ld (2\pi e\, \sigma_Z^2)$$

$\boxed{H(\Psi) \geq \frac{1}{2} \cdot ld (2\pi e)\sigma_Z^2}$ ← OVA TREBA DA SE DOKAZE !!!

$$H(\Psi) \leq H(V,Z) = H(Z) + H(V|Z) = H(Z) + H(V) \quad (\$)$$

$$I(X;Y) = H(Y) - H(\Psi) \geq H(Y) - H(Z) - H(V|Z)$$

$$I(X;Y) \geq I(X;Y|V) + H(Z) - H(Z) - H(V|Z) \geq I(X;Y|V) - H(V|Z)$$

$(*)\ \boxed{I(X;Y|V) \leq I(X;Y) + H(V|Z)} \quad = I(X;Y) + H(V) \quad \boxed{V, Z \text{ INDEP.}}$

$$I(V;\Psi) = H(\Psi) - H(\Psi|V) = H(\Psi) - H(Z)$$

$I(V;\Psi) = -H(V;\Psi) + H(V) + H(\Psi)$

$= -H(\Psi|V) \quad H(V;\Psi) = H(V) + H(\Psi|V)$

$\boxed{\Psi = V + Z}$

$I(V;\Psi) = H(V) - H(Z) \geq 0 \quad \Big| \quad H(V) \geq H(Z)$

$I(V;\Psi) = H(V) + H(\Psi) - H(V;\Psi)$

$-H(Z) = H(V) + H(\Psi) - H(V;\Psi)$

$Z = X + T$ $\quad I(X;Z) = H(Z) - H(T) \geq 0 \quad H(Z) \geq H(T)$

$H(Z X) = H(Z) + H(X|Z) = H(X) + H(Z|X)$, $H(Z|X) = H(T)$

$I(XT;Z) = I(X;Z) - I(T;Z|X)$

$I(T;Z) = H(Z) - H(Z|T) = H(Z) - H(X) \geq 0 \quad H(Z) \geq H(X)$

$\star \Rightarrow I(X;T|V) \geq I(X;T) \quad \boxed{H(V|Z) = H(V) \geq 0}$

$\Psi = f(X,Z) \qquad \Psi = f(\xi)$

$H(\Psi,\xi) = H(\Psi) + H(\xi|\Psi) = H(\xi) + H(\ldots)$

$\boxed{H(\xi) \geq H(\Psi) \geq H(f(\Psi))}$ } FUNCTION DECREASES ENTROPY

$I(X;T|V) = H(T) - H(Z) \qquad I(X;T) = H(T) - H(V+Z)$

$(V+Z) = f(V,Z) \qquad H(f(V,Z)) \leq H(V,Z)$

$I(X;T) = H(T) - H(f(V,Z)) \geq H(T) - H(V,Z) =$

$= H(T) - H(V) - H(Z|V) = H(T) - H(V) - H(Z)$

$I(X;T|V) = H(T) - H(Z)$
$I(X;T) \geq H(T) - H(V) - \ldots$ } DOKAZANO NOVE ... OČIGLEDNO DOKAZOT!!!

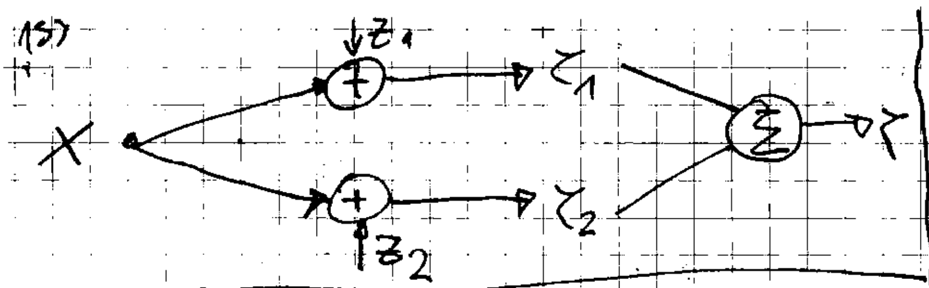
$H(V+Z,\Psi) = H(\Psi) + H(V+Z|\Psi) = H(V+Z) + H(\Psi|V+Z)$, $H(\Psi|V+Z) = 0$

? DALI E 0

...PATH GAUSSIAN CHANNEL ...

GAUSSIAN NOISE CHANNEL WITH POWER CONSTRAINT P, WHERE THE SIGNAL TAKES TWO DIFFERENT PATHS AND THE RECEIVED NOISY SIGNALS ARE ADDED TOGETHER AT THE ANTENNA.

15)



(a) FIND THE CAPACITY OF THIS CHANNEL IF Z_1 AND Z_2 ARE JOINTLY NORMAL WITH COVARIANCE MATRIX: $K_Z = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix}$

(b) WHAT IS THE CAPACITY FOR $\rho = 0$, $\rho = 1$ AND $\rho = -1$

$$K_Y = \begin{bmatrix} P+N & N\rho \\ N\rho & P+N \end{bmatrix} = (P+N)^2 - N^2\rho = P^2 + 2NP + N^2 - N^2\rho = N^2(1-\rho) + 2P(P+N)$$

$$K_Y = \begin{bmatrix} P+N & P+N\rho \\ P+N\rho & P+N \end{bmatrix} = (P+N)^2 - (P+N\rho)^2 = P^2 + 2PN + N^2 - P^2 - 2PN\rho + N^2\rho^2$$

$$= 2PN - 2PN\rho + N^2(1-\rho^2) = 2PN(1-\rho) + N^2(1-\rho^2)$$

$$Y = Y_1 + Y_2 = X + Z_1 + X + Z_2 = 2X + Z_1 + Z_2$$ $E[Y^2] = 4P + E[(Z_1+Z_2)^2]$

$$C = \frac{1}{2} \operatorname{ld} \frac{(2\pi e)^2 |K_Y|}{(2\pi e)^2 |K_Z|}$$

$I(X;Y) = H(Y) - H(Y|X)$

$H(f(X)) \leq H(X)$

$$I(X;Y) = H(Y_1 + Y_2) - H(Y|X) \leq H(Y_1 Y_2) - H(Y|X)$$

$$H(Y|X) = H(Y_1 + Y_2 | X) = H(2X + Z_1 + Z_2 | X) = H(Z_1 + Z_2)$$

$$I(X;Y) \leq H(Y_1 Y_2) - H(Z_1 + Z_2)$$

$\xi = Z_1 + Z_2$
$\eta = Z_1$

TEORIJA INFORMACIJA BOOKLET.

$$\begin{bmatrix} \frac{\partial(z_1+z_2)}{\partial \xi} & \frac{\partial z_1}{\partial \xi} \\ \frac{\partial(z_1+z_2)}{\partial \eta} & \frac{\partial z_1}{\partial \eta} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = 1$$

$$p(z_1 z_2) = |J|\, p(\xi, \eta) \qquad p(\xi, \eta) = \frac{p(z_1 z_2)}{|J|}$$

$z_1 = \eta$
$z_2 = \xi - z_1$

$$p(\xi,\eta) = p(\eta, \xi - \eta)$$

$$p(\xi) = \int_{-\infty}^{\infty} p(\xi,\eta)\, d\eta =$$

$$= \int_{-\infty}^{\infty} p(\eta, \xi-\eta)\, d\eta = \int_{-\infty}^{\infty} p(\eta) \cdot p(\xi - \eta)\, d\eta \qquad p(\xi) = p(\eta) * p(\xi)$$

STAT NEZAVISNI

$$p(Z_1+Z_2) = \int_{-\infty}^{\infty} p(z_1, z_2 - z_1)\, dz_1 = \frac{e^{-\frac{1}{2}\frac{z_2^2}{N(5+4\rho)}}}{\sqrt{2\pi N(5+4\rho)}}$$ GAUSSIAN ~~MANCE~~

$$p(Z_1+Z_2) \sim \mathcal{N}(0, N(5+4\rho))$$ $N = \sigma^2$

$h(Z_1+Z_2) = \frac{1}{2} ld\, 2\pi e\, N(5+4\rho)$

$\circledast \rightarrow I(X;Z) \le h(Z_1,Z_2) - h(Z_1+Z_2) = h(Z_1,Z_2) - \frac{1}{2} ld(2\pi e) N(5+4\rho)$

$I(X;Z) \le \frac{1}{2} ld\, 2\pi e\left(4P + N(5+4\rho)\right) - \frac{1}{2} ld(2\pi e) N(5+4\rho)$

$I(X;Z) \le \frac{1}{2} ld\left(1 + \frac{4P}{N(5+4\rho)}\right)$ $\quad C = \frac{1}{2} ld\left[1 + \frac{4P}{N(5+4\rho)}\right]$

(b) $C(\rho=0) = \frac{1}{2} ld\left[1 + \frac{4P}{5N}\right]$; $C(\rho=1) = \frac{1}{2} ld\left[1 + \frac{4P}{9N}\right]$

$C(\rho=-1) = \frac{1}{2} ld\left[1 + \frac{4P}{N}\right]$

■ UNIVERSITY OF TORONTO SOLUTION (sol5.pdf)

(a) $I(X;Z) = h(Z) - h(Z|X) = h(2X + Z_1 + Z_2) - h(Z_1+Z_2)$

GAUSSIAN DISTRO OF Z MAXIMIZES ENTROPY

$C = \frac{1}{2} ld(2\pi e) E[(2X+Z)^2] - \frac{1}{2} ld(2\pi e) E[Z^2]$

$E[(2X+Z)^2] = 4E[X^2] + E[Z^2] = 4 \cdot P + E[(Z_1+Z_2)^2]$

$E[(Z_1+Z_2)^2] = E[Z_1^2 + 2Z_1Z_2 + Z_2^2] = E[Z_1^2] + 2E[Z_1Z_2] + E[Z_2^2]$

VS MATLE SE ПОКАЗУВА РЕКА:

$E[Z_1Z_2] = \rho N$; $E[Z_1^2] = N$; $E[Z_2^2] = N$; $\quad E[Z^2] = E[(Z_1+Z_2)^2]$

$E[(Z_1+Z_2)^2] = N + 2\rho N + N = 2N + 2\rho N = 2N(1+\rho)$

$C = \frac{1}{2} ld(2\pi e)\left[4P + 2N(1+\rho)\right] - \frac{1}{2} ld(2\pi e)\left[2N(1+\rho)\right]$

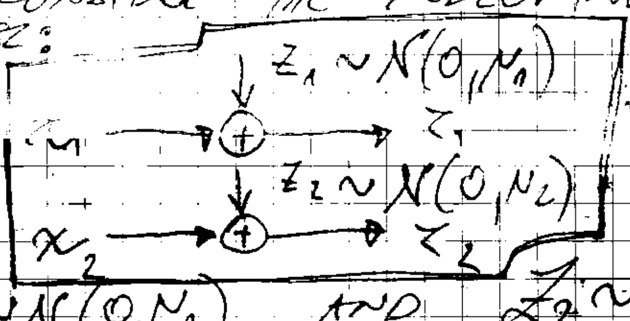
$C = \frac{1}{2} ld\left(1 + \frac{4P}{2N(1+\rho)}\right)$ $\quad N = \sigma^2$

(b) $\rho = 0$; $C = \frac{1}{2} ld\left(1 + \frac{2P}{N}\right)$; $\rho = 1$ $\quad C = \frac{1}{2} ld\left(1 + \frac{P}{N}\right)$

$\rho = -1$ $\quad C = \frac{1}{2} ld\left(1 + \frac{2P}{0}\right) \rightarrow \infty$ NOISE CANCEL ITSELF $Z_1 = -Z_2$

IN $\rho = 0$ THE NOISE POWER IS DOUBLED AND IN CASE $\rho = 1$ THE NOISE POWER IS 4 TIMES LARGER THEN IN THE CASE OF SINGLE NOISE CHANNEL.

PROBLEM 9.8 CONSIDER THE FOLLOWING PARALLEL GAUSSIAN CHANNEL:



WHERE $Z_1 \sim N(0,N_1)$ AND $Z_2 \sim N(0,N_2)$ ARE INDEPENDENT GAUSSIAN RANDOM VARIABLES AND:

(1??) $Y_i = X_i + Z_i$

We wish to allocate power to the two parallel channels. Let β_1 and β_2 be fixed. Consider total cost constraint $\beta_1 P_1 + \beta_2 P_2 \le \beta$ where P_i is power allocated to the ith channel and β_i is the cost per unit power in that channel. Thus $P_1 \ge 0$ and $P_2 \ge 0$ can be chosen subject to the cost constraint β.

(a) For what value of β does the channel stop acting like pair of channels?

(b) Evaluate the capacity and find P_1 and P_2 that achieve capacity for $\beta_1 = 1$, $\beta_2 = 2$, $N_1 = 3$, $N_2 = 2$ and $\beta = 10$.

RECALL

$$I(X;Y_1) = h(X_1^n) - h(Y_1^n | X_1^n) = h(Y_1^n) - h(Z_1^n) \le$$

$$\le \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Z_i) \le \sum_{i=1}^{n} \frac{1}{2} \operatorname{ld}\left(1 + \frac{P_i}{N_i}\right)$$

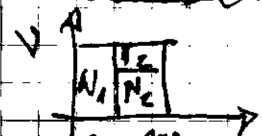
$$C = \sum_{i=1}^{n} \frac{1}{2} \operatorname{ld}\left(1 + \frac{P_i}{N_i}\right) \qquad \sum_{i}^{n} P_i \le P$$

$$\frac{dJ}{dP_i} = 0 ; \qquad J = \sum_{i=1}^{n} \frac{1}{2} \ln\left(1 + \frac{P_i}{N_i}\right) + \lambda \sum_{i=1}^{n} P_i$$

$$\frac{1}{2} \frac{1}{1 + \frac{P_i}{N_i}} \frac{1}{N_i} + \lambda = 0 \qquad \frac{1}{2} \frac{1}{P_i + N_i} + \lambda = 0$$

$$\frac{1}{P_i + N_i} = -2\lambda \qquad P_i + N_i = -\frac{1}{2\lambda} \qquad P_i = -\frac{1}{2\lambda} - N_i$$

$$P_i = V - N_i \qquad P_1 + P_2 \le P \qquad N_1 - N_2 \le P$$



$$P + N_1 = V \qquad N_1 = V$$

$$V = P_2 + N_2 \qquad P + N_2 \ge N_1 \qquad P \ge N_1 - N_2$$

[illegible]

VO OVOJ MOMENT KANALOT POČNUVA DA SE ODNESUVA KAKO PAR OD KANALI

$$C = \sum_{i=1}^{n} \frac{1}{2} \ln\left(1 + \frac{P_i}{N_i}\right) \qquad \sum_{i=1}^{n} \beta_i P_i \le \beta \qquad P_i = \frac{1}{2\lambda\beta_i} - N_i$$

$$\frac{1}{2} \frac{1}{P_i + N_i} - \lambda \cdot \beta_i = 0 \qquad \frac{1}{P_i + N_i} = 2\lambda \cdot \beta_i$$

$$P_i + N_i = \frac{1}{2\lambda \cdot \beta_i} \; (\Delta) \qquad N_1 = \frac{1}{2\lambda\beta_1} \qquad \lambda = \frac{1}{2N_1\beta_1}$$

$$P_2 + N_2 = \frac{1}{2\lambda \cdot \beta_2} \; (**) \qquad P_2 + N_2 = \frac{1}{\frac{1}{2N_1\beta_1} \cdot \beta_2} = N_1 \cdot \frac{\beta_1}{\beta_2}$$

$$P_2 \le \frac{\beta_1}{\beta_2} \cdot N_1 - N_2 \Rightarrow$$ channel stops acting like a pair of channels.

$p_1 \cdot P_1 + p_2 P_2 \leq P$

$p_1 P_1 + p_1 N_1 - p_2 N_2 \leq P$

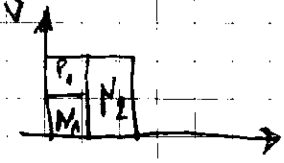
(6) $P_1, P_2 = ?$ $N_1 = 3$ $N_2 = 2$ $p_1 = 1$ $p_2 = 2$ $P = 10$

$P_1 = \frac{1}{p_1}(P - p_1 N_1 + p_2 N_2) = (10 - 3 + 2 \cdot 2) = 10 + 1 = 11$

$P_2 = \frac{p_1}{p_2} N_1 - N_2 = \frac{3}{2} - 2 = 1.5 - 2 = -0.5$

$1 \cdot (11) + 1 \cdot 3 - 2 \cdot 2 = 11 + 3 - 4 = 11 - 1 = 10$

- OBLATO CE IZBEGAM DE DVA NEGATIVNA SNAGA.



$P_1 + N_1 = \frac{1}{2\lambda p_1}$ $N_2 = \frac{1}{2\lambda p_2}$

$\lambda = \frac{1}{2 N_2 p_2}$ $P_1 + N_1 = \frac{1}{2\lambda \cdot p_1}$

$P_1 + N_1 = \frac{1}{2 \cdot \frac{1}{2 N_2 p_2} \cdot p_1} = \frac{p_2}{p_1} \cdot N_2$

$P_1 = \frac{p_2}{p_1} N_2 - N_1$

$P_1 = 2N_2 - N_1$ $P_1 = 4 - 3 = 1$

$P_2 = \frac{1}{p_2}(P - p_1 P_1) = \frac{1}{p_2}(P - p_2 N_2 + p_1 N_1)$

$p_1 P_1 + p_2 P_2 \leq P$ $p_1 \cdot 1 + 2 \cdot P_2 = 10$ $P_2 = \frac{10-1}{2} = 4.5$

$P_1 = 1$ $P_2 = 4.5$ $p_2 N_2 - p_1 N_1 + p_2 P_2 = P$

$p_1 P_1 + p_2 P_2 = 1 \cdot 1 + 2 \cdot 4.5 = 1 + 9 = 10 = P$

$C = \frac{1}{2} ld\left(1 + \frac{P_1}{N_1}\right) + \frac{1}{2} ld\left(1 + \frac{P_2}{N_2}\right) = \frac{1}{2} ld\left(1 + \frac{P_1}{N_1}\right)\left(1 + \frac{P_2}{N_2}\right)$

$= \frac{1}{2} ld\left(1 + \frac{1}{3}\right) \cdot \left(1 + \frac{4.5}{2}\right) = \frac{1}{2} ld\left(\frac{4}{3}\right) \cdot \left(\frac{2+4.5}{2}\right)$

$C = \frac{1}{2} ld\left(\frac{4 \cdot 6.5}{6}\right) = \frac{1}{2} ld\left(\frac{26,0}{6}\right) = \frac{1}{2} ld\left(\frac{13}{3}\right)$

$C = 0.5\, ld(4.33) = 1.058$ bits PER USE

$C = \frac{1}{2} ld\left(1 + \frac{5.5}{3}\right)\left(1 + \frac{2.25}{2}\right) = \frac{1}{2} ld\left(\frac{8.5 \cdot 4,25}{6}\right) = 1.27$

- UNIVERSITY OF ILINOIS SOLUTION (HW8s.PDF)

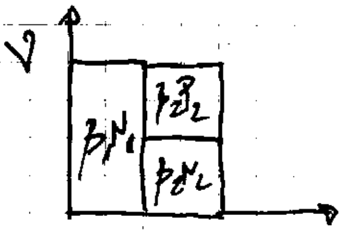
$P_1 = 5.5$ $P_2 = 2.25$

- ZA DA GO KORISTIS WATER FILLING IACTA ZA SE SVEDE NA (*) PA ZATOA (4) GO PISUVAS VO SLEDNAVA FORMA:

$p_i P_i + p_i N_i = \frac{1}{2\lambda} = \nu$ $\nu = \text{CONST}$

(**) MOZE PRISTAP NE E DOBAR ZOSTO (*) NE E CONST.

157



$p_1N_1 = V \qquad p_2P_2 + p_2N_2 = V$

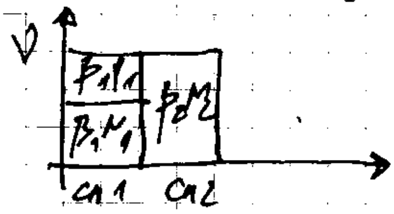
$p_2P_2 + p_2N_2 = p_1N_1$

$p_2P_2 = p_1N_1 - p_2N_2$

$p_2P_2 \geq p_1N_1 - p_2N_2$ } VO OVOJ SLUČAJ IMAME DVA KANALI

(b) $p_1N_1 = 1\cdot 3 = 3 \qquad p_2\cdot N_2 = 2\cdot 2 = 4$

ZNAČI OBRATNO TREBA:



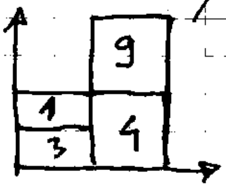
$p_1P_1 \geq p_2N_2 - p_1N_1$ } 2 CHANN.

$p_1P_1 \geq p_2N_2 - p_1N_1$

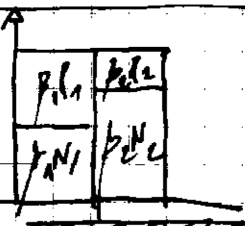
$\underbrace{p_2N_2 - p_1N_1}_{p_1P_1} + p_2P_2 = P$

$\underbrace{4-3}_{P_1} + p_2P_2 = 10 \qquad p_2P_2 = 10 - 1 = 9$

$P_2 = \frac{9}{p_2} = \frac{9}{2} = 4.5 \qquad P_1 = 1$ MINIMAL p_1P_1



} OVAKA ĆE ISPLEDA DEKA GO DOBIVME GO MINIMALNATA VREDNOST NA $p_1P_1 = 1$ NO NE GO ISPOLNUVA TAKA USLOVOT SUMATA NA SNAGATA NA ŠUMOT I SIGNALOT DA BIDE KONSTANTA ZA OVAKA KANALI.



I $\boxed{p_1P_1 + p_1N_1 = p_2P_2 + p_2N_2}$ OVAA ZAVISNA MI FALEŠE!

$P_1 = (p_2P_2 + p_2N_2 - p_1N_1)/p_1$

II $\boxed{p_1P_1 + p_2P_2 = P}$

$P_2 = \frac{1}{p_2}(P - p_1P_1) = \frac{P}{p_2} - \frac{p_1}{p_2}P_1$

$P_1 = (P - p_1P_1 + p_2N_2 - p_1N_1)\frac{1}{p_1} = \frac{P}{p_1} - P_1 + \frac{p_2}{p_1}N_2 - N_1$

$2P_1 = \left(\frac{P}{p_1} + \frac{p_2}{p_1}N_2 - N_1\right) \qquad P_1 = \frac{1}{2}\left(\frac{P}{p_1} + \frac{p_2}{p_1}N_2 - N_1\right)$

$P_1 = \frac{1}{2}\left(\frac{10}{1} + \frac{2}{1}\cdot 2 - 3\right) = \frac{1}{2}(10 + 4 - 3) = \frac{11}{2} = 5.5$

$P_2 = \frac{1}{2}(10 - 5.5) = \frac{1}{2}(4.5) = 2.25$

$10 = 1\cdot 5.5 + 2\cdot 2.25 = 5.5 + 4.5 = 10$ ⟵ SOLUTION

PROBLEM 9.11 (CONTINUE FROM N16n) $X \to Y \to Z$

$I(X,Y;Z) = H(Z) - H(Z|XY) = H(Z) - H(Z|Y)$

$= \frac{1}{2}\,ld(2\pi e)\sigma_2^2 - H(Z|Y), \qquad H(Z|Y) = ?$

$$I(X;Y;Z) = I(X;Z) - I(Y;Z|X) = I(Y;Z)$$ GREŠKA!!! [illegible] $H(Z|Y)$

$$I(Y;Z|X) = H(Z|X) - H(Z|XY) = H(Z|X) - H(Z|X) = 0$$

$$I(Y;Z) = \frac{1}{2}\, ld(2\pi e)\,\sigma_z^2 - H(Z|Y)$$

$H(Y;Z) = H(Y) - H(Z|Y)$

$$K = \begin{bmatrix} \sigma^2 & \rho_1\sigma^2 & 0 \\ \rho_1\sigma^2 & \sigma^2 & \rho_2\sigma^2 \\ 0 & \rho_2\sigma^2 & \sigma^2 \end{bmatrix} = \begin{bmatrix} P & \rho_1 P & 0 \\ \rho_1 P & P & \rho_2 P \\ 0 & \rho_2 P & P \end{bmatrix}$$

$H(Z|Y) = ?$ ①

$$f(Z|Y) = \frac{f(Y,Z)}{f(Y)} = \frac{1}{\sqrt{2\pi(1-\rho_2^2)P}}\, e^{-\frac{z^2 - 2yz\rho_2 + 2y^2\rho_2^2 - y^2}{2P(1-\rho_2^2)}}$$

① $= \dfrac{z^2 - 2yz\rho_2 - y^2(1 - 2\rho_2^2)}{2P(1-\rho_2^2)}$

1 KOMAD GREŠKA [illegible] EVEDUVANJE [illegible] MNOGO [illegible] PODIV IZLAZ ZA $h(Z|Y)$

$h(Y) - h(Z|Y) = I(Y;Z) - h(Y|\ldots)$
$+ h(Z|Y)$ | $I(Y;Z) = h(Z) - 2h(Z|Y) + h(\ldots)$

$$h(Z|Y) = -\iint f(Y,Z)\, \ln f(Z|Y)\, dy\, dz$$

$$h(Z|Y) = \frac{-(1-\rho_2^2) \ln(2\pi P(1-\rho_2^2))}{2(\rho_2^2 - 1)} = \frac{1}{2}\ln(2\pi)(1-\rho_2^2)P + \frac{1}{2}$$

$$\begin{bmatrix} P & \rho_2 P \\ \rho_2 P & P \end{bmatrix} = P^2 - \rho_2^2 P^2 = P^2(1-\rho_2^2) = |K_{YZ}|$$

$$I(X;Z) = \frac{1}{2}\, ld(2\pi e)\,\sigma_z^2 - \frac{1}{2}\ln 2\pi(1-\rho_2^2)P - \frac{1}{2}\ln e$$

$$I(X;Z) = \frac{1}{2}\, ld\, \frac{e}{e}\left(\frac{1}{1-\rho_2^2}\right) = -\frac{1}{2}\, ld\,(1-\rho_2^2)$$ bits

$$I(X;Z) = H(\ldots) - H(Z|Y) = H(Z) - (H(YZ) + H(Y)) =$$
$$= H(Z) - H(Y) + H(YZ) = \frac{1}{2}\, ld(2\pi e)\,\sigma_z^2 - \frac{1}{2}\, ld(2\pi e)\,\sigma_y^2 +$$
$$+ \frac{1}{2}\, ld(2\pi e)|K_{YZ}| = \frac{1}{2}\, ld(2\pi e)\,\sigma_z^2 - \frac{1}{2}\, ld(2\pi e)\,\sigma_y^2 + \frac{1}{2}\, ld(2\pi e)^2 P^2(1-\rho_2^2)$$

UIC PROBLEM 8.9 SOLUTION

$$I(X;Z) = h(X) - h(X|Z) \qquad h(X,Z) = h(Z) + h(X|Z)$$

$$h(X|Z) = h(X,Z) - h(Z) \qquad I(X;Z) = h(X) + h(Z) - h(X,Z)$$

$$I(X;Z) = 0.5\log(2\pi e\sigma_x^2) + 0.5\log(2\pi e\sigma_z^2) - \frac{1}{2}\ln(2\pi e)|K_{XZ}|$$

$$K_{XZ} = \begin{bmatrix} \sigma_x^2 & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xz}\sigma_x\sigma_z & \sigma_z^2 \end{bmatrix} = \sigma_x^2\sigma_z^2 - \rho_{xz}^2\sigma_x^2\sigma_z^2 = \sigma_x^2\sigma_z^2(1-\rho_{xz}^2)$$

$$I(X;Z) = 0.5\, ld(2\pi e\sigma_x^2) + 0.5\, ld(2\pi e\sigma_z^2) - \frac{1}{2}\, ld(2\pi e)^2\sigma_x^2\sigma_z^2(1-\rho_{xz}^2)$$

159

$$p(xz|y) = p(x|y)\cdot p(z|y) \qquad E[xz] = \rho_{xz}\sigma_x\sigma_z$$

$$H(xz|y) = H(x|y) + H(z|xy)$$

$$\rho_{xz} = \frac{E[xz]}{\sigma_x\sigma_z} = \frac{E_{xz}[xz]}{\sigma_x\sigma_z} = \frac{E_x[E[xz|y]]}{\sigma_x\sigma_z} = \frac{E_y[E[x|y]\cdot E[z|y]]}{\sigma_x\sigma_z}$$

$$E_z[z|y] = \int_{-\infty}^{\infty} z\, p(z|y)\, dz \overset{\text{MARKO}}{=} \underline{y\cdot\rho_2} \quad (\text{IF } \sigma_y = \sigma_z)$$

MARKO

IF $\sigma_y \neq \sigma_z$

$$E_z[z|y] = \rho_2\frac{y\cdot\sigma_z}{\sigma_y} \qquad E_x[x|y] = \frac{\rho_1 y\cdot\sigma_x}{\sigma_y}$$

$$E_y\left[\rho_2\frac{y\cdot\sigma_z}{\sigma_y}\cdot\frac{\rho_1 y\cdot\sigma_x}{\sigma_y}\right] = \frac{\rho_1\rho_2\sigma_x\sigma_z}{\sigma_y^2}E_y[y^2] = \rho_1\rho_2\sigma_x\sigma_z$$

$$\boxed{\rho_{xz} = \rho_1\rho_2} \qquad I(x;z) = 0.5\,ld\,(2\pi e)^2\sigma_x^2\sigma_z^2 - \frac{1}{2}ld(2\pi e)^2\sigma_x^2\sigma_z^2(1-\rho_{xz}^2)$$

$$I(x;z) = \frac{1}{2}ld\frac{(2\pi e)^2\sigma_x^2\sigma_z^2}{(2\pi e)^2\sigma_x^2\sigma_z^2(1-\rho_1^2\rho_2^2)} = -0.5\,ld\,(1-\rho_1^2\rho_2^2)$$

= OD ZADACAVA DIRETCI CE MI TEKA OD REVA DOLU ZA REŠIV. VO GLAVA 8 (MODELOT 8.9) NADOLU DOLA:

$$h(z|y) = \frac{1}{2} + \frac{1}{2}\ln 2\pi(1-\rho_2^2)P = \frac{1}{2}\ln(2\pi e)(1-\rho_2^2)P$$

$$h(y|x) = \frac{1}{2}\ln(2\pi e)(1-\rho_1^2)P$$

$$I(x;yz) = I(x;y) + I(x;z|y) = I(x;z) + I(x;y|z)$$

$$I(x;z|y) = H(z|y) - H(z|yx) = 0$$

$$I(x;z) = I(x;y) - I(x;y|z) = I(x;y) - H(x|z) + H(x|yz)$$

$$= I(x;y) - H(y|z) + H(y|xz) = I(x;y) - H(y|z) + H(y|x)$$

$$H(y|xz) = H(y|x) \quad \} \text{ OVA NE E DOVOLNO}$$

$$h(y|z) = \frac{1}{2}\ln 2\pi P(1-\rho_2^2) + 1/2 \qquad h(y|x) = \frac{1}{2}\ln 2\pi(1-\rho_1^2) + 1/2$$

$$I(x;z) = \underbrace{\frac{1}{2}ld\,(2\pi e)^2P^2(1-\rho_1^2)}_{\text{NE E TOČNO!!!}} - \underbrace{\frac{1}{2}\ln 2\pi P(1-\rho_2^2) + \frac{1}{2}\ln(2\pi e)(1-\rho_1^2)}_{= +\frac{1}{2}\ln\frac{(1-\rho_1^2)}{(1-\rho_2^2)}}$$

$$I(x;z) = \frac{1}{2}ld\,(2\pi e)^2\frac{P^2(1-\rho_1^2)(1-\rho_1^2)}{(1-\rho_2^2)}$$

$$I(x;y) = H(y) - H(y|x) \qquad I(x;z) = H(y) - H(y|x) - H(y|z) + H(y|x)$$

$$h(Y,Z) = \frac{1}{2}\,ld\,(2\pi e)^2 P^2 (1-\rho_z^2) = H(Z) + H(Y|Z)$$

$$h(Y|Z) = +h(Y,Z) - h(Z) = -\frac{1}{2}\,ld\,(2\pi e)\,P + h(Y,Z)$$

$$= -\frac{1}{2}\,ld(2\pi e)\,P + \frac{1}{2}\,ld\,(2\pi e)^2 \cdot P^2(1-\rho_z^2) =$$

$$= -\frac{1}{2}\,ld\,\frac{1}{(2\pi e)P(1-\rho_z^2)} = +\frac{1}{2}\,ld\,(2\pi e)P(1-\rho_z^2)$$

→ NESTO STO OPET ...

9.12 TIME-VARYING CHANNEL. A TRAIN PULLS OUT OF THE STATION AT THE CONSTANT VELOCITY. THE RECEIVED SIGNAL ENERGY THUS FALLS OFF WITH TIME AS $1/i^2$. THE TOTAL RECEIVED SIGNAL AT TIME i IS:

$$Y_i = \frac{1}{i} x_i + Z_i$$

3090311

WHERE $Z_1, Z_2, \ldots$ ARE I.I.D $\sim N(0,N)$. THE TRANSMITTER CONSTRAINT FOR BLOCK LENGTH n IS

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2(w) \le P \qquad w \in \{1, 2, \ldots, 2^{nR}\}$$

USING FANO INEQUALITY, SHOW THAT THE CAPACITY C IS EQUAL TO ZERO FOR THIS CHANNEL.

$$I(X^n; Y^n) = H(Y_1^n) - H(Y_1^n | X_1^n) \stackrel{?}{=} H(Y_1^n) - H(Z_1^n)$$

$$H(W|\hat{W}) \le 1 - P_e\, ld\, 2^{nR} \le 1 - P_e \cdot nR$$

$$E = \begin{cases} 1 & w \ne \hat{w} \quad P_e \\ 0 & w = \hat{w} \quad 1-P_e \end{cases} \quad \text{ERROR}$$

$$H(E, W|\hat{W}) = H(W|\hat{W}) + H(E|W\hat{W}) = H(W|\hat{W})$$

($H(E|W\hat{W}) = 0$, $= H(P_e)$)

$$= H(E|\hat{W}) + H(W|\hat{W}, E) \le H(E) + P_e \cdot H(W)$$

$$H(W|\hat{W}, E) = P(E=0)\,H(W|\hat{W}, E=0) + P(E=1)\,H(W|\hat{W}, E=1)$$

($H(W|\hat{W},E=0)=0$, $P(E=1)=P_e$, $\le H(W)$)

$$H(W|\hat{W}) \le H(P_e) + P_e H(W) \le H(P_e) + P_e\, ld\,|W|$$

$$[Y_1 Y_2] = [1, \tfrac{1}{2}] \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

$$E[Y_i^2] = \frac{1}{i^2} E[X_i^2] + E[Z_i^2]$$

$$C = \sum_{i=1}^{n} \frac{1}{2}\, ld \left(\frac{\frac{1}{i^2}\sigma_x^2 + \sigma_z^2}{\sigma_z^2} \right) = \sum_{1}^{n} \frac{1}{2}\, ld \left(1 + \frac{1}{i^2}\frac{\sigma_x^2}{\sigma_z^2} \right)$$

(16?) $Y_i = \frac{1}{i} X_i + Z_i$

$I(X_1^n; Y_1^n) = H(Y_1^n) - H(Y_1^n | X_1^n)$

$n = 3$

$H(Y_1^3|X_1^3) = H(Y_1|X_1 X_2 X_3) + H(Y_2|X_1^3 Y_1) + H(Y_3|X_1^3 Y_1^2)$

$= H(Y_1|X_1) + H(Y_2|X_2) + H(Y_3|X_3)$

$H(Y_1|X_1) = H(\frac{1}{1}X_1 + Z_1 | X_1) = H(Z_1)$

$H(Y_2|X_2) = H(\frac{1}{2}X_2 + Z_2 | X_2) = H(Z_2)$

$I(X_1^n; Y_1^n) = H(Y_1^n) - H(Z_1^n) \overset{\circledast}{=} H(Y_1^n) - \sum_{i=1}^{n} H(Z_i)$

$H(Y_1^n) = ?$ $H(Y_1^3) = H(Y_1) + H(Y_2|Y_1) + H(Y_3|Y_2 Y_1)$

$H(W|\hat{W}) = H(P_e) + P_e \cdot ld|W| = H(P_e) + P_e \cdot nR$

$\circledast$ $I(X_1^n; Y_1^n) = H(Y_1^n) - \sum_{i=1}^{n} H(Z_i) \leq \sum_{i=1}^{n} [H(Y_i) - H(Z_i)]$

$\leq \sum_{i=1}^{n} [H(Y_i) - ldN]$

$[Y_1 Y_2]$ $Y_1 = X_1 + Z_1$ $Y_2 = \frac{1}{2}X_2 + Z_2$

$X_2' = \frac{1}{2}X_2$ $p(X_2') = \frac{p(X_2)}{\frac{dX_2'}{dX_2}} \Big|_{X_2 = 2X_2'} = 2p(2X_2')$

$p(X_2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{X_2^2}{2\sigma^2}}$ $\frac{2}{\sigma\sqrt{2\pi}} e^{-\frac{X_2^2}{2\sigma^2}} \Big|_{X_2 = 2X_2'}$

$p(X_2') = \frac{2}{\sigma\sqrt{2\pi}} e^{-\frac{4X_2'^2}{2\sigma^2}} = \frac{1}{\frac{\sigma}{2}\sqrt{2\pi}} e^{-\frac{X_2'^2}{2(\frac{\sigma}{2})^2}}$

$\sigma_{X_2'} = \frac{\sigma_{X_2}}{2}$ $\boxed{\sigma_{X_2'}^2 = \frac{\sigma_{X_2}^2}{4}}$ $\sigma_{Y_2}^2 = \frac{\sigma_{X_2}^2}{4} + \sigma_{Z_2}^2$

$\sigma_{Y_i}^2 = \frac{\sigma_{X_i}^2}{i^2} + \sigma_{Z_i}^2$ $I(X_1^n; Y_1^n) \leq \sum_{i=1}^{n} \frac{1}{2} ld \frac{\frac{\sigma_{X_i}^2}{i^2} + \sigma_{Z_i}^2}{\sigma_{Z_i}^2}$

$= \sum_{i=1}^{n} \frac{1}{2} ld\left(1 + \frac{\sigma_{X_i}^2}{i^2 \sigma_{Z_i}^2}\right) = \frac{1}{2} ld\left(1 + \frac{\sigma_{X_i}^2}{\sigma_{Z_i}^2}\right)$

$\boxed{J = \sum_{i=1}^{n} \frac{1}{2} ld\left(1 + \frac{P_i}{i^2 N}\right) + \lambda \sum P_i = 0}$

$\frac{dJ}{dP_i} = 0$ $\frac{1}{1 + \frac{P_i}{i^2 N}} \cdot \frac{1}{i^2 N} - \lambda = 0$ $\boxed{\lambda = \frac{1}{i^2 N + P_i}}$

$i^2 N + P_i = \frac{1}{\lambda}$ $\quad i^2 N + P_i = \nu$

$P_i = \nu - i^2 N$

$n=2$ $\quad 1^2 N + P_1 = \nu \quad + \quad 2^2 N + P_2 = \nu$

$N + P_1 = 4N + P_2 \qquad P_1 + P_2 = 2P \qquad P_1 = 2P - P_2$

$N + 2P - P_2 = 4N + P_2 \qquad 2P = 3N + 2P_2 \qquad P = \frac{3}{2}N + P_2$

$P_2 = P - \frac{3N}{2}$ $\qquad P_1 = 2P - P + \frac{3N}{2} = P + \frac{3N}{2}$

$P_1 = P + \frac{3N}{2}$

$$I(X_1^2; Y_1^2) = \frac{1}{2} ld(2\pi e)\left(1 + \frac{P + \frac{3N}{2}}{N}\right) + \frac{1}{2} ld(2\pi e)\left(1 + \frac{P - \frac{3N}{2}}{N}\right)$$

$$= \frac{1}{2} ld(2\pi e)\frac{\left(N + P + \frac{3N}{2}\right)\left(N + P - \frac{3N}{2}\right)}{N} = \frac{1}{2} ld(2\pi e)\frac{\left(\frac{5N}{2} + P\right)\left(P - \frac{N}{2}\right)}{N}$$

$$\frac{1}{2} ld(2\pi e)\frac{\frac{5NP}{2} + P^2 - \frac{5N^2}{4} - \frac{NP}{2}}{N} = ld(2\pi e)\left(\frac{\frac{4NP}{2} + P^2 - \frac{5N^2}{4}}{N}\right)$$

$$= \frac{1}{2} ld(2\pi e)\left(\frac{4P}{2_n} + P^2 - \frac{5N}{4_n}\right)$$

$$I(X_1^n; Y_1^n) = \sum_i H(Y_i) - \sum_i H(Y_i|X_i) \geq \sum_i^n H(Y_i) - \sum_i^n (1 - P_e nR)$$

$$H(Y_i|X_i) \leq H(P_e) - P_e \cdot n \cdot R \leq 1 - P_e n \cdot R$$

UNIVERSITY OF TORONTO SOLUTIONS

$$E[Y_i^2] = \frac{1}{i^2} E[X_i^2] + E[Z_i^2] = \frac{P_i}{i^2} + N$$

$$h(Y_i) \leq \frac{1}{2} ld(2\pi e)\left(\frac{P_i}{i^2} + N\right) \qquad E[f(x)] \leq f(E[x]), \ f(x) \text{ convex } \vee$$

$$R \leq \frac{1}{n}\sum_{i=1}^n \frac{1}{2} ld\left(1 + \frac{P_i}{i^2 N}\right) \leq \Big| \text{ JENSEN'S INEQUALITY } \Big| \leq$$

$$\leq \frac{1}{2} ld\left(1 + \frac{1}{n}\sum_{i=1}^n \frac{P_i}{i^2 N}\right)$$

$A = \frac{P_i}{i^2}$ — $P_i = A i^2$ KLJUČNA PRETPOSTAVKA!

$$\frac{1}{n}\sum_{i=1}^n \frac{A i^2}{i^2 N} = \frac{A}{nN} \cdot n = \frac{A}{N} \qquad \frac{1}{n}\sum P_i \leq P \qquad \frac{A}{n}\sum i^2 \leq P$$

$$\sum i^2 = \frac{n(2n+1)(n+1)}{6} \qquad \frac{A}{n}\frac{n(2n+1)(n+1)}{6} \leq P$$

$$A \leq \frac{6P}{(2n+1)(n+1)} \qquad R \leq \frac{1}{2} ld\left(1 + \frac{6P}{N(2n+1)(n+1)}\right) + \epsilon_n$$

$n \to \infty \quad RHS \to 0$

PROBLEM 9.17 FEEDBACK CAPACITY. LET $(Z_1, Z_2) \sim N(0, K)$ $K = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$

FIND MAXIMUM OF $\frac{1}{2}\log\frac{|K_{X+Z}|}{|K_Z|}$ WITH AND WITHOUT FEEDBACK GIVEN A TRACE (POWER) CONSTRAINT $tr(K_X) \leq 2P$.

GAUSSIAN CHANNEL WITH FEEDBACK (RECALL)

$Y_i = X_i + Z_i$ $\quad Z_i \sim N(0, K_Z^{(n)})$

$X_i(W, Y^{i-1})$ $\quad W \in \{1, 2, \dots 2^{nR}\}$

$E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2(\omega, Y^{i-1})\right] \leq P$, $\quad \omega \in \{1, 2, \dots 2^{nR}\}$

$$C_{n,FB} = \max_{\frac{1}{n}tr(K_X^{(n)}) \leq P} \frac{1}{2n}\log\frac{|K_{X+Z}^{(n)}|}{|K_Z^{(n)}|} \quad (*A)$$

$X_i = \sum_{j=1}^{i-1} b_{ij} Z_j + V_i$ $\quad i = 1, 2, \dots n$

$X = BZ + V$ $\quad Y = X + Z$

$$C_{n,FB} = \max \frac{1}{2n}\log\frac{|(B+I)K_Z^{(n)}(B+I)^t + K_V|}{|K_Z^{(n)}|}$$

$tr(BK_Z^{(n)}B^t + K_V) \leq nP$ $\quad$ $B = 0$ IF FEEDBACK IS NOT ALLOWED

- WITHOUT FEEDBACK

$$C_n = \max_{\frac{1}{n}tr(K_X^{(n)}) \leq P} \frac{1}{2n}\log\frac{|K_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|}$$

$$C_n = \frac{1}{2n}\sum_{i=1}^{n}\log\left(1 + \frac{(\lambda - \lambda_i^{(n)})^+}{\lambda_i^{(n)}}\right) \quad (y)^+ = \max(y, 0)$$

$\sum_{i=1}^{n}(\lambda - \lambda_i^{(n)})^+ = n \cdot P$

THEOREM 9.6.1 FOR A GAUSSIAN CHANNEL WITH FEEDBACK, THE RATE R_n FOR ANY SEQUENCE OF $(2^{nR_n}, n)$ CODES WITH $P_e^{(n)} \to 0$ SATISFIES

$R_n \leq C_{n,FB} + \varepsilon_n$ $\quad \varepsilon_n \to 0$ AS $n \to \infty$

WHERE $C_{n,FB}$ IS DEFINED WITH $(*A)$

$H(W|\hat{W}) \leq 1 + nR_n \cdot P_e^{(n)} = n\varepsilon_n$

$nR_n = H(W) = I(W;\hat{W}) + H(W|\hat{W}) \leq I(W;\hat{W}) + n\varepsilon_n$

$$\leq I(W; Z^n) + n\varepsilon_n = \sum_{i=1}^{n} I(W; Z_i | Z_1^{i-1}) + n\varepsilon_n$$

$$\overset{(1)}{=} \sum \left(h(Z_i | Z_1^{i-1}) - h(Z_i | W, Z_1^{i-1}) \right) + n\varepsilon_n =$$

$$\overset{(2)}{=} \sum \left(h(Z_i | Z_1^{i-1}) - h(Z_i | W, Z_1^{i-1}, X_i, X^{i-1}, Z^{i-1}) \right) + n\varepsilon_n$$

$$W = f(X_i, Z_1^{i-1}) \qquad Z^{i-1} = Z_1^{i-1} - X^{i-1}$$

$$\overset{(3)}{=} \sum \left(h(Z_i | Z_1^{i-1}) - h(Z_i | Z^{i-1}) \right) + n\varepsilon_n =$$

$$= h(Z^n) - \sum h(Z_i | Z^{i-1}) + n\varepsilon_n = h(Z^n) - h(Z^n) + n\varepsilon_n$$

$$nR_n \leq h(Z^n) - h(Z^n) + n\varepsilon_n \qquad R_n \leq \frac{1}{n} h(Z^n) - \frac{1}{n} h(Z^n) + \varepsilon_n$$

$$\boxed{R_n \leq \frac{1}{2n} \log \frac{|K_Z^{(n)}|}{|K_Z^{(n)}|} + \varepsilon_n} \leq C_{n,FB} + \varepsilon_n$$

LEMA 9.6.1 $\quad \boxed{K_{X+Z} + K_{X-Z} = 2K_X + 2K_Z}$

$$K_{X+Z} = E\left[(X+Z)(X+Z)^T\right] = E\left[(X+Z)(X^T + Z^T)\right] =$$

$$= E\left[XX^T + XZ^T + ZX^T + ZZ^T\right] = K_X + K_{XZ} + K_{ZX} + K_Z$$

$$K_{X-Z} = E\left[(X-Z)(X-Z)^T\right] = E\left[X^2 - XZ - ZX + Z^2\right] =$$

$$= K_X - K_{XZ} - K_{ZX} + K_Z$$

$$K_{X+Z} + K_{X-Z} = 2K_X + 2K_Z$$

LEMA 9.6.2 $\quad A - B$ nonnegative definite

$$|A| \geq |B|$$

$C = A - B$ $\qquad X_1 \sim N(0, B) \quad X_2 \sim N(0, C)$

$\boxed{Z = X_1 + X_2}$ $\qquad h(Z) \geq h(Z | X_2) = h(X_1 | X_2) = h(X_1)$

$A = B + C$ $\qquad \frac{1}{2} \log (2\pi e)^n |A| \geq \frac{1}{2} \log (2\pi e)^n |B|$

$\Rightarrow |A| \geq |B|$

LEMA 9.6.3 $\quad |K_{X+Z}| \leq 2^n |K_X + K_Z|$

$$2K_X + 2K_Z - K_{X+Z} = K_{X-Z} \geq 0$$

$$|2K_X + 2K_Z| \geq |K_{X+Z}| \qquad |2(K_X + K_Z)| = 2^n |K_X + K_Z|$$

$$2^n |K_X + K_Z| \geq |K_{X+Z}|$$

LEMA 9.6.4 $\quad A, B \geq 0 \qquad 0 \leq \lambda \leq 1$

$$|\lambda A + (1-\lambda) B| \geq |A|^\lambda |B|^{1-\lambda}$$

167

PROOF: $X \sim N(0,A)$ $\quad Y \sim N(0,B)$

$$Z = \begin{cases} X & \theta = 1 \\ Y & \theta = 2 \end{cases} \qquad \theta = \begin{cases} 1 & p = \lambda \\ 2 & p = 1-\lambda \end{cases}$$

$K_Z = \lambda A + (1-\lambda) B$

$K_Z = E[Z^2]$ $\qquad H(Z,\theta) = H(Z) + H(\theta|Z) = H(\theta) + H(Z|\theta)$

$$H(Z|\theta) = p(\theta=1)\, H(Z|\theta=1) + p(\theta=2)\, H(Z|\theta=2)$$

$$\lambda\, H(X) + (1-\lambda)\, H(Y)$$

$E[Z] = \int p(z) \cdot z \, dz$

$$E_Z[E_\theta[Z|\theta]] = \int p(\theta) \int f(z|\theta) \cdot z \, dz =$$

$$= \sum_{\theta=1}^{2} p(\theta) \int f(z|\theta)\, z \, dz = p(\theta=1) \cdot E[X] + p(\theta=2) \cdot E[Y]$$

$E_Z[Z] = \lambda E[X] + (1-\lambda) E[Y]$

ANALOGNO: $\quad E_Z[Z^2] = \lambda E[X^2] + (1-\lambda) E[Y^2]$

$$\boxed{K_Z = \lambda K_X + (1-\lambda) K_Y} \qquad K_Z = \lambda A + (1-\lambda) B$$

$$\frac{1}{2} \ln (2\pi e)^n |\lambda A + (1-\lambda) B| \geq h(Z) \geq h(Z|\theta) =$$

$$= \lambda h(X) + (1-\lambda) h(Y) = \frac{\lambda}{2} \log (2\pi e)^n |K_X| + \frac{(1-\lambda)}{2} \log (2\pi e)^n |K_Y|$$

$$= \frac{1}{2} \log (2\pi e)^n |K_X|^\lambda |K_Y|^{1-\lambda} = \frac{1}{2} \log (2\pi e)^n |A|^\lambda |B|^{1-\lambda}$$

$$|\lambda A + (1-\lambda) B| \geq |A|^\lambda |B|^{1-\lambda}$$

LEMMA 9.6.5 IF X^n AND Z^n ARE CAUSALLY CONNECTED THEN:

$$\boxed{h(X^n - Z^n) \geq h(Z^n)} \qquad \boxed{|K_{X-Z}| \geq |K_Z|}$$

PROOF: $h(X^n - Z^n) = \sum_{i=1}^{n} h(X_i - Z_i | X^{i-1} - Z^{i-1}) \geq$

$$\overset{(b)}{\geq} \sum_{i=1}^{n} h(X_i - Z_i | X^{i-1}, Z^{i-1}, X_i) \overset{(d)}{=} \sum_{i=1}^{n} h(Z_i | Z^{i-1}) =$$

$$\overset{(c)}{=} \sum_{i=1}^{n} h(Z_i | X^{i-1}, Z^{i-1}, X_i)$$

$= h(Z^n)$

$h(X^n - Z^n) \geq h(Z^n)$

$$\frac{1}{2} \ln (2\pi e)^n |K_{X-Z}| = h(\tilde{X}^n - \tilde{Z}^n) \geq h(\tilde{Z}^n) = \frac{1}{2} \ln (2\pi e)^n |K_Z|$$

$$|K_{X-Z}| \geq |K_Z|$$

THEOREM 9.6.2 $C_{n,FB} \le C_n + \frac{1}{2}$

$$C_{n,FB} \le \max_{tr(K_X)\le nP} \frac{1}{2n} \log \frac{|K_Y|}{|K_Z|} \le$$

$$\le \max_{tr(K_X)\le nP} \frac{1}{2n} \log \frac{2^n |K_X + K_Z|}{|K_Z|} = \max_{tr(K_X)\le nP} \frac{1}{2n} \log \frac{|K_X+K_Z|}{|K_Z|}$$

$$+ \frac{1}{2n} \cdot \log 2^n = \max_{tr(K_X)\le nP} \frac{1}{2n} \log \frac{|K_X+K_Z|}{|K_Z|} + \frac{1}{2} \le C_n + \frac{1}{2}$$

THEOREM 9.6.3 $C_{n,FB} \le 2C_n$

$$\frac{1}{2}\frac{1}{2n} \log \frac{|K_{X+Z}|}{|K_Z|} \le \frac{1}{2n} \log \frac{|K_X+K_Z|}{|K_Z|}$$

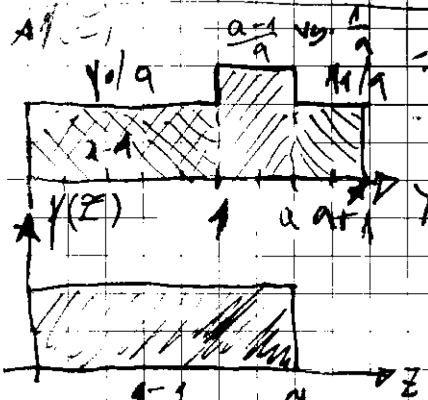
$$\frac{1}{2n} \log \frac{|K_X+K_Z|}{|K_Z|} \overset{(G)}{=} \frac{1}{2n} \log \frac{|\frac{1}{2}K_{X+Z} + \frac{1}{2}K_{X-Z}|}{|K_Z|} \ge$$

$$\overset{(e)}{\ge} \frac{1}{2n} \log \frac{|K_{X+Z}|^{\frac{1}{2}} |K_{X-Z}|^{1/2}}{|K_Z|} \ge \frac{1}{2n} \log \frac{|K_{X+Z}|^{1/2} |K_Z|^{\frac{1}{2}}}{|K_Z|} =$$

$$\frac{1}{2}\frac{1}{2n} \log \frac{|K_{X+Z}|}{|K_Z|} \qquad \frac{1}{2n} \log \frac{|K_X+K_Z|}{|K_Z|} \ge \frac{1}{2}\frac{1}{2n} \log \frac{|K_{X+Z}|}{|K_Z|}$$

$2C_n \ge C_{n,FB}$

... FROM NN6K | JOVAN STOSIĆ



$$p(0 \le y \le 1) = p(x=0) \cdot p(\ldots) = p_0 \cdot \frac{1}{a}$$

$$p(1 < y \le a) = p(\ldots) \cdot p(\ldots \le z \le a-\ldots) + p(x=0) \cdot p(1 \le z \le a) =$$

$$= p_1 \cdot \frac{a-1}{a} + p_0 \cdot \frac{a-1}{a} = \frac{a-1}{a}$$

$$\int_0^{a-1} \frac{1}{a} dz = \frac{1}{a}(a-1) \qquad \int_1^a \frac{1}{a} dz = \frac{1}{a}(a-1) \qquad \int_{a-1}^{a} \frac{1}{a} dz = \frac{1}{a}$$

$$p(a \le y \le a+1) = p(x=1) \cdot p(a \le z \le a+1) = p_1 \cdot 1/a$$

OVA SE VEROVATNO ... A NE PDF-I MMV!!!

CONCORDIA SOLUTION

$$\frac{p_0}{a} + \frac{a-1}{a} + \frac{p_1}{a} = \frac{p_0}{a} + 1 - \frac{1}{a} + \frac{p_1}{a} = \frac{1}{a} + 1 - \frac{1}{a} = 1$$ } MOŽDA DRUGAČIJE.

$$\frac{p_0}{a} + \frac{1}{a} + \frac{p_1}{a} = \frac{1}{a} + \frac{1}{a} = \frac{2}{a}$$

INTEGRALNO: $\frac{p_0}{a} + \frac{a-1}{a}(a-1) + \frac{p_1}{a}$ vs. $\frac{p_0}{a} + \frac{1}{a}(a-1) + \frac{p_1}{a} = 1$

(166) PROBLEM 8.8 RECALL:

$$\frac{1}{2}a_{-2} + \frac{1}{2}a_{-2} + \frac{1}{2}a_{-1} + \frac{1}{2}a_{-1} + \frac{1}{2}a_0 + \frac{1}{2}a_0 + \frac{1}{2}a_1 + \frac{1}{2}a_1 + \frac{1}{2}a_2$$
$$+ \frac{1}{2}a_2 = a_{-2} + a_{-1} + a_0 + a_1 + a_2 = 1$$

$$I(X;Y) = H(X) - H(X|Y) \qquad H(X) = H(p)$$

$$p(x,y) = p(x,y) \qquad p(x)\cdot p(y|x) = p(y)\cdot p(x|y)$$

BAYES CRITERION:

$$p(x|y) = \frac{p(x=1)p(y|x=1)}{p(y)} = \frac{p(x=1)p(y|x=1)}{p(x=0)p(y|x=0) + p(x=1)p(y|x=1)}$$

$$p(x=1|y) = \frac{p\cdot\frac{1}{a}}{(1-p)\frac{1}{a} + p\frac{1}{a}} = p$$

[MMV]

$$H(X|Y) = P(1 \le y \le a)\, H(q) = \frac{a-1}{a} H(q)$$

$x=0$ SIGURNO!!

$$H(X|Y) = P(0\le y\le 1)\cdot H(X|0\le y\le 1) + P(1\le y\le a)\cdot H(X|1\le y\le a) + P(a\le y\le a+1)\cdot H(X|a\le y\le a+1)$$

NE SE ZNAE KAKVO E X (IMA NEIZVESNOST)

$x=1$ SIGURNO!!!

$$H(X|1\le y\le a+1) = H(p)$$

$$H(X|Y) = P(1\le y\le a)\cdot H(X|1\le y\le a) = \frac{a-1}{a}\cdot H(p)$$

$$I(X;Y) = H(p) - \frac{a-1}{a}H(p) = H(p) - H(p) + \frac{1}{a}H(p)$$

$$I(X;Y) = \frac{H(p)}{a}$$

(b)

$$f(y) = \begin{cases} (1-p)\frac{1}{a} & 0\le y\le 1 \\ \frac{1}{a} & 1\le y\le a \\ p\,\frac{1}{a} & a\le y\le a+1 \end{cases}$$

$$h(Y) = -\int_0^1 (1-p)\frac{1}{a}\,ld\left((1-p)\frac{1}{a}\right) dy + \int_1^a \frac{1}{a}\,ld\,a\cdot dy - \int_a^{a+1} \frac{p}{a}\,ld\frac{p}{a}\,dy =$$

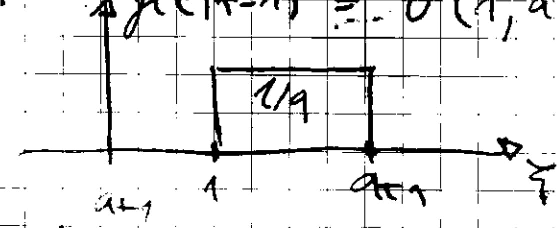
$$= -\frac{(1-p)}{a}\,ld\frac{(1-p)}{a}\,(1-0) + \frac{1}{a}\,ld\,a\,(a-1) - \frac{p}{a}\,ld\frac{p}{a}\,[a+1-a]$$

$$= -\frac{1-p}{a}\,ld(1-p) + \frac{(1-p)}{a}\,ld(a) + \frac{a-1}{a}\,ld(a) - \frac{p}{a}\,ld\,p + \frac{p}{a}\,ld\,a$$

$$= -\frac{1}{a}\left[p\,ld\,p + (1-p)\,ld(1-p)\right] + \frac{1}{a}ld(a) + ld(a) - \frac{1}{a}ld(a) = \frac{H(p)}{a} + ld\,a$$

$f(y|x=0) \sim U(0,a)$ $h(Y|X=0) = ld(a)$

$(Y|X=1)$ $f(y|x=1) = U(1, a+1)$



$h(Y|X=1) = \int \frac{1}{a}\, ld\, a\, dy = \frac{1}{a}\, ld\, (a+1-1) = ld\, a$

$h(Y) = P(X=0)\cdot h(Y|X=0) + P(X=1)\cdot h(Y|X=1) =$
$= (1-p)\, ld\, a + p\, ld\, a = ld\, a$

$I(X;Y) = h(Y) - h(Y|X) = \frac{H(p)}{a} + ld\, a - ld\, a = \frac{H(p)}{a}$

(d) THE MUTUAL INFORMATION IS MAXIMIZED FOR $p = \frac{1}{2}$ (UNIFORM DISTRIBUTION MAXIMIZES ENTROPY)

$C = \frac{1}{a} H\left(\frac{1}{2}\right) = \frac{1}{a}$ → [GOTO 8.9]

PROBLEM 9.16 GAUSSIAN MUTUAL INFORMATION

SUPPOSE THAT (X, Y, Z) ARE JOINTLY GAUSSIAN AND THAT $X \to Y \to Z$ FORMS A MARKOV CHAIN. LET X AND Y HAVE CORRELATION COEFFICIENT ρ_1 AND LET Y AND Z HAVE CORRELATION COEFFICIENT ρ_2. FIND $I(X;Z)$

$I(X,Y;Z) = I(X;Z) + I(Y;Z|X) = I(Y;Z) + I(X;Z|Y)$

$I(Y;Z|X) = H(Y|X) - H(Y|XZ) = H(Z|X) - H(Z|XY) =$
$= H(Z|X) - H(Z|Y)$

$I(X;Z|Y) = H(X|Y) - H(X|YZ) = H(Z|Y) - H(Z|XY) = H(Z|Y) - H(Z|Y) = 0$

$I(X;Y,Z) = I(X;Y) + I(X;Z|Y) = I(X;Z) + I(X;Y|Z)$

$I(X;Z|Y) = H(X|Y) - H(X|ZY) = H(Z|Y) - H(Z|XY) = 0$

$I(X,Y;Z) = H(XY) - H(XY|Z) = ld(2\pi e)^2 \sigma_X^2\sigma_Y^2(1-\rho_1^2) - H(XY|Z)$

$H(XZ) = H(Z) + H(X|Z)$

$H(X|Z) = H(X,Z) - H(Z)$

$E[(X-\bar{X})(Y-\bar{Y})] = \rho_1 \sigma_X^2 \sigma_Y^2$

$$\begin{bmatrix} \sigma_X^2 & \rho_1\sigma_X\sigma_Y \\ \rho_1\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

$$\begin{bmatrix} \sigma_X^2 & \rho_2\sigma_X\sigma_Z \\ \rho_2\sigma_X\sigma_Z & \sigma_Z^2 \end{bmatrix}$$

$I(X;Z) = H(X) - H(X|Z) = H(X) - H(X,Z) + H(Z)$
$= H(X) + H(Z) - H(X,Z) =$
$= \frac{1}{2} ld(2\pi e)\sigma_X^2 + \frac{1}{2} ld(2\pi e)\sigma_Z^2 - \frac{1}{2} ld(2\pi e)^2 |K_{XZ}| = \frac{1}{2} ld (2\pi e) \frac{\sigma_X^2\sigma_Z^2 - (\ldots)^2}{\sigma_X^2\sigma_Z^2(1-\rho_2^2)}$

169

(11) $I(x;z) = \frac{1}{2} ld \frac{1}{(2\pi e)} \frac{1}{(1-\rho_z^2)}$

$\begin{bmatrix} \sigma_x^2 & \rho_1 \sigma_x \sigma_y \\ \rho_1 \sigma_x \sigma_y & \sigma_y^2 \end{bmatrix}$ $\begin{bmatrix} \sigma_y^2 & \rho_2 \sigma_y \sigma_z \\ \rho_2 \sigma_y \sigma_z & \sigma_z^2 \end{bmatrix}$

$p(x,z|y) = p(x|y) \cdot p(z|y)$

$p(x,z|y) = \frac{p(xyz)}{p(y)} = \frac{p(xy) \cdot p(z|xy)}{p(y)} = \frac{p(x|y) p(z|y)}{p(y)}$ ✓

$p(x,z|y) = p(x|y) \cdot p(z|y)$

$E_{xz}[xz] = E_y \left[E_{xz}[xz|y] \right] = E_y \left[E_{xz}[x|y] \cdot E_{xz}[z|y] \right] = E[x] \cdot E[z]$

$\int x \cdot p(x|y)\, dx \qquad E_y[xz|y] = \int$

$E[R] = P(E) \cdot E_R[R|E] + P(\bar{E}) \cdot E_R[R|\bar{E}] =$

$= \sum_{i=E}^{\bar{E}} P(i)\, E_R[R|i]$

$E_R[R|E] = \sum_R R\, p(R|E) = 1 \cdot p(1|E) + 2\, p(2|E) +$

$+ 3\, p(3|E) + 4 \cdot p(4|E) + 5 \cdot p(5|E) + 6 \cdot p(6|E)$ (with $p(1|E)=0$, $p(3|E)=0$, $p(4|E)=1/3$, $p(5|E)=0$, $p(6|E)=1/3$, $p(2|E)=1/3$)

$E_R[R|E] = 2 \cdot \frac{1}{3} + 4 \cdot \frac{1}{3} + 6 \cdot \frac{1}{3} = \frac{12}{3} = 4$

$E_R[R|\bar{E}] = 1 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} + 5 \cdot \frac{1}{3} = \frac{9}{3} = 3$

$E[R] = \frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 3 = \frac{7}{2} = \sum_{i=1}^{6} \frac{i}{6} = \frac{1+2+3+4+5+6}{6}$

$= \frac{6+4+11}{6} = \frac{21}{6} = \frac{7}{2}$

$E_{xz}[x|y] = \int x\, p(xz|y)\, dx = E[x|y]$ PPNO (5)

$G_{xz}[xz] = \rho_z \cdot \sigma_x \cdot \sigma_z = \sigma_x \cdot \sigma_z \cdot \rho_1 \rho_2$

$\int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} x\, p(x|y) \cdot p(z|y)\, dx \right\} p(y)\, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, p(xy)\, p(z|y)\, dx\, dy$

$\begin{bmatrix} \sigma_x^2 & \rho_z \sigma_x \sigma_z \\ \rho_z \sigma_x \sigma_z & \sigma_z^2 \end{bmatrix}$

$\rho_z \sigma_x \sigma_z = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xz\, p(xz)\, dx\, dz = E_z[xz]$

$= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xz\, p(xz|y)\, dx\, dz \right\} p(y)\, dy =$

$$= \int_{-\infty}^{\infty}\int\int x z \, \gamma(x|y) \cdot \gamma(z|y)\, \gamma(y)\, dx\, dz\, dy = ⊛$$

$$\gamma(x|y) = \frac{\gamma(xy)}{\gamma(y)} \quad E_x[x|y] = \int x \cdot \gamma(x|y)\, dx = \sigma_x \cdot y \cdot \rho_1 / \sigma_y$$

$$E_z[x|y] = \int z \cdot \gamma(z|y)\, dz = \sigma_z \cdot y \cdot \rho_2 / \sigma_y$$

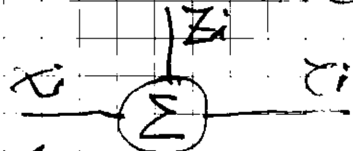
$$\gamma(z|y) = \frac{\gamma(zy)}{\gamma(y)}$$

$$⊛ = \frac{\sigma_x \sigma_z \rho_1 \rho_2}{\sigma_y^2} \int_{-\infty}^{\infty} y^2 \gamma(y)\, dy = \sigma_x \sigma_z \cdot \rho_1 \rho_2 \quad (\#)$$

$$\boxed{\rho_z = \rho_1 \rho_2}$$

p. 169 (A) $\Rightarrow$ $I(x;z) = -\frac{1}{2}\, ld (2\pi e)(1-\rho_1^2\rho_2^2)$

PROBLEM 9.17 IMPULSE POWER. CONSIDER THE ADDITIVE WHITE GAUSSIAN CHANNEL



WITH $Z_i \sim N(0, N)$ AND THE INPUT SIGNAL HAS AVERAGE CONSTRAINT P.

(a) SUPPOSE THAT WE USE ALL OUR POWER AT TIME 1 (i.e. $E x_1^2 = nP$, $E x_i^2 = 0$ FOR $i = 2, 3, \ldots n$). FIND

$$\max_{f(x^n)} \frac{I(x^n; y^n)}{n}$$

WHERE THE MAXIMIZATION IS OVER ALL DISTRIBUTIONS $f(x^n)$ SUBJECT TO THE CONSTRAINT: $E[x_1^2] = nP$, $E x_i^2 = 0$, $i = 2, \ldots n$

(b) FIND:

$$\max \frac{1}{n} I(x^n; y^n) \qquad c: E\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] \le P$$

AND COMPARE TO PART (a).

$$I(x^n; y^n) = h(y^n) - h(y^n|x^n) = h(y^n) - h(z^n) =$$

$$= h(y^n) - \frac{1}{2}\, ld\, (2\pi e)^n |K| = h(y^n) - \frac{1}{2}\, ld (2\pi e)^n N^n$$

$$\le \sum_{i=1}^{n} \frac{1}{2}\, ld\, (P_i + N_i)(2\pi e)^n - \sum_{i=1}^{n} \frac{1}{2}\, ld\, (2\pi e)^n \cdot N_i$$

$$= \sum_{i=1}^{n} \frac{1}{2}\, ld\, \frac{P_i + N_i}{N_i} (2\pi e)^n = \sum_{i=1}^{n} \frac{1}{2}\, ld \left(1 + \frac{P_i}{N_i}\right)$$

$$\sum P_i \le nP \qquad \frac{1}{1+\frac{P_i}{N_i}} \cdot \frac{1}{N_i} - \lambda = 0$$

$$\frac{1}{N_i + P_i} - \lambda = 0 \qquad N_i + P_i = \frac{1}{\lambda} \qquad \nu = N_1 + P_1 = N_2 + P_2 = N_n + P_n$$

171 JENSEN'S INEQUALITY: $E[f(X)] \geq f(E[X])$ f – CONVEX ✓

$E[f(X)] \leq f(E[X])$ CONCAVE

(b) $$\frac{I(X^n;Y^n)}{n} \leq \frac{1}{2n}\sum_{i=1}^{n} ld\left(1+\frac{P_i}{N}\right) \leq \frac{1}{2} ld\, \frac{1}{n}\sum_{i=1}^{n}\left(1+\frac{P_i}{N}\right)$$

$$= \frac{1}{2} ld\left(1+\frac{\frac{1}{n}\sum_{i=1}^{n} P_i}{N}\right) \leq \frac{1}{2} ld\left(1+\frac{P}{N}\right)$$

(a) $$\frac{I(X^n;Y^n)}{n} \leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2} ld\left(1+\frac{P_i}{N}\right) = \frac{1}{n} ld\left(1+\frac{nP}{N}\right)$$

$P_1 = nP \quad P_i = 0 \quad i = 2, \dots, n$

$$f(x^n) = \frac{1}{\sqrt{2\pi}\,|K_n|} e^{-\frac{X \cdot K^{-1} \cdot X^T}{2}}$$

$$K = \begin{bmatrix} nP & 0 & 0 & 0 \\ 0 & 0 & & \\ 0 & & \ddots & 0 \end{bmatrix}$$

$$f(x^n) = \frac{1}{nP\sqrt{2\pi}} e^{-\frac{x^2}{2nP}}$$

$$\frac{1}{2} ld\left(1+\frac{P}{N}\right) \quad \text{vs.} \quad \frac{1}{n} ld\left(1+\frac{nP}{N}\right)$$

$\geq$

"=" FOR "$n = 1$"

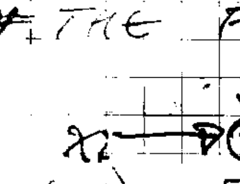
→ PROOF GRAPHICALLY IN MATLAB (FOR DIFFERENT "n")

$$\ln(1+x) \doteq 1 + \underbrace{\frac{f'(x)}{1!}\cdot x}_{O(x)} + \frac{f''(x)}{2!} x^2 \qquad f''(\ln x) = \left(\frac{1}{x}\right)' = -\frac{1}{x^2}$$

$$O(1) = \frac{1}{1+x}\cdot x \qquad \ln(1+x) \geq \ln x$$

$$\ln(x) = 1 + \left.\frac{1}{x}\right|_{x=1} \cdot x - \frac{1}{2x}\cdot x^2 \leq 1+x \qquad 0 \leq x \leq 0$$

PROBLEM 9.18 GAUSSIAN CHANNEL WITH TIME-VARYING MEAN. FIND THE CAPACITY OF THE FOLLOWING GAUSSIAN CHANNEL:



LET $Z_1, Z_2, \dots$ BE INDEPENDENT AND LET THERE BE A POWER CONSTRAINT, P ON $x^n(W)$. FIND THE CAPACITY WHEN:

(a) $\mu_i = 0$ FOR ALL i

(b) $\mu_i = e^i$ $i = 1, 2, \dots$

ASSUME THAT μ_i IS KNOWN TO TRANSMITTER AND RECEIVER

→ μ_i UNKNOWN, BUT μ_i i.i.d. $\sim \mathcal{N}(\ , 1)$ FOR ALL "i"?

$$C = \frac{1}{4}\sum_{i=1}^{4} \frac{1}{2} ld\left(1+\frac{P_i}{N_i}\right) \leq \frac{1}{2} ld\left(1+\sum \frac{P_i}{N_i}\right)$$

$N_i = E[Z_i^2] - \mu_i^2$ $\qquad$ ⊛ $= E\left[\frac{P_i}{N_i}\right]$

$$C = \frac{1}{4}\sum_{i=1}^{4} \frac{1}{2} ld\left(1+\frac{P_i}{E[Z_i^2]}\right)$$

$$C = \frac{1}{4}\sum_{i=1}^{4} \frac{1}{2} ld\left(1+\frac{P_i}{E[Z_i]^2 - e^i}\right) =$$

$$\frac{1}{4}\sum_{i=1}^{4} \frac{1}{2} ld\left(\frac{P_i}{E[Z_i^2]-e^i}\right) = \frac{1}{4}\sum_{i=1}^{4} \frac{1}{2} ld\, P_i - \frac{1}{4}\sum_{i=1}^{4} ld\left[E[Z_i^2]-e^i\right]$$

$$\leq \frac{1}{2\cdot 4}\sum_{i=1}^{4} ld\left(1+\frac{P_i}{E[Z_i^2] - 0}\right)$$

$E[\mu_i] = 0$

b) (REVISITED)

$\nu = P_i + N_i$ $\qquad$ $P_1 + N_1 = P_2 + N_2 = \dots P_4 + N_4$

$P_i > \nu - N_i$

UNIVERSITY OF CHICAGO SOLUTION

a) WHEN $\mu_i = 0$ $\qquad$ $C = 0.5 \log\left(1+\frac{P}{N}\right)$

b) WHEN $\mu_i = e^i$ SINCE IT IS KNOWN [illegible] TRANSMITTER AND THE RECEIVER, THE [illegible] WILL BE ABLE TO SUBSTRACT IT, WHILE [illegible] AND THEN WE WILL GO BACK TO THE ZERO CASE! $C = 0.5 \log\left(1+\frac{P}{N}\right)$

(c) WHEN $\mu_i \sim \mathcal{N}(0, N_1)$ $\qquad$ $Z' = Z +$

$\sigma_{Z'}^2 = \sigma_Z^2 + \sigma_\mu^2 = N + N_1$

$$C = 0.5 \log\left(1+\frac{P}{N+N_1}\right)$$

PROBLEM 9.19 PARAMETRIC FORM FOR [illegible] CAPACITY. CONSIDER "m" PARALLEL GAUSSIAN CHANNELS $Y_i = X_i + Z_i$ WHERE $Z_i \sim \mathcal{N}(0, \lambda_i)$ AND X_i ARE INDEPENDENT RANDOM VARIABLES.

$$C = \sum_{i=1}^{m} \frac{1}{2} \log\left(1+\frac{(\lambda-\lambda_i)^+}{\lambda_i}\right) \quad \text{WHERE} \quad \sum_{i=1}^{m}(\lambda-\lambda_i)^+ = P$$

177 SHOW THAT THIS CAN BE REWRITTEN AS:

$$P(\lambda) = \sum_{i:\lambda_i \leq \lambda} (\lambda - \lambda_i) \qquad C(\lambda) = \sum_{i:\lambda_i \leq \lambda} \frac{1}{2} \log \frac{\lambda}{\lambda_i}$$

WHERE $P(\lambda)$ IS PIECEWISE LINEAR AND $C(\lambda)$ IS PIECEWISE LOGARITHMIC IN λ.

$$C = \sum_{i=1}^{m} \frac{1}{2} \log\left(1 + \frac{(\lambda - \lambda_i)^+}{\lambda_i}\right) \qquad \sum_{i=1}^{m} (\lambda - \lambda_i)^+ = P$$

$$\hookrightarrow C(\lambda) = \sum_{i=1}^{m} \frac{1}{2} \log\left(1 + \left(\frac{\lambda}{\lambda_i} - 1\right)^+\right) = \sum_{\substack{i=1 \\ \lambda_i \leq \lambda}}^{m} \frac{1}{2} \log\left(1 + \frac{\lambda}{\lambda_i} - 1\right)$$

$$= \sum_{i:\lambda_i \leq \lambda} \frac{1}{2} \log\left(\frac{\lambda}{\lambda_i}\right)$$

$$P(\lambda) = \sum_{i=1}^{m} (\lambda - \lambda_i)^+ = \sum_{i:\lambda_i \leq \lambda} (\lambda - \lambda_i)$$

PROBLEM 9.20 (STAEDLER PAGE 14) ROBUST DECODING. CONSIDER AN ADDITIVE NOISE CHANNEL WHOSE OUTPUT Y IS GIVEN WITH:

$$Y = X + Z$$

WHERE THE CHANNEL INPUT X IS AVERAGE POWER LIMITED: $E[X^2] \leq P$ AND THE NOISE PROCESS $\{Z_k\}_{k=-\infty}^{\infty}$ IS I.I.D. WITH MARGINAL DISTRIBUTION $P_Z(z)$ (NON NECESSARY GAUSSIAN) OF POWER „N", $E[Z^2] = N$.

(a) SHOW THAT THE CHANNEL CAPACITY, $C = \max_{E[X^2] \leq P} I(X, Y)$ IS <u>LOWER BOUNDED</u> BY C_G WHERE

$$C_G = \frac{1}{2} \log\left(1 + \frac{P}{N}\right)$$

(I.E. THE CAPACITY C_G CORRESPONDING TO WHITE GAUSSIAN NOISE.)

(b) DECODING THE RECEIVED VECTOR TO THE CODEWORD THAT IS CLOSEST TO IT IN EUCLIDEAN DISTANCE IS IN GENERAL SUBOPTIMAL IF THE NOISE IS NON-GAUSSIAN. SHOW, HOWEVER, THAT THE RATE C_G IS ACHIEVABLE EVEN IF ONE INSIST ON PERFORMING NEAREST-NEIGHBOR DECODING (MINIMUM EUCLIDEAN DISTANCE DECODING) RATHER THAN THE OPTIMAL MAXIMUM-LIKELIHOOD OR JOINT TYPICALITY DECODING (WITH RESPECT TO THE TRUE NOISE DISTRIBUTION).

(c) EXTEND THE RESULT TO THE CASE WHERE NOISE IS NOT I.I.D BUT IS STATIONARY ERGODIC WITH POWER N.

HINT FOR (b) AND (c): CONSIDER A SIZE 2^{nR} RANDOM CODEBOOK WHOSE CODEWORDS ARE DRAWN INDEPENDENTLY OF EACH OTHER ACCORDING TO A UNIFORM DISTRIBUTION OVER THE n-DIMENSIONAL SPHERE OF

RADIUS $\sqrt{nP}$.)

(a) USING A SYMMETRY ARGUMENT, SHOW THAT CONDITIONED ON THE CODE VECTOR THE CONDITIONAL AVERAGE PROBABILITY OF ERROR DEPENDS ON THE [redacted] NOISE VECTOR ONLY VIA ITS EUCLIDEAN NORM $\|Z\|$.

(b) USE A GEOMETRIC ARGUMENT TO SHOW THAT THIS DEPENDENCE IS MONOTONIC.

(c) GIVEN A RATE $R < C_G$, CHOOSE SOME $N' > N$ SUCH THAT

$$R < \frac{1}{2}\log\left(1+\frac{P}{N'}\right)$$

COMPARE THE CASE WHERE NOISE IS I.I.D. $\mathcal{N}(0,N')$ TO THE CASE AT HAND.

CONCLUDE THE PROOF USING THE FACT THAT THE ... ENSEMBLE OF CODEBOOKS CAN ACHIEVE THE ... OF GAUSSIAN CHANNEL (NO NEED TO PROVE)

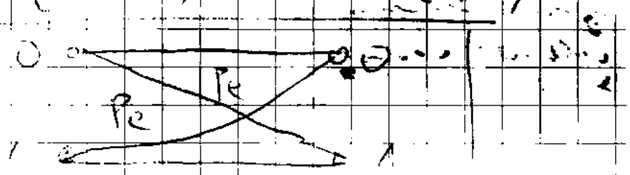
$(x_1, x_2, \ldots, x_n)$

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \le P$$

$$P_e = \frac{1}{2}P_r(Y>0 \mid X=-\sqrt{P}) + \frac{1}{2}P_r(Y<0 \mid X=\sqrt{P}) = \frac{1}{2}P_r(Z \ge \sqrt{P} \mid X=-\sqrt{P})$$

$$+\frac{1}{2}P_r(Z \le \sqrt{P} \mid X=+\sqrt{P}) = P_r(Z \ge \sqrt{P}) = 1-\Phi(\sqrt{P})$$

$$\Phi = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$$



1) GENERATION OF THE CODEBOOK.

$$\frac{1}{n}\sum x_i^2 \to P-\varepsilon \qquad x_i(w),\ i=1,2,\ldots,n;\ w=1,2,\ldots,2^{nR} \quad \sim \mathcal{N}(0, P-\varepsilon)$$

$$x^n(1), x^n(2), \ldots, x^n(2^{nR}) \in R^n$$

$$Y^n = X^n(1) + Z^n \qquad E_0 = \left\{\frac{1}{n}\sum_{i=1}^{n} x_i^2(1) > P\right\}$$

$$E_i = \left\{(X^n(i), Y^n) \in A_\varepsilon^{(n)}\right\}$$

$$P_r(\mathcal{E} \mid W=1) = P(\mathcal{E}) = P\left(E_0 \cup E_1^c \cup E_2 \cup \cdots \cup E_{2^{nR}}\right)$$

$$\le P(E_0) + P(E_1^c) + \sum_{i=2}^{2^{nR}} P(E_i) \qquad P(E_1^c) \le \varepsilon \quad n\uparrow$$

$$P_r(\mathcal{E}) = \frac{1}{2^{nR}}\sum \lambda_i = P_e^{(n)} \qquad P_e^{(n)} = P_r(\mathcal{E}) = P(E_0) + P(E_1^c) + \sum_{i=2}^{2^{nR}} P(E_i)$$

$$\le \varepsilon + \varepsilon + \sum_{i=2}^{2^{nR}} 2^{-n(I(X;Y)-3\varepsilon)} = 2\varepsilon + (2^{nR}-1)\,2^{-n(I(X;Y)-3\varepsilon)}$$

Berkley Solutions

(a) The proof depends on the entropy power inequality

$$2^{2h(X+Z)} \geq 2^{2h(X)} + 2^{2h(Z)} \quad \Big/ \; 2^{2h(Z)}$$

$$2^{2(h(X+Z)-2h(Z))} \geq 2^{2h(X)-2h(Z)} + 1$$

$E[Z^2]=N \qquad h(Z) \leq \frac{1}{2}\log(2\pi e)N \qquad h(X) = \frac{1}{2}\log(2\pi e)P$

$$2^{2(h(X+Z)-h(Z))} \geq 1 + \frac{2^{\log(2\pi e)P}}{2^{\log(2\pi e)N}} = \frac{2\pi e P}{2\pi e N} + 1$$

$$2^{2(h(X+Z)-h(Z))} \geq \frac{P}{N}+1 \qquad 2\left(h(X+Z)-h(Z)\right) \geq ld\ \frac{P}{N}+1$$

$$\underbrace{h(X+Z)-h(Z)}_{I(X;Y)} \geq \frac{1}{2}\, ld\left(1+\frac{P}{N}\right) \qquad I(X;Y) \geq \frac{1}{2}\, ld\left(1+\frac{P}{N}\right)$$

Thus we have found a specific input distribution (the Gaussian distribution with mean 0 and variance P) for which $I(X;Y)$ is at least as big as $\frac{1}{2}\log\left(1+\frac{P}{N}\right)$. This proves that capacity of the channel is bounded below by $\frac{1}{2} ld\left(1+\frac{P}{N}\right)$

Problem 9.22 (continue from N/6u) $\quad Y^n = X^n + Z^n$

$Z_i \sim N(0,N) \quad i=1,2,\dots,n \qquad \frac{1}{n}\sum_{i=1}^{n} x_i^2 \leq P$

$x^n = (0,0 \dots 0) \qquad Y^n = Z^n$

$R > 0$

$$x^n \longrightarrow (+) \longrightarrow Y^n \longrightarrow \hat{Z}^n(Y^n) \qquad P_r\{\hat{Z}^n \neq Z^n\} \to 0 \quad n \to \infty$$

(with Z^n entering the adder from above)

$$nR = H(W) = I(W;\hat{W}) + H(W|\hat{W}) \leq I(X^n;Y^n) + n\varepsilon_n =$$

$$= H(Y^n) - H(Z^n) + n\varepsilon_n = \sum_{i=1}^{n} \frac{1}{2}\, ld(2\pi e)\left(1+\frac{P_i}{N}\right) + n\varepsilon_n$$

$$R \leq \frac{1}{n}\sum_{i=1}^{n} \frac{1}{2}\, ld(2\pi e)^n\left(1+\frac{P_i}{N}\right) + \varepsilon_n \leq \frac{1}{2}\, ld(2\pi e)^n\left(1+\frac{P}{N}\right) + \varepsilon_n$$

$$R \leq \frac{1}{n}\,\frac{1}{2}\, ld(2\pi e)^n\left(1+\frac{P_1}{N}\right) \qquad P_2, P_3, \dots P_n = 0$$

$$R \leq \frac{1}{2n}\, ld(2\pi e)^n + \frac{1}{2n}\, ld\left(1+\frac{P_1}{N}\right) = ld\sqrt{2\pi e} + ld\left(1+\frac{P_1}{N}\right)^{\frac{1}{2n}}$$

$$n \to \infty \quad \Rightarrow \quad R \leq ld\sqrt{2\pi e} + 0 = ld\sqrt{2\pi e}$$

O/C SOLUTION

WE KNOW THAT: $\sup R = C = C(P/N)$

IF $R<C$ FROM THE ACHIEVABILITY PROOF OF THE CHANNEL CODING THEOREM 2^{nR} DIFFERENT X^n SEQUENCES CAN BE RECOVERED CORRECTLY WITH ARBITRARILY SMALL ERROR FOR n LARGE ENOUGH. ONCE X^n IS DETERMINED, Z^n CAN BE EASILY COMPUTED AS $Y^n - X^n$.

WE SHOW THAT THIS IS OPTIMAL BY USING PROOF BY CONTRADICTION. ASSUME THAT THERE IS SOME $R>C$ SUCH THAT Z^n CAN BE RECOVERED WITH $\Pr\{\hat{Z}^n \neq Z^n\} \to 0$ AS $n \to \infty$. BUT THIS IMPLIES THAT $\hat{X}^n = Y^n - \hat{Z}^n$ CAN BE DETERMINED WITH ARBITRARY PRECISION. THAT IS, THERE IS A CODEBOOK $X^n(W)$, $W = 1, 2, \ldots, 2^{nR}$ WITH $R>C$ AND $\Pr\{\hat{X}^n \neq X^n\} = P_e^{(n)} \to 0$ AS $n \to \infty$. AS WE SAW IN THE CONVERSE PROOF OF THE CHANNEL CODING THEOREM, THIS IS IMPOSSIBLE. HENCE WE HAVE CONTRADICTION AND R CANNOT BE GREATER THAN C.

$$y(t) = (x(t) + z(t)) * h(t)$$

BAND LIMITED CHANNEL

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega) e^{i\omega t}\, d\omega = \frac{1}{2\pi}\int_{-2\pi W}^{2\pi W} F(\omega) e^{i\omega t}\, d\omega$$

$$f\left(\frac{n}{2W}\right) = \frac{1}{2\pi}\int_{-2\pi W}^{2\pi W} F(\omega) e^{i\omega \frac{n}{2W}}\, d\omega$$

$$\operatorname{sinc}(t) = \frac{\sin(2\pi W t)}{(2\pi W t)}$$

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right) \operatorname{sinc}\left(t - \frac{n}{2W}\right)$$

IF THE NOISE HAS POWER SPECTRAL DENSITY $N_0/2$ W/Hz AND BANDWIDTH W HERTZ, THE NOISE HAS POWER $\frac{N_0}{2} 2W = N_0 W$ AND EACH OF THE $2WT$ NOISE SAMPLES IN TIME T HAS VARIANCE $N_0 W T / 2WT = N_0/2$

$$C = \frac{1}{2}\log\left(1 + \frac{P}{N}\right) \quad \text{BITS/TRANSMISSION}$$

THE ENERGY PER SAMPLE OVER THE TIME INTERVAL $[0,T]$

IN THIS CASE, THE ENERGY PER SAMPLE IS: $\frac{PT}{2WT} = \frac{P}{2W}$

THE NOISE VARIANCE PER SAMPLE IS: $\frac{N_0}{2} 2W \frac{T}{2WT} = \frac{N_0 W T}{2WT} = N_0/2$

$$C = \frac{1}{2}\log\left(1 + \frac{\frac{P}{2W}}{\frac{N_0}{2}}\right) = \frac{1}{2}\log\left(1 + \frac{P}{WN_0}\right) \quad \text{BITS/TRANSM}$$

I.E. SAMPLE

(177)

$$C = W \log\left(1 + \frac{P}{N_0 W}\right) = 2W \cdot \frac{1}{2} \log\left(1 + \frac{P}{N_0 W}\right)$$

SINCE THERE ARE 2W SAMPLES PER SECOND

THIS EQUATION IS ONE OF THE MOST FAMOUS FORMULAS OF INFORMATION THEORY. IT GIVES THE CAPACITY OF BANDLIMITED GAUSSIAN CHANNEL WITH NOISE SPECTRAL DENSITY $N_0/2$ WATTS/Hz AND POWER P WATTS.

$W \to \infty \Rightarrow C = \lim_{W \to \infty} \frac{\frac{1}{2} \log\left(1 + \frac{P}{N_0 W}\right)}{\frac{1}{2W}} =$

$= \lim_{W \to \infty} \frac{\frac{1}{2} \frac{1}{1 + \frac{P}{N_0 W}} \cdot \frac{P}{N_0 W^2} \cdot \frac{1}{\ln 2}}{\frac{1}{2W^2}} = \lim_{W \to \infty} \frac{N_0 W}{W N_0 + P} \cdot \frac{P}{N_0} \cdot \frac{1}{\ln 2}$

$= \frac{P}{N_0} \cdot \frac{\ln e}{\ln 2} = \frac{P}{N_0} \cdot \mathrm{ld}\, e$

$\log e = \mathrm{ld}\, e = \frac{\ln e}{\ln 2}$

$$C = \frac{P}{N_0} \cdot \mathrm{ld}\, e$$

FOR CHANNEL WITH INFINITE BANDWIDTH THE CAPACITY GROWS LINEARLY WITH THE POWER

$W \,\mathrm{ld}\left(1 + \frac{P}{W N_0}\right) = 3300 \cdot \mathrm{ld}(1 + 2000) = 36200$ kbps

EIGENVALUES AND EIGENVECTORS (WIKIPEDIA)

$A v = \lambda \cdot v$ — λ - EIGENVALUE, v - EIGENVECTOR

IN THREEDIMENSIONAL SPACE EIGENVECTOR v IS AN ARROW WHOSE DIRECTION IS EITHER PRESERVED OR EXACTLY REVERSED AFTER MULTIPLICATION BY A.

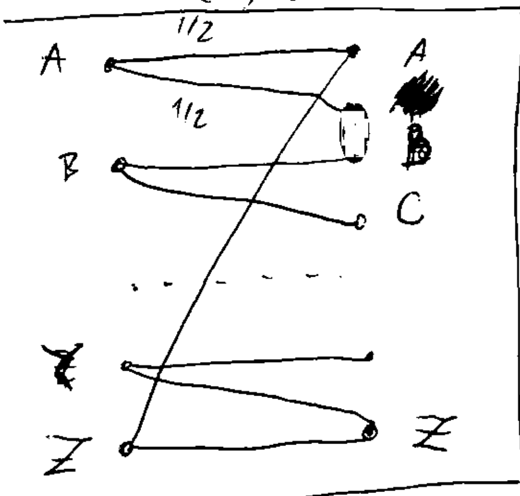
1B

7.6 **NOISY TYPEWRITER.** CONSIDER 26-KEY TYPEWRITER.

(a) IF PUSHING THE KEY RESULTS IN PRINTING THE ASSOCIATED LETTER, WHAT IS THE CAPACITY C IN BITS?

(b) NOW SUPPOSE THAT PUSHING A KEY RESULTS IN PRINTING THE LETTER OR NEXT (WITH EQUAL PROBABILITY). THUS $A \to A$ OR B, ..., $Z \to Z$ OR A. WHAT IS THE CAPACITY?

(c) WHAT IS THE HIGHEST RATE CODE WITH BLOCK LENGTH ONE THAT YOU CAN FIND THAT ACHIEVES ZERO PROBABILITY OF ERROR FOR THE CHANNEL IN PART (b)?



(b) $I(X;Y) = H(Y) - H(Y|X) =$

$= H(X) - H(X|Y)$

$H(Y) = 26 \cdot \frac{1}{26} \cdot ld\, 26 = ld\, 26$ — 1

$H(Y|X) = \sum_x P(x) \cdot H(Y|X=x)$

$P(A) \cdot H(Y|X=A) = \frac{1}{26} \cdot \Big[P(Y=A|X=A) \cdot ld \frac{1}{P(Y=A|A)} + P(Y=B|A) \cdot ld \frac{1}{P(Y=B|A)} \Big] = \frac{1}{26} \Big[\underbrace{\frac{1}{2} ld\, 2 + \frac{1}{2} ld\, 2}_{1} \Big]$

$H(Y|X) = \sum_x P(x) \cdot 1 = 1$ $\quad$ $\boxed{I(X;Y) = H(Y) - 1}$

$C = \max_{P(x)} \big[H(Y) - 1 \big] = ld\, 26 - 1 = ld\, 26 - ld\, 2 = \underline{ld\, 13}$

(a) $I(X;Y) = \underline{H(Y)} - H(Y|X) = ld\, 26 - \sum_x P(x) \big[\underbrace{1 \cdot ld\, 1}_{0} \big] =$

$= \underline{ld\, 26}$

$W \to X^n \to Y^n \to$

(c) $R = \frac{ld M}{n} = \frac{ld\, 2^{nR}}{n}$ $\quad$ $W \in \{1, 2, \ldots, 2^{nR}\}$

$H(W) = 2^{nR} \cdot \frac{1}{2^{nR}}\, ld\, 2^{nR} = n \cdot R = H(W|\hat{W}) + I(W;\hat{W}) \leq$

$\leq 1 + P_e \cdot ld\, 2^{nR} + \underline{I(W;\hat{W})} \leq 1 + P_e\, nR + \underline{I(X^n;Y^n)}$

$I(X^n;Y^n) = H(Y^n) - H(Y^n|X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1^{i-1} X_1^n$

[redacted]

$$I(X^n;Y^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}, X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i)$$

$$= \sum_{i=1}^{n} H(Y_i|Y_1^{i}) - \sum_{i=1}^{n} H(Y_i|X_i) \le \sum_{i=1}^{n} H(Y_i) - H(Y_i|X_i) =$$

$$= \sum_{i=1}^{n} I(X_i;Y_i) \le n\cdot C$$

$$n\cdot R \le 1 + P_e \cdot nR + nC \qquad R \le \frac{1}{n} + P_e\cdot R + C$$

$$n\to\infty \Rightarrow \boxed{R \le C}$$

$\boxed{n=1}$ $\quad R \le 1 + P_e\cdot R + C \qquad P_e = 0$

$$R \le 1 + C \qquad \max(R) = C+1$$

$$\boxed{R - C - 1 \le P_e R} \qquad 1 - \frac{C}{R} - \frac{1}{R} \le 0$$

$$\frac{R - C - 1}{R} \le 0 \qquad \boxed{R \le C+1}$$

$$P_e = \frac{1}{2^{nR}} \sum_{i=1}^{n} \lambda(i)$$

$$\lambda_i = P(\hat{W} \ne W | W) = P_r\left(g(Y^n) \ne i \,|\, X^n = x^n(i)\right) = \sum_{y^n} p(y^n|x^n(i)) \cdot I(g(y^n)\ne i)$$

- ZERO ERROR CODES:

$$n\cdot R = \underbrace{H(W|\hat{W})}_{0 \text{ zero error}} + I(W;\hat{W}) \le I(X^n;Y^n) \le n\,C$$

$\boxed{R \le C}$ $\quad \boxed{R \le ld\,13}$ $\quad \boxed{R_{max} = ld\,13}$

- EDITION 1 SOLUTION

SIMPLE ZERO ERROR BLOCK LENGTH ONE CODE IS ONE THAT USES EVERY ALTERNATIVE LETTER: A, C, E ... IN THIS CASE NONE OF THE CODEWORDS WILL BE CONFUSED. „A" WILL PRODUCE EITHER „A" OR „B", „C" WILL PRODUCE „C" OR „D" AND SO ON.

[redacted]

$$R = \frac{ld(\#\text{ CODEWORDS})}{\text{BLOCK LENGTH}} = \frac{ld\,13}{1} = ld\,13$$

IN THIS CASE WE CAN ACHIEVE THE CAPACITY WITH SIMPLE CODE WITH ZERO ERRORS.

PROBLEM 7.7 CASCADE OF BINARY SIMMETRIC CHANNELS.

SHOW THAT A CASCADE OF „n" IDENTICAL INDEPENDANT BINARY SIMETRIC CHANNELS,

$$x_0 \to \boxed{BSC} \to x_1 \to \cdots \to x_{n-1} \to \boxed{BSC} \to x_n .$$

EACH WITH RAW ERROR PROBABILITY p IS EQUIVALENT TO A SINGLE BSC WITH ERROR PROBABILITY

$$\frac{1}{2}\left(1-(1-2p)^n\right)$$

AND HENCE THAT $\lim_{n\to\infty} I(x_0; x_n) = 0$ IF $p \neq 0,1$.

NO ENCODING OR DECODING TAKES PLACE AT THE INTERMEDIATE TERMINALS $x_1, \ldots, x_{n-1}$. THUS, THE CAPACITY OF CASCADE TENDS TO ZERO.

$$\begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \end{bmatrix} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}^T \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (1-p)x_1 + p x_2 \\ p x_1 + (1-p)x_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1^2 \\ y_2^2 \end{bmatrix} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \begin{bmatrix} (1-p)x_1 + p x_2 \\ p x_1 + (1-p)x_2 \end{bmatrix} = \begin{bmatrix} (1-p)^2 x_1 + (1-p)p x_2 + p^2 x_1 + (1-p)p x_2 \\ p(1-p)x_1 + p^2 x_2 + p(1-p)x_1 + (1-p)^2 x_2 \end{bmatrix}$$

$$= \begin{bmatrix} \left[(1-p)^2 + p^2\right] x_1 + 2(1-p)p\, x_2 \\ 2(1-p)p\, x_1 + \left[(1-p)^2 + p^2\right] x_2 \end{bmatrix}$$

$$C = 1 - H(p) = H(Y) - H(Y|X) = 1 - H(p)$$

$$H(Y|X) = \sum p(x) \cdot H(Y|X=x) = p(x_1) H(Y|X=x_1) + p(x_2) H(Y|X=x_2) = \frac{1}{2} H(p) + \frac{1}{2} H(p) = H(p)$$

[redacted]

$$\begin{bmatrix} Y_1^{(2)} \\ Y_2^{(2)} \end{bmatrix} = \begin{bmatrix} [(1-\gamma)^2+\gamma^2]x_1 + 2[(1-\gamma)\gamma]x_2 \\ 2(1-\gamma)\gamma\, x_1 + [(1-\gamma)^2+\gamma^2]x_2 \end{bmatrix}$$

$$I(X; Y^{(2)}) = H(Y^{(2)}) - H(Y^{(2)}|X) = ld\,2 - H(Y^{(2)}|X)$$

$$H(Y^{(2)}|X) = \sum_x p(x)\, H(Y^{(2)}|X=x) = \frac{1}{2}[(1-\gamma)^2+\gamma^2]\, ld \frac{1}{(1-\gamma)^2+\gamma^2} + \frac{1}{2}\, 2(1-\gamma)\gamma\, ld\frac{1}{2(1-\gamma)\gamma} =$$

$$= \Big/ (1-\gamma)^2+\gamma^2 = 1-2\gamma+\gamma^2+\gamma^2 = 1-2\gamma+2\gamma^2,\quad 2(1-\gamma)\gamma = 2\gamma-2\gamma^2 \Big/ =$$

$$= \frac{1}{2}\, ld\frac{1}{(1-\gamma)^2+\gamma^2} - \gamma\, ld\frac{1}{(1-\gamma)^2+\gamma^2} + \gamma^2\, ld\frac{1}{(1-\gamma)^2+\gamma^2} + \gamma\, ld\frac{1}{2(1-\gamma)\gamma} - \gamma^2\, ld\frac{1}{2(1-\gamma)\gamma}$$

$$= \frac{1}{2}\, ld\frac{1}{(1-\gamma)^2+\gamma^2} + \gamma\, ld\frac{(1-\gamma)^2+\gamma^2}{2(1-\gamma)\cdot\gamma} + \gamma^2\, ld\frac{2(1-\gamma)\gamma}{(1-\gamma)^2+\gamma^2}$$

$$\frac{1}{2}\left(1-(1-2\gamma)^2\right) = \frac{1}{2}\left(1-1+4\gamma-4\gamma^2\right) = \frac{4}{2}(\gamma-\gamma^2)$$

$$= 2\gamma(1-\gamma)$$

$$P_e = \frac{1}{2}\sum_i \lambda(i) = \frac{1}{2}\Big[P_r(g(Y^n)=X_2 \mid X=x_1) + P_r(g(Y^n)=x_1 \mid X=x_2)\Big] = \frac{1}{2}\Big[2(1-\gamma)\gamma + 2(1-\gamma)\gamma\Big] = 2(1-\gamma)\gamma$$

$$\begin{bmatrix} Y_1^{(3)} \\ Y_2^{(3)} \end{bmatrix} = \begin{bmatrix} \sim & \gamma_{12} \\ \gamma_{21} & \sim \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\frac{1}{2}\left(1-(1-2\gamma)^2\right) = \frac{1}{2}\left(1-1+4\gamma-4\gamma^2\right) = \frac{1}{2}\cdot 4(\gamma-\gamma^2) =$$

$$= 2\gamma(1-\gamma)$$

[redacted]

$$\begin{bmatrix} (1-\gamma) & \gamma \\ \gamma & (1-\gamma) \end{bmatrix} \begin{bmatrix} [(1-\gamma)^2+\gamma^2]x_1 + 2(1-\gamma)\gamma\, x_2 \\ 2(1-\gamma)\gamma\, x_1 + [(1-\gamma)^2+\gamma^2]x_2 \end{bmatrix}$$

$$\gamma_{12}^{(3)} = 2(1-\gamma)^2\gamma + [(1-\gamma)^2+\gamma^2]\gamma = \underbrace{3(1-\gamma)^2\gamma + \gamma^3}_{\circledast}$$

$$\frac{1}{2}\left(1-(1-2\gamma)^3\right) = \frac{1}{2}\left(1-(1-6\gamma+12\gamma^2-8\gamma^3)\right)$$

$$= \frac{1}{2}\left(1-1+6\gamma-12\gamma^2+8\gamma^3\right) = (3-6\gamma+4\gamma^2)\cdot\gamma$$

$$\circledast = 3(1-2\gamma+\gamma^2)\gamma + \gamma^3 = 3\gamma - 6\gamma^2 + 3\gamma^3 + \gamma^3 =$$

$$= 3\gamma - 6\gamma^2 + 4\gamma^3 = (3-6\gamma+4\gamma^2)\gamma$$

$$\gamma_{11}^{(3)} = (1-\gamma)^3 + \gamma^2(1-\gamma) + 2(1-\gamma)\gamma^2 = (1-\gamma)^3 + 3\gamma^2(1-\gamma)$$

$$= (1-\gamma)\left[(1-\gamma)^2 + 3\gamma^2\right] = (1-\gamma)\left[1-2\gamma+\gamma^2+3\gamma^2\right] =$$

$$= (1-\gamma)\left[1-2\gamma+4\gamma^2\right] = 1-2\gamma+4\gamma^2-\gamma+2\gamma^2-4\gamma^3 =$$

$$= 1-3\gamma+6\gamma^2-4\gamma^3 = (1-2\gamma)^3$$

$$\boxed{\gamma_{12} = 1-\gamma_{11} = 1-(1-2\gamma)^3}$$

$$\Pi^{(3)} = \begin{bmatrix} (1-2\gamma)^3 & 1-\gamma_{11} \\ 1-\gamma_{22} & (1-2\gamma)^3 \end{bmatrix}$$

$$\Pi^{(2)} = \begin{bmatrix} (1-\gamma)^2+\gamma^2 & 2(1-\gamma)\gamma \\ 2(1-\gamma)\gamma & (1-\gamma)^2+\gamma^2 \end{bmatrix} =$$

$$1-\frac{1}{2}\left(1-(1-2\gamma)^2\right) = 1-\frac{1}{2}+\frac{1}{2}(1-2\gamma)^2 = \frac{1}{2}+\frac{1}{2}(1-2\gamma)^2 = \frac{1}{2}(1+(1-2\gamma)^2)$$

$$\gamma_{11}^{(2)} = (1-\gamma)^2+\gamma^2 = 1-2\gamma+\gamma^2+\gamma^2 = (1-2\gamma+2\gamma^2) = \frac{1}{2}\left(1+(1-2\gamma)^2\right)$$

$$\frac{1}{2}\left[1+(1-2\gamma)^2\right] = \left[1+1-4\gamma+4\gamma^2\right]\frac{1}{2} = 1-2\gamma+2\gamma^2$$

[redacted]

$$\begin{bmatrix} 1-\gamma & \gamma \\ \gamma & 1-\gamma \end{bmatrix}\begin{bmatrix} 1-\gamma & \gamma \\ \gamma & 1-\gamma \end{bmatrix}\begin{bmatrix} 1-\gamma & \gamma \\ \gamma & 1-\gamma \end{bmatrix}\begin{bmatrix} 1-\gamma & \gamma \\ \gamma & 1-\gamma \end{bmatrix}$$

$$\begin{bmatrix} (1-\gamma)^2+\gamma^2 \,;\, 2(1-\gamma)\gamma \\ 2(1-\gamma)\gamma \,;\, (1-\gamma)^2+\gamma^2 \end{bmatrix}\begin{bmatrix} (1-\gamma)^2+\gamma^2 \,;\, 2(1-\gamma)\gamma \\ 2(1-\gamma)\gamma \,;\, (1-\gamma)^2+\gamma^2 \end{bmatrix}$$

$$\begin{bmatrix} 1-\gamma & \gamma \\ \gamma & 1-\gamma \end{bmatrix}\begin{bmatrix} (1-\gamma)^2+\gamma^2 & 2(1-\gamma)\gamma \\ 2(1-\gamma)\gamma & (1-\gamma)^2+\gamma^2 \end{bmatrix} =$$

$$= \begin{bmatrix} (1-\gamma)^3+\gamma^2(1-\gamma)+2\gamma^2(1-\gamma) \,;\, 2(1-\gamma)^2\gamma+\gamma(1-\gamma)^2+\gamma^3 \\ 2(1-\gamma)^2\gamma+\gamma(1-\gamma)^2+\gamma^3 \,;\, (1-\gamma)^3+\gamma^2(1-\gamma)+2\gamma^2(1-\gamma) \end{bmatrix}$$

$$(1-\gamma-\gamma)^2 = (1-\gamma)^2 + 2\gamma(1-\gamma) + \gamma^2$$

$$(1-\gamma-\gamma)^3 = (1-\gamma)^3 - \binom{3}{1}(1-\gamma)\cdot\gamma^2 + \binom{3}{2}(1-\gamma)^2\gamma - \gamma^3$$

$$= (1-\gamma)^3 - 3(1-\gamma)\gamma^2 + 3(1-\gamma)^2\gamma - \gamma^3$$

$$1-(1-\gamma-\gamma)^3 = \cancel{1} - \cancel{\ldots} + \gamma^3 + 3(1-\gamma)\gamma^2 - 3(1-\gamma)^2\gamma + \gamma^3$$

$$= 3(1-\gamma)\gamma^2 - 3(1-\gamma)^2\gamma + 2\gamma^3 =$$

$$= 3(1-\gamma)\gamma^2 - 5(1-\gamma)^2\gamma + 2(1-\gamma)^2\gamma + 2\gamma^3 =$$

$$= 3(1-\gamma)\gamma^2 - 5(1-\gamma)^2\gamma + 2(1-\gamma)^2\gamma = 3\gamma^2-3\gamma^3-5\gamma+10\gamma^2-5\gamma^3$$

$$(*) = 3\gamma^2 - 3\gamma^3 - 5\gamma(1-2\gamma+\gamma^2) = 3\gamma^2-3\gamma^3-5\gamma+10\gamma^2-5\gamma^3$$

CONTINUE FROM N16a. 74

$$P_e^{(n)} = \frac{1}{2^{nR}}\sum_{i=1}^{n}\binom{n}{i}2^{-i}$$

$$P_e^{(n)} = \frac{1}{2^{nR}}\left(\left(\frac{3}{2}\right)^n - 1\right)$$

$$\left(\frac{3}{2}\right)^n = 1 \Rightarrow n \cdot ld\left(\frac{3}{2}\right) = 0$$

$$\boxed{R_{min} = 0.6}$$

$$\lim_{n\to\infty} \frac{1}{2^{nR}}\left(\left(\frac{3}{2}\right)^n - 1\right) = ? \Rightarrow P_e = 0 \text{ FOR } R \in [0.6, 1]$$

[redacted]

14.09.13

SUŠTINATA NA ZADAČAVA 7-E NA MOETO REŠENIE E OVAA:

$$\lambda_i = P_r\left(g(Y^n)\neq i \mid X^n = x^n(i)\right) = \sum_{Y^n} P\left(Y^n \mid X^n = x^n(i)\right) I\left(g(Y^n)\neq i\right)$$

$P(Y^n \mid X^n = x^n(i))$ → OVA E VAŽNO DA SE NAJDE

- PRETHODNO VAKA GO IMAM REŠENO:

i=1: $$\lambda_i = \binom{n}{1} \cdot P\left(Y^n \mid X^n(i=1)\right) = \binom{n}{1}\frac{1}{2^1}$$ — SAMO EDNA EDINICA IMA OVAA SEKVENCA

i=2: $$\lambda_i = \binom{n}{2}\frac{1}{2^2}$$ — VEROJATNOST DVE EDINICI DA BIDAT ZGREŠENI

$\binom{n}{2}$: BROJ NA SEKVENCI SO „2" EDINICI

AMA NE MORA 2 EDINICI DA BIDAT ZGREŠENI, MOŽE EDNA DA BIDE ZGREŠENA ILI „2" DA SE ZGREŠENI PA ZATOA:

$$\lambda_i = \binom{n}{2}\left(\frac{1}{2}+\frac{1}{2^2}\right)$$

- VO GENERALEN SLUČAJ

i=k: $$\lambda_k = \binom{n}{k}\left(\frac{1}{2}+\frac{1}{2^2}+\dots+\frac{1}{2^k}\right) = \binom{n}{k}\frac{1}{2}\underbrace{\left(1+\frac{1}{2}+\dots\frac{1}{2^{k-1}}\right)}_{S}$$

$$S = \frac{1-\frac{1}{2^k}}{1-\frac{1}{2}} = 2\left(1-\frac{1}{2^k}\right); \quad \boxed{\lambda_k = \binom{n}{k}\left(1-\frac{1}{2^k}\right)}$$

$$P_e^{(n)} = \frac{1}{2^{nR}}\sum_{i=1}^{n}\binom{n}{i}\left(1-\frac{1}{2^i}\right) = \frac{1}{2^{nR}}\left[\sum_{i=1}^{n}\binom{n}{i} - \sum_{i=1}^{n}\binom{n}{i}\frac{1}{2^i}\right]$$

$$= \frac{1}{2^{nR}}\left[2^n - \left(\frac{3}{2}\right)^n\right] = \frac{2^n}{2^{nR}} - \frac{3^n}{2^{n+nR}} = \frac{2^{2n}-3^n}{2^{nR+n}}$$

$$= \frac{1}{2^{nR-n}} - \frac{3^n}{2^{nR+n}}$$

$$\lim_{n\to\infty}\frac{2^{2n}-3^n}{2^{nR+n}} = 0 \Rightarrow R = 1.000..1 = 1+\varepsilon$$

$$n\cdot R \le 1 + P_e \cdot nR + nC \qquad R \le \frac{1}{n} + P_e\cdot R + C$$

$$n\to\infty \qquad R \le C$$

[redacted]

EDITION 1 SOLUTION

$x \in [0, 1]$

$$\begin{bmatrix} p(y_1) \\ p(y_2) \end{bmatrix} = \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \end{bmatrix} \qquad // \quad p(x) = \left\{ \frac{1}{2}, \frac{1}{2} \right\}$$

$$H(Y) = -\left[p(Y=y_1) \cdot ld\, p(Y=y_1) + p(Y=y_2)\, ld\, p(Y=y_2) \right]$$

$$H(Y|X) = \sum_x p(x) \cdot H(Y|X=x)$$

$$p(Y=y_1) = p(x_1) + \frac{1}{2} p(x_2) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

$$p(Y=y_2) = \frac{1}{2} \cdot p(x_2) = \frac{1}{4}$$

$$H(Y) = + \left(\frac{3}{4}\, ld \frac{4}{3} + \frac{1}{4}\, ld 4 \right) = + \left[\frac{3}{4}\, ld 4 - \frac{3}{4}\, ld 3 + \frac{1}{2} \right] =$$

$$= + \left[\frac{3}{2} + \frac{1}{2} - \frac{3}{4}\, ld 3 \right] = 2 - \frac{3}{4}\, ld 3$$

$$H(Y|X) = \underbrace{\frac{1}{2}}_{p(X=x_1)} \cdot \underbrace{\left[1 \cdot ld 1 \right]}_{H(Y|X=x_1)} + \underbrace{p(X=x_2)}_{1/2} \cdot \underbrace{\left[\frac{1}{2}\, ld 2 + \frac{1}{2}\, ld 2 \right]}_{H(Y|X=x_2)} =$$

$$= \frac{1}{2} \left[\frac{1}{2} + \frac{1}{2} \right] = 1/2$$

$$I(X;Y) = H(Y) - H(Y|X) = 2 - \frac{3}{4}\, ld 3 - \frac{1}{2} = \frac{3}{2} - \frac{3}{4}\, ld 3$$

$$= \underbrace{0.31128}_{RATE} \le \underbrace{0.322}_{CAPACITY}$$

$$R = \frac{2^{nR}}{n}$$

RESENIETO VO PROBLEM 8 DOAN DOBIV ? $p(x_2) = \frac{2}{5}$, $p(x_1) = \frac{3}{5}$ SE POSTIGNUVA CEGA $p(x_1) = p(x_2)$

UMA CAPACITETOT NA KANALOT $= \frac{1}{2}$ IA ~~ZATOA~~ NE SE POSTIGNUVA KAPACITETOT NA KANALOT.

!!! (MMV)

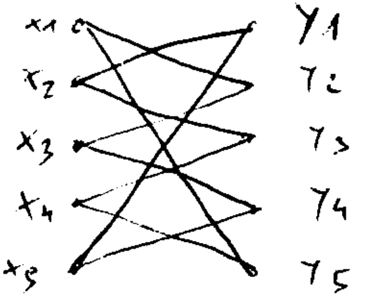
FROM THE PROOF OF THE CHANNEL CODING THEOREM, IT FOLLOWS THAT USING A RANDOM CODE WITH CODEWORDS GENERATED ACCORDING TO PROBABILITY $p(x)$ WE CAN SEND INFORMATION AT RATE $I(X;Y)$ CORRESPONDING TO THAT $p(x)$ WITH AN ARBITRARY PROBABILITY OF ERROR.

JAS GA REŠAVAV ZADAČATA KAKO DA NE GA ZNAM CHANNEL CODING THEOREMATA

[redacted]

CONTINUE FROM (80a)

$X^4 = \{14, 20, 31, 42, 03\}$



$(x_2 x_2) \to Z_4 = (y_1 \vee y_3, y_1 \vee y_3) = y_1y_1, y_1y_3, y_3y_1, y_3y_3$

$(x_4 x_5) \to Z_4 = (y_3 \vee y_5, y_1 \vee y_4) = y_3y_1, y_3y_4, y_5y_1, y_5y_4$

DOKOLKU BI GI NAJDE IDEALNITE KOMBINACII NA $x_i x_j$ TAKA ŠTO NE SE JAVI GREŠKA:

$H(Y^4|X^4) = \sum p(x_i x_j)\, H(Y^4|x_i x_j) = 5 \cdot \frac{1}{5} \cdot \frac{1}{4} \operatorname{ld} 4 = \frac{1}{2}$

$C = \max_{p(x)} I(X^4, Y^4) = H(Y^4) - H(Y^4|X^4) = \operatorname{ld} 5 - \frac{1}{2} = 1.82$

$(x_3 x_5) \to (y_2 \vee y_4, y_1 \vee y_4) = y_2y_1, y_2y_4, y_4y_1, y_4y_4$

$(x_1 x_3) \to (y_2 \vee y_5, y_2 \vee y_4) = y_2y_2, y_2y_4, y_5y_2, y_5y_4$

$(x_2 x_5) \to (y_1 \vee y_3, y_1 \vee y_4) = y_1y_1, y_1y_4, y_3y_1, y_3y_4$

$(x_4 x_2) \to (y_3 \vee y_5, y_1 \vee y_3) = y_3y_1, y_3y_3, y_5y_1, y_5y_3$

$(x_1 x_4) \to (y_2 \vee y_5, y_3 \vee y_5) = y_2y_3, y_2y_5, y_5y_3, y_5y_5$

$(x_2 x_4) \to (y_1 \vee y_3, y_3 \vee y_5) = y_1y_3, y_1y_5, y_3y_3, y_3y_5$

$(x_5 x_3) \to (y_1 \vee y_4, y_2 \vee y_4) = y_1y_2, y_1y_4, y_4y_2, y_4y_4$

$(x_5 x_2) \to (y_1 \vee y_4, y_1 \vee y_3) = y_1y_1, y_1y_3, y_4y_1, y_4y_3$

$(x_4 x_2) \to (y_3 \vee y_5, y_1 \vee y_3) = y_3y_1, y_3y_3, y_5y_1, y_5y_3$

$(x_4 x_1) \to (y_3 \vee y_5, y_2 \vee y_5) = y_3y_2, y_3y_5, y_5y_2, y_5y_5$

$(x_3 x_1) \to (y_2 \vee y_4, y_2 \vee y_5) = y_2y_2, y_2y_5, y_4y_2, y_4y_5$

$= \frac{n!}{(n-k)!\,k!}$ 1 2 3 / 12 13 23 $\binom{n}{2}$ } KOMBINACII

12	13	23
21	31	32

$P = \frac{n!}{(n-k)!} = \frac{6}{1} = 6$

$(x_2 x_5) \to (y_1 \vee y_3, y_1 \vee y_4) = y_1y_1, y_1y_4, y_3y_1, y_3y_4$

$(x_3 x_1) \to (y_2 \vee y_4, y_2 \vee y_5) = y_2y_2, y_2y_5, y_4y_2, y_4y_5$

$(x_4 x_2) \to (y_3 \vee y_5, y_1 \vee y_3) = y_3y_1, y_3y_3, y_5y_1, y_5y_3$

$(x_5 x_3) \to (y_1 \vee y_4, y_2 \vee y_4) = y_1y_2, y_1y_4, y_4y_2, y_4y_4$

$(x_1 x_4) \to (y_2 \vee y_5, y_3 \vee y_5) = y_2y_3, y_2y_5, y_5y_3, y_5y_5$

[redacted]

$X^4 = \{11, \underset{(a)}{20}, \underset{(c)}{31}, 42, 03\}$

SAMOSTOJNO DOPOLN PO OVIE VODNI ZBOROVI. (a) GI POTPRV ZA DA NEMAM 2X „0" T.E. „1" NA PRVIOT DIGIT OD KODNIOT ZBOR.
SOTAK VARVIOD KOD NE BEZGREŠEN ZATO ŠTO VO VCORITE ZBOROVI

$(x_3 x_1) \to (y_2 V y_4, y_2 V y_5) = (y_2 y_2, \underline{\underline{y_2 y_5}}, y_4 y_2, y_4 y_5)$

$(x_1 x_4) \to (y_2 V y_5, y_3 V y_5) = (y_2 y_3, \underline{\underline{y_2 y_5}}, y_5 y_2, y_5 y_5)$

IMA PO EDNA MOŽNA KOMBINACIJA KOJA MOŽE DA VODI KON GREŠKA T.E.

$H(Y^4|X^4) = ?$ $\quad H(Y^4|X^4) = \sum_{x^4} P(x^4)\, H(Y^4|X^4 = x^4)$

ČLENOVITE ZA: $x^4 = (x_2 x_5, x_4 x_2, x_5 x_3)$ VODAT KON UNIKATNI Y^4 T.E. NE SE JAVUVA GREŠKA.

$\hat{W} = (0, 1, 2, 3, 4)$ $\quad Y^4(\hat{W}=0) = Y_1^4 = (\underbrace{y_1 y_1, y_1 y_4, y_3 y_1, y_3 y_4})$

$Y^4 = (Y_1^4, Y_2^4, Y_3^4, Y_4^4, Y_5^4)$

BILO KOJA KOMBINACIJA OD OVIE IZLEZNI KODNI ZBOROVI ISKLUČIVO PODRAZBIRA DEKA BIL ISPRATEN $(x_2 x_5) = x_1^4 = x^4(W=0)$

ISTOTO VAŽI I ZA: $x^4(W=2) = (x_4 x_2)$

$\hat{W} = 2 = g(Y^4)$ $\quad Y^4(\hat{W}=2) = (\underbrace{y_5 y_1, y_5 y_3, y_5 y_1, y_5 y_3})$

BILO KOJ OD OVIE KOMBINACII OD KODNI ZBOROVI DA SE DOBIJAT NA IZLEZ KJE SE DEKODIRA $\hat{W} = 2$

$H(Y^4|X^4) = \sum_{x_m = x_3 x_1, x_1 x_4} P(x^4)\, H(Y^4|X^4 = x_m^4) = \frac{1}{5} H(Y^4|X^4 = x_3 x_1) +$ **MMV**

$+ \frac{1}{5} H(Y^4|X^4 = x_1 x_4) = (*)$

AKO PRATIŠ $(x_3 x_1)$: $g(Y^4) = 1$ ILI $g(Y^4) = 4$

AKO PRATIŠ $(x_1 x_4)$: $g(Y^4) = 4$ ILI $g(Y^4) = 1$

$(*) = \frac{1}{5}\Big[\underbrace{\frac{3}{4} ld \frac{4}{3} + \frac{1}{4} ld \frac{4}{1}}\Big] + \frac{1}{5}\Big[\frac{3}{4} ld \frac{4}{3} + \frac{1}{4} ld \frac{4}{1}\Big] =$

$= \frac{2}{5}\Big[\overset{g(Y^4)=1}{\frac{3}{4} ld \frac{4}{3}} + \overset{g(Y^4=4)}{\frac{1}{4_2} 2}\Big] = \frac{2}{5}\Big[\frac{3}{4} \cdot 2 - \frac{3}{4} ld 3 + \frac{1}{2}\Big] =$

$= \frac{2}{5}\Big[2 - \frac{3}{4} ld 3\Big] = 0.32$, $\quad$ $C_2 = H(Y^4) - H(Y^4|X^4) = ld 5 - 0.32 = 1.99742$; $C_1 = \frac{C_2}{2} \doteq 1$

[redacted]

CONTINUE FROM 87a

CHANNEL WITH DEPENDANCE BETWEEN LETTERS.

$00 \to 01 \quad 01 \to 10 \quad 10 \to 11 \quad 11 \to 00$

(a) $I(x_1x_2; z_1z_2) = H(z_1z_2) - \underbrace{H(z_1z_2|x_1x_2)}_{0}$

$p(x_1x_2) = \{p_1, p_2, p_3, p_4\}$

$H(y_1y_2|x_1x_2) = \sum_{x_1x_2} p(x_1x_2)\, \underbrace{H(z_1z_2|x_1x_2 = x_1x_2)}_{0} = 0$

$H(z_1z_2) = ?$ $\quad P(z_1z_2 = y_1y_2) = p(x_1x_2) \cdot \underbrace{p(z_1z_2 = y_1y_2|x_1x_2)}_{1}$

$p(y_1y_2) = p(x_1x_2) = \{p_1, p_2, p_3, p_4\}$

$H(z_1z_2) = H(x_1x_2) = H(p_1, p_2, p_3, p_4)$

e.g. IF $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$

$H(z_1z_2) = 4 \cdot \frac{1}{4} \, ld\, 4 = 4 \cdot 1/2 = 2$

(b) $C = \max_{p(x_1x_2)} I(x_1x_2; y_1y_2) = \max_{p(x_1x_2)\ \text{UNIFORM}} [H(z_1z_2)] = \left(\frac{1}{4}\, ld\, 4\right) 4 = ld\, 4 = 2$ bits

(c) $I(x_1; z_1) = 0$

$p(x_1x_2) = \left\{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right\}$

$p(x_1x_2)$

$x_1 \backslash x_2$	0	1	$p(x_1)$
0	1/4	1/4	1/2
1	1/4	1/4	1/2
$p(x_2)$	$\frac{1}{2}$	$\frac{1}{2}$	

$p(x_1) = \sum_{x_2} p(x_1x_2 = x_2)$

$p(y_1y_2) = \left\{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right\}$

$y_1y_2 = \{01, 10, 11, 00\}$

$p(y_1) = \left\{\frac{1}{2}, \frac{1}{2}\right\} \quad p(z_2) = \left\{\frac{1}{2}, \frac{1}{2}\right\}$

$I(x_1; z_1) = H(z_1) - H(z_1|x_1) = 2 \cdot \frac{1}{2}\, ld\, 2 - H(z_1|x_1)$

$= 1 - H(z_1|x_1)$; $H(z_1|x_1) = ?$

$H(z_1|x_1) = \sum_{x_1} p(x_1) \cdot H(z_1|x_1 = x_1) = \frac{1}{2} \cdot H(z_1|x_1 = 0) + \frac{1}{2} H(z_1|x_1 = 1) = -\frac{1}{2}\left[P(z_1 = 0|x_1 = 0) \cdot ld\, P(z_1 = 0|x_1 = 0) + P(z_1 = 1|x_1 = 0) \cdot ld\, P(z_1 = 1|x_1 = 0)\right] - \frac{1}{2}\left[P(z_1 = 0|x_1 = 1) \cdot ld\, P(z_1 = 0|x_1 = 1) + P(z_1 = 1|x_1 = 1) \cdot ld\, P(z_1 = 1|x_1 = 1)\right]$

[redacted]

12b

$H(Y_1|X_1) = +\frac{1}{2}\left[\frac{1}{2} ld 2 + \frac{1}{2} ld 2\right] + \frac{1}{2}\left[\frac{1}{2} ld 2 + \frac{1}{2} ld 2\right] =$

$= 2 \cdot \frac{1}{2}\left[\frac{1}{2} + \frac{1}{2}\right] = 1$

$I(X_1; Y_1) = H(Y_1) - H(Y_1|X_1) = 1 - 1 = 0$

$\max_{p(x_1 x_2)} I(X_1 X_2; Y_1 Y_2) = 2$

12e

CONTINUE FROM 919

(a) $H(X) = \left(\frac{1}{2} ld 2\right) \cdot 2 = 1$ $\quad H(Y) = \left(\frac{1}{2} ld 2\right) \cdot 2 = 1$

$H(X,Y) = -2(0.45\, ld\, 0.45) - 2(0.05\, ld\, 0.05) = 1.469$

$H(Y|X) = ?$

$I(X;Y) = H(X) - H(Y|X)$

X\Y	0	1
0	0.9	0.1
1	0.1	0.9

$P(Y|X)$

$P(Y|X) = \frac{P(X,Y)}{P(X)}$

$H(Y|X) = -2(0.9\, ld\, 0.9) - 2(0.1\, ld\, 0.1)$
$= 0.94 - 0.73799$ (?)

?

$I(X;Y) = 1 - 0.94 \doteq 0.06$

$H(X,Y) = H(X) + H(Y|X) = 1 + 0.937?$

$H(Y|X) = \sum_{xy} p(x,y) \cdot ld \frac{1}{P(Y|X)} =$

$= (0.45 \cdot ld\, 0.9) \cdot 2 + 2(0.05\, ld\, 0.1) = 0.469$

$H(X;Y) = H(X) + H(Y|X) = 1 + 0.469 = 1.469$ ✓

$I(X;Y) = 1 - 0.469 = 0.531$

(b) X_1^n Bernoulli$\left(\frac{1}{2}\right)$ $\quad \left| -H(X) - \frac{1}{n} ld\, p(x_1^n) \right| \le \varepsilon$

$2^{-n(H+\varepsilon)} \le p(x_1^n) \le 2^{-n(H-\varepsilon)}$

$H(X) = \left(\frac{1}{2} ld 2\right) \cdot 2 = 1$ $\quad H - \varepsilon \le -\frac{1}{n} ld (x_1^n) \le H + \varepsilon$

$1 - 0.2 \le \frac{1}{n} ld(p(x_1^n)) \le 1.2$ $\quad 0.8 \le \frac{1}{n} ld\, p(x_1^n) \le 1.2$

[redacted]

$$0.8 \le -\frac{1}{n} ld\, p(x_1^n) \le 1.2$$

$$p(x_1^n) = p(x_1)\, p(x_2) \cdots p(x_n)$$

$$\frac{1}{n} ld\, p(x_1^n) = \frac{1}{n} \sum_{i=1}^{n} ld\, p(x_i)$$

$k=1 \quad \binom{n}{k} p^k \cdot (1-p)^{n-k}$ } VEROJATNOS DA SE JAVI SEKVENCA OD „n" BITA KADE 1 BIT E „1", n-1 SE „0"

$$P(A_\varepsilon^n) = \sum_{k=r}^{l} \binom{n}{k} p^k (1-p)^{n-k} \ge 1-\varepsilon = 1-0.2 = \underline{0.8}$$

$$p = \frac{1}{2} \Rightarrow \sum_{k=r}^{l} \binom{n}{k} \frac{1}{2^k} \frac{1}{2^{n-k}} = \frac{1}{2^n} \sum_{k=r}^{l} \binom{n}{k}$$

EXAMPLE! $n = 10$

$$\frac{1}{2^{10}} = 0.00098$$

$$-\frac{1}{n} ld \frac{1}{2^n} = +\frac{1}{n} ld\, 2^n = +\frac{1}{n} n = 1$$

$p(x_1^n) = \frac{1}{2^n}$ } ZA BILO KOJA SEKVENCA $x_1 x_2 \ldots x_n$ ~~ZATOA~~ ŠTO $p(x_i = 0) = p(x_i = 1) = \frac{1}{2}$

ZATOA ZA BILO KOJA SEKVENCA x_1^n

$$0.8 \le -\frac{1}{n} ld\, p(x_1^n) = -\frac{1}{n} \cdot (-n) = 1 \le 1.2 \quad \forall x_1, x_2 \ldots x_n$$

ZNAČI SITE 2^n MOŽNI SEKVENCI SE TIPIČNI !!!

$$p(z_1^n) = p(z_1) \cdot p(z_2) \cdots p(z_n)$$

$$p(z_i) = \begin{cases} 1/2 & z_i = 0 \\ 1/2 & z_i = 1 \end{cases} \qquad 0.8 \le -\frac{1}{n} ld\, P(Z_1^n) \le 1.2$$

$$P(Z_1^n = z_1, z_2 \ldots z_n) = \frac{1}{2^n} \qquad 0.8 \le -\frac{1}{n} \cdot (-n) \le 1.2$$

$0.8 \le 1 \le 1.2$ } ALL 2^n z_1^n SEQUENCES ARE MEMBERS OF THE TYPICAL SET.

[redacted]

JOINT AEP

1.) $\Pr\left((X^n, Y^n) \in A_\varepsilon^n\right) \to 1$ as $n \to \infty$

2.) $|A_\varepsilon^n| \leq 2^{n(H(X,Y)+\varepsilon)}$

3.) $(\tilde{X}^n, \tilde{Y}^n) \sim p(x^n)\, p(y^n)$

$$\Pr\left[(\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^n\right] \leq 2^{-n(I(X,Y)-3\varepsilon)}$$

$$\Pr\left[(\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^n\right] \geq (1-\varepsilon)\, 2^{-n(I(X,Y)+3\varepsilon)}$$

$$1 = \sum_{(x^n,y^n)} p(x^n, y^n) \geq \sum_{(x^n,y^n)\in A_\varepsilon^n} p(x^n, y^n) \geq |A_\varepsilon^n|\, 2^{-n(H(X,Y)+\varepsilon)}$$

$$\boxed{|A_\varepsilon^n| \leq 2^{n(H(X,Y)+\varepsilon)}}$$

$$(\tilde{X}^n, \tilde{Y}^n) \quad \Pr\left((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^n\right) = \ldots \leq \sum_{(\tilde{x}^n,\tilde{y}^n)\in A_\varepsilon^n} p(\tilde{x}^n)\, p(\tilde{y}^n)$$

$$\leq |A_\varepsilon^n| \cdot 2^{-n(H(X)-\varepsilon)}\, 2^{-n(H(Y)-\varepsilon)} \leq 2^{n(H(X,Y)+\varepsilon)}\, 2^{-n(H(X)-\varepsilon)} \cdot 2^{-n(H(Y)-\varepsilon)}$$

$$= 2^{n H(Y|X) + n\varepsilon + n\varepsilon - nH(Y) + \varepsilon} = 2^{-n(I-3\varepsilon)}$$

$$= 2^{-n(I(X;Y)-3\varepsilon)} \Rightarrow \underline{\Pr\left[(\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^n\right] \leq 2^{-n(I(X;Y)-3\varepsilon)}}$$

$$(1-\varepsilon) \leq \sum_{(x^n,y^n)\in A_\varepsilon^n} p(x^n, y^n) \leq |A_\varepsilon^n|\, 2^{-n(H(X,Y)-\varepsilon)}$$

$$\boxed{|A_\varepsilon^n| \geq (1-\varepsilon)\, 2^{n(H(X,Y)-\varepsilon)}}$$

$$\Pr\left((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^n\right) = \sum_{(\tilde{x}^n,\tilde{y}^n)\in A_\varepsilon^n} p(\tilde{x}^n)\cdot p(\tilde{y}^n) \geq |A_\varepsilon^n|\, 2^{-n(H(X)+\varepsilon)} \cdot 2^{-n(H(Y)+\varepsilon)}$$

$$\geq (1-\varepsilon)\, 2^{n(H(X,Y)-\varepsilon) - nH(X) - n\varepsilon - nH(Y) - n\varepsilon}$$

$$= (1-\varepsilon)\, 2^{-n(I(X,Y)) - 3n\varepsilon} = \underline{\underline{(1-\varepsilon)\, 2^{-n(I+3\varepsilon)}}}$$

[redacted]

$$(1-\varepsilon)2^{-n(H(X,Y)+3\varepsilon)} \le P_r\left[(\tilde{x}^n, \tilde{y}^n) \in A_\varepsilon^{(n)}\right] \le 2^{-n(I(X,Y)-3\varepsilon)}$$

(i) $$p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i) = (0.45)^{n-k} \cdot (0.05)^k = \left(\frac{1}{2}\right)^n (1-p)^{n-k} \cdot p^k$$

k - NUMBER OF PLACES IN WHICH x^n DIFFERS FROM y^n, $Z = Y \oplus X$

$$p(x^n, y^n) = p(x^n) \cdot p(y^n | x^n) = p(x^n) \cdot p(z^n | x^n) =$$

$$= p(x^n)\, p(z^n) = \left(\frac{1}{2}\right)^n \cdot (1-p)^{n-k}\, p^k$$

SADA DA KAŽE: KADA ŠTO $Z=1$ TAMO $Y \neq X$ T-e

KADA ŠTO $Z=0$ $Y = X$

$Z = (y - x) \bmod 2$

$X = (Y - Z) \bmod 2$

$Y = X \oplus Z$

$I(X;Y) = ?$

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \sum_{x \in \mathcal{X}} p(x) \cdot H(Y|x)$$

$$= H(Y) - \sum_{x \in \mathcal{X}} p(x)\, H(x + Z | x) = H(Y) - H(Z)$$

$$p(x + Z | x) = p(Z)$$

X	Z	Y	(x−y) mod 2
0	0	0	0
0	1	1	1 — (−1) mod 2 = 1
1	0	1	0
1	1	0	1

$$H(X,Y) = H(X) + H(Y|X) = H(X) + H(Z)$$

$$2^{-n(H(X,Y)+\varepsilon)} \le p(x^n, y^n) \le 2^{-n(H(X,Y)-\varepsilon)}$$

$$p(x^n, y^n) = p(x^n) \cdot p(z^n)$$

$$2^{-n(H(X)+\varepsilon)} \le p(x^n) \le 2^{-n(H(X)-\varepsilon)} \quad \cdot\; p(z^n)$$

$$2^{-n(H(Z)+\varepsilon)} \le p(z^n) \le 2^{-n(H(Z)-\varepsilon)}$$

$$2^{-n(H(X)+H(Z)+2\varepsilon)} \le p(x^n) \cdot p(z^n) \le 2^{-n(H(X)+H(Z)-2\varepsilon)}$$

$$2^{-n(H(X,Y)+2\varepsilon)} \le p(x^n y^n) \le 2^{-n(H(X,Y)-2\varepsilon)}$$

[redacted]

(d) $\binom{n}{k} p^k (1-p)^{n-k} = \left| k=0 \right| = 1 \cdot 1 \cdot (1-p)^{25} = 0.071790$

$(1-p) = \sqrt[25]{0.071790} = 0.9$ $\quad (p = 1-0.9 = 0.1)$ $\quad p = p(Y=0|X=1) + p(Y=1|X=0)$

$\Pr(Z=1) = p = 0.1 = p(Y=1|X=0) + p(Y=0|X=1)$ (MMV)

NA PRETHODNATA STRANA GO DOKAZAV ANALITIČKI!!!

$-\frac{1}{n} \operatorname{ld} p(x^n) = -\frac{1}{25} \operatorname{ld}\left(\frac{1}{2^{25}}\right) = 1$

$\Pr(Z) = \{0.1, 0.9\}$ $\quad p(Z=1) = 0.1$, $\quad p(Z=0) = 0.9$

$H(Z) = -[0.1 \operatorname{ld}(0.1) + 0.9 \operatorname{ld}(0.9)] = H(Y|X) = 0.469$

. Neizvesnosta za greškata e ednakva na uslovnata neizvesnost $H(Y|X)$!!! MMV

$2^{-n(H(Z)+\varepsilon)} \leq p(Z) \leq 2^{-n(H(Z)-\varepsilon)}$

$\left| -\frac{1}{n} \operatorname{ld} p(Z) - H(Z) \right| \leq \varepsilon = 0.2$ } OVA E MOŽNO SAMO ZA: $k=2,3,4,5$

$|A_\varepsilon^{(n)}| = \sum_{k=1}^{4} \binom{n}{k} = \binom{25}{1} + \binom{25}{2} + \binom{25}{3} + \binom{25}{4} = 15275$

(e) $[x^n(1), Y^n]$ ARE JOINTLY TYPICAL IF Z^n IS TYPICAL (SINCE X^n IS ALWAYS TYPICAL)

$\Pr(Z^n \in A_\varepsilon^n) = \sum_{k=1}^{4} \binom{n}{k} p^k (1-p)^{n-k} = 0.89022$

CONTINUE FROM N16d

PROBLEM 7.26 CONSIDER THE CHANNEL WITH $X, Y \in \{0,1,2,3\}$ AND TRANSITION PROBABILITIES $p(Y|X)$ GIVEN AS:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \end{bmatrix}$$

(a) FIND THE CAPACITY OF THIS CHANNEL

(b) DEFINE THE RANDOM VARIABLE $Z = g(Y)$ WHERE

$g(Y) = \begin{cases} A & \text{IF } Y \in \{0,1\} \\ B & \text{IF } Y \in \{2,3\} \end{cases}$

[redacted]

FOR THE TWO PMFs FOR X COMPUTE $I(X;Z)$:

(i) $p(x) = \begin{cases} 1/2 & \text{IF } x \in \{1,3\} \\ 0 & \text{IF } x \in \{0,2\} \end{cases}$

(ii) $p(x) = \begin{cases} 0 & \text{IF } x \in \{1,3\} \\ 1/2 & \text{IF } x \in \{0,2\} \end{cases}$

(c) FIND THE CAPACITY OF THE CHANNEL BETWEEN X & Z SPECIFICALLY WHERE $x \in \{0,1,2,3\}$, $z \in \{A,B\}$ AND THE TRANSITIONAL PROBABILITIES $P(z|x)$ ARE GIVEN BY:

$$p(Z=z \mid X=x) = \sum_{g(y_0)=z} P(Y=y_0 \mid X=x)$$

(d) FOR THE X DISTRIBUTION OF PART (i) OF (b), DOES $X \to Z \to Y$ FORM MARKOV CHAIN?

(a) $C = ld\, m - 1 = ld\, 4 - 1 = ld \frac{4}{2} = ld\, 2 = 1$

$I(X;Y) = H(Y) - H(Y|X)$ $\quad H(Y)\big|_{max} = ld\, m$

$H(Y|X) = \sum_x p(x) H(Y|X=x) = \sum_x P(X=x_i) \cdot 1 = 1$

$-H(Y|X=1) = -H(Y|X=0) = P(Y|X=0) \cdot ld\, P(Y|X=0) + P(Y|X=1) \cdot ld\, P(Y|X=1)$

$= (1/2\; ld\, 2) \cdot 2 = 1$

$I(X;Y) = H(Y) - 1$ $\quad C = \max_{p(x)} I(X;Y) = ld(4) - 1 = 2 - 1 = 1$

(b) $Z = g(Y)$ $\quad I(X;Y) = ?$

$g(Y) = \begin{cases} A & Y = \{0,1\} \\ B & Y \in \{2,3\} \end{cases}$

(i) $p(x) = \begin{cases} 1/2 & x \in \{1,3\} \\ 0 & x \in \{0,2\} \end{cases}$

$I(X;Z) = H(X) - H(X|Z) = H(Z) - H(Z|X) = (\frac{1}{2} ld\, 2) \cdot 2 - H(Z|X)$

Z \ Y	0	1	2	3
A	y_1	y_2	0	0
B			y_3	y_4

$P(Z=A) = P(Z=A|Y=0) + P(Z=A|Y=1)$

$P(Z=B) = P(Z=B|Y=2) + P(Z=B|Y=3)$

X \ Y	0 (A)	1 (A)	2 (B)	3 (B)	P(X)
0	0	0	0	0	0
1	1/4	1/4	0	0	1/2
2	0	0	0	0	0
3	0	0	1/4	1/4	1/2
P(Y)	1/4	1/4	1/4	1/4	

X \ Y	0	1	2	3
0	1/2	0	0	1/2
1	1/2	1/2	0	0
2	0	1/2	1/2	0
3	0	0	1/2	1/2

$P(Y|X)$ I.E. TRANSITION MATRIX

$P(X) \cdot P(Y|X) = P(X,Y)$

$P(Z=A) = P(Y=0) + P(Y=1) = 1/2$

$P(Z=B) = P(Y=2) + P(Y=3) = 1/2$

[redacted]

P(X, Z)

X \ Z	A	B	P(X)
0	0	0	0
1	1/2	0	1/2
2	0	0	0
3	0	1/2	1/2
P(Z)	1/2	1/2	

$P(X=1, Z=A) = 1/2$ $\quad P(X=3, Z=B) = 1/2$

$P(X=0, Z=A) = \emptyset$ $\quad P(X=2, Z=B) = 0$

X \ Z	A	B	P(Z|X)
0	0	0	
1	1	0	
2	0	0	
3	0	1	

$H(Z|X) = -\sum_{X,Z} P(X,Z) \cdot ld\, P(Z|X) = -\left[\frac{1}{2} \cdot ld\, 1 + \frac{1}{2}\, ld\, 1\right] = 0$

$I(X;Z) = H(Z) - H(Z|X) = 1 - 0 = 1$

$H(X,Z) = H(X) + H(Z|X) = H(Z) + H(X|Z)$

$Y = g(X)$

$H(X) \geq H(Y)$

(11)

P(Z)	A		B		
X \ Y	0	1	2	3	P(X)
0	1/4	1/4	0	0	1/2
1	0	0	0	0	0
2	0	0	1/4	1/4	1/2
3	0	0	0	0	0
P(Y)	1/4	1/4	1/4	1/4	

X \ Z	A	B	P(X)
0	1/2	0	1/2
1	0	0	0
2	0	1/2	1/2
3	0	0	0
P(Z)	1/2	1/2	

P(X, Z)

X \ Z	A	B
0	1	0
1	0	0
2	0	1
3	0	0

P(Z|X)

$H(Z|X) = -\sum_{X,Z} P(X,Z) \cdot ld\, P(Z|X) = -\left[\frac{1}{2}\, ld\, 1 + \frac{1}{2}\, ld\, 1\right] = 0$

$I(X;Z) = H(Z) - H(Z|X) = ld\, 2 - 0 = 1$

PROBLEM 8.6 VARIATION INEQUALITY. VERIFY FOR POSITIVE RANDOM VARIABLES X THAT

$ld\, E_P(X) = \sup_Q \left[E_Q\, ld(X) - D(Q\|P)\right]$,

WHERE $E_P(X) = \sum_x x\, P(x)$ AND $D(Q\|P) = \sum Q\, ld \frac{Q(x)}{P(x)}$

AND THE SUPREMUM IS OVER ALL $Q(x) \geq 0$, $\sum Q(x) = 1$.

IT IS ENOUGH TO EXTREMIZE $J(Q) = E_Q[\ln X] - D(Q\|P) + \lambda\left(\sum Q(x) - 1\right)$

[redacted]

LAGRANGE MULTIPLIERS

$J(Q) = E_Q[\ln x] - D(Q\|P) + \lambda\left(\sum Q(x) - 1\right)$

$f_x(x,y,z) = \lambda\, g_x(x,y,z) \qquad f_y(x,y,z) = \lambda\, g_y(x,y,z)$

$f_z(x,y,z) = \lambda\, g_z(x,y,z)$

$\frac{\partial J(Q(x_1))}{\partial Q(x_1)}, \frac{\partial J(Q(x_2))}{\partial Q(x_2)} \ldots \frac{\partial J(Q(x_n))}{\partial Q(x_n)}$

$$\frac{\partial J(Q)}{\partial Q(x_i)} = \frac{d}{dQ(x_i)}\left[E_Q[\ln x]\right] - \frac{d}{dQ(x_i)} D(Q\|P) + \lambda \frac{d}{dQ(x_i)}\left[\sum Q(x) - 1\right]$$

$E_Q[\ln x] = \sum_x \ln(x) Q(x) \qquad \frac{d}{dQ_i}\left[E_Q(\ln x)\right] = \sum_x \ln(x) \frac{dQ(x)}{dQ_i} = \ln(x_i)$

$D(Q\|P) = \sum_x Q \operatorname{ld} \frac{Q}{P} \qquad \frac{\partial D(Q\|P)}{\partial Q(x_i)} = \frac{\partial}{\partial Q(x_i)}\left[\sum_{x_i} Q(x_i) \operatorname{ld} \frac{Q(x_i)}{P(x_i)}\right]$

$= \operatorname{ld}\frac{Q(x_i)}{P(x_i)} + Q(x_i) \frac{1}{\left(\frac{Q(x_i)}{P(x_i)}\right)\ln 2} \frac{1}{P(x_i)} = \operatorname{ld}\frac{Q(x_i)}{P(x_i)} + \frac{\ln e}{\ln 2} =$

$= \operatorname{ld}\left[\frac{Q(x_i)}{P(x_i)} \cdot e\right]$

$$\frac{\partial J(Q(x_i))}{\partial Q(x_i)} = \ln(x_i) - \operatorname{ld}\frac{Q(x_i)}{P(x_i)} - \operatorname{ld} e + \lambda = 0$$

DEFINITION: THE MUTUAL INFORMATION BETWEEN TWO RANDOM VARIABLES X AND Y IS GIVEN AS:

$$I(X;Y) = \sup_{P,Q} I\left([X]_P, [Y]_Q\right)$$

$$I(X;Y) = \sum_{x,y} P(x,y) \operatorname{ld} \frac{P(x,y)}{P(x)\,P(y)} = D\left(P(x,y)\,\|\,P(x)P(y)\right)$$

$\operatorname{ld} E_P(x) = \operatorname{ld}\left[\sum_x P(x)\right]$

JENSEN INEQUALITY: $E[f(x)] \geq f(E[x])$ $\quad f$ – CONVEX

ld – CONCAVE FUNCTION $\quad \operatorname{ld} E_P(x) \geq E_P[\operatorname{ld}(x)] = H_P(x)$

$\operatorname{ld}\left[e \frac{Q(x_i)}{P(x_i)}\right] = +\left(\ln(x_i) + \lambda\right) = +\left[\ln x_i + \ln e^{\lambda}\right]$

$\operatorname{ld}\left(e \frac{Q(x_i)}{P(x_i)}\right) = +\ln\left(x_i e^{\lambda}\right) \quad \frac{1}{\ln 2} \ln\left(\frac{P(x_i)}{e\,Q(x_i)}\right)^{-1} = \ln\left(x_i e^{\lambda}\right)$

[redacted]

$$\ln[E_P(X)] = \sup_Q \left[E_Q \ln(X) - D(Q\|P)\right]$$

$$J(Q) = E_Q[\ln X] - D(Q\|P) + \lambda\left(\sum Q(x) - 1\right)$$

$$\operatorname{ld}\left[e \frac{Q(x_i)}{P(x_i)}\right] = \ln x_i e^{\lambda} = \ln 2 \cdot \operatorname{ld}(x_i e^{\lambda})$$

$$\operatorname{ld}\left[e \frac{Q(x_i)}{P(x_i)}\right] = \operatorname{ld} x_i^{\ln 2} e^{\lambda \ln 2}; \quad \boxed{e \frac{Q(x_i)}{P(x_i)} = x_i^{\ln 2} e^{\lambda \ln 2}}$$

$$\boxed{f(\lambda_i) = \frac{e\, Q(x_i)}{\lambda_i^{\ln 2} e^{\lambda \ln 2}}} \qquad \sum P(x_i) = \sum \frac{Q(x_i)}{x_i^{\ln 2} e^{\lambda \ln 2 - 1}}$$

$$\sum_x P(x_i) = \frac{1}{e^{\lambda \ln 2 - 1}} \sum \frac{Q(x_i)}{x_i^{\ln 2}} = 1$$

$$\sum_{x_i} \frac{Q(x_i)}{x_i^{\ln 2}} = e^{\lambda \ln 2 - 1}$$

If: $\boxed{D(Q\|P) = \sum Q \ln \frac{Q}{P}}$

$$\frac{dJ(Q)}{dQ_i} = \ln x_i - \ln \frac{Q(x_i)}{P(x_i)} - 1 + \lambda = 0$$

$$\ln(x_i e^{\lambda}) = \ln \frac{Q(x_i)}{P(x_i)} \cdot e$$

$$\ln x_i + \lambda = \ln \frac{Q}{P} \cdot e$$

$$x_i e^{\lambda} = \frac{Q(x_i)}{P(x_i)} \cdot e \qquad \frac{Q(x)}{P(x)} = x_i e^{\lambda - 1} \qquad \boxed{P(x_i) = \frac{Q(x)}{x_i e^{\lambda - 1}}}$$

$$\sum P(x_i) = \sum_{x_i} \frac{Q(x_i)}{x_i e^{\lambda - 1}} = e^{1-\lambda} \sum_{x_i} \frac{Q(x_i)}{x_i}$$

$$\boxed{\sum_{x_i} \frac{Q(x_i)}{x_i} = e^{1-\lambda}} \qquad Q(x_i) = P(x_i)\, x_i\, e^{\lambda - 1} \qquad \sum_{x_i} Q(x_i) = 1$$

$$\sum_{x_i} x_i e^{\lambda - 1} P(x_i) = e^{\lambda - 1} \sum_{x_i} x_i P(x_i) = e^{\lambda - 1} E_P(X) = 1$$

$$E_P(X) = \frac{1}{e^{\lambda - 1}} = e^{1-\lambda}$$

MMV

[redacted]

$$D(Q\|P) = \sum_{x_i} Q_i \ln \frac{Q}{P} = \sum_{x_i} Q \ln(x_i e^{\lambda-1}) = \sum_{x_i} Q(x_i) \ln x_i + \sum_{x_i} Q(x_i)(\lambda-1) = E_Q[\ln X] + (\lambda-1)$$

$$\ln[E_P(X)] = \sup_Q [E_Q \ln(X) - D(Q\|P)]$$

$$\ln[E_P(X)] = (1-\lambda) \qquad \sup[E_Q[\ln X] - D(Q\|P)] = E_Q[\ln X] - E_Q[\ln X] - (\lambda \ldots)$$

$$\sup[E_Q[\ln X] - D(Q\|P)] = 1-\lambda = \ln[E_P(X)]$$

SIMPLER SOLUTION

$$\ln x_i - \ln \frac{Q}{P} - 1 + \lambda = 0$$

$$\ln \frac{Q}{P} + 1 = \lambda + \ln x_i \quad : \quad \ln \frac{Q}{P} = (\lambda-1) + \ln x_i$$

$$\ln \frac{Q}{P} = \ln(e^{\lambda-1} x_i) \qquad \frac{Q}{P} = K x_i \qquad Q(x_i) = K x_i P(x_i)$$

$$\sum_{x_i} Q(x_i) = K \sum_{x_i} x_i P(x_i) = K \cdot E_P(X) = 1 \qquad K = \frac{1}{E_P(X)}$$

$$Q(x_i) = \frac{x_i P(x_i)}{E_P(X)} \Rightarrow \sum_x Q(x) \ln(x) - \sum Q(x) \ln \frac{Q(x)}{P(x)} + \lambda\left(\sum Q(x) - 1\right)$$

$$= \sum_x Q(x) \ln(x) - \sum_x Q(x) \ln \frac{x_i}{E_P(X)} + \lambda\left(\sum Q(x_i) - 1\right) =$$

$$= \sum_x Q(x) \ln(x) - \sum_x Q(x) \ln x + \sum_x Q(x) \ln E_P(X) + \lambda\left[\sum Q(x) - 1\right] =$$

$$= E_Q[\ln(x)] - E_Q[\ln x] + \ln E_P(X) + \lambda\left[\sum Q(x) - 1\right] = 0$$

$$\ln[E_P(X)] = \sup[E_Q[\ln(x)] - D(Q\|P)]$$

$$\ln[E_P(X)] = \ln[E_P(X)]$$

ZA MAKSIMIZACIJA VREDNOST I MAKU:

$$\ln[E_P(X)] \geq E_Q[\ln(X)] - D(Q\|P)$$

PROBLEM 8.7 DIFFERENTIAL ENTROPY BOUND ON DISCRETE ENTROPY. LET X BE A DISCRETE RANDOM VARIABLE ON THE SET $\mathcal{X} = \{a_1, a_2, \ldots\}$ WITH $\Pr(X = a_i) = p_i$. SHOW THAT:

$$H(p_1, p_2, \ldots) \leq \frac{1}{2} \log(2\pi e)\left(\sum_{i=1}^{\infty} p_i i^2 - \left(\sum_{i=1}^{\infty} i p_i\right)^2 + \frac{1}{12}\right)$$

[redacted]

MOREOVER FOR EVERY PERMUTATION σ,

$$H(p_1, p_2, \dots) \leq \frac{1}{2} \operatorname{ld}(2\pi e)\left(\sum_{i=1} p_{\sigma(i)} i^2 - \left(\sum_{i=1} i\, p_{\sigma(i)}\right)^2 + \frac{1}{12}\right)$$

HINT: CONSTRUCT A RANDOM VARIABLE X' SUCH THAT $\Pr(X'=i)=p_i$. LET U BE A UNIFORM $(0,1]$ RANDOM VARIABLE AND LET $Y = X' + U$, WHERE X' AND U ARE INDEPENDENT. USE THE MAXIMUM ENTROPY BOUND ON Y TO OBTAIN BOUNDS IN THE PROBLEM.

$$H(p_1, p_2, \dots) \leq \frac{1}{2} \operatorname{ld}(2\pi e)\left(\sum_{i=1}^{\infty} p_i i^2 - \left(\sum_{i=1}^{\infty} i\, p_i\right)^2 + \frac{1}{12}\right)$$

$Y = X' + U$ $\quad h(Y) \leq \frac{1}{2} \operatorname{ld}(2\pi e)\sigma^2$

$$p(U) = \begin{cases} 1 & u \in 0..1 \\ 0 & \text{OTHERWISE} \end{cases}$$

$$I(X';Y) = h(X') - h(X'|Y) = h(Y) - H(Y|X')$$

$$h(Y|X') = \int_x p(x)\, h(Y|X'=x)\, dx = \int_x p(x)\, h(x+U)\, dx$$

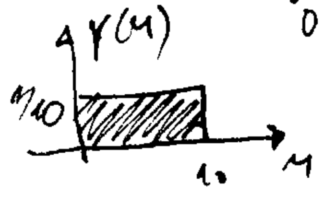
$$= \int_x p(x)\, h(x+U)\, dx = h(U)$$

$$I(X';Y) = H(Y) - H(U)$$

$H(X^4) = h(X) - \operatorname{ld} 4$ $\quad 2^{?}$

$= h(Y) - 4$

$$h(U) = \int_0^1 1 \operatorname{ld} 1\, du = 0 = \frac{\operatorname{ld} 1}{10}$$



$$h(U) = \int_0 \frac{1}{10} \operatorname{ld} 10\, du = \frac{1}{10} \cdot 10 \cdot \operatorname{ld} 10 = \operatorname{ld} 10$$

$$I(X';Y) = H(Y) = h(X') - h(X'|Y)$$

$$H(X'|Y) = \sum p(Y=y)\, H(y - U | y) = H(U)$$

$\Pr(X'=i) = p_i$ $\quad \Pr(X'=1) = p_1$ $\quad \Pr(X'=2) = p_2$, $\Pr(X'=3) = p_3$

$$E[X'] = \sum_{i=1}^{\infty} i\, p_i \qquad E[X'^2] = \sum_{i=1} i^2 p_i$$

[redacted]

$$\sigma^2 = \sum_{i=1}^{\infty} (X-\bar{x})^2 p_i = \sum_{i=1}^{\infty} \left(i - \sum_{j=1}^{\infty} j p_j\right)^2 p_i$$

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

$1 + 2 + 3 + \dots + n$

$n + n-1 + n-2 + \dots + 1$ $\quad +$

$(n+1) + (n+1) + (n+1) + \dots (n+1) = n \cdot (n+1)$

$$\left(i - \sum_{i=1}^{\infty} i p_i\right)^2 = i^2 - 2i \sum_{j=1}^{\infty} j p_j + \left(\sum_{j=1}^{\infty} j p_j\right)^2$$

$$\sum_{i=0}^{\infty} i \frac{1}{2^i} = \frac{\frac{1}{2}}{\left(1-\frac{1}{2}\right)^2} = 2$$

$$\sum_{i=0}^{\infty} \frac{1}{2^i} = \frac{1}{1-\frac{1}{2}}$$

$S = 1 + \frac{1}{2} + \frac{1}{2^2} + \dots \frac{1}{2^n}$

$\frac{S}{2} = \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n+1}}$

$1 - \frac{1}{2^{n+1}} = S - \frac{S}{2}$

$S\left(1 - \frac{1}{2}\right) = 1 - \frac{1}{2^n}$

$S = \frac{1 - 2^n}{1-2} = /n \to \infty/ = \frac{1}{1-2}$

$$S = \sum_{n=0}^{\infty} x^n \qquad dS/dx = \sum_{n=0}^{\infty} n x^{n-1}$$

$$S = \frac{1}{1-x} \qquad \left(\frac{1}{1-x}\right)' = \sum_{n=0}^{\infty} n x^{n-1} \qquad \frac{1}{(1-x)^2} = \sum_{n=0}^{\infty} n x^{n-1}$$

$$\left(\frac{1}{1-x}\right)' = \left[(1-x)^{-1}\right]' = -1 \cdot (1-x)^{-1-1} \cdot (-1) = \frac{1}{(1-x)^2}$$

$$\frac{1}{(1-x)} = \frac{1}{x} \sum_{n=0}^{\infty} n x^n \qquad \sum_{n=0}^{\infty} n x^n = \frac{x}{(1-x)^2}$$

$p(z) = p(x, v) \qquad K = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 1 \end{bmatrix}$

$h(z) \leq \frac{1}{2} \, ld \, (2\pi e) |K| = \frac{1}{2} \, ld \, (2\pi e) \, \sigma^2$

$H(W) = nR = H(W, \hat{W}) = H(W) + H(\hat{W} | W) = H(\hat{W}) + H(W | \hat{W}) \qquad W \in [1, 2, \dots 2^{nR}]$, $H(\hat{W}|W) = 0$

$H(\hat{W}) = 1 + P_e \, ld \, 2^{nR} \leq 1 + P_e \cdot n \cdot R$

$W \to X^n \to Y^n \to \hat{W}$

$I(W, \hat{W}) \leq I(W, Y^n)$

$I(W, \hat{W}) \leq I(X^n, Y^n)$

$$H(W) = n \cdot R = \underbrace{I(W;\hat{W})}_{H(W) - H(W|\hat{W})} + H(W|\hat{W}) \leq 1 + P_e \, n \cdot R + I(W;\hat{W}) \leq 1 + P_e \, n R + I(X^n, Y^n)$$

$$n \cdot R \leq 1 + P_e \, n R + n \cdot C \qquad R \leq \frac{1}{n} + P_e \cdot R + C$$

$$\boxed{R \leq C}$$

CONTINUE FROM N. 16K

8.8 CHANNEL WITH UNIFORMLY DISTRIBUTED NOISE

CONSIDER A ADDITIVE CHANNEL WHOSE INPUT ALPHABET $\mathcal{X} = \{0, \pm 1, \pm 2\}$ AND WHOSE OUTPUT $Y = X + Z$, WHERE Z IS DISTRIBUTED UNIFORMLY OVER THE INTERVAL $[-1, 1]$. THUS THE INPUT OF THE CHANNEL IS A DISCRETE RANDOM VARIABLE, WHEREAS THE OUTPUT IS CONTINUOUS. CALCULATE $C = \max_{p(x)} I(X;Y)$ OF THIS CHANNEL.

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

FROM 8.7

$$H(X) = -\sum_{i=1}^{\infty} p_i \, ld \, p_i = -\sum_{i=1}^{\infty} \left(\int_i^{i+1} p(y)\,dy \right) \left(ld \int_i^{i+1} p(y)\,dy \right) =$$

$$= \sum_{i=1}^{\infty} \int_i^{i+1} p(y) \, ld \, y \, dy = h(Y)$$

$$f(x_i) \cdot \Delta = \int_{i\Delta}^{(i+1)\Delta} f(x)\,dx$$

$$\mathcal{X} = \{-2, -1, 0, 1, 2\}$$

$$p(x) = \{p_1 \; p_2 \; p_3 \; p_4 \; p_5\}$$

$$H(X) = -\sum_{i=1}^{5} p_i \, ld \, p_i = -\sum_{i=1}^{5} \left(\int_{i-1}^{i+1} p(y)\,dy \right) \left(ld \int_{i-1}^{i+1} p(y)\,dy \right)$$

$$= -\sum_{i=1}^{5} \int_{i-1}^{i+1} p(y) \, ld \, p(y)\,dy = -\left[\int_0^2 p(y) \, ld \, p(y)\,dy + \int_1^3 p(y) \, ld \, p(y)\,dy + \dots + \int_4^6 p(y) \, ld \, p(y)\,dy \right] = -\left[\int_0^1 p(y) \, ld \, p(y)\,dy + \int_1^2 p(y) \, ld \, p(y)\,dy + \dots + \int_4^5 p(y) \, ld \, p(y)\,dy \right] - \left[\int_1^2 p(y) \, ld \, p(y)\,dy + \int_2^3 p(y) \, ld \, p(y)\,dy + \dots + \int_5^6 p(y) \, ld \, p(y) \right]$$

[redacted]

$$= -\int_1^5 p(y)\, ld\, p(y)\, dy - \int_2^6 p(y)\, ld\, p(y)\, dy$$

$x \in [-2, -1, 0, 1, 2]$ $\qquad p(x) = \{p_{-2}, p_{-1}, p_0, p_1, p_2\}$

$y \in [(-3, -1); (-2, 0); (-1, 1); (0, 2); (1, 3)]$

x \ y	-3	-2	-1	0	1	2	3	p(x)
-2								p_{-2}
-1								p_{-1}
0								p_0
1								p_1
2								p_2
p(y)								

$$H(x) = -\sum_{i=-2}^{2} p_i\, ld\, p_i =$$

$$= -\sum_{i=-2}^{2} \int_{i-1}^{i+1} p(y)\, ld\, p(y)\, dy =$$

$$= \int^{3} p(y)\, ld\, p(y)\, dy = h(y)$$

$$\int_{-3}^{-1};\ \int_{-2}^{0};\ \int_{-1}^{1};\ \int_{0}^{2};\ \int_{1}^{3} \qquad \int_{-3}^{-2};\ \int_{-2}^{-1};\ \int_{-1}^{0};\ \int_{0}^{1};\ \int_{1}^{2};\ \int_{2}^{3} \oplus$$

$$\oplus \int_{-2}^{-1};\ \int_{-1}^{0};\ \int_{0}^{1};\ \int_{1}^{2};\ \int_{2}^{3} \qquad H(x) = -\left[\int_{-3}^{2} p(y)\, ld\, p(y)\, dy + \int_{-2}^{3} p(y)\, ld\, p(y)\, dy\right]$$

$$H(x) = -\sum_{i=-2}^{2} p_i\, ld\, p_i \qquad H(x|y) \qquad x = y - z$$

$$H(x|y) = \int p(y)\, H(y - z|y)\, dy = h(z)$$

$$I(x;y) = H(x) - h(z) \qquad h(z) = \int^{1} \frac{1}{2}\, dz = \frac{1}{2} \cdot 2 = 1$$

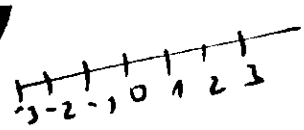
$$I(x;y) = H(x) - 1 \qquad p(x) = \left\{\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right\}$$

$$I(x;y) = ld\,5 - 1 = ld\,\frac{5}{2} = ld\,(2.5)$$

$$f_Y(y) = \begin{cases} \frac{1}{2} p_{-2} & y \in (-3, -2) \\ \frac{1}{2}(p_{-2} + p_{-1}) & y \in (-2, -1) \\ \frac{1}{2}(p_{-1} + p_0) & y \in (-1, 0) \\ \frac{1}{2}(p_0 + p_1) & y \in (0, 1) \\ \frac{1}{2}(p_1 + p_2) & y \in (1, 2) \\ 1/2\, p_2 & y \in (2, 3) \end{cases}$$

SIRMOV VO SOLUTIONS TA PUSLE PROBLEM

[redacted]



$$h(Y) = -\Big[\int_{-3}^{-2} \tfrac{1}{2} P_{-2} \,ld\, \frac{P_{-2}}{2}\,dy + \int_{-2}^{-1} \frac{P_{-2}+P_{-1}}{2} \,ld\, \frac{P_{-2}+P_{-1}}{2}\,dy + \int_{-1}^{0} \frac{P_{-1}+P_0}{2} \,ld\, \frac{P_{-1}+P_0}{2}\,dy$$

$$+ \int_{0}^{1} \frac{P_0+P_1}{2} \,ld\, \frac{P_0+P_1}{2}\,dy + \int_{1}^{2} \frac{P_1+P_2}{2} \,ld\, \frac{P_1+P_2}{2}\,dy + \int_{2}^{3} \frac{P_2}{2} \,ld\, \frac{P_2}{2}\,dy\Big] =$$

$$= -\Big[\frac{P_{-2}}{2} \,ld\, \frac{P_{-2}}{2} + \frac{P_{-2}+P_{-1}}{2} \,ld\, \frac{P_{-2}+P_{-1}}{2} + \frac{P_{-1}+P_0}{2} \,ld\, \frac{P_{-1}+P_0}{2} + \frac{P_0+P_1}{2} \,ld\, \frac{P_0+P_1}{2}$$

$$+ \frac{P_1+P_2}{2} \,ld\, \frac{P_1+P_2}{2} + \frac{P_2}{2} \,ld\, \frac{P_2}{2}\Big]$$

PRILOGY IZGLEDA KAKO GRADOVI ZA DISCRETNA PROMJENLJIVA !!!

$$= -\Big[\frac{P_{-2}}{2} \,ld\, \frac{P_{-2}}{2}\cdot\frac{P_{-2}+P_{-1}}{2} + \frac{P_{-1}}{2} \,ld\Big(\frac{P_{-2}+P_{-1}}{2}\cdot\frac{P_{-1}+P_0}{2}\Big) + \frac{P_0}{2} \,ld\, \frac{P_{-1}+P_0}{2}\cdot\frac{P_0+P_1}{2} +$$

$$+ \frac{P_1}{2} \,ld\Big(\frac{P_0+P_1}{2}\cdot\frac{P_1+P_2}{2}\Big) + \frac{P_2}{2} \,ld\, \frac{P_1+P_2}{2}\,\frac{P_2}{2}\Big]$$

$+ I_2 + I_{-1} + I_0 + I_1 + I_2$

$$-H(X|Y) = \int f(Y)\cdot H(x|Y)\,dy = \int_{-3}^{-2} \frac{P_{-2}}{2}\cdot H(x|Y\in(-3:-2))\,dy$$

$x = -2 \Rightarrow H = 0$

$$\Rightarrow I_2 = \int_{-2}^{-1} \frac{P_{-2}+P_{-1}}{2}\cdot H(x|Y\in(-2:-1))\,dy = \frac{P_{-2}+P_{-1}}{2}$$

$x = -2 \quad P = \frac{1}{2}$

$x = -1 \quad P = \frac{1}{2}$

$$H(x|Y\in(-2:-1)) = P(x=-2|Y\in-2:-1)\,ld\,P(x=-2|Y\in-2:-1) +$$

$$+ P(x=-1|Y\in-2:-1)\cdot ld\,P(x=-1|Y\in-2:-1) = \Big(\frac{1}{2}\,ld\,2\Big)\cdot 2 = 1$$

$$I_{-1} = \frac{P_{-1}+P_0}{2} \qquad I_0 = \frac{P_0+P_1}{2} \qquad I_1 = \frac{P_1+P_2}{2} \qquad I_2 = 0$$

$$-H(x|Y) = \frac{P_{-2}+P_{-1}+P_{-1}+P_0+P_0+P_1+P_1+P_2}{2} = \frac{P_{-2}+2P_{-1}+2P_0+2P_1+P_2}{2}$$

$$-H(X|Y) = \sum_{i=-1}^{1} \frac{P_i+P_{i+1}}{2} = \frac{P_{-1}+P_0}{2} + \frac{P_0+P_1}{2} + \frac{P_1+P_2}{2}$$

$$-h(Y) = \sum_{i=-2}^{3} \frac{P_{i-1}+P_i}{2} \,ld\, \frac{P_{i-1}+P_i}{2} = \frac{P_{-2}}{2} \,ld\, \frac{P_{-2}}{2} + \frac{P_{-2}+P_{-1}}{2} \,ld\, \frac{P_{-2}+P_{-1}}{2} +$$

$$+ \dots + \frac{P_1+P_2}{2} \,ld\, \frac{P_1+P_2}{2} + \frac{P_2}{2} \,ld\, \frac{P_2}{2}$$

$$I(X,Y) = -\sum_{i=-2}^{3} \frac{P_{i-1}+P_i}{2} \,ld\, \frac{P_{i-1}+P_i}{2} + \sum_{i=-1}^{1} \frac{P_i+P_{i+1}}{2}$$

[redacted]

$$f(P) = -\sum_{i=-2}^{3} \frac{P_{i-1}+P_i}{2} \operatorname{ld} \frac{P_{i-1}+P_i}{2} + \sum_{i=-1}^{1} \frac{P_i + P_{i+1}}{2} + \lambda \left(\sum_{i=-2}^{2} P_i - 1 \right)$$

$$\frac{d f(P_i)}{d P_i} = -\frac{1}{2} \operatorname{ld} \frac{P_{i-1}+P_i}{2} - \underbrace{\frac{P_{i-1}+P_i}{2} \cdot \frac{2}{P_{i-1}+P_i} \frac{1}{2}}_{1} + \frac{1}{2} + \lambda = 0$$

$$\lambda = \frac{1}{2} \operatorname{ld} \frac{P_{i-1}+P_i}{2} \qquad \lambda = \frac{1}{2}\left(1 + \operatorname{ld} \frac{P_{i-1}+P_i}{2}\right)$$

$$\lambda = \frac{1}{2} \operatorname{ld} \frac{(P_{i-1}+P_i)}{2} \qquad 2 \cdot 2^{2\lambda} = (P_{i-1}+P_i) \qquad P_i = 2^{2\lambda+1} - P_{i-1}$$

$$\sum_{i=-2}^{2} P_i = 1 \qquad \sum_{i=-2}^{2} \left(2^{2\lambda+1} - P_{i-1}\right) = 1 \qquad 5 \cdot 2^{2\lambda} - \sum_{i=-2}^{2} P_{i-1} = 1$$

$$5 \cdot 2^{2\lambda+1} - \sum_{i=-1}^{2} P_{i-1} = 1 \qquad \boxed{5 \cdot 2^{2\lambda+1} - \sum_{j=-2}^{1} P_j = 1}$$

$j = i-1$; $i=-1$ $j=-2$; $i=2$ $j=1$

$$\sum_{j=-2}^{1} P_j = 1 - P_2 \qquad 5 \cdot 2^{2\lambda+1} - 1 + P_2 = 1 \qquad \boxed{P_2 = 2 - 5 \cdot 2^{2\lambda+1}}$$

$$5 \cdot 2^{2\lambda+1} = 2 - P_2 \qquad 2^{2\lambda+1} = \frac{2-P_2}{5} \qquad 2\lambda+1 = \operatorname{ld} \frac{2-P_2}{5}$$

$$\lambda = -\frac{1}{2} + \frac{1}{2} \operatorname{ld} \frac{2-P_2}{5} = \frac{1}{2} \operatorname{ld} \frac{2-P_2}{10} \qquad \frac{1}{2} \operatorname{ld} \frac{2-P_2}{10} = \frac{1}{2} \operatorname{ld} (P_{i-1}+P_i)$$

$$\boxed{P_i + P_{i-1} = (2-P_2)/10}$$

$$\left\{ \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16} \right\}$$

$$\frac{1}{2} + \frac{1}{4} = \frac{3}{4} \qquad \left(2 - \frac{1}{16}\right)\frac{1}{10} = \frac{32-1}{16} = \frac{31}{16} \cdot \frac{1}{10} = \frac{31}{160}$$

$$\left\{ \frac{1}{16}, \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2} \right\} \qquad \frac{1}{8} \qquad \left(2 - \frac{1}{2}\right)\frac{1}{10} = \frac{1}{10} \frac{4-1}{2} = \frac{3}{20}$$

$$C = -\sum_{i=-2}^{3} \frac{2-P_2}{20} \operatorname{ld} \frac{2-P_2}{20} + \sum_{i=-1}^{1} \frac{2-P_2}{20}$$

$$C = 6 \cdot \frac{2-P_2}{20} \operatorname{ld} \frac{2-P_2}{20} + 3 \cdot \frac{2-P_2}{20} = \frac{6 P_2}{10} \operatorname{ld} \frac{2-P_2}{20} + 3 \frac{2-P_2}{20}$$

$C \nearrow$ AS $P_2 \searrow$ FOR $\boxed{P_2 = 0}$ $\boxed{C = 0.3}$

- BERKLEY SOLUTION $P_2(Y) = \left\{ \frac{P_{-2}}{2}, \frac{P_{-2}+P_{-1}}{2}, \frac{P_{-1}+P_0}{2}, \frac{P_0+P_1}{2}, \frac{P_1+P_2}{2}, \frac{P_2}{2} \right\}$

(4) [11.26] $\Rightarrow H(Y)\big|_{max} = \operatorname{ld} 6$ → 0

$$\frac{P_{-2}}{2} = \frac{P_2}{2} = \frac{1}{6} \qquad \frac{P_{-2}+P_{-1}}{2} = \frac{1}{6}; \qquad \frac{P_1+P_2}{2} = \frac{1}{6}; \qquad \frac{P_0}{2} = \frac{1}{6} \qquad \boxed{P_0 = \frac{1}{3}}$$

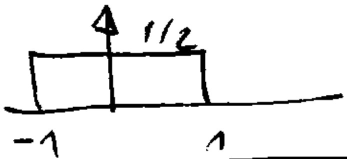
$P_{-2} = P_2 = P_0 = \frac{1}{3}$ $\quad P_{-1} = P_1 = 0 \quad$ $Y = X + Z$

$I(X;Y) = H(Y) - H(Y|X)$

$h(Y|X) = \sum_x p(x) \cdot H(Y|X=x) = \sum_x p(x) \cdot H(x+Z|x)$

$h(Y|X) = h(Z)$

$h(Z) = \int_{-1}^{1} \frac{1}{2}\, ld\, 2 \, dz = \frac{1}{2}(1+1) = 1$



$h(Y|X) = P_{-2} H(Y|X=-2) + P_{-1} H(Y|X=-1) + P_0 H(Y|X=0) +$
$+ P_1 H(Y|X=1) + P_2 H(Y|X=2) =$

$P_{-2} H(Y|X=-2) = -P_{-2} \int_{-3}^{-1} f(y)\, ld\, f(y)\, dy = -P_{-2} \int_{-3}^{-2} \frac{P_{-2}}{2}\, ld \frac{P_{-2}}{2}\, dy$
$+ \int_{-2}^{-1} \frac{P_{-2}+P_{-1}}{2}\, ld \frac{P_{-2}+P_{-1}}{2}\, dy = -P_{-2}\left[\frac{P_{-2}}{2}\, ld \frac{P_{-2}}{2} + \frac{P_{-2}+P_{-1}}{2}\, ld \frac{P_{-2}+P_{-1}}{2}\right]$
$= -P_2 \left[\frac{1}{6}\, ld \frac{1}{6} + \frac{1}{6}\, ld \frac{1}{6}\right] = -\frac{1}{6} \cdot \frac{1}{3}\, ld \frac{1}{6} = \frac{1}{24} \cdot ld\, 6$

$P_{-1} H(Y|X=-1) = \frac{1}{24}\, ld\, 6$!

$h_Y(y) = \int_{-3}^{-2} \frac{P_{-2}}{2}\, ld \frac{P_{-2}}{2}\, dy = \frac{1}{6}\, ld\, 6 \qquad h_Y(y) = \left(\frac{1}{6}\, ld\, 6\right) 6 = ld\, 6$

— MENE MI IZLEGUVA DOLLA

$C = ld\, 6 - h(Z) = ld\, 6 - 1 = ld\, 6 - ld\, 2 = ld\, 3$

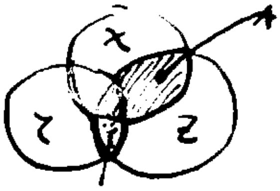
TORONTO U.

POTVRDENO OVAA GODINA VO sol4.pdf ∞

PROBLEM 8.9 GAUSSIAN MUTUAL INFORMATION. Suppose that (X,Y,Z) are jointly Gaussian and that $X \to Y \to Z$ forms a Markov chain. Let X and Y have correlation coefficient ρ_1 and let Y and Z have correlation coefficient ρ_2. Find $I(X;Z)$.

$I(X,Y;Z) = I(X;Z) + I(Y;Z|X) = I(X;Z) + I(Y;Z|X)$
$= I(Y;Z) + I(X;Z|Y)$, with $I(X;Z|Y) = 0$

[redacted]



$$I(x,y;z) = I(x;z) + I(y;z|x)$$

$$I(x;z|y) = H(z|y) - H(z|x,y)$$
$$= H(z|y) - H(z|y) = 0$$

$$K_1 = \begin{bmatrix} \sigma_x^2 & \rho_1 \\ \rho_1 & \sigma_y^2 \end{bmatrix}$$

$$K_2 = \begin{bmatrix} \sigma_y^2 & \rho_2 \\ \rho_2 & \sigma_z^2 \end{bmatrix}$$

$$K = \begin{bmatrix} \sigma_x^2 & \rho_1 & 0 \\ \rho_1 & \sigma_y^2 & \rho_2 \\ 0 & \rho_2 & \sigma_z^2 \end{bmatrix}$$

$$p(x,y) = \frac{1}{(\sqrt{2\pi})^2 |K_1|^{1/2}} e^{-[x,y] K_1^{-1} \begin{bmatrix} x \\ y \end{bmatrix}} ; \quad p(x,y,z) = \frac{1}{(\sqrt{2\pi})^3 \sqrt{K}} e^{-\begin{bmatrix} x \\ y \\ z \end{bmatrix} K^{-1} [x\,y\,z]}$$

$$I(x;z) = H(x) - H(x|z)$$

$$H(x,y,z) = H(x,y) + H(z|x,y) = H(x,y) + \sout{H(z|y)}$$

$$h(x,y,z) = \frac{1}{2} \operatorname{ld} (2\pi e)^3 |K| \qquad h(x,y) = \frac{1}{2} \operatorname{ld} (2\pi e)^2 |K_1|$$

$$p(x) = \int_{-\infty}^{\infty} \frac{1}{(\sqrt{2\pi})^2 |K_1|^{1/2}} e^{-[x\,y] \cdot K_1^{-1} \begin{bmatrix} x \\ y \end{bmatrix}} dy$$

$$p(x) = \frac{1}{2\pi \sqrt{|K_1|}} \int_{-\infty}^{\infty} e^{+\frac{x^2 \sigma_y^2 - 2xy\rho_1 + y^2 \sigma_x^2}{\sigma_x^2 \sigma_y^2 - \rho_1^2}} dy$$

$$p(x) = \frac{1}{2\pi \sqrt{(\sigma_x^2 \sigma_y^2 - \rho_1^2)}} \cdot \frac{e^{-\frac{x^2 (\sigma_x^2 - 1)}{(\sigma_x^2 \sigma_y^2 - \rho_1^2)\sigma_x^2}}}{\sqrt{\sigma_x^2}} \cdot \sqrt{\sigma_x^2 \sigma_y^2 - \rho_1^2}$$

$$p(x) = \frac{1}{2\pi\, \sigma_x} e^{-\frac{x^2 (\sigma_x^2 - 1)}{(\sigma_x^2 \sigma_y^2 - \rho_1^2)\sigma_x^2}}$$

• PROBLEM 8.11 CONTINUE FROM N16K

(e) EVEN THOUGH THE PROCESS IS NOT ERGODIC, IT IS STATIONARY, AND IT DOES HAVE AN AEP, BECAUSE

$$-\frac{1}{n} \ln f(x^n) = -\frac{1}{n} \ln \frac{1}{(2\pi)^{n/2} \sqrt{|K_{x^n}|}} e^{-x^t K^{-1} x/2} =$$

$$= +\frac{1}{n} \ln (2\pi)^{n/2} + \frac{1}{n} \ln \sqrt{|K_{x^n}|} + \frac{1}{2n} x^t K^{-1} X =$$

$= \frac{1}{2n}\ln(2\pi)^n + \frac{1}{2n}\ln|K_{x^n}| + \frac{1}{2n}x^t K^{-1} x =$

$= \frac{1}{2n}\ln(2\pi)^n|K_{x^n}| + \frac{1}{2n}x^t K^{-1}x + \frac{1}{2n}\ln e^n - \frac{1}{2n}\ln e^n$

$= \frac{1}{2n}\ln(2\pi e)^n|K_{x^n}| + \frac{1}{2n}x^t K^{-1}x - \frac{1}{2} = \frac{1}{n}h(x^n) - \frac{1}{2} + \frac{1}{2n}x^t K^{-1} x$

SINCE: $x \sim N(0,K)$ WE ~~CAN~~ WRITE $x = K^{1/2}W$ WHERE $W = N(0,I)$. THEN:

$x^t K^{-1} x = W^T K^{\frac{1}{2}T}\cdot K^{-1}\cdot K^{1/2}W = W^T K^{\frac{1}{2}}K^{-1}K^{\frac{1}{2}}W = W^t W$

K IS SIMMETRICA

$\boxed{x^t K^{-1} x = W^t W} = \sum W_i^2$ THEREFORE $x^t K^{-1} x$ HAS χ^2 (chi-squared) DISTRIBUTION WITH n DEGREES OF FREEDOM. THE DENSITY OF CHI-SQUARED DISTRIBUTION IS:

$$f(x) = \frac{x^{\frac{n}{2}-1}e^{-\frac{x}{2}}}{\Gamma\left(\frac{n}{2}\right)2^{n/2}} \qquad M(s) = \frac{1}{\sqrt{(1-2t)^k}}$$

$$M(s) = \int_{-\infty}^{\infty} f(x)e^{sx}\,dx$$

BY CHERNOFF BOUND:

$$Pr\left\{\frac{1}{n}\sum W_i^2 > 1+\epsilon\right\} \le \min_s e^{-s(1+\epsilon)}(1-2s)^{-\frac{n}{2}}$$

$$\le e^{-\frac{n}{2}(\epsilon - \ln(1+\epsilon))}$$

$s = \frac{\epsilon}{2(1+\epsilon)}$; $\quad e^{-\frac{\epsilon}{2(1+\epsilon)}\cdot(1+\epsilon)}\left(1 - \frac{\epsilon}{(1+\epsilon)}\right)^{-\frac{n}{2}} =$

$= e^{-\frac{\epsilon}{2}}\left(\frac{2+2\epsilon-\epsilon}{2(1+\epsilon)}\right)^{-\frac{n}{2}} = e^{-\frac{\epsilon}{2}}\left(\frac{1}{(1+\epsilon)}\right)^{-\frac{n}{2}} = e^{-\frac{\epsilon}{2}}(1+\epsilon)^{\frac{n}{2}}$

$= e^{-\frac{\epsilon}{2}}e^{\frac{n}{2}\ln(1+\epsilon)} = e^{-\frac{1}{2}[\epsilon - n\ln(1+\epsilon)]}$

$-\frac{1}{n}\ln[f(x^n)] - \frac{1}{n}h(x^n) = -\frac{1}{2} + \frac{1}{2n}\sum W_i^2$

$$Pr\left\{\left|-\frac{1}{n}\ln[f(x^n)] - \frac{1}{n}h(x^n)\right| > \epsilon\right\} = Pr\left\{\left|\frac{1}{2n}\sum W_i^2 - \frac{1}{2}\right| > \epsilon\right\} \le e^{-\frac{n}{2}(\epsilon - \ln(1+\epsilon))}$$

THE BOUND GOES TO 0 AS $n \to \infty$ AND THEREFORE BY BOREL CANTELLI LEMMA: $-\frac{1}{n}\ln f(x^n) - h(x^n) \to 0$ WITH PROBABILITY 1.

[redacted]

SO z_i SATISFIES AEP EVEN THOUGH IT IS NOT ERGODIC.

LEMMA 11.9.1 (CHERNOFF LEMMA) LET Y BE ANY RANDOM VARIABLE AND LET $\psi(s)$ BE THE MOMENT GENERATING FUNCTION Y.

$$\psi(s) = E[e^{sY}]$$

THEN FOR ALL $s \geq 0$: $\quad \Pr(Y \geq a) \leq e^{-sa}\psi(s)$

AND THUS:

$$\Pr(Y \geq a) \leq \min_{s \geq 0} e^{-sa}\psi(s)$$

MARKOV INEQUALITY:

$$\Pr\{p(x) \leq d\} \log \frac{1}{d} \leq H(X)$$

$$\Pr\{x \geq a\} \leq \frac{E[x]}{a}$$

$X_i = V + Z_i$

$i = n+1 \quad z_i = V$

$$f(V|z^n) \sim N\left(\frac{S}{nS+N}\sum z_i \, , \, \frac{SN}{nS+N}\right)$$

$$f(V|z^n) = \frac{f(V, z^n)}{f(z^n)} \qquad \frac{f(x,y)}{f(y)} = f(x|y)$$

$$\int x \cdot f(x|y)\, dx = E[x|y]$$

$$E[x] = \sum_y p(y)\, E[x|y] = \sum_y p(y) \sum_x x \cdot p(x|y)$$

$$f(V|z^n) = \frac{f(V, z^n)}{f(z^n)} \longrightarrow K_{(V,z^n)} = \begin{bmatrix} S+N & S & S \\ S & S+N & S \\ S & S & S \end{bmatrix}$$

MMV

MAILE: $\quad f(V|z^n) = \dfrac{1}{\sqrt{2\pi}\cdot\left(\frac{NS}{2S+N}\right)^{1/2}} e^{-\frac{1}{2}\frac{1}{SN(2S+N)}\left[VN - S(x_1+x_2+2V)\right]^2}$

$(*)$, σ

$$(*) = -\frac{1}{2}\frac{N^2}{SN(2S+N)}\left[V - \frac{S}{N}\sum_{i=1}^{2} x_i + 2V\cdot\frac{S}{N}\right]^2 =$$

$$= -\frac{1}{2}\frac{N^2}{SN(2S+N)}\left[V\left(\frac{N+2S}{N}\right) - \frac{S}{N}\sum_{i=1}^{2} x_i\right]^2 = -\frac{N}{S(2S+N)}\left[V\left(\frac{N+2S}{N}\right) - \frac{S}{N}\sum_{i=1}^{2} x_i\right]^2$$

$$= -\frac{N}{S(2S+N)}\left(\frac{N+2S}{N}\right)^2\left[V - \frac{N}{N+2S}\cdot\frac{S}{N}\sum_{i=1}^{2} x_i\right]^2 = -\frac{N+2S}{NS}\left[V - \frac{S}{N+2S}\sum_{1}^{2} x_i\right]^2$$

$1/\sigma^2$

$$M = \frac{S}{N+2S}\sum_{i=1}^{2} x_i \qquad \sigma^2 = \frac{NS}{N+2S}$$

$\frac{Q(x)}{P(x)} = k \cdot x$ $\quad \sum_x Q(x) = 1 \quad Q(x) = k \cdot x \, P(x)$

$\sum k \cdot x \, P(x) = 1 \quad k = \frac{1}{\sum x P(x)} = \frac{1}{G(x)}$

$Q(x) = \frac{x P(x)}{G(x)} \quad \log G(x) = \log \frac{x P(x)}{Q(x)}$

$\sum_x Q(x) \log(x) - \sum Q(x) \log \frac{x P(x)}{G(x) \cdot P(x)} = 0$

$\sum_x Q(x) \log(x) - \sum Q(x) \log x - \sum Q(x) \log \frac{1}{G(x)} = 0$

$\sum_x Q(x) \log(x) - E_Q[\log x] = \log \frac{1}{G(x)}$

$E_Q[\log x] - \sum_x Q(x) \log x$

$J(Q,\lambda) = \sum_x Q(x) \ln(x) - \sum Q(x) \ln \frac{Q(x)}{P(x)} + \lambda \left(\sum_x Q(x) - 1 \right)$

$\frac{dJ(Q,\lambda)}{dQ} = \ln x - \ln \frac{Q(x)}{P(x)} - 1 + \lambda = 0$

$\ln \frac{Q(x)}{P(x)} = \ln x - 1 + \lambda \quad e^{\ln x - 1 + \lambda} = \frac{Q(x)}{P(x)}$

$Q(x) = P(x)\, e^{\ln x - 1 + \lambda} = P(x) \cdot \underbrace{e^{\lambda - 1}}_{k} x = k \cdot P(x) \cdot x$

$\sum_x Q(x) = 1 \quad k \cdot \sum_x x P(x) = 1 \quad k = \frac{1}{G(x)}$

$e^{\lambda - 1} = \frac{1}{G(x)} \quad G(x) = e^{1-\lambda} \quad 1 - \lambda = \ln G(x)$

$\lambda = 1 - \ln G(x) \quad Q(x) = \frac{P(x)\, x}{G(x)}$

$J(Q,\lambda)\big|_{max} = E_Q(\ln x) - \sum Q(x) \ln \frac{P(x)\, x}{P(x)\, G(x)} - 0 =$

$= E_Q(\ln x) - \sum Q(x) \ln x - \sum Q(x) \ln \frac{1}{G(x)} = \ln G(x)$

$J(Q,\lambda) \le J(Q,\lambda)\big|_{max} = \ln G(x) \quad = \lambda \left(\sum_x Q(x) - 1 \right)$

$\sum_x Q(x) \ln(x) - \sum Q(x) \ln \frac{Q(x)}{P(x)} - 0 \le \ln G(x)$

$$E[X|Y] = \frac{Y \cdot \rho \sigma_1}{\sqrt{\sigma_1^2 \sigma_2^2 (1-\rho^2)} \sqrt{\frac{1}{\sigma_1^2(1-\rho^2)}}} = \frac{Y \rho \sigma_1}{\sigma_2}$$

CHAPTER 9.1 GAUSSIAN CHANNEL: DEFINITIONS

WE NOW DEFINE THE (INFORMATION) CAPACITY OF THE CHANNEL AS THE MAXIMUM OF THE MUTUAL INFORMATION BETWEEN THE INPUT AND OUTPUT OVER ALL DISTRIBUTIONS ON THE INPUT THAT SATISFY THE POWER CONSTRAINT

DEFINITION THE INFORMATION CAPACITY OF GAUSSIAN CHANNEL AS MAXIMUM OF THE MUTUAL INFORMATION BETWEEN THE INPUT AND OUTPUT OVER ALL DISTRIBUTIONS ON THE INPUT THAT SATISFY THE POWER CONSTRAINT IS:

$$C = \max_{f(x): EX^2 \le P} I(X;Y)$$

$I(X;Y) = h(Y) - h(Y|X) = h(Y) - h(X+Z|X) =$
$h(Y) - h(Z)$; $h(Z) = \frac{1}{2} \log 2\pi e N$

$E[Y^2] = E[(X+Z)^2] = E[X^2 + 2XZ + Z^2] = E[X^2] + 2E[XZ] + E[Z^2]$, with $2E[XZ] = 0$

$E[Y^2] = E[X^2] + E[Z^2] = P + N$

$$h(Y) \le \frac{1}{2} ld (2\pi e) \cdot (P+N)$$

$I(X;Y) = h(Y) - h(Z) \le \frac{1}{2} ld (2\pi e)(P+N) - \frac{1}{2} ld(2\pi e N)$

$$I(X;Y) \le \frac{1}{2} ld (2\pi e)\left(1 + \frac{P}{N}\right)$$

HENCE THE INFORMATION CAPACITY OF GAUSSIAN CHANNEL IS:

$$C = \max_{E[X^2] \le P} I(X;Y) = \frac{1}{2} ld (2\pi e)\left(1 + \frac{P}{N}\right)$$

THE MAXIMUM IS ATTAINED WHEN $X \sim N(0,P)$.

- WE WILL NOW SHOW THAT THIS CAPACITY IS ALSO THE SUPREMUM OF THE RATES ACHIEVABLE FOR THE CHANNEL.

DEFINITION: AN (M, n) CODE FOR THE GAUSSIAN CHANNEL WITH POWER CONSTRAINT P CONSIST OF THE FOLLOWING:

1. AN INDEX SET $\{1, 2, \dots M\}$
2. AN ENCODING FUNCTION $X: \{1, 2, \dots M\} \to \mathcal{X}^n$

[redacted]

YIELDING CODEWORDS: $x^n(1), x^n(2) \dots x^n(M)$, SATISFYING THE POWER CONSTRAINT P; THAT IS, FOR EVERY CODEWORD:

$$\boxed{\sum_{i=1}^{n} x_i^2(w) \leq nP} \qquad w = 1, 2, \dots M$$

3. A DECODING FUNCTION

$$g: \mathcal{Y}^n \to \{1, 2, \dots M\}$$

- THE ARITMETIC AVERAGE OF PROBABILITY OF ERROR IS DEFINED BY:

$$P_e^{(n)} = \frac{1}{2^{nR}} \sum \lambda_i$$

$$\lambda_i = \Pr\left(g(Y^n) \neq i \mid X^n = x^n(i)\right) = \sum_{y^n} P(y^n | x^n(i)) \cdot I\left(P(y^n|x^n(i)) \neq 0\right)$$

i.e. $g(y^n) \neq i$

$$\lambda_i = \Pr\left(g(Y^n) \neq i \mid X^n = x^n(i)\right) = \sum_{y^n} P(y^n | x^n(i)) \cdot I\left(g(y^n) \neq i\right)$$

$$\boxed{\lambda^{(n)} = \max_{i \in [1 \dots M]} \lambda_i} \qquad P_e^{(n)} = \frac{1}{2^{nR}} \sum_{i=1}^{M} \lambda_i$$

($2^{nR} = M$)

DEFINITION A RATE „R" IS SAID TO BE ACHIEVABLE FOR GAUSSIAN CHANNEL WITH POWER CONSTRAINT P IF THERE EXIST A SEQUENCE OF $(2^{nR}, n)$ CODES WITH CODEWORDS SATISFYING THE POWER CONSTRAINT SUCH THAT THE MAXIMAL PROBABILITY OF ERROR $\lambda^{(n)}$ TENDS TO ZERO. THE CAPACITY OF THE CHANNEL IS THE SUPREMUM OF THE ACHIEVABLE RATES.

THEOREM 9.1.1 THE CAPACITY OF A GAUSSIAN CHANNEL WITH POWER CONSTRAINT P AND NOISE VARIANCE „N" IS:

$$\boxed{C = \frac{1}{2} \mathrm{ld}\left(1 + \frac{P}{N}\right)} \quad \text{BITS PER TRANSMISSION}$$

PROOF (ACHIEVABILITY)

1. GENERATION OF THE CODEBOOK. WE WISH TO GENERATE CODEBOOK IN WHICH ALL THE CODEWORDS SATISFY THE POWER CONSTRAINT. TO ENSURE THIS, WE GENERATE THE CODEWORDS WITH EACH ELEMENT i.i.d ACCORDING TO

A NORMAL DISTRIBUTION WITH VARIANCE $P-\varepsilon$. SINCE FOR LARGE n, $\frac{1}{n}\sum x_i^2 \to P-\varepsilon$ THE PROBABILITY THAT CODEWORD DOESN'T SATISFY THE POWER CONSTRAINT IS VERY SMALL.

LET $X_i(w)$, $i=1,2,\dots n$, $w=1,2,\dots 2^{nR}$ BE I.I.D. $\sim N(0,P-\varepsilon)$, FORMING CODEWORDS $X^n(1), X^n(2), \dots X^n(2^{nR}) \in R^n$.

(2) <u>ENCODING.</u> AFTER GENERATION OF THE CODEBOOK, THE CODEBOOK IS REVEALED TO BOTH SENDER AND RECEIVER. <u>TO SEND THE MESSAGE INDEX</u> W, THE TRANSMITTER SENDS THE W-TH CODEWORD $X^n(W)$ IN THE CODEBOOK.

(3) DECODING. THE RECEIVER LOOKS DOWN THE LIST OF CODEWORDS $\{X^n(w)\}$ AND SEARCHES FOR ONE THAT IS JOINTLY TYPICAL WITH RECEIVED VECTOR. IF THERE IS ONE AND ONLY ONE SUCH CODEWORD $X^n(w)$, THE RECEIVER DECLARES $\hat{W}=w$ TO BE THE TRANSMITTED CODEWORD. OTHERWISE THE RECEIVER DECLARES AN ERROR. THE RECEIVER ALSO DECLARES AN ERROR IF THE CHOSEN CODEWORD DOESN'T SATISFY THE POWER CONSTRAINT.

(4) WITHOUT LOSS OF GENERALITY ASSUME THAT CODEWORD 1 WAS SENT. THUS $Y^n = X^n(1) + Z^n$. DEFINE THE FOLLOWING EVENTS:

$$E_0 = \left\{ \frac{1}{n}\sum_{i=1}^{n} X_i(1) > P \right\}$$ AND

$$E_i = \left\{ (X^n(i), Y^n) \text{ IS IN } A_\varepsilon^{(n)} \right\}$$

THEN AN ERROR OCCURS IF E_0 OCCURS (THE POWER CONSTRAINT IS VIOLATED) OR E_1^c OCCURS (THE TRANSMITTED CODEWORD AND THE RECEIVED SEQUENCE ARE NOT JOINTLY TYPICAL) OR $E_2 \cup E_3 \cup \dots \cup E_{2^{nR}}$ OCCURS (SOME WRONG CODEWORD IS JOINTLY TYPICAL WITH RECEIVED SEQUENCE). LET "E" DENOTE THE EVENT $\hat{W} \neq W$ AND LET "P" DENOTE THE CONDITIONAL PROBABILITY GIVEN THAT $W=1$. HENCE,

$$\Pr(E|W=1) = P(E) = P(E_0 \cup E_1^c \cup E_2 \cup E_3 \cup \dots \cup E_{2^{nR}}) \leq$$

$$\leq P(E_0) + P(E_1^c) + \sum_{i=2}^{2^{nR}} P(E_i)$$

- BY THE LAW OF LARGE NUMBERS $P(E_0) \to 0$ AS $n \to \infty$
- $P(E_1^c) \to 0$ hence $P(E_1^c) \leq \varepsilon$ FOR SUFFICIENTLY LARGE n

[redacted]

PROBABILITY THAT $X^n(i)$ AND Y^n ARE JOINTLY TYPICAL IS $\le 2^{-n(I(X;Y)-3\varepsilon)}$

- LET W BE UNIFORMLY DISTRIBUTED $\{1,2,\dots,2^{nR}\}$

$$Pr(\mathcal{E}) = \frac{1}{2^{nR}}\sum_{i=1}^{2^{nR}} \lambda_i = P_e^{(n)}$$

$$P_e^{(n)} = Pr(\mathcal{E}) = Pr(\mathcal{E}|W=1) \le P(E_0)+P(E_1^c)+\sum_{i=2}^{2^{nR}} P(E_i) \le$$

$$\le \varepsilon+\varepsilon+\sum_{i=2}^{2^{nR}} 2^{-n(I(X;Y)-3\varepsilon)} = 2\varepsilon + (2^{nR}-1)\,2^{-n(I(X;Y)-3\varepsilon)}$$

~~$= 2\varepsilon + 2^{nR-n(I(X;Y)-3\varepsilon)} - 2^{-n(I(X;Y)-3\varepsilon)} =$~~

~~$= 2\varepsilon + 2^{-n(I(X;Y)-3\varepsilon-nR)}$~~

$$\le 2\varepsilon + 2^{nR}\cdot 2^{-n(I(X;Y)-3\varepsilon)} = 2\varepsilon + 2^{3n\varepsilon - n(I(X;Y)-R)} \le 3\varepsilon$$

RECALL

$$Pr\left((\tilde X^n,\tilde Y^n)\in A_\varepsilon^{(n)}\right) = \sum_{(\tilde x^n,\tilde y^n)\in A_\varepsilon^{(n)}} P(\tilde x^n,\tilde y^n) = \sum_{(\tilde x^n,\tilde y^n)\in A_\varepsilon^{(n)}} P(\tilde x^n)\cdot P(\tilde y^n) \ge (*)$$

$$\ge \sum_{(\tilde x^n,\tilde y^n)\in A_\varepsilon^{n}} 2^{-n(H(X)+\varepsilon)}\cdot 2^{-n(H(Y)+\varepsilon)} \ge (1-\varepsilon)\, 2^{+n(H(X,Y)-\varepsilon)}\, 2^{-nH(X)-n\varepsilon}\, 2^{-nH(Y)-n\varepsilon}$$

$$|A_\varepsilon^{(n)}| \ge (1-\varepsilon)\, 2^{n(H(X,Y)-\varepsilon)} \qquad // \qquad |A_\varepsilon^n| \le 2^{n(H(X,Y)+\varepsilon)}$$

$Pr(\tilde X^n,\tilde Y^n \in A_\varepsilon^{(n)}) \ge (1-\varepsilon)\, 2^{-n(I+3\varepsilon)}$

$$nH(X,Y) - n\varepsilon - nH(X) - n\varepsilon - nH(Y) - n\varepsilon =$$

$$= n(\underbrace{H(X,Y)-H(X)}_{H(Y|X)}-H(Y)) - 3n\varepsilon = -n\cdot I(X;Y) - 3n\varepsilon$$

$$1 = \sum_{(x^n,y^n)} P(x^n,y^n) \ge \sum_{(x^n,y^n)\in A_\varepsilon^n} P(x^n,y^n) \ge |A_\varepsilon^n|\cdot 2^{-n(H(X,Y)+\varepsilon)}$$

$$|A_\varepsilon^n| \le 2^{+n(H(X,Y)+\varepsilon)}$$

$$(*) \le \sum_{(\tilde x^n,\tilde y^n)\in A_\varepsilon^n} P(x^n)\cdot P(y^n) \le |A_\varepsilon^n|\, 2^{-n(H(X)-\varepsilon)}\, 2^{-n(H(Y)-\varepsilon)} \le (\#)$$

$$(1-\varepsilon) \le \sum_{x^n y^n\in A_\varepsilon^n} P(x^n,y^n) \le |A_\varepsilon^{(n)}|\cdot 2^{-n(H(X,Y)-\varepsilon)} \qquad A_\varepsilon^{(n)} \ge (1-\varepsilon)\, 2^{+n(H(X,Y)-\varepsilon)}$$

[redacted]

(II) $\leq$ $2^{n(H(X,Y)+\varepsilon) - nH(X) + n\varepsilon - nH(Y) + n\varepsilon} \leq 2^{n(H(X,Y)-H(X)-H(Y)+3\varepsilon)}$

$$P_r\left((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^{(n)}\right) \leq 2^{-n(I(X;Y)-3\varepsilon)}$$

$$\boxed{(1-\varepsilon)\, 2^{-n(I+3\varepsilon)} \leq P_r\left(\tilde{X}^n, \tilde{Y}^n \in A_\varepsilon^{(n)}\right) \leq 2^{-n(I(X;Y)-3\varepsilon)}}$$

$$2^{[3\varepsilon - I(X;Y) + R]\cdot n}$$

$$-[I(X;Y) - 3\varepsilon] + R < 0$$

$I(X;Y) - 3\varepsilon \geq R \leftarrow R \leq I - 3\varepsilon$; $-I + 3\varepsilon + R \leq 0$

ТРЕБА ДА Е ПОМАЛО ОД НУЛА ЗА ДА ЗА ДОВОЛНО ГОЛЕМО n $2^{n\cdot x} \to 0$

9.2 CONVERSE TO THE CODING THEOREM FOR GAUSSIAN CHANNELS

IN THIS SECTION WE COMPLETE THE PROOF THAT CAPACITY OF A GAUSSIAN CHANNEL IS: $C = \frac{1}{2} ld\left(1 + \frac{P}{N}\right)$ BY PROVING THAT RATES $R > C$ ARE NOT ACHIEVABLE. THE PROOF PARALLELS THE PROOF FOR DISCRETE CHANNEL. THE MAIN NEW INGREDIENT IS THE POWER CONSTRAINT.

$$\boxed{R \leq C = \frac{1}{2}\, ld\left(1 + \frac{P}{N}\right)}$$

CONSIDER $(2^{nR}, n)$ CODE THAT SATISFIES THE POWER CONSTRAINT: $\frac{1}{n}\sum_{i=1}^{n} x_i^2(w) \leq P \qquad w = 1, 2, \ldots 2^{nR}$

$$\boxed{W \to x^n(w) \to Y^n \to \hat{W}}$$

$$H(W|\hat{W}) = 1 + P_e^{n}\cdot nR = n\cdot\varepsilon_n \qquad \varepsilon_n \to 0 \text{ as } P_e^{(n)} \to 0$$

$$nR = H(W) = H(W|\hat{W}) + \underbrace{I(W;\hat{W})}_{H(W)-H(W|\hat{W})} \leq 1 + P_e^{(n)}\cdot nR + I(X^n;Y^n)$$

$$= 1 + P_e^{(n)} nR + h(Y^n) - h(Y^n|X^n) \leq n\cdot\varepsilon_n + \sum_{i=1}^{n} h(Y_i) - h(Z^n)$$

$$= n\cdot\varepsilon_n + \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Z_i) = n\cdot\varepsilon_n + \sum_{i=1}^{n} I(X_i;Y_i)$$

NOW LET P_i BE THE AVERAGE POWER OF THE i-TH COLUMN OF THE CODEBOOK, THAT IS:

$$P_i = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} x_i^2(w)$$

$Y_i = X_i + Z_i$

X_i & Z_i ARE INDEPENDENT

$$E[Y_i^2] = E[X_i^2] + E[Z_i^2] = P_i + N$$

ENTROPY IS MAXIMIZED BY THE NORMAL DISTRIBUTION:

$$h(Y_i) \le \frac{1}{2}\operatorname{ld}(2\pi e)\cdot(P_i+N)$$

$$nR \le \sum\left(h(Y_i) - h(Z_i)\right) + n\varepsilon_n \le \sum\left(\frac{1}{2}\operatorname{ld} 2\pi e(P_i+N) - \frac{1}{2}\operatorname{ld} 2\pi e N\right) + n\varepsilon_n = \sum \frac{1}{2}\operatorname{ld}\left(1+\frac{P_i}{N}\right) + n\varepsilon_n \quad (*)$$

SINCE EACH OF THE CODEWORDS SATISFIES THE POWER CONSTRAINT, SO DOES THEIR AVERAGE, AND HENCE:

$$\frac{1}{n}\sum_i P_i \le P$$

SINCE $\frac{1}{2}\operatorname{ld}(1+x)$ IS CONCAVE FUNCTION OF x, WE CAN APPLY JENSEN'S INEQUALITY TO OBTAIN:

$$\underbrace{\frac{1}{n}\sum \frac{1}{2}\operatorname{ld}\left(1+\frac{P_i}{N}\right)}_{E[f(x)]} \le \underbrace{\frac{1}{2}\operatorname{ld}\left(n\cdot\frac{1}{n} + \frac{1}{n}\sum_{i=1}^{n}\frac{P_i}{N}\right)}_{f[E[x]]} \le \frac{1}{2}\operatorname{ld}\left(1+\frac{P}{N}\right)$$

$$(*) \quad R \le \frac{1}{n}\sum\frac{1}{2}\operatorname{ld}\left(1+\frac{P_i}{N}\right) + \varepsilon_n$$

$$nR \le n\cdot\frac{1}{n}\sum\frac{1}{2}\operatorname{ld}\left(1+\frac{P_i}{N}\right) + n\varepsilon_n$$

$$R \le \frac{1}{2}\operatorname{ld}\left(1+\frac{P}{N}\right) + \varepsilon_n \qquad \varepsilon_n \to 0$$

9.3 BANDLIMITED CHANNELS

A COMMON MODEL FOR COMMUNICATION OVER A RADIO NETWORK OR A TELEPHONE LINE IS A BANDLIMITED CHANNEL WITH WHITE NOISE.

$$Y(t) = \left(X(t) + Z(t)\right) * h(t)$$

$X(t)$ – SIGNAL WAVEFORM

$Z(t)$ – IS THE WAVEFORM OF WHITE GAUSSIAN NOISE

$h(t)$ – IMPULSE RESPONSE OF IDEAL BANDPASS FILTER WHICH CUTS ALL FREQUENCIES GREATER THEN W.

THEOREM 9.3.1 SUPPOSE THAT A FUNCTION $f(t)$ IS BANDLIMITED TO W, NAMELY, THE SPECTRUM OF THE FUNCTION IS 0 FOR ALL FREQUENCIES GREATER THEN W.

THEN THE FUNCTION IS COMPLETELY DETERMINED BY SAMPLES OF THE FUNCTION SPACED $1/2W$ SECONDS APART.

<u>PROOF</u>: $F(\omega) = \mathcal{F}\{f(t)\}$

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega = \frac{1}{2\pi}\int_{-2\pi W}^{2\pi W} F(\omega) e^{j\omega t} d\omega$$

VALUES OF THE SIGNAL AT THE SAMPLE POINTS IS:

$$f\left(\frac{n}{2W}\right) = \frac{1}{2\pi}\int_{-2\pi W}^{2\pi W} F(\omega) e^{j\omega \frac{n}{2W}} d\omega$$

$$\operatorname{sinc}(t) = \frac{\sin(2\pi W t)}{2\pi W t}$$

THE SPECTRUM OF THIS FUNCTION IS CONSTANT IN THE BAND $(-W, W)$ AND 0 OUTSIDE.

$$g(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right) \operatorname{sinc}\left(t - \frac{n}{2W}\right) \qquad g(t) = f(t)$$

$g(t)$ REPRESENTS $f(t)$ IN TERMS OF ITS SAMPLES.

A GENERAL FUNCTION HAS AN INFINITE NUMBER OF DEGREES OF FREEDOM - THE VALUE OF THE FUNCTION AT EVERY POINT CAN BE CHOSEN INDEPENDENTLY. THE NYQUIST-SHANNON SAMPLING THEORY SHOWS THAT BANDLIMITED FUNCTION HAS ONLY $2W$ DEGREES OF FREEDOM PER SECOND.

$$f(t) = \sum_{n=-\infty}^{\infty} C_n e^{+j\omega_0 n t} \qquad C_n = \int f(t) e^{-j\omega_0 n t} dt$$

$$f(t) = \sum_{n=-\infty}^{\infty} a_n \cos(\omega_0 n t) + b_n \sin(\omega_0 n t)$$

THERE ARE ABOUT <u>$2TW$ ORTHONORMAL BASIS FUNCTIONS</u> FOR THE SET OF ALMOST TIME-LIMITED, ALMOST BANDLIMITED FUNCTIONS, AND WE CAN DESCRIBE ANY FUNCTION WITHIN THE SET BY ITS COORDINATES IN THIS BASIS. MOREOVER THE PROJECTION OF WHITE NOISE ON THIS BASIS VECTORS FORMS AN I.I.D. GAUSSIAN PROCESS. <u>THE ABOVE ARGUMENTS ENABLE US TO VIEW THE BANDLIMITED TIME LIMITED FUNCTIONS AS VECTORS IN VECTOR SPACE OF $2TW$ DIMENSIONS.</u>

ASSUMING THAT CHANNEL HAS BANDWIDTH W, WE CAN REPRESENT BOTH THE INPUT AND THE OUTPUT BY SAMPLES TAKEN $1/2W$ SECONDS APART. EACH OF THE INPUT SAMPLES

IS CORRUPTED BY NOISE TO PRODUCE THE CORRESPONDING OUTPUT SAMPLE. SINCE THE NOISE IS WHITE AND GAUSSIAN, IT CAN BE SHOWN THAT EACH NOISE SAMPLE IS AN INDEPENDENT, IDENTICALLY DISTRIBUTED GAUSSIAN RANDOM VARIABLE.

IF THE NOISE HAS POWER SPECTRAL DENSITY $\frac{N_0}{2}\ \frac{W}{Hz}$ AND THE BANDWIDTH OF W Hz, THE NOISE HAS POWER

$$\frac{N_0}{2}\cdot 2W = N_0\cdot W \quad \}①$$

SAMO DA KAŽE DEKA SEKOJ SAMPLE IMA VARIJANSA ① SEKI, RAMNO JUSTINA (VO SEKOJA SEKUNDA SPAGA)

EACH OF THE $2WT$ NOISE SAMPLES IN TIME T HAS VARIANCE $N_0WT/2WT = N_0/2$

LOOKING AT THE INPUT AS A VECTOR IN THE $2TW$-DIMENSIONAL SPACE WE SEE THAT THE RECEIVED SIGNAL IS SPHERICALLY NORMALLY DISTRIBUTED ABOUT THIS POINT WITH COVARIANCE $\frac{N_0}{2}I$.

- CAPACITY OF DISCRETE GAUSSIAN CHANNEL IS:

$$C=\frac{1}{2}\,ld\left(1+\frac{P}{N}\right)$$

- LET THE CHANNEL BE USED OVER THE INTERVAL: $[0,T]$

THE ENERGY PER SAMPLE IS $PT/2WT = \boxed{P/2W}$

- NOISE VARIANCE PER SAMPLE IS:

$$\frac{N_0}{2}2W\cdot\frac{T}{2WT} = \boxed{\frac{N_0}{2}}$$

HENCE THE CAPACITY PER SAMPLE IS:

$$\boxed{C=\frac{1}{2}\,ld\left(1+\frac{\frac{P}{2W}}{\frac{N_0}{2}}\right)=\frac{1}{2}\,ld\left(1+\frac{P}{N_0\cdot W}\right)}\ \text{bits per sample}$$

SINCE THERE ARE $2W$ SAMPLES EACH SECOND THE CAPACITY OF THE CHANNEL IS:

$$\boxed{C=W\,ld\left(1+\frac{P}{N_0\cdot W}\right)}\ \text{BITS PER SECOND}$$

HMV

CAPACITY OF A BANDLIMITED GAUSSIAN CHANNEL WITH NOISE SPECTRAL DENSITY $N_0/2$ WATTS/Hz AND POWER P WATTS.

$$C=\lim_{W\to\infty}\frac{ld\left(1+\frac{P}{N_0 W}\right)}{\frac{1}{W}}=\lim_{W\to\infty}\frac{\frac{1}{\ln 2}\frac{1}{1+\frac{P}{N_0 W}}\cdot\frac{P}{N_0}\cancel{\left(\frac{-1}{W^2}\right)}}{\cancel{\left(\frac{-1}{W^2}\right)}}=\frac{P}{N_0}\cdot\frac{1}{\ln 2}$$

$$C=\frac{P}{N_0}\cdot\frac{\ln e}{\ln 2} \qquad \boxed{C=\frac{P}{N_0}\cdot ld\,e}$$

OTTUKA ZA KANALI SO BESKONEČEN OPSEG, KAPACITETOT RASTE LINEARNO SO MOĆNOSTA NA KANALOT.

[redacted]

EXAMPLE 9.3.1 (TELEPHONE LINE) TO ALLOW MULTIPLEXING OF MANY CHANNELS, TELEPHONE SIGNALS ARE BANDLIMITED TO 3300 Hz. USING A BANDWIDTH OF 3300 Hz AND SNR OF 33 dB (i.e. $P/N_0W = 2000$)

$N_0 \,(=)\, \frac{\text{Watt}}{\text{Hz}}$

$$10 \log \frac{P}{N_0 W} = 33 \qquad \frac{P}{N_0 W} = 10^{3.3} = 1995$$

WE FIND THE CAPACITY OF THE CHANNEL TO BE:

$$C = 3300 \log\left(1 + 1995\right) = 36.178,00 \text{ bits/sec}$$

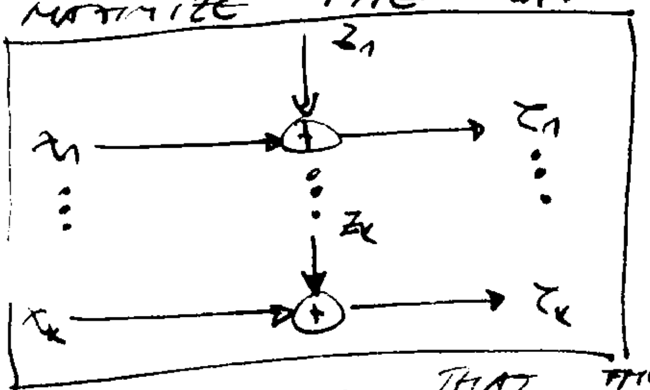
IN REAL TELEPHONE CHANNELS, THERE ARE OTHER FACTORS, SUCH AS CROSSTALK, INTERFERENCE, ECHOES, AND NONFLAT CHANNELS WHICH MUST BE COMPENSATED FOR TO ACHIEVE THIS CAPACITY.

V90 MODEMS ACHIEVE 56 kbps (> 36 kbps) OVER TELEPHONE CHANNEL, ACHIEVE THIS RATE IN ONE DIRECTION TAKING ADVANTAGE OF PURELY DIGITAL CHANNEL FROM THE SERVER TO THAT TELEPHONE SWITCH IN THE NETWORK.

15 24
12
27

9.4 PARALLEL GAUSSIAN CHANNELS

IN THIS SECTION WE CONSIDER "k" INDEPENDENT GAUSSIAN CHANNELS IN PARALLEL WITH A COMMON POWER CONSTRAINT. THE OBJECTIVE IS TO DISTRIBUTE THE TOTAL POWER AMONG THE CHANNELS SO AS TO MAXIMIZE THE CAPACITY.



THE OUTPUT OF EACH IS SUM OF THE INPUT AND THE GAUSSIAN NOISE. FOR CHANNEL "j":

$$Y_j = X_j + Z_j \qquad j = 1, 2, \dots k,$$

$$Z_j \sim \mathcal{N}(0, N_j)$$

AND THE NOISE IS ASSUMED TO BE INDEPENDENT FROM CHANNEL TO CHANNEL.

WE ASSUME THAT THERE IS A COMMON POWER CONSTRAINT ON THE TOTAL POWER USED i.e:

$$E\left[\sum_{j=1} X_j^2\right] \leq P$$

WE WISH TO DISTRIBUTE THE POWER AMONG THE VARIOUS CHANNELS SO AS TO MAXIMIZE THE TOTAL CAPACITY.

$$C = \max_{f(x_1, x_2, \dots x_k) : \sum E[X_i^2] \leq P} I(X_1 X_2 \dots X_k ; Y_1, Y_2, \dots Y_k)$$

$$I(X_1^k ; Y_1^k) = h(Y_1^k) - h(Y_1^k | X_1^k) \leq \sum_{i=1}^{k} h(Y_i) - h(Z_1^k)$$

[redacted]

$$= \sum_{i=1}^{k} h(Y_i) - \sum_{i=1}^{k} h(Z_i) = \sum_{i=1}^{k} I(X_i;Y_i) \le \sum_i \frac{1}{2} ld\left(1+\frac{P_i}{N_i}\right)$$

$P_i = E[X_i^2]$ $\qquad P = \sum P_i$

$$(X_1^k) \sim N\left(0, \begin{bmatrix} P_1 & 0 & 0 & \cdots & 0 \\ 0 & P_2 & 0 & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & 0 & P_k \end{bmatrix}\right)$$

$$\boxed{\sum_i \frac{1}{2} ld\left(1+\frac{V-N_i}{N_i}\right)}$$

SO THE PROBLEM IS REDUCED TO FINDING THE POWER ALLOTMENT THAT MAXIMIZES THE CAPACITY SUBJECT TO CONSTRAINT THAT $\sum P_i = P$.

$$J(P_1, P_2, \dots P_k) = \sum \frac{1}{2} ld\left(1+\frac{P_i}{N_i}\right) + \lambda\left(\sum_{i=1}^{k} P_i - P\right)$$

BITS

$$\frac{\partial J(P_i)}{\partial P_i} = 0 \qquad \frac{1}{\ln 2}\frac{1}{2}\cdot\frac{1}{1+\frac{P_i}{N_i}}\cdot\frac{1}{N_i} + \lambda = 0$$

$$\frac{1}{2}\cdot\frac{1}{N_i+P_i} + \lambda = 0 \text{ (NATS)} \qquad \frac{1}{2(N_i+P_i)} = -\lambda$$

$$2(N_i+P_i) = -\frac{1}{\lambda} \qquad (*)\ \boxed{P_i = -\frac{1}{2\lambda} - N_i = V - N_i}$$

P_i MUST BE POSITIVE (NONNEGATIVE) → KUHN-TUCKER CONDITIONS ⇒

$$P_i = (V-N_i)^+ \qquad \boxed{\sum (V-N_i)^+ = P}$$

$$(x)^+ = \begin{cases} x & \text{IF } x \ge 0 \\ 0 & \text{IF } x < 0 \end{cases}$$

!!! V E CONSTANTA →

(*) E SUŠTINSKI IZRAZ !!! E GORNATA NA ~~[illegible]~~ SNAGATA VO i-OT KANAL ... $N_i \downarrow$ $P_i \uparrow$ I OBRATNO. OVA $V-N_i$ T.E. AKO ... E VSUŠNOST WATER FILLING ALGORITAM.

9.5 CHANNELS WITH COLORED GAUSSIAN NOISE

IN 9.4. WE CONSIDERED A SET OF PARALLEL INDEPENDENT GAUSSIAN CHANNELS IN WHICH THE NOISE SAMPLES FROM DIFFERENT CHANNELS WERE INDEPENDENT. NOW WE WILL CONSIDER THE CASE WHERE THE NOISE IS DEPENDENT. FOR CHANNELS WITH MEMORY WE CAN CONSIDER A BLOCK OF "n" CONSECUTIVE USES OF THE CHANNEL AS "n" CHANNELS IN PARALLEL WITH DEPENDENT NOISE.

LET K_Z BE THE COVARIANCE MATRIX OF THE NOISE, AND LET K_X BE THE INPUT COVARIANCE MATRIX.

[redacted]

TRACE = (SUM OF THE DIAGONAL)

THE POWER CONSTRAINT ON THE INPUT CAN THEN BE WRITTEN AS:

$$\frac{1}{n}\sum_i \sigma x_i^2 \le P \quad \text{OR EQUIVALENTLY:} \quad \frac{1}{n}\,tr(K_x) \le P$$

UNLIKE SECTION 9.4 THE POWER CONSTRAINT HERE DEPENDS ON „n", THE CAPACITY WILL HAVE TO BE CALCULATED FOR EACH „n".

$$I(x_1^n; y_1^n) = h(Y_1^n) - h(Z_1^n)$$

$h(Z_1^n)$ IS DETERMINED ONLY BY THE DISTRIBUTION OF THE NOISE AND IS NOT DEPENDENT ON THE CHOICE OF THE INPUT DISTRIBUTION. SO FINDING CAPACITY AMOUNTS TO MAXIMIZING $h(Y_1, Y_2, \ldots Y_n)$. THE ENTROPY OF THE OUTPUT IS MAXIMIZED WHEN Y IS NORMAL, WHICH IS ACHIEVED WHEN INPUT IS NORMAL. SINCE THE INPUT AND NOISE ARE INDEPENDENT THE COVARIANCE OF THE OUTPUT IS: $K_Y = K_X + K_Z$ AND:

$$h(Y_1^n) = \frac{1}{2}\, ld\left((2\pi e)^n |K_x + K_z|\right)$$

NOW THE PROBLEM IS REDUCED TO CHOSING K_X SO AS TO MAXIMIZE $|K_X + K_Z|$, SUBJECT TO A TRACE CONSTRAINT ON K_X. TO DO THIS, WE DECOMPOSE K_Z INTO DIAGONAL FORM,

$$K_Z = Q \Lambda Q^t \quad \text{WHERE} \quad QQ^T = I$$

$$|K_x + K_Z| = |K_x + Q\Lambda Q^T| = |Q(Q^T K_x Q + \Lambda)Q^T| =$$

$$= |Q| \cdot |Q^t K_x Q + \Lambda| |Q^T| = |A + \Lambda|$$

MMV

- DISTRIBUTIVEN ZAKON VELJA ZA DETERMINANTI

$$|AB| = |A| \cdot |B| \qquad |I| = |AA^{-1}| = |A| \cdot |A^{-1}| = 1$$

$$A = Q^t K_x Q$$

- THE MATRICES IN THE TRACE OF THE PRODUCT CAN BE SWITCHED: $tr(BC) = tr(CB)$

$$tr(A) = tr(\underbrace{Q^t K_x}_{B} \underbrace{Q}_{C}) = tr(QQ^t K_x) = tr(K_x)$$

NOW THE PROBLEM IS REDUCED TO MAXIMIZING $|A+\Lambda|$ SUBJECT TO TRACE CONSTRAINT: $tr(A) \le nP$.
HADAMARD'S INEQUALITY STATES THAT THE DETERMINANT OF ANY POSITIVE DEFINITIVE MATRIX K IS LESS THEN THE PRODUCT OF ITS DIAGONAL ELEMENTS, THAT IS:

$|K| \le \prod_i K_{ii}$ WITH EQUALITY IF THE MATRIX IS DIAGONAL.

$|A+\Lambda| \le \prod_i (A_{ii}+\lambda_i)$ $\frac{1}{n}\sum_i A_{ii} \le P$

MAXIMUM VALUE OF $\prod_i (A_{ii}+\lambda_i)$ IS ATTAINED WHEN

$A_{ii}+\lambda_i = \nu$

- KUHN-TUCKER CONDITIONS

$$A_{ii} = (\nu - \lambda_i)^+$$

WHERE WATER LEVEL ν IS CHOSEN SO THAT $\sum A_{ii} = nP$

- FOR CHANNELS IN WHICH NOISE FORMS A STATIONARY STOCHASTIC PROCESS, THE INPUT SIGNAL SHOULD BE CHOSEN TO BE GAUSSIAN PROCESS WITH SPECTRUM THAT IS LARGE AT FREQUENCES WHERE THE NOISE SPECTRUM IS SMALL.

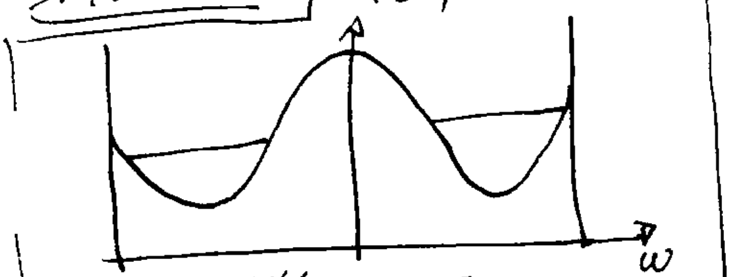


FIG. 9.5 WATER FILLING IN SPECTRAL DOMAIN

THE CAPACITY OF AN ADDITIVE GAUSSIAN NOISE CHANNEL WITH POWER SPECTRUM $N(f)$ IS:

$$C = \int_{-\pi}^{\pi} \frac{1}{2} \log\left(1 + \frac{(\nu - N(f))^+}{N(f)}\right) df$$

WHERE ν IS CHOSEN SO THAT: $\int (\nu - N(f))^+ df = P$

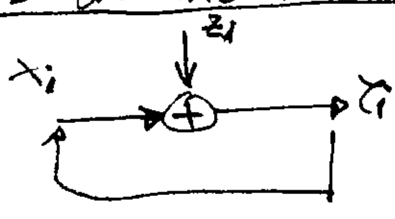
9.6 GAUSSIAN CHANNELS WITH FEEDBACKS

FOR CHANNELS WITH MEMORY, WHERE THE NOISE IS CORRELATED FROM TIME INSTANT TO TIME INSTANT, FEEDBACK DOES INCREASE CAPACITY.

$H(f(x), x) = \underline{H(x)} + \underbrace{H(f(x)|x)}_{0} = H(x) = \underline{H(f(x))} + H(x|f(x))$

$$\boxed{H(x) \ge H(f(x))}$$

IN THIS SECTION WE DESCRIBE AN EXPRESSION FOR THE CAPACITY OF THE CHANNEL WITH FEEDBACK IN TERMS OF THE COVARIANCE MATRIX OF THE NOISE Z. WE PROVE A CONVERSE FOR THIS EXPRESSION FOR CAPACITY. WE THEN DERIVE A SIMPLE BOUND ON THE INCREASE IN CAPACITY DUE TO FEEDBACK.



THE OUTPUT OF THE CHANNEL Y_i IS:

$$Y_i = X_i + Z_i \qquad Z_i \sim N(0, K_Z^{(n)})$$

THE FEEDBACK ALLOWS THE INPUT OF THE CHANNEL TO DEPEND ON THE PAST VALUES OF THE OUTPUT.

[redacted]

A $(2^{nR}, n)$ code for Gaussian channel with feedback consist of a sequence of mappings $x_i(W, Y^{i-1})$ where $W \in \{1, 2, \ldots 2^{nR}\}$ is the input message and Y^{i-1} is sequence of past values of the output. Thus $x(W, \cdot)$ is a code function rather than a codeword. In addition:

$$E\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2(w, Y^{i-1})\right] \le P \qquad w \in \{1, 2, \ldots 2^{nR}\}$$

Expectation is over all possible noise sequences. Because of the feedback X^n and Z^n are not independent; X_i depends casually on the past values of Z.

1.) <u>With feedback</u>. The capacity $C_{n,FB}$ in bits per transmition of the time-varying Gaussian channel with feedback is:

$$C_{n,FB} = \max_{\frac{1}{n} tr(K_X^{(n)}) \le P} \frac{1}{2n} ld \frac{|K_{X+Z}^{(n)}|}{|K_Z^{(n)}|} \qquad (1)$$

where maximization is taken over all X^n of the form

$$X_i = \sum_{j=1}^{i-1} b_{ij} Z_j + V_i \qquad i = 1, 2, \ldots, n \qquad (2)$$

and V_n is independent of Z_n.
Recasting (1) & (2) using $X = BZ + V$ and $Y = X + Z$ we can write:

$$C_{n,FB} = \max \frac{1}{2n} ld \frac{|(B+I) K_Z^{(n)} (B+I)^t + K_V|}{|K_Z^{(n)}|} \qquad (3)$$

where maximum is taken over all nonnegative definite K_V and strictly triangular B such that:

$$tr(B K_Z^{(n)} B^t + K_V) \le nP \qquad (4)$$

. Note that $B = \emptyset$ if feedback is not allowed.

(2) <u>Without feedback</u>. The capacity C_n of the time-varying Gaussian channel without feedback is given by:

$$C_n = \max_{\frac{1}{n} tr(K_X^{(n)}) \le P} \frac{1}{2n} ld \frac{|K_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|} \qquad (5)$$

This reduces to water-filling on the eigenvalues: $\{\lambda_i^{(n)}\}$ of $K_Z^{(n)}$:

$$C_n = \frac{1}{2n} \sum_{i=1}^{n} ld \left(1 + \frac{(\lambda - \lambda_i^{(n)})^+}{\lambda_i^{(n)}}\right) \qquad (6)$$

WHERE $(y)^+ = \max\{y, 0\}$ AND WHERE λ IS CHOSEN:

$$\sum_{i=1}^{n} \left(\lambda - \lambda_i^{(n)}\right)^+ = nP$$

- WE NOW PROVE AN UPPER BOUND FOR THE CAPACITY OF THE GAUSSIAN CHANNEL WITH FEEDBACK.

THEOREM 9.6.1 FOR A GAUSSIAN CHANNEL WITH FEEDBACK, THE RATE R_n FOR ANY SEQUENCE $(2^{nR_n}, n)$ CODES WITH $P_e^{(n)} \to 0$ SATISFIES:

$$R_n \leq C_{n,FB} + \varepsilon_n$$

WITH $\varepsilon_n \to 0$ AS $n \to \infty$, WHERE $C_{n,FB}$ IS (1).

PROOF: LET W BE UNIFORM OVER 2^{nR}, AND THEREFORE THE PROBABILITY OF ERROR $P_e^{(n)}$ IS BOUNDED BY FANO'S INEQUALITY

$$H(W|\hat{W}) \leq 1 + nR_n P_e^{(n)} = n\varepsilon_n$$

WHERE $\varepsilon_n \to 0$ AS $P_e^{(n)} \to 0$

WE CAN THEN BOUND THE RATE AS FOLLOWS:

$$P(x_1 x_2 \dots x_n) = P(x_1) \cdot P(x_2|x_1) \cdot P(x_3|x_1 x_2) \cdots P(x_n|x_1^{n-1})$$

$$nR_n = H(W) = I(W;\hat{W}) + H(W|\hat{W}) \leq I(W;\hat{W}) + n\varepsilon_n \leq$$

$$\leq I(W;Y^n) + n\varepsilon_n = \sum I(W;Y_i|Y_1^{i-1}) + n\varepsilon_n =$$

$$= \sum h(Y_i|Y_1^{i-1}) - h(Y_i|Y_1^{i-1}, W) + n\varepsilon_n =$$

$h(X+Z|X) = h(Z|X)$

$$= \sum h(Y_i|Y_1^{i-1}) - h(Z_i|Y_1^{i-1}, W, X_i, X^{i-1}, Z^{i-1}) + n\varepsilon_n$$

$X_i = f(Y_1^{i-1}, W)$ $\qquad Z^{i-1} = Y^{i-1} - X^{i-1}$

$$= \sum \left[h(Y_i|Y_1^{i-1}) - h(Z_i|Z^{i-1})\right] + n\varepsilon_n = h(Y) - h(Z) + n\varepsilon_n$$

$$R_n \leq \frac{1}{n}\left[h(Y) - h(Z)\right] + \varepsilon_n \leq \frac{1}{2n}\left[\log |K_Y|(2\pi e)^n - \log (2\pi e)^n |K_Z|\right]$$

$$R_n \leq \frac{1}{2n} \log \frac{|K_Y|}{|K_Z|} + \varepsilon_n \qquad R_n \leq C_{n,FB} + \varepsilon_n$$

WE HAVE PROVED AN UPPER BOUND ON THE CAPACITY OF THE GAUSSIAN CHANNEL WITH FEEDBACK IN TERMS OF COVARIANCE MATRIX: $K_{X+Z}^{(n)}$. WE NOW DERIVE BOUNDS ON THE CAPACITY WITH FEEDBACK IN TERMS OF $K_X^{(n)}$ AND $K_Z^{(n)}$, WHICH WILL THEN BE USED TO DERIVE BOUNDS IN TERMS OF THE CAPACITY WITHOUT FEEDBACK.

LEMMA 9.6.1 LET X AND Z BE n-DIMENSIONAL RANDOM VECTORS. THEN: (★)

$$K_{X+Z} + K_{X-Z} = 2K_X + 2K_Z$$

PROOF: $K_{X+Z} = E[(X+Z)\cdot(X+Z)^T] = E[X X^T + X Z^T + Z X^T + Z Z^T]$

$= K_X + K_{XZ} + K_{ZX} + K_Z$

$K_{X-Z} = E[(X-Z)(X-Z)^T] = E[X X^T] - E[X Z^T] - E[Z X^T] + E[Z Z^T]$

$= K_X - K_{XZ} - K_{ZX} + K_Z$

$K_{X+Z} + K_{X-Z} = K_X + \cancel{K_{XZ}} + \cancel{K_{ZX}} + K_Z + K_X - \cancel{K_{XZ}} - \cancel{K_{ZX}} + K_Z$

$= 2K_X + 2K_Z$

LEMMA 9.6.2 FOR TWO $n \times n$ NONNEGATIVE DEFINITE MATRICES A & B, IF $A-B$ IS NONNEGATIVE DEFINITE, THEN: $|A| \geq |B|$

PROOF: $C = A - B$ B, C — NONNEGATIVE DEFINITE HENCE WE CAN CONSIDER THEM AS COVARIANCE MATRICES.

$X_1 \sim \mathcal{N}(0, B)$; $X_2 \sim \mathcal{N}(0, C)$. LET $Y = X_1 + X_2$

$h(Y) \geq h(Y|X_2) = h(X_1|X_2) = h(X_1)$ SINCE X_1 & X_2 ARE INDEPENDENT

$\frac{1}{2}\log(2\pi e)^n |K_Y| \geq \frac{1}{2}\log(2\pi e)^n |K_{X_1}|$ $\log(2\pi e)^n|A| \geq \log(2\pi e)^n|B|$

$\Rightarrow |A| \geq |B|$

LEMMA 9.6.3 FOR TWO n-DIMENSIONAL VECTORS X & Z,

$$|K_{X+Z}| \leq 2^n |K_X + K_Z|$$

PROOF: $\underbrace{2(K_X + K_Z)}_{A} - \underbrace{K_{X+Z}}_{B} = K_{X-Z} \succeq 0$ (★) $\Big\}$ $A - B \succeq 0$

WHERE "$\succeq$" MEANS $A \succeq 0$ A IS NONNEGATIVE DEFINITE. HENCE UTILIZING LEMMA 9.6.2, WE HAVE:

$$|K_{X+Z}| \leq |2(K_X + K_Z)| = 2^n |K_X + K_Z|$$

FROM PROPERTIES OF DETERMINANTS

$|c \cdot A| = |c \cdot I \cdot A| = |cI| \cdot |A| = c^n \cdot |A|$ $\Big\}$ IF A IS $n \times n$ MATRIX

LEMMA 9.6.4 FOR A, B NONNEGATIVE DEFINITE MATRICES AND $0 \leq \lambda \leq 1$

$$|\lambda A + (1-\lambda) B| \geq |A|^{\lambda} |B|^{1-\lambda}$$

[redacted]

PROOF: LET $X \sim N_n(0,A)$ AND $Y \sim N_n(0,B)$. LET Z BE THE MIXTURE RANDOM VECTOR:

$$Z = \begin{cases} X & \text{IF } \theta = 1 \\ Y & \text{IF } \theta = 2 \end{cases} \qquad \text{WHERE } \theta = \begin{cases} 1 & \text{WITH } P=\lambda \\ 2 & \text{WITH } P=1-\lambda \end{cases}$$

LET X, Y AND θ BE INDEPENDENT

$$K_Z = \lambda A + (1-\lambda) B$$

$$H(Z|\theta) = P(\theta=1)\cdot H(Z|\theta=1) + P(\theta=2)\, H(Z|\theta=2) =$$

$$= \lambda \cdot H(X) + (1-\lambda)\, H(Y)$$

$$H(Z) \qquad ld(2\pi e)^n A \qquad (1-\lambda)\; ld(2\pi e)^n B \qquad H(Z|\theta)$$

$$ld(2\pi e)^n |K_Z| \geq ld(2\pi e)^{n\lambda}|A|^\lambda + ld(2\pi e)^{n(1-\lambda)}|B|^{1-\lambda}$$

$$= ld(2\pi e)^{n\lambda + n - n\lambda} A^\lambda \cdot B^{1-\lambda} = ld(2\pi e)^n |A|^\lambda \cdot |B|^{1-\lambda}$$

$$|K_Z| \geq |A|^\lambda \cdot |B|^{1-\lambda} \qquad (\$)$$

$$H(Z,\theta) = H(\theta) + H(Z|\theta) = H(Z) + H(\theta|Z)$$

$$H(Z) = H(\lambda) + \lambda H(X) + (1-\lambda) H(Y) \qquad H(Z) \geq H(Z|\theta)$$

$$E[Z] = E_\theta[E[Z|\theta]] = \sum_\theta P(\theta)\, E[Z|\theta] = \lambda \cdot E[Z|\theta=1] + (1-\lambda)\, E[Z|\theta=2] = \lambda \cdot E[X] + (1-\lambda)\, E[Y]$$

АНАЛОГИЧНО ЗА ВТОРИЯ МОМЕНТ:

$$K_Z = E[Z^2] = E_\theta[E[Z^2|\theta]] = \sum_\theta P(\theta)\, E[Z^2|\theta] = \lambda E[X^2] + (1-\lambda)\, E[Y^2] = \lambda K_X + (1-\lambda) K_Y$$

MMV

$$K_Z = \lambda K_X + (1-\lambda) K_Y \quad (\$) \Rightarrow |K_Z| \geq |A|^\lambda |B|^{1-\lambda} \Rightarrow$$

$$|\lambda K_X + (1-\lambda) K_Y| \geq |A|^\lambda |B|^{1-\lambda} \qquad |\lambda A + (1-\lambda) B| \geq |A|^\lambda |B|^{1-\lambda}$$

- CORRECT SOLUTION: $\frac{1}{2} ld(2\pi e)^n |\lambda A + (1-\lambda) B| \geq h(Z) \geq h(Z|\theta) =$

$$= \lambda h(X) + (1-\lambda) h(Y) = \lambda \tfrac{1}{2} ld(2\pi e)^n |A| + (1-\lambda) \tfrac{1}{2} ld(2\pi e)^n |B| =$$

$$= ld(2\pi e)^{n\lambda} |A|^\lambda (2\pi e)^{n(1-\lambda)} |B|^{1-\lambda} = \tfrac{1}{2} ld(2\pi e)^n |A|^\lambda |B|^{1-\lambda}$$

$$\Rightarrow |\lambda A + (1-\lambda) B| \geq |A|^\lambda |B|^{1-\lambda}$$

DEFINITION: WE SAY THAT A RANDOM VECTOR X^n IS CAUSALLY RELATED TO Z^n IF:

$$f(x^n, z^n) = f(z^n)\prod_{i=1}^{n} f(x_i \mid x^{i-1}, z^{i-1}) \qquad (A)$$

$$f(x^2, z^2) = f(z_1^2)\cdot f(x_1)\cdot f(x_2 \mid x_1, z_1) = f(z_1 z_2) f(x_1) f(x_2 \mid z_1 x_1)$$

$$f(x_1 x_2 \mid z_1)\ (?) \qquad f(x_1 x_2 \mid z_1) = f(x_1 \mid z_1) f(x_2 \mid x_1 z_1)$$

$$f(x^3, z^3) = f(z_1^3)\, f(x_1)\cdot f(x_2 \mid x_1, z_1)\cdot f(x_3 \mid x_1^2\, z_1^2)$$

LEMMA 9.6.5 IF X^n AND Z^n ARE CAUSALLY RELATED THEN:

$$h(X^n - Z^n) \geq h(Z^n) \quad \text{AND} \quad |K_{X-Z}| \geq |K_Z|$$

WHERE K_{X-Z} AND K_Z ARE COVARIANCE MATRICES OF $X^n - Z^n$ AND Z^n RESPECTIVELY.

PROOF: $h(X^n - Z^n) \overset{(a)}{=} \sum_{i=1}^{n} h(X_i - Z_i \mid X^{i-1}, Z^{i-1}) \geq$

$\overset{(b)}{\geq} \sum_{i=1}^{n} h(X_i - Z_i \mid X^{i-1}, Z^{i-1}, X_i) \overset{(c)}{=} \sum_{i=1}^{n} h(Z_i \mid X^{i-1}, Z^{i-1}, X_i) \overset{(d)}{=}$ (#★)

$\overset{(d)}{=} \sum_{i=1}^{n} h(Z_i \mid Z^{i-1}) \overset{(e)}{=} h(Z^n)$

$$f(x^{i-1} z^{i-1}) = f(z^{i-1})\cdot \prod_{j=1}^{i-1} f(x_j \mid x^{j-1} z^{j-1})$$

$$f(x^{i-1}, x_i, z^{i-1}) = f(x^i, z^{i-1}) = f(z^{i-1})\cdot \prod_{j=1}^{i} f(x_j \mid x^{j-1} z^{j-1})$$

(#★) ⟶ KAUZALNA POVRZANOST PODRAZBIRA X ZAVISI OD Z NO NE Z OD X (TAKA GO TOLKUVAM (A))

$$\frac{1}{2}\log(2\pi e)|K_{X-Z}| \geq h(\tilde{Z}^n) = \frac{1}{2}\log(2\pi e)|K_Z| \Rightarrow$$

$$= h(\tilde{X}^n - \tilde{Z}^n)$$

$$|K_{X-Z}| \geq |K_Z|$$

$\tilde{X}^n, \tilde{Z}^n$ } MULTIVARIATE NORMAL (CAUSALLY RELATED)

WE ARE NOW IN POSITION TO PROVE THAT FEEDBACK INCREASES THE CAPACITY OF NONWHITE GAUSSIAN ADDITIVE NOISE CHANNEL BY AT MOST HALF A BIT.

[redacted]

THEOREM 9.6.2 $C_{n,FB} \le \max_{tr(K_x)\le nP} \frac{1}{2n} ld \frac{|K_z|}{|K_z|} \le$

$$\le \max_{tr(K_x)\le nP} \frac{1}{2n} ld \frac{2^n |K_x+K_z|}{|K_z|} = \max_{tr(K_x)\le nP}\left\{\frac{ld\, 2^n}{2n} + \frac{1}{2n} ld \frac{|K_x+K_z|}{|K_z|}\right\}$$

$$= \max_{tr(K_x)\le nP}\left[\frac{1}{2} + \frac{1}{2n} ld \frac{|K_x+K_z|}{|K_z|}\right] = \underbrace{\max_{tr(K_x)\le nP} \frac{1}{2n} ld \frac{|K_x+K_z|}{|K_z|}}_{\text{N16n.45.(5)}} + \frac{1}{2}$$

$$= C_n + \frac{1}{2} \qquad \boxed{C_{n,FB} \le C_n + \frac{1}{2}}$$

- WE NOW PROVE PINSKER'S STATEMENT THAT FEEDBACK CAN AT MOST DOUBLE THE CAPACITY OF COLORED NOISE CHANNEL.

THEOREM 9.6.3 $\boxed{C_{n,FB} \le 2C_n}$

IT IS ENOUGH TO SHOW THAT:

$$\frac{1}{2}\frac{1}{2n} ld \frac{|K_{x+z}|}{|K_z|} \le \frac{1}{2n} ld \frac{|K_x+K_z|}{|K_z|}$$

$$\Rightarrow \frac{1}{2} C_{n,FB} \le C \quad \Rightarrow \quad \boxed{C_{n,FB} \le 2C}$$

$$\frac{1}{2n} ld \frac{|K_x+K_z|}{|K_z|} \overset{(a)}{=} \frac{1}{2n} ld \frac{|\frac{1}{2}K_{x+z} + \frac{1}{2}K_{x-z}|}{|K_z|} \overset{(b)}{\ge} \frac{1}{2n} ld\left(\frac{|K_{x+z}|^{\frac{1}{2}}|K_{x-z}|^{1/2}}{|K_z|}\right) \overset{(c)}{\ge} \frac{1}{2n} ld\left(\frac{|K_{x+z}|^{1/2}|K_z|^{1/2}}{|K_z|}\right) = \frac{1}{2}\frac{1}{2n} ld \frac{|K_{x+z}|}{|K_z|}$$

$$\frac{1}{2n} ld \frac{|K_x+K_z|}{|K_z|} \ge \frac{1}{2}\frac{1}{2n} ld \frac{|K_{x+z}|}{|K_z|}$$

PROVED

THUS, WE HAVE SHOWN THAT GAUSSIAN CHANNEL CAPACITY IS NOT INCREASED BY MORE THAN A HALF BIT OR BY MORE THAN A FACTOR 2 WHEN WE HAVE FEEDBACK. FEEDBACK HELPS BUT NOT MUCH.

SUMMARY

- MAXIMUM ENTROPY $\max_{E(X^2)=\alpha} h(X) = \frac{1}{2} ld(2\pi e)\alpha$
- GAUSSIAN CHANNEL: $Y_i = X_i + Z_i$, $Z_i \sim N(0,N)$; POWER CONSTRAINT: $\frac{1}{n}\sum_{i=1}^{n} x_i^2 \le P$; $C = \frac{1}{2} ld\left(1+\frac{P}{N}\right)$ bits per transmition

[redacted]

446.231.564

■ BANDLIMITED ADDITIVE WHITE GAUSSIAN NOISE CHANNEL

BANDWIDTH W; TWO-SIDED POWER SPECTRAL DENSITY $N_0/2$; SIGNAL POWER P; AND:

$$C = W \, ld \left(1 + \frac{P}{N_0 W}\right) \quad \text{BITS PER SECOND}$$

■ WATER-FILLING (k PARALLEL GAUSSIAN CHANNELS).

$Y_j = X_j + Z_j$, $j = 1, 2, \ldots, k$; $Z_j \sim N(0, N_j)$; $\sum_{j=1}^{k} x_j^2 \le P$

$$C = \sum_{i=1}^{k} \frac{1}{2} ld \left(1 + \frac{(\nu - N_i)^+}{N_i}\right)$$

WHERE ν IS CHOSEN SO THAT: $\sum (\nu - N_i)^+ = P$.

□ ADDITIVE NONWHITE GAUSSIAN NOISE CHANNEL.

$Y_i = X_i + Z_i$; $Z^n \sim N(0, K_Z)$; AND

$$C = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} ld \left(1 + \frac{(\nu - \lambda_i)^+}{\lambda_i}\right)$$

WHERE $\lambda_1, \lambda_2, \ldots \lambda_n$ ARE THE EIGENVALUES OF K_Z AND ν IS CHOSEN SO THAT $\sum_i (\nu - \lambda_i)^+ = nP$

■ CAPACITY WITHOUT FEEDBACK

$$C_n = \max_{tr(K_X) \le nP} \frac{1}{2n} ld \frac{|K_X + K_Z|}{|K_Z|}$$

■ CAPACITY WITH FEEDBACK

$$C_{n,FB} = \max_{tr(K_X) \le nP} \frac{1}{2n} ld \frac{|K_{X+Z}|}{|K_Z|}.$$

■ FEEDBACKS BOUNDS:

$$C_{n,FB} \le C_n + \frac{1}{2} \qquad C_{n,FB} \le 2 C_n$$

PROBLEMS

PROBLEM 9.1 CHANNEL WITH TWO INDEPENDENT LOOKS AT Y.

LET Y_1 AND Y_2 BE CONDITIONALLY INDEPENDENT AND CONDITIONALLY IDENTICALLY DISTRIBUTED GIVEN X.

(a) SHOW: $I(X; Y_1 Y_2) = 2 I(X; Y_1) - I(Y_1; Y_2)$

(c) CONCLUDE THAT CAPACITY OF THE CHANNEL:

(Δ) $X \rightarrow [\ ] \rightarrow (Z_1; Z_2)$

IS LESS THAN TWICE THE CAPACITY OF THE CHANNEL

(□) $X \rightarrow [\ ] \rightarrow Z_1$ $\boxed{I(Z_1;Z_2)}$

(a) $I(X;Z_1Z_2) = I(X;Z_1) + I(X;Z_2|Z_1)$

$I(X;Z_2|Z_1) = H(Z_2|Z_1) - H(Z_2|Z_1X) = H(Z_2|Z_1) - H(Z_2|X) =$

$= \underbrace{-H(Z_2) + H(Z_2|Z_1)}_{-I(Z_1;Z_2)} + \underbrace{H(Z_2) - H(Z_2|X)}_{I(X;Z_2)} = I(X;Z_2) - I(Z_1;Z_2)$

$\boxed{H(Z_1|X) = H(Z_2|X)}$ CONDITIONALLY IDENT. DIST.

$I(X;Z_1) = I(X;Z_2)$ Z_1 & Z_2 ARE IDENTICALLY DISTRIBUTED !!

(b) $I(X;Z_1Z_2) = I(X;Z_1) + I(X;Z_1) - I(Z_1;Z_2) = \underline{2I(X;Z_1) - I(Z_1;Z_2)}$

(□) $C_□ = \max_{p(x)} I(X;Z_1)$

(Δ) $C_Δ = \max_{p(x)} I(X;Z_1Z_2) = \max_{p(x)} \left[2I(X;Z_1) - I(Z_1;Z_2)\right]$

$= \max_{p(x)} \left[2I(X;Z_1)\right] - I(Z_1;Z_2) = 2 \underbrace{\max_{p(x)} I(X;Z_1)}_{C_□} - I(Z_1;Z_2)$

$\boxed{C_Δ = 2C_□ - I(Z_1;Z_2)}$ $2C_□ = C_Δ + I(Z_1;Z_2)$

$C_□ = \frac{1}{2}\left[C_Δ + I(Z_1;Z_2)\right]$ $\boxed{2C_□ \geq C_Δ}$

→ PROVED $C_Δ$ IS LESS THEN TWICE THEN CAPACITY OF $C_□$

• REVISITED (CONDITIONALLY INDEPENDENT Z_1 & Z_2)

$I(X;Z_1Z_2) = H(Z_1Z_2) - H(Z_1Z_2|X) = H(Z_1) + H(Z_2|Z_1) -$

$- H(Z_1|X) - H(Z_2|XZ_1) = \underline{H(Z_1)} + \underline{H(Z_2)} - H(Z_1|X) - \underline{H(Z_2|X)} =$

$= I(X;Z_1) + I(X;Z_2)$ (AND BECAUSE $H(Z_2|Z_1) = H(Z_2)$)

$=$| VO SPLOŠNEM |$= \underbrace{H(Z_1) - H(Z_1|X)}_{I(X;Z_1)} + H(Z_2|Z_1) - H(Z_2|X) =$

$= I(X;Z_1) + \underbrace{H(Z_2) - H(Z_2|X)}_{I(X;Z_2)} \underbrace{- H(Z_2) + H(Z_2|Z_1)}_{-I(Z_1;Z_2)}$

[redacted]

$$I(X;Y_1Y_2) = I(X;Y_1) + I(X;Y_2) - I(Y_1;Y_2)$$

$$\left.\begin{aligned} I(X;Y_1) &= H(Y_1) - H(Y_1|X) \\ I(X;Y_2) &= H(Y_2) - H(Y_2|X) \end{aligned}\right\}$$

$H(Y_1|X) = H(Y_2|X)$ CONDITIONALLY IDENTICAL DISTRIBUTED GIVEN X

$H(Y_1) = H(Y_2) \Rightarrow$ SINCE Y_1 & Y_2 ARE IDENTICALLY DISTRIBUTED

$$I(X;Y_1Y_2) = 2I(X;Y_1) - I(Y_1;Y_2)$$

PROBLEM 9.2 TWO LOOK GAUSSIAN CHANNEL

CONSIDER THE ORDINARY GAUSSIAN CHANNEL WITH TWO CORRELATED LOOKS AT X, THAT IS: $Y = (Y_1, Y_2)$ WHERE:

$$Y_1 = X + Z_1 \qquad Y_2 = X + Z_2$$

WITH POWER CONSTRAINT P ON X, AND $(Z_1, Z_2) \sim N_2(0, K)$

$$K = \begin{bmatrix} N & N\rho \\ N\rho & N \end{bmatrix}$$

FIND CAPACITY FOR:

(a) $\rho = 1$ (b) $\rho = 0$ (c) $\rho = -1$

$Y = X + Z_\theta$

$$\theta = \begin{cases} 1 & Z = Z_1 \\ 2 & Z = Z_2 \end{cases}$$

$$I(X;Y_1Y_2) = 2I(X;Y_1) - I(Y_1;Y_2)$$

$$I(Y_1;Y_2) = H(Y_1) - H(Y_2|Y_1)$$

IF $Y = \begin{cases} Y_1 & \theta = 1 \\ Y_2 & \theta = 2 \end{cases}$

$$H(Y,\theta) = H(Y) + \underbrace{H(\theta|Y)}_{0} = H(\theta) + H(Y|\theta) = H(\theta) + P(\theta=1)\cdot H(Y_1) + P(\theta=2)\cdot H(Y_2)$$

(b) $K = \begin{bmatrix} N & 0 \\ 0 & N \end{bmatrix}$ $\quad p(Z_1, Z_2) = \dfrac{1}{(2\pi)^{n}|K|^{1/2}} e^{-\frac{z\cdot K^{-1}\cdot z^T}{2}}$

$$p(Z_1, Z_2) = p(Z_1)\cdot p(Z_2) = \frac{1}{\sqrt{2\pi N_1}} e^{-\frac{z_1^2}{2N_1}} \cdot \frac{1}{\sqrt{2\pi N_2}} e^{-\frac{z_2^2}{2N_2}}$$

[redacted]

$$I(X;Y_1Y_2) = 2I(X;Y_1) - I(Y_1;Y_2)$$

$$C = \max_{p(x)} \left[I(X;Y_1) \right] - \underline{I(Y_1;Y_2)}$$

$$I(X;Y_1) = H(Y_1) - H(Y_1|X) = H(Y_1) - H(X+Z_1|X)$$

$$= H(Y_1) - H(Z_1) \leq \frac{1}{2}\mathrm{ld}\, 2\pi e\,(P+N) - \frac{1}{2}\mathrm{ld}(2\pi e)N =$$

$$= \frac{1}{2}\,\mathrm{ld}\left(1+\frac{N}{P}\right)$$

$$C = \frac{1}{2}\,\mathrm{ld}\left(1+\frac{N}{P}\right) - I(Y_1;Y_2)$$

$$I(Y_1;Y_2) = H(Y_1) - H(Y_1|Y_2) = H(Y_1) - H(X+Z_1|X+Z_2)$$

$$= H(Y_1) - \underline{H(Z_1|Z_2)}$$

$$H(X+Z_1|X+Z_2) = \sum_{xz_1z_2} p(x\, z_1\, z_2)\cdot \mathrm{ld}\,\frac{1}{p(x+z_1|x+z_2)} =$$

X indep. Z_1, Z_2

$$= \sum_x p(x) \sum_{z_1z_2} p(z_1z_2)\,\mathrm{ld}\,\frac{1}{p(x=x, z_1|z_2\, x=x)} = H(Z_1|Z_2)$$

ALTERNATIVELY.

$$H(Y_1|Y_2) = \sum_x p(x)\, H(Y_1|Y_2 = x+Z_2) =$$

$$= \sum_x p(x)\cdot H(x+Z_1|x+Z_2) = h(Z_1|Z_2)$$

(b) $\begin{bmatrix} N & 0 \\ 0 & N \end{bmatrix}$ Z_1 INDEPENDENT Z_2

$$h(Z_1|Z_2) = h(Z_1)$$

$$I(Y_1;Y_2) = h(Y_1) - h(Z_1) =$$

$$= \frac{1}{2}\mathrm{ld}\left(1+\frac{N}{P}\right) - \left[h(Y_1) - h(Z_1)\right]$$

- A DA GO TESTIRAM KAKO PORACUNAM GAUSOV KANAL

$$I(X;Y_1Y_2) = h(Y_1Y_2) - h(Y_1Y_2|X) = h(Y_1Y_2) -$$

$$- \sum_x p(x)\, H(x+Z_1, x+Z_2|X=x) = h(Y_1Y_2) - h(Z_1Z_2)$$

$$I(X;Y_1Y_2) \le \sum_{i=1}^{2} h(Y_i) - \sum_{i=1}^{2} h(Z_i) \le \sum_{i=1}^{2}\left\{\frac{1}{2}\operatorname{ld}(2\pi e)(P+N_i) - \frac{1}{2}\operatorname{ld}(2\pi e)N_i\right\} = \sum_{i=1}^{2}\frac{1}{2}\operatorname{ld}\left(\frac{P+N_i}{N_i}\right) = \frac{1}{2}\sum_{i=1}^{2}\operatorname{ld}\left(1+\frac{P}{N_i}\right)$$

$$J(P_1,P_2) = \sum_{i=1}^{2}\frac{1}{2}\operatorname{ld}\left(1+\frac{P_i}{N_i}\right) + \lambda\left(\sum P_i\right)$$

$$\frac{\partial J}{\partial P_i} = \frac{1}{2}\cdot\frac{1}{1+\frac{P_i}{N_i}}\cdot\frac{1}{N_i} + \lambda = 0 \qquad \frac{1}{2}\frac{1}{P_i+N_i} + \lambda = 0$$

$$\frac{1}{2(P_i+N_i)} = -\lambda \qquad \frac{1}{P_i+N_i} = -2\lambda \qquad P_i+N_i = -\frac{1}{2\lambda}$$

$$P_i = -N_i - \frac{1}{2\lambda} = \underline{\nu - N_i}$$

(b) $P_1 = P_2 = P \qquad N_1 = N_2 = N \qquad K_Z = \begin{bmatrix} N & 0 \\ 0 & N \end{bmatrix}$

ZNAČI ZA (b)

$$\boxed{C = \frac{1}{2}\sum_{i=1}^{2}\operatorname{ld}\left(1+\frac{\nu-N}{N}\right)} \qquad \sum(\nu - N_i) = P \quad (k=1)$$

$$p(z_1|z_2) = \frac{p(z_1 z_2)}{p(z_2)} = \frac{\frac{1}{2\pi|K|^{1/2}}e^{-\frac{z^T K^{-1} z}{2}}}{\frac{1}{\sqrt{2\pi N}}e^{-\frac{z_1^2}{2N}}}$$

↓ MY SOLUTION BASED ON 8.11

(a) $E[Y_1Y_2] = E[(X+Z_i)(X+Z_j)] = E[X^2] + \underline{E[Z_iZ_j]} =$

$$= \begin{cases} P & i\ne j \\ P+N & i=j \end{cases} \qquad K_Y = \begin{bmatrix} P+N & P \\ P & P+N \end{bmatrix} \qquad K_Z = \begin{bmatrix} N & 0 \\ 0 & N \end{bmatrix}$$

$$I(X;Y_1Y_2) = H(Y_1Y_2) - H(Y_1Y_2|X) = \underline{h(Y_1Y_2) - h(Z_1,Z_2)}$$

$$\le \frac{1}{2}\operatorname{ld}(2\pi e)^2|K_Y| - \frac{1}{2}\operatorname{ld}(2\pi e)^2|K_Z| = \frac{1}{2}\operatorname{ld}\frac{|K_Y|}{|K_Z|}$$

↑ GAUSSIAN DISTRIBUTION MAXIMIZES ENTROPY

$$C = \frac{1}{2}\operatorname{ld}\frac{(P+N)^2 - P^2}{N^2} = \frac{1}{2}\operatorname{ld}\frac{P^2+2PN+N^2-P^2}{N^2} = \frac{1}{2}\operatorname{ld}\frac{N(N+2P)}{N^2} = \frac{1}{2}\operatorname{ld}\left(1+\frac{2P}{N}\right)$$

(a) $E[\zeta_1\zeta_2] = E[x^2] + E[Z_i Z_j] = \begin{cases} P+N & i=j \\ P+N & i\neq j \end{cases}$

$K_Z = \begin{bmatrix} N & N \\ N & N \end{bmatrix} \qquad K_\zeta = \begin{bmatrix} P+N & P+N \\ P+N & P+N \end{bmatrix}$

$I(x;\zeta_1\zeta_2) = h(\zeta_1\zeta_2) - h(Z_1Z_2) = \frac{1}{2}\,ld\,(2\pi e)^2|K_\zeta| - \frac{1}{2}\,ld\,(2\pi e)^2|K_Z|$

$C = \frac{1}{2}\,ld\,\frac{|K_\zeta|}{|K_Z|} = \frac{1}{2}\,ld\,\frac{(P+N)^2-(P+N)^2}{N^2-N^2} = 0$

(c) $K_Z = \begin{bmatrix} N & -N \\ -N & N \end{bmatrix} \qquad K_\zeta = \begin{bmatrix} P+N & P-N \\ P-N & P+N \end{bmatrix}$

$C = \frac{1}{2}\,ld\,\frac{(P+N)^2-(P-N)^2}{N^2-N^2} = \frac{1}{2}\,ld\,\frac{P^2+2PN+N^2-P^2+2PN-N^2}{0}$

$= \frac{1}{2}\,ld\,\frac{4PN}{0} = \infty$ — SOLIDRADA SE SVEDU- NA 1/4 PLESMEDRA NA KORELACIONA MATRICA ZA N = ZA DA GO KALKULIRAM

• DA GI PROVERAM SO 3 HW8c.PDF (SOGLASNO) OK DOBIO LIMESOT

$K_\zeta = \begin{bmatrix} P & P \\ P & P \end{bmatrix} + K_Z \;;\; K_Z = \begin{bmatrix} N & \rho N \\ \rho N & N \end{bmatrix} \qquad |K_Z| = N^2 - \rho^2 N^2 = N^2(1-\rho^2)$

$K_\zeta = \begin{bmatrix} P+N \;;\; P+\rho N \\ P+\rho N \;;\; P+N \end{bmatrix} \qquad |K_\zeta| = (P+N)^2 - (P+\rho N)^2$

$|K_\zeta| = P^2 + 2NP + N^2 - P^2 - 2\rho PN - \rho^2 N^2 = N^2(1-\rho^2) + 2NP(1-\rho)$

$C = \frac{1}{2}\,ld\,\frac{N^2(1-\rho^2) + 2NP(1-\rho)}{N^2(1-\rho^2)} = \frac{1}{2}\,ld\left(1 + \frac{2P}{N(1+\rho)}\right)$

(a) $\boxed{\rho = 1}$ $\quad C = \frac{1}{2}\,ld\left(1 + \frac{2P}{N(1+1)}\right) = \frac{1}{2}\,ld\left(1 + \frac{2P}{2N}\right)$

$C = \frac{1}{2}\,ld\left(1 + \frac{P}{N}\right)$

(b) $\boxed{\rho = 0}$ $\quad C = \frac{1}{2}\,ld\left(1 + \frac{2P}{N}\right)$

(c) $(\rho = -1)$ $\quad C = \infty$

[redacted]

PROBLEM 9.3 OUTPUT POWER CONSTRAINT. CONSIDER AN ADDITIVE GAUSSIAN CHANNEL WITH AN EXPECTED OUTPUT POWER CONSTRAINT P. THUS, $Y = X + Z$, $Z \sim N(0, \sigma^2)$, Z IS INDEPENDENT OF X, AND $E[Y^2] \leq P$. FIND THE CAPACITY.

$$C = \max_{p(x)} I(X;Y) \qquad I(X;Y) = h(Y) - h(Z)$$

$$\frac{1}{n}\sum P_i = P \qquad P_i = y_i^2$$

$$J = h(Y) + \lambda \sum P_i$$

$$J = -\sum_{i=1}^{n} p(i) \operatorname{ld} p(i) + \lambda\, E[Y^2] =$$

$$= -\sum_{i=1}^{n} p(i) \operatorname{ld} p(i) + \lambda \sum_{i=1}^{n} p(i) \cdot \frac{y_i^2}{p_i}$$

$$\frac{dJ}{dP_i} = -\left(\operatorname{ld} p(i) + p(i) \cdot \frac{1}{p(i)} + \lambda y_i^2\right) = 0$$

$$\operatorname{ld} p(i) + \lambda \cdot P_i = 0 \qquad \operatorname{ld} p(i) = -\lambda P_i$$

$$p(i) = 2^{-\lambda P_i} \qquad \sum_{i=1}^{n} p(i) = 1 \qquad \sum_{i=1}^{n} \frac{1}{2^{\lambda P_i}} = 1$$

$$\sum_{i=1}^{n} \frac{1}{2^n} = \sum_{i=0}^{n} 2^n - 1 = \frac{1 - 2^{n+1}}{1-2} - 1 = \frac{1 - 2^{n+1} - 1 + 2}{1-2}$$

$$= \frac{\left(-2^{n+1} + 2\right)}{1-2} = 2\,\frac{(1 - 2^n)}{1-2} = 1 \qquad 2 - 2^{n+1} = 1 - 2$$

$$2 \cdot 2 - 2^{n+1} = 1 \qquad 2\left(2 - 2^n\right) = 1$$

$$2^{\lambda P_i} = 2^{i \cdot \lambda P_i} \qquad \lambda P_i = i \qquad \boxed{P_i = \frac{1}{\lambda}} \; (?)$$

$$I(X;Y) = h(Y) - h(Z) \leq \frac{1}{2} \operatorname{ld} \frac{P_Y}{N}$$

$$J = \frac{1}{2} \operatorname{ld} \frac{P_Y}{N} + \lambda \sum_{i=1}^{n} P_i \qquad E(Y^2) = P_Y$$

$$J = \frac{1}{2} \operatorname{ld} \left[\frac{\sum_{i=1}^{n} P_i \cdot p_i}{N}\right] + \lambda \sum_{i=1}^{n} P_i \; ; \quad \frac{dJ}{dP_i} = 0$$

[redacted]

$$J = \frac{1}{2}\, ld\, \frac{\sum_{i=1}^{n} p_i P_i}{N} + \lambda \sum_{i=1}^{n} P_i$$

$$\frac{dJ}{dP_i} = -\frac{1}{2} ld N + \left[\frac{1}{2} ld \sum_{i=1}^{n} p_i P_i\right]' + \frac{d}{dP_i}\left[\lambda \sum_{i=1}^{n} P_i\right]$$

JENSEN'S INEQUALITY:

$E[f(x)] \geq f(E[x])$ — $f(x) \Rightarrow$ CONVEX

$E[f(x)] \leq f(E[x])$ — $f(x) \rightarrow$ CONCAVE

$$\frac{1}{2} ld \left(\underbrace{\sum p_i P_i}_{E[x]}\right) \geq \frac{1}{2}\sum p_i \, ld P_i$$

$$J \geq -\frac{1}{2} ld N + \frac{1}{2}\sum p_i \, ld P_i + \lambda \sum P_i$$

$$\frac{dJ}{dP_i} = 0 \qquad \frac{p_i}{P_i}\frac{1}{\ln 2} - \lambda \cdot 1 = 0 \qquad \lambda = +\frac{p_i}{P_i}\cdot\frac{1}{\ln 2}$$

$$\boxed{P_i = \frac{p_i}{\lambda}\cdot ld e} \qquad p_i = \frac{P_i \cdot \lambda}{ld e}$$

$$\sum p_i = 1 \qquad \sum_{i=1}^{n} \frac{P_i \lambda}{ld e} = 1 \qquad \frac{\lambda}{ld e}\underbrace{\sum_{i=1}^{n} P_i}_{P} = 1$$

$$\boxed{\lambda = \frac{ld e}{P}} \qquad \boxed{P_i = \frac{p_i}{\frac{ld e}{P}}\cdot ld e = P \cdot p_i}$$

$$P_r = \sum_{i=1}^{n} P_i \cdot p_i = \sum_{i=1}^{n} P p_i^2 = P \sum_{i=1}^{n} p_i^2$$

$$\boxed{C = \frac{1}{2} ld \frac{P_x}{N} = \frac{1}{2} ld \frac{P}{N}} \qquad N = \sigma^2$$

- sols_5_229A spr07.pdf (Berkeley)

IF $P \leq \sigma^2$ IT IS IMPOSSIBLE TO MEET THE POWER CONSTRAINT, SO CAPACITY IN THIS CASE IS 0 TRUE.

$$C = \left(\frac{1}{2} ld \frac{P}{\sigma^2}\right)^+ \qquad (M)^+ = \max(M, 0)$$

THIS BOUND CAN BE ACHIEVED WHEN $p(x)$ IS GAUSSIAN

WITH: $\sigma_x^2 = P - N = P - \sigma^2 \qquad \sigma^2 = \sigma_z^2$

$\sigma_y^2 = P = \sigma_x^2 + \sigma_z^2 = P - \sigma^2 + \sigma^2 = P$

PROBLEM 9.9 Vector Gaussian channel. Consider the vector Gaussian noise channel: $Y = X + Z$ where $X = (X_1 X_2 X_3)$, $Z = (Z_1 Z_2 Z_3)$, $Y = (Y_1 Y_2 Y_3)$, $E[\|X\|^2] \le P$ and

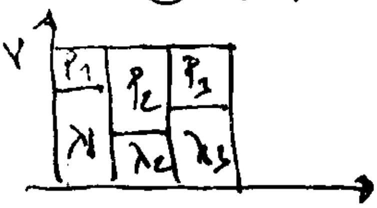
$$Z \sim N\left(0, \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\right)$$

Find the capacity. The answer may be surprising

$K_z = V \Lambda V \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \qquad \begin{matrix} \lambda_1 = 0 \\ \lambda_2 = 3 \\ \lambda_3 = 1 \end{matrix}$

$$C = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \operatorname{ld}\left(1 + \frac{(\nu - \lambda_i)^+}{\lambda_i}\right)$$

WE DEAL SO 3 TO TO $\sum P_i \le P$ (VENIA SPAGA)



$P_1 + \lambda_1 = P_2 + \lambda_2 = P_3 + \lambda_3$

$P_1 + P_2 + P_3 = P$

$P_1 + 0 = \nu \quad \Rightarrow \quad \nu = P_1$

$P_2 + \lambda_2 = P_1$

$P_3 + \lambda_3 = P_1$

$P_1 - P_2 = 3 \qquad P_2 = (P_1 - 3)^+$

$P_1 - P_3 = 1 \qquad P_3 = P_1 - 1$

$P_1 + P_2 + P_3 = P \qquad P_1 + P_1 - 3 + P_1 - 1 = P \qquad 3P_1 = P + 4 \qquad \boxed{P_1 = \frac{P+4}{3}}$

$P_2 = \frac{P+4}{3} - 3 = \frac{P+4-9}{3} = \frac{P-5}{3} \rightarrow$ AND $P < 5 \Rightarrow P_2 = 0$!!

$P_3 = \frac{P+4}{3} - 1 = \frac{P+4-3}{3} = \frac{P+1}{3}$

$P_1 + P_2 + P_3 = \frac{P+4+P-5+P+1}{3} = \frac{3P + 5 - 5}{3} = P$

$$C = \left[\frac{1}{2}\operatorname{ld}\left(1 + \frac{P+4}{3 \cdot 0}\right) + \frac{1}{2}\operatorname{ld}\left(1 + \frac{P-5}{3 \cdot 3}\right) + \frac{1}{2}\operatorname{ld}\left(1 + \frac{P+1}{3 \cdot 1}\right)\right]$$

(first term $\to \infty$; remaining terms: $\frac{1}{2}\operatorname{ld}\left(1 + \frac{P-5}{9}\right) + \frac{1}{2}\operatorname{ld}\left(1 + \frac{P+1}{3}\right)$)

$C \to \infty$

$\frac{1}{n} \sum_i E[X_i^2] \le P \qquad \frac{1}{n} \operatorname{tr}(K_x) \le P$

$I(X_1^n; Y_1^n) = h(Y_1^n) - h(Z_1^n) \qquad K_Y = K_X + K_Z$

$h(Y_1^n) = \frac{1}{2} \operatorname{ld}\{|K_X + K_Z| (2\pi e)^n\}$

$K_Z = Q \cdot \Lambda \cdot Q^t \quad (Q \cdot Q^T = I) \qquad |K_X + K_Z| = |K_X + Q \cdot \Lambda \cdot Q^T| = |Q| \cdot |Q^T K_X Q + \Lambda| \cdot |Q^T| = |A + \Lambda|$

[redacted]

$A = Q^t \cdot K_x \cdot Q$ $\quad tr(BC) = tr(CB)$

$tr(A) = tr(\underbrace{Q^t \cdot K_x}_{B} \cdot \underbrace{Q}_{C}) = tr(\underbrace{Q \cdot Q^t}_{I} \cdot K_x) = tr(K_x)$

$\max: |A + \Lambda|$ s.t. $tr(A) \leq nP$

$K_z = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

$|K| \leq \prod_i K_{ii}$ $\quad |A+\Lambda| \leq \prod_i (A_{ii} + \lambda_i)$

$\frac{1}{n} \sum_i A_{ii} \leq P$

MAXIMAL VALUE OF $\prod_i (A_{ii} + \lambda_i)$ IS ATTAINED WHEN

$A_{ii} = (\nu - \lambda_i)^+$ i.e. $A_{ii} + \lambda_i = \nu$

$\sum A_{ii} = nP$

$C = \int_{-\pi}^{\pi} \frac{1}{2} \log\left(1 + \frac{(\nu - N(f))^+}{N(f)}\right) df$ $\quad \int (\nu - N(f))^+ df = P$.

PROBLEM 9.9 (UIC SOLUTIONS)

- CALCULATION OF EIGEN VALUES:

$|K_z - xI| = \left| \begin{bmatrix} 1-x & 0 & 1 \\ 0 & 1-x & 1 \\ 1 & 1 & 2-x \end{bmatrix} \right| = (1-x)\left[(1-x)(2-x) - 1\right] + (x-1) = (1-x)\left[(1-x)(2-x) - 1 - 1\right]$

$= (1-x)\left[2 - x - 2x + x^2 - 2\right] = (1-x)\,x\,[x-3]$ $\quad x = [1, 0, 3]$

$|K_z| = 0$ THIS MEANS THAT COLUMNS (OR ROWS SINCE MATRIX IS SIMMETRIC) ARE LINEARLY DEPENDENT: IN PARTICULAR WE CAN SEE THAT LAST COLUMN (ROW) CAN BE COMPUTED AS THE SUMMATION OF THE FIRST TWO COLUMNS (ROWS).

- SINCE K_Z IS COVARIANCE MATRIX, THIS MAKE THAT Z_3 CAN BE REWRITTEN AS A FUNCTION OF Z_1 AND Z_2: $Z_3 = Z_1 + Z_2$. THIS FACT CAN BE EXPLOITED TO NULLIFY THE EFFECT OF NOISE. IN PARTICULAR IF THE SAME $x = x_1 = x_2 = x_3$ SIGNAL IS SENT OVER THE THREE CHANNELS ITS VALUE CAN PERFECTLY DERIVED BY SUBTRACTING Y_3 TO THE SUMMATION OF Y_1 AND Y_2:

$Y_1 + Y_2 - Y_3 = x + Z_1 + x + Z_2 - x - Z_3 = x + Z_1 + Z_2 - Z_1 - Z_2 = x$

THUS, AS IN EXCERCISE 9.2, CAPACITY OF THE CHANNEL CAN BE CONSIDERED TO BE INFINITE. (NO PROBABILITY OF ERROR) i.e $P_e^{(n)} = 0$

- UNIVERSITY OF TORONTO:

$C = \frac{1}{2} \log (2\pi e)^n \frac{|K_x + K_z|}{|K_z|}$ $\quad |K_z| = 0 \Rightarrow C \to \infty$

[redacted]

$$P(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$ POISSON

• The optimal coding scheme is to set $x_1 = x_2 = 0$ from which receiver detects values of Z_1 and Z_2 then uses them to cancel the noise in Z_3, thus providing noise free channel from x_3 to Z_3, which clearly permits an infinite rate of error free communication for any $\rho \neq 0$.

Problem 9.10 Capacity of photographic film. Here is a problem with a nice answer that takes a little time. We're interested in the capacity of photographic film. The film consist of silver iodide crystals Poison distributed, with a density of λ particles per square inch. The film is illuminated without knowledge of the position of the silver iodide particles. It is then developed and the receiver sees only the silver iodide particles that have been illuminated. It is assumed that light incident on the cell exposes the grain if it is there and otherwise results in a blank response. Silver iodide particles that are not illuminated and vacant portions of the film remain blank. What is the capacity of this film?

We make following assumptions. We grid the film very finely into cells of area dA. It is assumed that there is at most one silver iodide particle per cell and no silver iodide particle is intersected at cell boundaries. Thus the film can be considered to be large number of binary asymmetric channels with crossover probability $1 - \lambda dA$. By calculating the capacity of this binary channel (asymetric ch.) to first order in dA (making the necessary approxima-tions) one can calculate the capacity of the film in bits per square inch. It is, of course proportional to λ. What is the multiplicative constant? The answer will be λ bits per unit area if both the illuminator and the receiver knew the positions of the crystals.

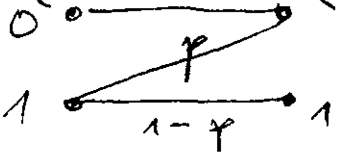
$$\bar{K} = \sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = e^{-\lambda}\left(\frac{\lambda}{0!} + \frac{\lambda^2}{1!} + \dots\right) = e^{-\lambda} \cdot \lambda \left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots\right)$$

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots \frac{x^n}{n!} + \dots = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad e^{\lambda}$$

$$\boxed{\bar{K} = \lambda}$$

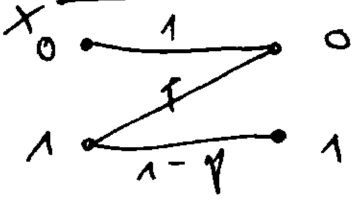
$$\sum_{k=0}^{\infty} P(k,\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda}\underbrace{\left(1 + \frac{\lambda^1}{1} + \frac{\lambda^2}{2!} + \dots\right)}_{e^{\lambda}} = 1$$

$E[X] = \lambda \quad (X \sim P(k,\lambda))$



$\boxed{p = 1 - \lambda dA}$

BINARY ARITMETHIC CHANNEL



$$I(X;Y) = H(Y) - H(Y|X) = H(X) - H(X|Y)$$

$$H(Y|X) = P(X=0)\cdot H(Y|X=0) + P(X=1)\cdot H(Y|X=1)$$

$$H(Y|X=0) = -\left[P(Y=0|X=0)\cdot ld\, P(Y=0|X=0) + P(Y=0|X=1)\cdot ld\, P(Y=1|X=0)\right]$$

(under the terms: 1, 1, 0)

$H(Y|X=0) = 0$ — nema neizvesnost.

$$H(Y|X=1) = -\left(P(Y=0|X=1)\, ld\, P(Y=0|X=1) + P(Y=1|X=1)\, ld\, P(Y=1|X=1)\right)$$

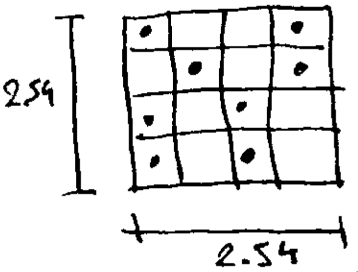
(under the terms: γ, γ, $(1-\gamma)$, $(1-\gamma)$)

$H(Y|X=1) = + H(\gamma)$ //

$$\boxed{H(Y|X) = P(X=1)\cdot H(\gamma)} \quad (A)$$

$$\boxed{\gamma = 1 - \lambda\cdot dA} \qquad (1-\gamma) = \lambda\cdot dA$$

$$\boxed{\lambda = 8 \text{ particles}/\text{inch}^2} \qquad \boxed{\lambda = \frac{8}{16} = 50\ \%} \quad (\star)$$



Sega da kaže deka na eden kvadraten inch imaš vo prosek λ iodide kristalni čestički. Brojot na kristalni čestički vo kvadraten inch e raspredelen vo Puasonova raspredelba.

Dokaz – varijansa na Puasonova:

$$E[k^2] = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k e^{-\lambda}}{k!}$$

$$E[k^2] = \sum_{k=1}^{\infty} k \frac{\lambda^k e^{-\lambda}}{(k-1)!} = \underbrace{\sum_{k=1}^{\infty} (k-1) \frac{\lambda^k e^{-\lambda}}{(k-1)!}}_{(*)} + \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-1)!}}_{(**)}$$

$m = k-1 \quad k=1 \quad m=0 \,/\quad k = m+1$

$$(*) = \sum_{m=0}^{\infty} m \frac{\lambda^{m+1} e^{-\lambda}}{m!} = \lambda \cdot \underbrace{\sum_{m=0}^{\infty} m \frac{\lambda^m e^{-\lambda}}{m!}}_{\lambda} = \lambda^2$$

$$(**) = \sum_{m=0}^{\infty} \frac{\lambda^{m+1} e^{-\lambda}}{m!} = \lambda \sum_{m=0}^{\infty} \frac{\lambda^m e^{-\lambda}}{m!} = \lambda$$

$$E[(k-\bar{k})^2] = E[k^2] - E^2[k] = \lambda^2 + \lambda - \lambda^2 = \lambda$$

$$P_X(k, A) = \frac{(\lambda\cdot A)^k}{k!}\cdot e^{-\lambda\cdot A}$$

$$P(X=1) = \lambda\cdot dA$$
$$P(X=0) = 1 - \lambda\cdot dA$$

[redacted]

(*) ⇒ 33% OD KLETKITE VO 1 KVADRATEN INCH IMAAT IODIDNE KRISTALNA ČESTIČKA.

ZNAČI, DEKA VEROJATNOSTA EDNA KLETKA DA IMA KRISTALNA ČESTIČKA E:

$$\boxed{P(\lambda=1) = \lambda \overset{(*)}{=} \frac{1}{2} = 0.5}$$ (KAKO NA SLIKATA ZA SLUČAJ (*))

$$\boxed{p = 1-\lambda}$$

- ZNAČI OD (1) SLEDI:

$$H(Z|X) = \lambda \cdot dA \cdot H(p) \doteq \lambda \cdot H(p) = \lambda \cdot H(\lambda)$$

$$I(X;Z) = H(Z) - \lambda \cdot \left[(1-\lambda)\, ld\frac{1}{1-\lambda} + \lambda\, ld\frac{1}{\lambda}\right]$$

GAUSSIAN DISTRIBUTION MAXIMIZES ENTROPY ⇒

$$h(Y)\Big|_{max} = \frac{1}{2}\, ld(2\pi e)\cdot\lambda \qquad \boxed{C = \frac{1}{2} ld(2\pi e)\lambda - \lambda\cdot H(\lambda)}$$

$$C = \frac{1}{2} ld(2\pi e)\cdot\lambda - \lambda\left[(1-\lambda)\, ld\frac{1}{(1-\lambda)} + \lambda\, ld\frac{1}{\lambda}\right] =$$

$$= \frac{1}{2} ld(2\pi e) + \frac{1}{2} ld\lambda - \lambda\, ld\frac{1}{(1-\lambda)} + \lambda^2 ld\frac{1}{(1-\lambda)} - \lambda^2 ld\frac{1}{\lambda}$$

$$= \frac{1}{2} ld(2\pi e) + \frac{1}{2} ld\lambda - \lambda\, ld\frac{1}{(1-\lambda)} + \lambda^2 ld\frac{\lambda}{(1-\lambda)}$$

$$= \frac{1}{2} ld(2\pi e) + ld\sqrt{\lambda} + \lambda\, ld(1-\lambda) + \lambda^2 ld\frac{\lambda}{(1-\lambda)} =$$

$$= \frac{1}{2} ld(2\pi e) + ld\sqrt{\lambda}\cdot(1-\lambda)^{\lambda} + \lambda^2 ld\frac{\lambda}{(1-\lambda)} = \frac{1}{2} ld(2\pi e) + \frac{1}{2} ld\lambda\cdot(1-\lambda)^{2\lambda} + \lambda^2 ld\frac{\lambda}{(1-\lambda)}$$

$$= \frac{1}{2} ld(2\pi e)\lambda(1-\lambda)^{2\lambda} + \lambda^2 ld\left(\frac{\lambda}{(1-\lambda)}\right) = \frac{1}{2} ld(2\pi e)\lambda(1-\lambda)^{2\lambda} + \frac{1}{2} ld\left(\frac{\lambda}{1-\lambda}\right)^{\lambda^2\cdot 2}$$

$$= \frac{1}{2} ld(2\pi e)\boxed{\lambda\cdot(1-\lambda)^{2\lambda}\cdot\left(\frac{\lambda}{1-\lambda}\right)^{2\lambda^2}}^{(*)} = \frac{1}{2} ld(2\pi e)\,\lambda^{2\lambda^2+1}(1-\lambda)^{2\lambda(1-\lambda)}$$

$$(*) = \lambda\cdot(1-\lambda)^{2\lambda}\cdot\frac{\lambda^{2\lambda^2}}{(1-\lambda)^{2\lambda\cdot\lambda-2\lambda}} = \frac{\lambda^{2\lambda^2+1}}{(1-\lambda)^{2\lambda(\lambda-1)}} = \frac{\lambda^{2\lambda^2+1}}{(1-\lambda)^{-2\lambda(1-\lambda)}}$$

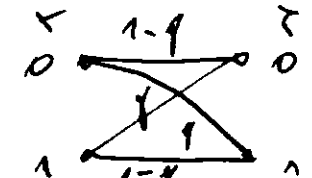
$$C = \frac{1}{2} ld(2\pi e)\,\lambda(1-\lambda)^{2\lambda}\left(\frac{\lambda}{1-\lambda}\right)^{2\lambda^2} \quad \frac{BITS}{CELL}$$

- FOR DISCRETE VARIABLES UNIFORM DISTRIBUTION MAXIMIZES ENTROPY HENCE: (1) ⇒

$$\boxed{C = 1 - \lambda\cdot H(\lambda)} = 1 - \lambda\left((1-\lambda)\, ld\frac{1}{1-\lambda} + \lambda\, ld\frac{1}{\lambda}\right) \quad \frac{bits}{CELL}$$

$1\ SQ\ INCH = 25{,}4 \times 25{,}4\ mm^2 = 645{,}16\ mm^2$ → 1 CELL = $1\ mm^2$

$$C_A = C\cdot 645{,}16 = C(0.8)\cdot 645{,}16 = 272{,}55\ bits/INCH$$



$I(X;Y) = H(Y) - H(Y|X) = \log 2 - H(Y|X)$

$H(Y|X) = P(X=0)\cdot \underbrace{H(Y|X=0)}_{H(q)} + P(X=1)\cdot \underbrace{H(Y|X=1)}_{H(q)}$

$= H(q)$

$I(X;Y)\Big|_{max} = 1 - H(q)$

$C = \operatorname{ld}(m) - \lambda\cdot A \; H(\lambda\cdot A)$ A – SIZE OF THE FILM IN SQINCH

e.g. $\lambda = 1000/\text{SQINCH}$ $A = 0.64 \text{ INCH}^2$

APS-C $25.1\times 16.7\ mm = \underline{419.17\ mm^2} = \frac{419.17}{645.16} = 0.64\ inch^2$

$C =$ $dA = \frac{419.17}{1000}$ $\lambda = \frac{1000}{645.16} \frac{\text{KRISTALS}}{mm^2}$

$\lambda = 1.55 \frac{\text{KRISTALS}}{mm^2}$ $1 \text{ KRISTAL} = 0.5\ mm^2 = 1 \text{ UNIT} = dA$

$\lambda = \frac{1.55}{2} \frac{\text{KRISTALS}}{\text{UNIT}} = 0.725 \frac{\text{KRISTALS}}{\text{UNIT}}$

$m = A = \frac{645\ mm^2}{0.5\ mm^2} = 2\cdot 645 = \underline{1290}$

$C = \operatorname{ld} 1290 - 0.725\ H(0.725) =$

$\frac{A}{\lambda\cdot dA} \Rightarrow$ NUMBER OF KRISTALS

$C = \operatorname{ld}\frac{A}{\lambda dA} - \lambda\cdot dA\ H(\lambda\cdot dA)$

$\lambda = \frac{N}{645.16} = 0.0155\cdot N \text{ KRISTALS}/1\,mm^2$ $\lambda (=) \frac{\text{KRISTALS}}{mm^2}$

$dA = (\text{UNIT AREA})\ mm^2$ A – TOTAL AREA OF FILM (SENSOR)

$C(A,\lambda,dA) = \operatorname{ld}\left(\frac{A}{dA}\right) - \lambda\cdot dA\ H(\lambda\cdot dA)$

$C = \lambda\frac{A}{dA}\left(1 - \lambda\cdot dA\ H(\lambda\cdot dA)\right)$ bits / (IN AREA A)

λ – PARTICLES FOR CELL $0\le\lambda\le 1$
A – AREA IN mm^2
dA – CELL (UNIT AREA)

PROBLEM 9.11 SUPPOSE THAT (X,Y,Z) ARE JOINTLY GAUSSIAN AND THAT $X\to Y\to Z$ FORMS MARKOV CHAIN. LET X AND Y HAVE CORRELATION COEFFICIENT ρ_1 AND LET Y AND Z HAVE CORRELATION COEFFICIENT ρ_2. FIND $I(X;Z)$.

$I(X,Y;Z) = I(X;Z) + I(Y;Z|X) = I(Y;Z) + \cancel{I(X;Z|Y)}^{0}$

$I(X;Z|Y) = H(Z|Y) - H(Z|XY) = H(Z|Y) - H(Z|Y) = 0$

$$I(X;Z) = I(Y;Z) - \underline{I(Y;Z|X)}$$

$$I(Y;Z|X) = H(Z|X) - H(Z|XY) = \underline{H(Z|X)} - \underline{H(Z|Y)}$$

$$I(Y;Z) = H(Z) - H(Z|Y)$$

$$I(Y;Z) = H(Z) - \underline{H(Z|Y)} = \frac{1}{2} ld(2\pi e)G_Z^2 -$$

$$h(Z|Y) = \int_{-\infty}^{\infty} p(y) \cdot h(Z|Y=y)\, dy$$

$$p(Y,Z) = \frac{1}{(2\pi e)^2 |K_{YZ}|} e^{-\frac{[Y,Z]\cdot K_{YZ}^{-1}\cdot [Y,Z]^T}{2}}$$

$$h(Z|Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(y,z)\, ld \frac{1}{p(z|y)}\, dy\, dz$$

$$= \int_{-\infty}^{\infty} p(y) \int_{-\infty}^{\infty} p(z|y)\, ld \frac{1}{p(z|y)}\, dz$$

$$p(y) = \int_{-\infty}^{\infty} p(yz)\, dz$$

$$p(z|y) = \frac{p(yz)}{p(y)}$$

$$I(XY;Z) = H(Z) - H(Z|XY) \underset{\infty}{=} H(Z) - h(Z|Y)_{\infty}$$

$$G_{YZ}^2 = E_{zy}[Z^2] \qquad \int p(y)\,dy \int_{-\infty}^{\infty} z^2 \cdot p(z|y)\, dz = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z^2 p(xy)\, dx\, dy$$

$$E[z^2] = E_y[E[z^2|y]]$$

PROBLEM 7.13 (CONTINUE FROM N16K)

(a) $C_{2,FB} = \frac{1}{2n} \max_{tr[K_X] \leq 2P} ld \frac{|K_{X+Z}|}{|K_Z|}$

(b) $C_{2,NFB} = C_2 = \frac{1}{2n} \max_{tr[K_X] \leq 2P} ld \frac{|K_X + K_Z|}{|K_Z|}$

$K = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$

$(Z_1 Z_2) \sim N(0,K)$

$|K| = 1-\rho^2$

(a) $K_{X+Z} + K_{X-Z} = 2K_X + 2K_Z$

$= E[X^2] + E[XZ] + E[ZX] + E[Z^2]$

$|K_{X+Z}| \qquad |A+B| \leq |A|^{\frac{1}{2}}|B|^{\frac{1}{2}}$

$K_{X+Z} = E[(X+Z)(X+Z)] =$

$|A-B| > 0 \Rightarrow |A| > B$

$2K_X + 2K_Z - K_{X+Z} = K_{X-Z} \geq 0$

$\Rightarrow |2(K_X + K_Z)| \geq |K_{X+Z}|$

$\frac{1}{2} ln (2\pi e)^{2n} |K_{X+Z}| \leq \frac{1}{2} ln (2\pi e)^n (2^n |K_X + K_Z|)$

~~$C_{2,FB} = \frac{1}{2n} \max_{tr[K_X]\leq 2P} ld \frac{|K_X+K_Z|}{|K_Z|} \leq \frac{1}{2n} \max_{tr[K_X]\leq 2P} ld \frac{|K_X|^{\frac{1}{2}} |K_Z|^{\frac{1}{2}}}{|K_Z|}$~~

[redacted]

$\frac{1}{2n} ld \frac{|K_x|^{1/2}}{|K_z|} \leq \frac{1}{2} ld \frac{|K_x|}{|K_z|} \leq \frac{1}{8} ld \frac{P^2}{(1-\rho^2)}$

(b) $C_2 = \frac{1}{2n} \max_{tr[K_x] \leq 2P} \sum_{i=1}^{n} ld\left(1 + \frac{(\lambda - \lambda_i)^+}{\lambda_i}\right)$

$\boxed{\lambda_1 = 1 + \rho \qquad \lambda_2 = 1 - \rho}$

$P_i = \lambda - \lambda_i \qquad \lambda = P_i + \lambda_i$

$P_1 + \lambda_1 = P_2 + \lambda_2 \qquad P_1 + 1 + \rho = P_2 + 1 - \rho \qquad \boxed{P_1 = P_2 - 2\rho}$

$P_1 + P_2 = 2P \qquad P_2 - 2\rho + P_2 = 2P \qquad 2P_2 - 2\rho = 2P$

$\boxed{P_2 = P + \rho} \qquad \boxed{P_1 = P + \rho - 2\rho = P - \rho}$

$C_2 = \frac{1}{4}\left[ld\left(1 + \frac{1-\rho}{1+\rho}\right) + ld\left(1 + \frac{P+\rho}{(1-\rho)}\right)\right] =$

$= \frac{1}{4} ld \left[\frac{(1+\rho+P-\rho)}{(1+\rho)} \frac{(1-\rho+P+\rho)}{(1-\rho)}\right] = \frac{1}{4} ld \frac{(1+P)^2}{(1-\rho^2)} = \frac{1}{4} ld \frac{1+2P+P^2}{(1-\rho^2)}$

(a) $C_{FB,n} = \max_{tr[K_x] \leq 2P} \frac{1}{2n} ld \frac{|K_{x+z}|}{|K_z|} \leq \frac{1}{2n} ld\, 2^n + \frac{1}{2n} ld \frac{|K_x + K_z|}{|K_z|}$ (*)

$|K_{x+z}| \leq |2(K_x + K_z)| \leq 2^n |K_x + K_z| \leq$

$\leq \frac{1}{2} + \frac{1}{2n} ld \frac{|K_x|^{\frac{1}{2}} |K_z|^{\frac{1}{2}}}{|K_z|^{1/2}} = \frac{1}{2} + \frac{1}{2n} \frac{1}{2} ld \frac{|K_x|}{|K_z|}$

$= \frac{1}{2} + \frac{1}{8} ld \frac{P^2}{(1-\rho^2)} = \frac{1}{8} \cdot ld\, 16 + \frac{1}{8} ld \frac{P^2}{(1-\rho^2)} =$

$= \frac{1}{8} ld \frac{16 P^2}{(1-\rho^2)}$

(1) $K_{x+z} + K_{x-z} = 2K_x + 2K_z$
(2) $|A - B| \geq 0 \Rightarrow |A| \geq |B|$ (3) $|K_{x+z}| \leq 2^n |K_x + K_z|$
(4) $\frac{1}{2^n}|K_x + K_z| \geq |K_x|^{1/2} |K_z|^{1/2}$

$C_{FB,n} - C_2 = \frac{1}{8} ld \frac{16 P^2}{(1-\rho^2)} - \frac{1}{4} \cdot \frac{1}{2} ld \frac{(1+2P+P^2)^2}{(1-\rho^2)^2}$

$|\lambda K_x + (1-\lambda) K_z| \leq |K_x|^{\lambda} |K_z|^{1-\lambda}$

$|\frac{1}{2} K_x + \frac{1}{2} K_z| \leq |K_x|^{1/2} |K_z|^{1/2} \qquad \frac{1}{2^n} |K_x + K_z| \leq |K_x|^{1/2} |K_z|^{1/2}$

(*) $\rightarrow \frac{1}{2} + \frac{1}{4} ld \frac{|K_x + K_z|}{(1-\rho^2)} \leq \frac{1}{2} + \frac{1}{4} ld \frac{4 \cdot \frac{1}{4} |K_x + K_z|}{(1-\rho^2)} \leq$

$\leq \frac{1}{2} + \frac{1}{4} ld \frac{4 |K_x|^{1/2} \cdot |K_z|^{1/2}}{|K_z|} = \frac{1}{2} + \frac{1}{4} ld\, 4 + \frac{1}{4} ld \frac{|K_x|^{1/2}}{|K_z|^{1/2}}$

$= \frac{1}{2} + \frac{1}{2} + \frac{1}{8} ld \frac{|K_x|}{|K_z|} = \frac{1}{8} ld\, 256 + \frac{1}{8} ld \frac{P^2}{1-\rho^2} = \frac{1}{8} ld \frac{256 P^2}{1-\rho^2}$

$$\frac{1}{2} ld \frac{|K_{X+Z}|}{|K_Z|} = \left(\begin{smallmatrix}\text{WITHOUT}\\ \text{FEED}\end{smallmatrix} \right) \overset{?}{\geq} \frac{1}{2} ld \frac{|K_X|+|K_Z|}{|K_Z|} =$$

$$= \frac{1}{2} ld \left(1 + \frac{|K_X|}{|K_Z|}\right) = \frac{1}{2} ld \left(1 + \frac{P^2}{1-\rho^2}\right) = \frac{1}{2} ld \frac{1+P^2-\rho^2}{1-\rho^2}$$

SOLUTION

(b) $$\left| \underbrace{\begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix}}_{K_X} + \underbrace{\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}}_{K_Z} \right| = \begin{vmatrix} P+1 & \rho \\ \rho & P+1 \end{vmatrix} = (P+1)^2 - \rho^2$$

$$|K_{X+Z}| \leq 2^n |K_X + K_Z| = 2^n \left[(P+1)^2 - \rho^2 \right]$$

$$C_2 = \frac{1}{2} ld \frac{|K_X+K_Z|}{|K_Z|} = \frac{1}{2} ld \frac{(P+1)^2-\rho^2}{1-\rho^2}$$

(a) $$C_{FB,2} = \frac{1}{2} ld \frac{|K_{X+Z}|}{|K_Z|} \leq \frac{1}{2} ld \frac{2^n |K_X+K_Z|}{|K_Z|} = \frac{1}{2} ld\, 2^n + \frac{1}{2} ld \frac{|K_X+K_Z|}{|K_Z|}$$

$$= 1 + \frac{1}{2} ld \frac{(P+1)^2-\rho^2}{1-\rho^2} = \frac{1}{2} ld 4 + \frac{1}{2} ld \frac{(P+1)^2-\rho^2}{1-\rho^2} =$$

$$= \frac{1}{2} ld \frac{4\left[(P+1)^2-\rho^2\right]}{1-\rho^2}$$

$$C_{FB,2} - C_2 = \frac{1}{2} ld \frac{4\left[(P+1)^2-\rho^2\right]}{1-\rho^2} \cdot \frac{1-\rho^2}{(P+1)^2-\rho^2} = \underbrace{1}_{} \ \text{bit/tran}$$

NE È MOLTO SOLCASTO THEOREM 9.6.2 ZAVOA!

$$C_{FB,2} \leq C_2 + \frac{1}{2} \qquad C_{FB,2}\Big|_{max} = C_2 + \frac{1}{2} = \frac{1}{2} + \frac{1}{2} ld \frac{(P+1)^2-\rho^2}{1-\rho^2}$$

$$C_{FB,2}\Big|_{max} = C_2 + \frac{1}{2} :$$

PROBLEM 9.14 ADDITIVE NOISE CHANNEL. CONSIDER THE CHANNEL $Y = X + Z$ WHERE X IS THE TRANSMITTED SIGNAL WITH POWER CONSTRAINT P, Z IS INDEPENDENT ADDITIVE NOISE, AND Y IS RECEIVED SIGNAL. LET

$$Z = \begin{cases} 0 & \text{WITH PROBABILITY } 1/10 \\ Z^* & \text{WITH PROBABILITY } 9/10 \end{cases}$$

WHERE $Z^* \sim \mathcal{N}(0, N)$. THUS Z HAS MIXTURE DISTRIBUTION THAT IS THE MIXTURE OF GAUSSIAN DISTRIBUTION AND A DEGENERATE DISTRIBUTION WITH MASS 1 AT 0.

(a) WHAT IS THE CAPACITY OF THIS CHANNEL? THIS SHOULD BE A PLEASANT SURPRISE.

(b) HOW WOULD YOU SIGNAL TO ACHIEVE CAPACITY?

[redacted]

(a) $Y = X + Z$ $\quad Z = \begin{cases} 0 & p = 1/10 \\ Z^* & q = 9/10 \end{cases}$ $\quad Z^* \sim \mathcal{N}(0, N)$

$I(X;Y) = H(Y) - H(Y|X) = H(Y) - H(X+Z|X) = H(Y) - H(Z)$

$= H(Z, Z^*) = H(Z^*) + \underbrace{H(Z|Z^*)}_{\neq 0} = H(Z) + \underbrace{H(Z^*|Z)}_{\neq 0}$

$Z = f(Z)$

$X = X_1 + X_2 \quad X_1 \sim p_1 \quad X_2 \sim p_2$

RECALL

$\theta = \begin{cases} 1 & X = X_1 \quad \alpha \\ 2 & X = X_2 \quad 1-\alpha \end{cases} \quad \theta = f(X)$

$H(\theta, X) = H(\theta) + H(X|\theta) = H(X) + \underbrace{H(\theta|X)}_{0} = H(\alpha) + H(X|\theta)$

$H(X|\theta) = P(\theta = 1) \cdot \underbrace{H(X|\theta=1)}_{H(X_1)} + P(\theta = 2) \cdot \underbrace{H(X|\theta=2)}_{H(X_2)} =$

$= \alpha H(X_1) + (1-\alpha) H(X_2)$

$H(X) = H(\alpha) + \alpha H(X_1) + (1-\alpha) H(X_2)$

$H(Z^*) = H(Z) + H(Z^*|Z) = H\left(\frac{1}{10}\right) + H(Z^*|Z)$

$H(Z^*|Z) = \underbrace{P(Z=0)}_{1/10} \cdot \underbrace{H(Z^*|Z=0)}_{H(Z^*)} + P(Z = Z^*) \cdot \underbrace{H(Z^*|Z^*)}_{0}$

$H(Z^*) = H\left(\frac{1}{10}\right) + \frac{1}{10} H(Z^*) \qquad \frac{9}{10} H(Z^*) = H\left(\frac{1}{10}\right)$

$H(Z^*) = \frac{10}{9} H\left(\frac{1}{10}\right) \qquad f(X, \theta) = \begin{cases} 1 & Z = 0 \\ 0 & Z \neq 0 \end{cases}$

$Z = \begin{cases} Z^+ & f(Z^+, 0) \\ Z^* & \mathcal{N}(0, N) \end{cases} \qquad \theta = \begin{cases} 1 & \Rightarrow Z^+; \; p = \frac{1}{10} \\ 2 & Z = Z^*; \; p = \frac{9}{10} \end{cases}$

$\theta = f(Z)$

$H(Z, \theta) = H(Z) + \underbrace{H(\theta|Z)}_{0} = H(\theta) + H(Z|\theta) = H\left(\frac{1}{10}\right) + H(Z|\theta)$

$H(Z|\theta) = P(\theta = 1) H(Z|\theta = 1) + P(\theta = 2) \cdot H(Z|\theta = 2) =$

$= \frac{1}{10} \cdot \underbrace{H(Z^+)}_{0} + \frac{9}{10} \cdot H(Z^*) = \frac{9}{10} \cdot \frac{1}{2} ld\left[(2\pi e) \cdot N\right]$

$\rightarrow Z^+ \in \{0, \text{other}\} \quad P(Z^+) = \{1, 0\} \quad H(Z^+) = 1 \cdot ld\, 1 = 0$

$H(Z) = H\left(\frac{1}{10}\right) + H(Z|\theta) = H\left(\frac{1}{10}\right) + \frac{9}{20} ld(2\pi e) N = \frac{1}{10} ld\, 10 + \frac{9}{10} ld \frac{10}{9}$

$+ \frac{9}{20} ld(2\pi e) N = \frac{1}{10} ld\, 10 + \frac{9}{10}\left[ld \frac{10}{9} + ld\sqrt{2\pi e N}\right] = \frac{1}{10} ld\, 10 + \frac{9}{10} ld \frac{10}{9}\sqrt{2\pi e N}$

$H(Z) = \frac{1}{10}\left[ld\,10 + ld\left(\frac{10}{9}\right)^9 \cdot \sqrt{(2\pi e N)^9}\right] = \frac{1}{10}\, ld\, \frac{10^{10}}{9^9} (2\pi e N)^{\frac{9}{2}}$

$I(X;Y) = H(Y) - H(Z)$ $\quad C = \frac{1}{2} ld(2\pi e)P_Y - \frac{1}{10} ld \frac{10^{10}}{9^9}\sqrt{2\pi e N}$

$P_Y = E[X^2] + E[Z^2] = P + E_Z[Z^2]$

$E_Z[Z^2] = E_\Theta\left[E[Z^2/\Theta]\right] = \sum_{\Theta=1}^{2} P(\Theta)\, E_Z[Z^2 | \Theta] = \frac{1}{10} E[Z'^2] + \frac{9}{10} E[\ldots]$

$E_Z[Z^2] = \frac{9}{10} E[Z'^2] = \frac{9}{10} N$

$C = \frac{1}{2} ld(2\pi e)\left(P + \frac{9}{10}N\right) - \frac{1}{2} ld \frac{10^{\frac{10}{5}}}{9^{\frac{9}{5}}} \cdot (2\pi e)^{\frac{9}{10}} N^{\frac{9}{10}}$

$C = \frac{1}{2} ld \frac{(2\pi e)}{(2\pi e)^{9/10}} \frac{P + \frac{9}{10}N}{N^{9/10}} = \frac{1}{2} ld \sqrt[10]{2\pi e}\left(\frac{9}{10}\sqrt[10]{N} + \frac{P}{\sqrt[10]{N^9}}\right)$

• UIC SOLUTION (HW9c.PDF)

CAPACITY OF THIS CHANNEL IS INFINITE, SINCE AT THE TIMES THE NOISE IS 0 THE OUTPUT IS EXACTLY EQUAL TO THE INPUT, AND WE CAN SEND INFINITE NUMBER OF BITS. TO SEND INFORMATION THROUGH THIS CHANNEL JUST REPEAT SAME REAL NUMBER AT THE INPUT. WHEN WE HAVE THREE OR FOUR OUTPUTS THAT AGREE, THAT SHOULD CORRESPOND TO THE POINTS WHERE THE NOISE IS 0, AND WE CAN DECODE INFINITE NUMBER OF BITS.

PROBLEM 9.15 DISCRET INPUT, CONTINUOUS OUTPUT CHAN.

LET $Pr\{X=1\} = p$, $Pr\{X=0\} = 1-p$, AND LET $Y = X + Z$ WHERE Z IS UNIFORM OVER THE INTERVAL $[0,a]$ $a>1$, AND Z IS INDEPENDENT OF X.

(a) CALCULATE : $I(X;Y) = H(X) - H(X|Y)$

(b) NOW CALCULATE $I(X;Y)$ THE OTHER WAY BY:

$I(X;Y) = h(Y) - h(Y|X)$

(c) CALCULATE CAPACITY OF THE CHANNEL MAXIMIZING OVER p.

RECALL: (PROBLEM 8.7) $\quad Y = X' + Z \qquad Z \sim U[0,1]$

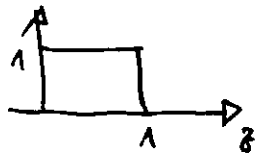
$X' \in [a_1, a_2, \ldots]$

$h(Y) = H(X')$ $\qquad P(X') = \{p_1, p_2 \ldots\}$

$h(Y) \leq \frac{1}{2} ld(2\pi e)\sigma_Y^2 = \frac{1}{2} ld(2\pi e)\left[\sigma_X^2 + \sigma_Z^2\right] =$

$= \frac{1}{2} ld(2\pi e)\left[\sum_{i=1}^{n} i^2 p(i) - \left(\sum_{i=1}^{n} i\, p(i)\right)^2 + \sigma_Z^2\right]$

$Z \sim U[0,1]$



$$E[z] = \int_0^1 z \cdot 1 \, dz = \frac{z^2}{2}\Big|_0^1 = \frac{1}{2}$$

$$E[z^2] = \int_0^1 z^2 \, dz = \frac{z^3}{3}\Big|_0^1 = \frac{1}{3}$$

$$\sigma_z^2 = E[z^2] - E^2[z] = \frac{1}{3} - \frac{1}{4} = \frac{4-3}{12} = \frac{1}{12}$$

$$h(Y) = \frac{1}{2}\, ld\,(2\pi e)\left[\sum_{i=1} i^2 p_i - \sum_{i=1} i\, p_i + \frac{1}{12}\right]$$

(a) $I(X;Y) = H(X) - H(Y|X)$ $\qquad Y = X + Z$

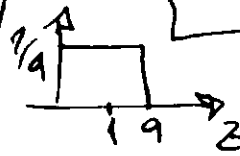
$X \in \{0, 1\}$

$H(X) = p\, ld\, \frac{1}{p} + (1-p)\, ld\, \frac{1}{(1-p)} = H(p)$

$p(X) = \{1-p, p\}$

$$H(Y|X) = \sum_{x=0}^{1} p(x)\, H(Y|x) = (1-p)\, \underbrace{H(Y|x=0)}_{H(z)} + p\, \underbrace{H(Y|x=1)}_{H(z)}$$

$H(Y|X) = H(z)$



$$h(z) = \int_0^a \frac{1}{a}\, ld\, a \, dz = \frac{1}{a}\, ld\, a\, [a - 0] = ld\, a$$

$a > 1$

$I(X;Y) = H(Y) - H(z)$.

$I(X;Y) = h(Y) - ld\, a$

(b) $I(X;Y) = H(X) - H(X|Y)$ $\qquad H(Y) = h(X) = \,?$

$$h(X) = \sum_{i=0}^{1} p_i\, ld\, \frac{1}{p_i} = -\sum_{i=0}^{1} p_i\, ld\, p_i$$

$Y = X + Z \in [X+0,\ X+a]$ $\qquad \in [0, 1+a]$

$0+0 \quad 0+a$
$1+0 \quad 1+a$

$Y_1 \in \{0, 1\}$
$Y_2 \in [a, a+1]$

$p(Y=0) = p$

Recall from 8.7 $\qquad X = \{a_1\ a_2\}$ $\qquad p_r\{X = a_i\} = p_i$

$Y = X + U \qquad U \sim U[0,1]$

$$H(X) = -\sum_{i=0}^{1} p_i\, ld\, p_i = -\left[p_0\, ld\, p_0 + p_1\, ld\, p_1\right] =$$

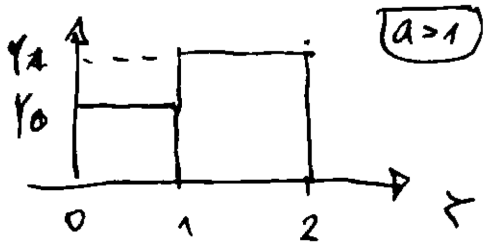
$$= -\sum_{i=0}^{1} \int_{i}^{i+1} f_Y(y)\, dy\ ld \int_{i}^{i+1} f_Y(y)\, dy = -\left(\int_0^1 f_Y(y)\, dy\right) ld \int_0^1 f_Y(y)\, dy -$$

$$- \left(\int_1^2 f_Y(y)\, dy\right) ld \left(\int_1^2 f_Y(y)\, dy\right)$$

[redacted]

$$\ln\left[\int_0^1 x\,dx\right] = \ln\left(\frac{x^2}{2}\Big|_0^1\right) = \ln\frac{1}{2}$$

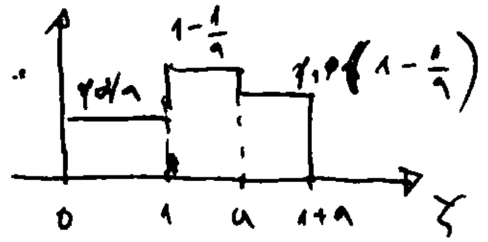
$$\int_0^1 \ln x\,dx = \frac{1}{x}\Big|_0^1 = (1-\infty)$$

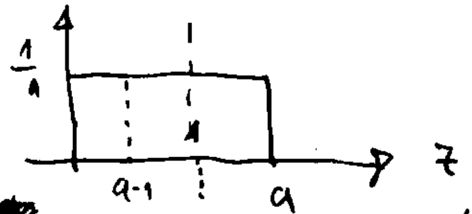


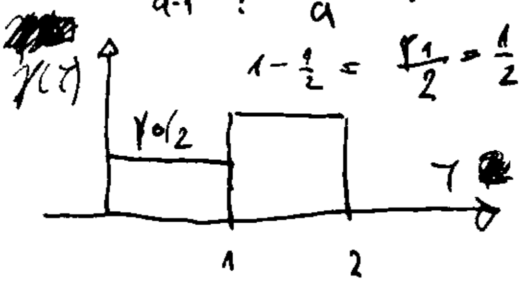
$a>1$ $\quad$ $\gamma_0+\gamma_1=1$ $\quad$ $\gamma_0 = 1-\gamma$, $\gamma_1=\gamma$

$$h(z) = \int_0^2 \gamma(z)\,\mathrm{ld}\,\gamma(z)\,dy =$$

$$= \int_0^1 \gamma_1\,\mathrm{ld}\,\gamma_1\,dy + \int_1^2 \gamma_2\,\mathrm{ld}\,\gamma_2\,dy = \sum_{i=0}^{1}\int_i^{i+1} \gamma_i\,\mathrm{ld}\,\gamma_i\,dy$$







$$\gamma(z\in 0..1) = \gamma(x=0)\cdot\gamma(z\le 1) = \gamma_0\cdot\frac{1}{a}$$

$$\gamma(z\le 1) = \int_0^1 \frac{1}{a}\,dz = \frac{1}{a}\cdot 1 = \frac{1}{a}$$

$$\gamma(z\ge 1) = \int_1^a \frac{1}{a}\,dz = \frac{1}{a}\cdot(a-1) = 1-\frac{1}{a}$$

$$\gamma(z\in 1..a) = \gamma(x=0)\cdot\gamma(z\ge 1) + \gamma(x=1)\cdot\gamma(z\le a-1) = \gamma_0\left(1-\frac{1}{a}\right) + \gamma_1\left(1-\frac{1}{a}\right) = 1-\frac{1}{a}$$

$$\gamma(z\le a-1) = \int_0^{a-1} \frac{1}{a}\,dz = \frac{1}{a}(a-1) = 1-\frac{1}{a} = \gamma_1\cdot\left(1-\frac{1}{a}\right)$$

?

$$\gamma(z\in a..a+1) = \gamma(x=1)\cdot\gamma(z\ge a-1) = \gamma_1\cdot 1/a = \gamma_1/a$$

[illegible]

$$\int_0^{1+a} \gamma(y)\,dy = \int_0^1 \frac{\gamma_0}{a}\,dy + \int_1^a \left(1-\frac{1}{a}\right)dy + \int_a^{a+1} \frac{\gamma_1}{a}\,dy = \frac{\gamma_0}{a} + \left(1-\frac{1}{a}\right)(a-1) + \frac{\gamma_1}{a}(a+1-a)\left(1-\frac{1}{a}\right)$$

$$= \frac{\gamma_0 \ldots}{a} + \left(\frac{a-1}{a}\right)(a-1) \ldots [illegible] =$$

$$= \ldots \frac{a^2-a \ldots a+1}{a} = \frac{1+a^2 \ldots}{a} = \frac{a^2-2a+2}{a} = 1$$

$$a^2-2a+2 = a \quad a^2-3a+2=0 \quad a_{1,2} = 2, 1$$

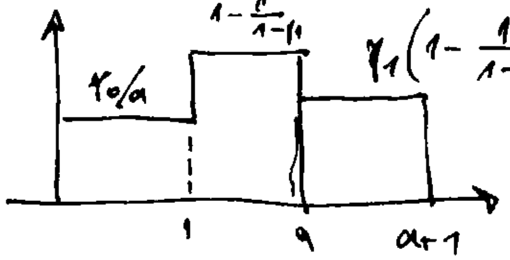
[redacted]

$$\int_0^{1+a} \gamma(y)\,dy = \frac{p_0}{a} + \frac{(a-1)}{a}(a-1) + \gamma_1 \frac{a-1}{a} = \frac{p_0 + a^2 - 2a + 1 + \gamma_1 a - \gamma_1}{a}$$

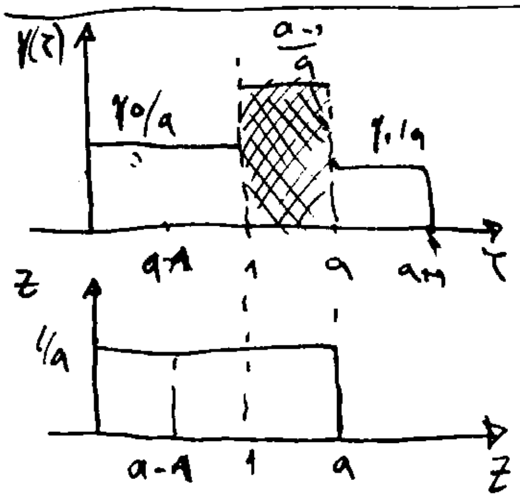
$$\gamma_0 + a^2 - 2a + 1 + \gamma_1 a - \gamma_1 = a \qquad \gamma_0 = 1 - \gamma_1$$

$$1 - \gamma_1 + a^2 - 3a + 1 + \gamma_1 a - \gamma_1 = 0 \qquad a^2 - 3a + 2 + \gamma_1 a - 2\gamma_1 = 0$$

$$a^2 + (\gamma_1 - 3)a + 2(1-\gamma_1) = 0 \quad \Rightarrow \quad \boxed{a = 1 - \gamma_1}$$



$$\gamma_1\left(\frac{1-\gamma_1 - 1}{1-\gamma_1}\right) = -\frac{\gamma_1^2}{1-\gamma_1}$$



- $\gamma(\tau \le 1) = p(x=0)\, p(0 \le z \le 1) = p_0 \cdot 1/a$

$$\gamma(z \le 1) = \int_0^1 \frac{1}{a}\,dz = \frac{1}{a};$$

- $\gamma(1 \le \tau \le a) = p(x=1)\cdot p(0 \le z \le a-1) + p(x=0)\cdot \gamma(1 \le z \le a) = p_1 \frac{a-1}{a} + p_0 \frac{a-1}{a}$

$$\int^{a-1} \frac{1}{a}\,dz = \frac{1}{a}(a-1) \qquad \int_1^a \frac{1}{a}\,dz = \frac{1}{a}(a-1)$$

$$\boxed{\gamma(1 \le \tau \le a) = \frac{a-1}{a}}$$

- $\gamma(a \le \tau \le a+1) = p(x=1)\cdot \gamma(z > a-1) =$
$= p_1 \cdot 1/a$

$$\gamma(z \ge a-1) = \int^a \frac{1}{a}\,dz = \frac{1}{a}(a - a + 1) = \frac{1}{a}$$

$$\frac{\gamma_0}{a} + \frac{a-1}{a}\cdot(a-1) + \frac{\gamma_1}{a} = \frac{1}{a} + \left(1 - \frac{1}{a}\right)(a-1) = \frac{1}{a} + a - 1 - 1 + \frac{1}{a}$$

$$-H(\tau) = \int_0^1 \frac{\gamma_0}{a} \operatorname{ld} \frac{\gamma_0}{a}\,dy + \int_1^a \frac{a-1}{a} \operatorname{ld} \frac{a-1}{a}\,dy + \int_a^{a+1} \frac{\gamma_1}{a} \operatorname{ld} \frac{\gamma_1}{a}\,dy =$$

$$= \frac{\gamma_0}{a} \operatorname{ld} \frac{\gamma_0}{a} + \left(\frac{a-1}{a} \operatorname{ld} \frac{a-1}{a}\right)(a-1) + \frac{\gamma_1}{a} \operatorname{ld} \frac{\gamma_1}{a}\cdot 1 =$$

$$= \frac{\gamma_0}{a} \operatorname{ld} \gamma_0 - \frac{\gamma_0}{a} \operatorname{ld} a + \frac{\gamma_1}{a} \operatorname{ld} \gamma_1 - \frac{\gamma_1}{a} \operatorname{ld} a + \frac{(a-1)^2}{a} \operatorname{ld} \frac{a-1}{a}$$

$$= \frac{1}{a}\left[-H(\gamma)\right] - \frac{\operatorname{ld} a}{a}\underbrace{(\gamma_0 + \gamma_1)}_{1} + \frac{(a-1)^2}{a} \operatorname{ld}(a-1) - \frac{(a-1)^2}{a} \operatorname{ld} a$$

$$-H(Y) = -\frac{H(q)}{a} - \frac{1}{a}\,ld\,a + \frac{(a-1)^2}{a}\,ld\,(a-1) - \frac{a^2-2a}{a}\,ld\,a =$$

$$= -\frac{H(q)}{a} - \frac{(a-1)^2}{a}\,ld(a-1) + \frac{2a-a^2}{a}\,ld\,a = -\frac{H(q)}{a} - \frac{a^2-2a+1}{a}\,ld(a-1)$$

$$= -\frac{H(q)}{a} - \frac{(a-1)^2}{a}\,ld(a-1) + \frac{(2-a)a}{a}\,ld\,a =$$

$$= -\frac{H(q)}{a} - \frac{a^2-2a+1}{a}\,ld(a-1) + 2\,ld\,a - a\,ld\,a =$$

$$= -\frac{H(q)}{a} + \frac{2a-a^2-1}{a}\,ld(a-1) + 2\,ld\,a - a\,ld\,a =$$

$$= -\frac{H(q)}{a} + (2-a)\,ld(a-1) - \frac{1}{a}\,ld(a-1) + (2-a)\,ld\,a =$$

$$= -\frac{H(q)}{a} + (2-a)\,ld\,a(a-1) - \frac{1}{a}\,ld(a-1)$$

$$\boxed{H(Y) = \frac{H(q)}{a} - (2-a)\,ld\,a(a-1) + \frac{1}{a}\,ld(a-1)}$$

$H(Y|X) = ?$ $\qquad I(X;Y) = H(q) - ld\,a = H(Y) - H(Y|X)$

$$H(Y|X) = -H(q) + ld\,a + \frac{H(q)}{a} - (2-a)\,ld\,a(a-1) + \frac{1}{a}\,ld(a-1) =$$

$$= H(q)\left(\frac{1}{a}-1\right) + ld\,a - (2-a)\,ld\,a - (2-a)\,ld\,(a-1) + \frac{1}{a}\,ld(a-1)$$

$$= H(q)\left(\frac{1}{a}-1\right) + ld\,a - 2\,ld\,a + a\,ld\,a + \left(\frac{1}{a} - 2 + a\right)ld(a-1) =$$

$$= H(q)\,\frac{1-a}{a} + (a-1)\,ld\,a + \frac{(1-2a+a^2)}{a}\,ld\,(a-1) =$$

$$= H(q)\,\frac{1-a}{a} + (a-1)\,ld\,a + \frac{(a-1)^2}{a}\,ld\,(a-1) =$$

$$= H(q)\,\frac{1-a}{a} + (a-1)\left[ld\,a + \frac{a-1}{a}\,ld\,(a-1)\right] =$$

$$= H(q)\,\frac{1-a}{a} + (a-1)\left[ld\,a + ld(a-1) - \frac{1}{a}\,ld\,(a-1)\right] =$$

$$= H(q)\,\frac{1-a}{a} + (a-1)\left[ld\,a(a-1) - \frac{1}{a}\,ld(a-1)\right]$$

(a) $I(X;Y) = H(X) - H(X|Y)$ $\qquad \boxed{Y = X + Z}$ $\qquad X = Y - Z$

$H(X) = H(q)$ $\qquad H(X|Y) = H(Y - Z|Y) = H(Z) = \ln(a)$

$$\boxed{I(X;Y) = H(q) - \ln(a)}$$

(b) $I(X;Y) = H(Y) - H(Y|X) = H(Y) - \ln(a)$

PROBLEM 8.8 REVISITED $\qquad x \in \{0, \pm 1, \pm 2\}$ $\qquad Z \sim [-1,1]$ $\qquad Y = X + Z$

$q(x) = [q_{-2}, q_{-1}, q_0, q_1, q_2]$ $\qquad Y \in [-3, -2, -1, 0, 1, 2, 3\}$

$$\gamma(r) = \begin{cases} q_{-2}/2 \\ (q_{-2}+q_{-1})/2 \\ (q_{-1}+q_0)/2 \\ (q_0+q_1)/2 \\ (q_1+q_2)/2 \\ q_2/2 \end{cases}$$

[redacted]

$x \in [0,1]$ $\quad y(x) \in \{y_0, y_1\}$ $\quad z \sim U[0..a]$

$\tau \in [0..1, 1..a, a..1+a]$

$y(\tau) = \left[\frac{y_0}{a}, 1 - \frac{1}{a}, \frac{y_1}{a} \right]$ $\qquad 1 - \frac{y_0 + y_1}{a} = 1 - \frac{1}{a}$

$h(\tau)|_{max} = ld\,3$ $\qquad \frac{y_0}{a} = \frac{1}{3} = \frac{y_1}{a}$ $\qquad y_0 = \frac{a}{3} = y_1$

$y_0 + y_1 = \frac{a}{3} + \frac{a}{3} = \frac{2a}{3} = 1$ $\qquad \boxed{a = \frac{3}{2}}$ $\qquad y_0 = \frac{1}{2} = y_1$

$y(\tau) = \left[\frac{a}{3}, 1 - \frac{2}{3}, \frac{1}{3} \right] = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right]$

$$h(\tau) = -\left[\frac{y_0}{a}\, ld\, \frac{y_0}{a} + \left[\left(1 - \frac{1}{a}\right) ld \left(1 - \frac{1}{a}\right) \right] \frac{(a-1)}{\int_1^a (\cdot)\, d\tau} + \frac{y_1}{a}\, ld\, \frac{y_1}{a} \right]$$

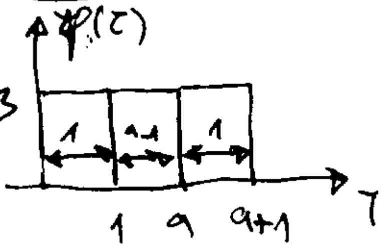
(c) $\quad C = \max_{y} I(x;\tau)$; $\quad I(x;\tau) = h(\tau) - h(\tau|x)$ $\quad h(z) = l_n a$

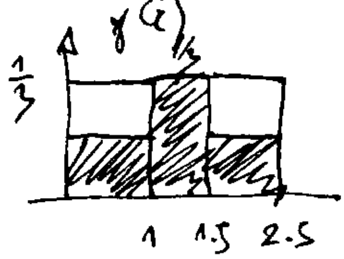
$h(\tau)|_{max} \Rightarrow$ FOR UNIFORM DISTRIBUTION

$h(\tau) = ld\,3$ $\qquad y(\tau) = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right]$

$\Rightarrow \frac{y_0}{a} = \frac{1}{3}$ $\qquad y_0 = \frac{a}{3} = y_1$ $\qquad \frac{a}{3} + \frac{a}{3} = 1$ $\qquad \frac{2a}{3} = 1$

$\boxed{a = \frac{3}{2}}$ $\qquad y_0 = \frac{1}{2} = y_1$ $\qquad \Rightarrow y(\tau) = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right]$





$\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3}$

$\frac{y_0}{a} = \left(1 - \frac{1}{a}\right)(a-1) = \frac{y_1}{a} = \frac{1}{3}$ $\qquad y_0 = a \left(\frac{a-1}{a} \right)(a-1) = (a-1)^2$

$y_1 = 1 - (a-1)^2$ $\qquad (a-1)^2 = \frac{a}{3}$ $\qquad a^2 - 2a + 1 = \frac{a}{3}$

$a^2 - 2a - \frac{a}{3} + 1 = 0$ $\qquad a^2 - \frac{6a + a}{3} + 1 = 0$ $\qquad a^2 - \frac{7a}{3} + 1 = 0$

$\boxed{a = \frac{7}{6} \oplus \frac{\sqrt{13}}{6}}$ $\qquad \boxed{a = 1.76759}$

$y_0 = (a-1)^2 =$

$C = ld\,3 - ld\,a = 0.76318$

[redacted]

PROBLEM 9.20 (CONTINUE FROM N 162) 75

THE ENTROPY POWER INEQUALITY

$$2^{2h(X+Z)} \geq 2^{2h(X)} + 2^{2h(Z)} \quad / \; 2^{2h(Z)}$$

$h(X) = \frac{1}{2} ld(2\pi e)P$

$h(Z) = \frac{1}{2} ld(2\pi e)N$

$$2^{2[h(X+Z) - h(Z)]} \geq 2^{2h(X) - 2h(Z)} + 1$$

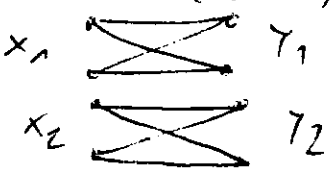
$$2^{2[h(X+Z) - h(Z)]} \geq 2^{2h(X) - 2h(Z)} + 1 = 2^{\frac{ld(2\pi e)P}{ld(2\pi e)N}} + 1 = \frac{2\pi e P}{2\pi e N} + 1$$

$$2^{2[h(X+Z) - h(Z)]} \geq \frac{P}{N} + 1 \qquad 2[h(X+Z) - h(Z)] \geq ld\left(\frac{P}{N} + 1\right)$$

$$I(Y;X) = I(X;Y) \geq \frac{1}{2} ld\left(1 + \frac{P}{N}\right)$$

Щом сега $I(X;Y)$ е по-голямо от $\frac{1}{2} ld\left(1+\frac{P}{N}\right)$, тогаз и $C = \max_{F(x)} I(X;Y)$ е по-голямо, т.е. $\frac{1}{2} ld\left(1+\frac{P}{N}\right)$ е долна граница на капацитета.

- Recall (using two channels at once) Problem 7.5



$$\Rightarrow \quad C \leq C_1 + C_2$$

$$\left. I(X;Y) \right|_{max} \leq \left. I(X_1;Y_1) \right|_{max} + \left. I(X_2;Y_2) \right|_{max}$$

- Entropy of disjoint mixture Problem 2.10

$$H(X) = H(\alpha) + \alpha H(X_1) + (1-\alpha) H(X_2) \quad / \; 2^x$$

$$2^{H(X)} = 2^{H(\alpha)} + 2^{\alpha H(X_1)} + 2^{(1-\alpha)H(X_2)} \leq 2 + 2^{\alpha H(X_1)} + 2^{(1-\alpha)H(X_2)}$$

$$= 2 + 2^{H(X_2)} 2^{-\alpha H(X_2)} + 2^{\alpha H(X_1)} \leq 2 + 2^{H(X_2)} + 2^{\alpha H(X_1)}$$

$$\frac{dH(X)}{d\alpha} = -\left[\alpha\, ld\, \alpha + (1-\alpha)\, ld(1-\alpha)\right]' + H(X_1) - H(X_2) = 0$$

$$-\left[ld\,\alpha + \alpha \cdot \frac{1}{\alpha} \frac{1}{\ln 2} + \left[-1 \cdot ld(1-\alpha)\right] + \frac{(1-\alpha)}{\ln 2} \frac{-1}{1-\alpha}\right] + H(X_1) - H(X_2) = 0$$

$$-\left[ld\,\alpha + ld\,e - ld(1-\alpha) - ld\,e\right] + H(X_1) - H(X_2) = 0$$

$$-ld \frac{\alpha}{1-\alpha} = H(X_2) - H(X_1) \qquad \frac{\alpha}{1-\alpha} = 2^{-H(X_2) + H(X_1)}$$

$$\alpha = 2^{-H_2 + H_1} - \alpha 2^{-H_2 + H_1} \qquad \alpha\left(1 + 2^{-H_2 + H_1}\right) = 2^{-H_2 + H_1}$$

(NOTATION 4.1 PP. 91)

$$\alpha = \frac{2^{-H_2 + H_1}}{1 + \frac{2^{H_1}}{2^{H_2}}} = \frac{2^{H_1}}{2^{H_1} + 2^{H_2}}; \quad \left. H(X) \right|_{max} = \left. H(X) \right|_{\alpha = \frac{2^{H_1}}{2^{H_1} + 2^{H_2}}}$$

$$2^{H_{max}} = 2^{H_1} + 2^{H_2} \qquad 2^{H} \leq 2^{H_{max}} \qquad 2^{H(X)} \leq 2^{H(X_1)} + 2^{H(X_2)}$$

[redacted]

(6) THE RANDOM CODEBOOK CAN BE LISTED AS AN ARRAY WITH 2^{nR} ROWS AND "n" COLUMNS, THE ENTRY $x_i(m)$ IN ROW "m" AND COLUMN "i" BEEING THE i-TH COORDINATE OF THE m-TH CODEWORD. EACH $x_i(m)$ IS GAUSSIAN WITH MEAN ZERO AND VARIANCE P AND ALL THIS' RANDOM VARIABLES ARE MUTUALY INDEPENDENT. LET W DENOTE THE MESSAGE TRANSMITTED, THIS IS UNIFORMLY DISTRIBUTED ON $\{1, 2, \ldots 2^{nR}\}$ AND IS INDEPENDENT OF THE RANDOM CODEBOOK. THE RECEIVED VECTOR $(r_1, \ldots r_n)$ THEN HAS COORDINATES GIVEN WITH: $r_i = x_i(W) + Z_i$, WHERE $(Z_1, \ldots, Z_n)$ ARE I.I.D WITH MARGINAL DISTRIBUTION $f_Z(z)$ AND THESE ARE INDEPENDENT OF W AND THE RANDOM CODEBOOK. LET $\hat{W}$ DENOTE THE DECODED MESSAGE. THIS IS DETERMINED BY THE RULE

$$\boxed{\hat{W} = \arg\min_m \sum_{i=1}^{n} (r_i - x_i(m))^2}$$

AND SINCE TIES OCCUR WITH PROBABILITY OF ZERO, WE CAN IGNORE THE EVENT THAT TIE-BREAKING IS NEEDED. THE PROBABILITY OF ERROR IS: $P(\hat{W} \neq W)$. NOTE THAT THIS IS AUTOMATICALY BEEING AVERAGED OVER THE TRANSMITTED MESSAGE AND THE RANDOM CODEBOOK.

(I) THE PROBABILITY OF ERROR CONDITIONED ON NOISE REALISATION BEING z CAN BE WRITEN AS $P(\hat{W} \neq W / Z = z)$ WHERE Z DENOTES $(Z_1, \ldots, Z_n)$. THE SUGESTED SYMETRY ARGUMENT WORKS AS FOLLOWS: LET U BE ANY ORTHOGONAL **n×n REAL MATRIX**. GIVEN RANDOM CODEBOOK, WE CAN CONSIDER ANOTHER ONE WITH THE SAME NUMBER OF CODEWORDS BUT WITH m-TH ROW BEING GIVEN BY: $(\tilde{x}_1(m), \ldots \tilde{x}_n(m))$ WHERE:

$$\begin{bmatrix} \tilde{x}_1(m) \\ \vdots \\ \tilde{x}_n(m) \end{bmatrix} = U \begin{bmatrix} x_1(m) \\ \vdots \\ x_n(m) \end{bmatrix} \qquad \boxed{\begin{matrix} U^T = U^{-1} \\ U^T \cdot U = I \end{matrix}}$$

BY INVARIANCE UNDER ROTATION, FOR THE NEW CODEBOOK THE PROBABILITY OF ERROR CONDITIONED ON THE NOISE REALIZATION BEING $(Uz^T)^T$ IS IDENTICAL TO THE PROBABILITY OF ERROR UNDER THE OLD CODEBOOK CONDITIONED ON THE NOISE REALIZATION BEING z. HOWEVER, THE NEW CODEBOOK IS STATISTICALY INDISTINGUISHABLE FROM OLD CODEBOOK, HENCE PROBABILITY OF ERROR CONDITIONED ON THE NOISE VECTOR BEING z DEPENDS ONLY ON THE EUCLIDIAN NORM $\|z\|$.

(II) ERROR OCCURS IF THE VECTOR $x(W)^T + Z^T$ IS CLOSER TO $x(m)^T$ THAN TO $x(W)^T$ FOR SOME $m \neq W$. THE CLOSEST SUCH m IS $\hat{W}$. LET $\alpha > 1$. BY CONSIDERING THE TRIANGLE DEFINED BY THREE POINTS $x(W)^T$, $x(\hat{W})^T$ AND $X(W)^T + Z^T$ IN THE TWO DIMENSIONAL PLANE THAT THEY SPAN, SIMPLE GEOMETRIC ARGUMENT SHOW THAT THIS IMPLIES THAT, IN EUCLIDIAN NORM, $x(W)^T + \alpha Z^T$ IS CLOSER TO $x(\hat{W})^T$ THEN IT IS TO $x(W)^T$. THIS

[redacted]

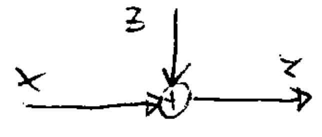
IMPLIES THAT PROBABILITY OF ERROR CONDITIONED ON EUCLIDEAN NORM OF THE NOISE VECTOR IS MONOTONICALLY INCREASING IN THIS EUCLIDEAN NORM.

(III) BY THE WEAK LAW OF LARGE NUMBERS THE PROBABILITY THAT THE NORM OF THE NOISE VECTOR Z EXCEEDS THAT OF GAUSSIAN NOISE VECTOR $(Z_1' \ldots Z_n')$ EACH OF WHOSE COORDINATES HAVE MEAN ZERO AND VARIANCE N', WHERE $N' > N$, GOES TO 0. BY THE GEOMETRIC ARGUMENT OF THE PRECEDING PART IT FOLLOWS THAT THE <u>PROBABILITY OF ERROR</u> IN THE ORIGINAL CASE IS ASYMPTOTICALLY NO BIGGER THAN THAT IN THE GAUSSIAN NOISE CASE (WHERE THE VARIANCE OF THE GAUSSIAN NOISE IS STRICTLY BIGGER THAN N)

(IV) SINCE THE PROBABILITY OF ERROR IN THE GAUSSIAN CASE (WITH NOISE VARIANCE N' STRICTLY BIGGER THEN N BUT SUCH AS $R < \frac{1}{2}\log\left(1+\frac{P}{N'}\right)$) ASYMPTOTICALLY GOES TO ZERO IT DOES SO IN THE CASE WE ARE CONSIDERING.

(V) EVERY STEP IN THE DISCUSSION ABOVE GOES THROUGH EVEN IF WE ONLY ASSUME THAT THE NOISE PROCESS IS STATIONARY AND ERGODIC. THE ONLY THING THAT NEEDS TO CHANGE IS THAT INSTEAD OF WEAK LAW OF LARGE NUMBERS FOR I.I.D NOISE VARIABLES Z_i, WE NEED TO APPLY AN ERGODIC THEOREM.

PROBLEM 9.21 <u>MUTUAL INFORMATION GAME</u>. CONSIDER THE FOLLOWING CHANNEL:



THROUGHOUT THIS PROBLEM WE SHALL CONSTRAIN THE SIGNAL POWER $E[X]=0 \quad EX^2 = P$ AND THE NOISE POWER $E[Z]=0 \quad EZ^2 = N$

AND ASSUME THAT X AND Z ARE INDEPENDENT. THE CHANNEL CAPACITY IS GIVEN BY: $I(X; X+Z)$.
NOW FOR THE GAME. THE NOISE PLAYER CHOOSES A DISTRIBUTION ON Z TO MINIMIZE $I(X;X+Z)$, WHILE THE SIGNAL PLAYER CHOOSES A DISTRIBUTION ON X TO MAXIMIZE $I(X;X+Z)$. SHOW THAT GAUSSIAN LETTING $X^* \sim N(0,P)$; $Z^* \sim N(0,N)$, CONDITIONS X^* AND Z^* SATISFY THE SADDLEPOINT

$$I(X;X+Z^*) \leq I(X^*;X^*+Z^*) \leq I(X^*;X^*+Z)$$

THUS:
$$\min_Z \max_X I(X;X+Z) = \max_X \min_Z I(X;X+Z) = \frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

AND THE GAME HAS A VALUE. IN PARTICULAR DEVIATION FROM NORMAL FOR EITHER PLAYER WORSENS THE ~~DEVIATION~~ MUTUAL INFORMATION FROM THAT PLAYER'S STANDPOINT. CAN YOU DISCUSS THE IMPLICATIONS OF THIS?

NOTE: PART OF THE PROOF HINGES ON ENTROPY POWER INEQUALITY FROM SECTION 17.8, WHICH STATES THAT IF X AND Y ARE <u>INDEPENDENT RANDOM n-VECTORS</u> WITH DENSITIES, THEN:

$$2^{\frac{2}{n}h(X+Y)} \geq 2^{\frac{2}{n}h(X)} + 2^{\frac{2}{n}h(Y)}$$

[redacted]

$$2^{\frac{2}{n}h(x+z)} \geq 2^{\frac{2}{n}h(x)} + 2^{\frac{2}{n}h(z)}$$

$$I(x; x+z) = H(x+z) - H(x+z|x) = H(x+z) - H(z)$$

$x = x^* \sim N(0,P)$ $\quad z = z^* \sim N(0,N)$

$$I(x^*; x^*+z^*) = \frac{1}{2} ld(2\pi e)(P+N) - \frac{1}{2} ld(2\pi e) N = \frac{1}{2} ld(2\pi e)\left(1+\frac{P}{N}\right)$$

$$I(x^*; x^*+z) = \frac{1}{2} H(x^*+z) - H(z) \;;\; I(x; x+z^*) = \frac{1}{2} H(x+z^*) - H(z^*)$$

$$I(x; x+z^*) = \frac{1}{2} H(x+z^*) - \frac{1}{2} ld(2\pi e)N \leq \underbrace{\frac{1}{2} H(x^*+z^*)}_{(F)} - \frac{1}{2} ld(2\pi e)N = I(x^*+z^*)$$

(F) GAUSSIAN DISTRIBUTION MAXIMIZES ENTROPY! (F) HENCE

$\Rightarrow I(x^*+z^*; z^*) \geq I(x; x+z^*)$ FIRST PART IS PROVEN

$$I(x^*; x^*+z) = \frac{1}{2} H(x^*+z) - H(z)$$

$$2^{2h(x^*+z)} \geq 2^{2h(x^*)} + 2^{2h(z)} = 2^{2\frac{1}{2} ld(2\pi e)P} + 2^{2h(z)}$$

$$2^{2h(x^*+z)} \geq (2\pi e)P + 2^{2h(z)} \,/\, 2^{2h(z)} \qquad h(z)_{max} = \frac{1}{2} ld(2\pi e N)$$

$$2^{2[h(x^*+z)-h(z)]} \geq \frac{(2\pi e)P}{2^{2h(z)}} + 1 \geq \left| \begin{array}{c}\text{GAUSSIAN}\\ \text{MAXIMIZES}\\ \text{ENTROPY}\end{array} \right| \geq \frac{(2\pi e)P}{(2\pi e)N} + 1$$

$$h(x^*+z) - h(z) = \underbrace{I(x^*+z; z) \geq \frac{1}{2} ld\left(1+\frac{P}{N}\right)}_{\text{SECOND PART OF THE PROOF}} = I(x^*+z^*; z^*)$$

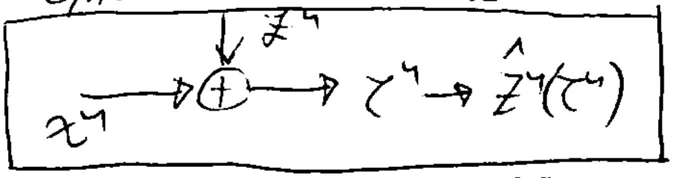
HENCE: $I(x^*+z; z) \geq I(x^*+z^*; z^*) \geq I(x; x+z^*)$ i.e.

$$I(x; x+z^*) \leq \underbrace{I(x^*+z^*; z^*) \leq I(x^*+z; z)}_{(\$)} \quad \text{PROVED !!!}$$

- REGARDING \$ THE STATEMENT IN BERKELEY SOLUTIONS sol5_229ASPR07.PDF: \$ WAS PROVED IN 9.20, WHERE WE ARGUED THAT THE MEMORYLESS ADDITIVE NOISE CHANNEL WITH NOISE VARIANCE N, WHEN THE INPUT DISTRIBUTION IS CHOSEN TO BE GAUSSIAN WITH MEAN ZERO AND VARIANCE "P" THE MUTUAL INFORMATION BETWEEN THE INPUT AND THE OUTPUT IS AT LEAST AS LARGE AS IN THE CASE OF GAUSSIAN NOISE OF VARIANCE N.

PROBLEM 9.22 RECOVERING THE NOISE. CONSIDER A STANDARD GAUSSIAN CHANNEL $Y^n = X^n + Z^n$, WHERE Z_i IS I.I.D. $\sim N(0,N)$ $i = 1, 2, \ldots n$ AND $\frac{1}{n}\sum_{i=1}^{n} x_i^2 \leq P$. HERE WE ARE INTERESTED IN RECOVERING THE NOISE Z^n AND WE DONT CARE ABOUT THE SIGNAL X^n. BY SENDING

$X^n = (0, 0, \dots, 0)$, THE RECEIVER GETS $Y^n = Z^n$ AND CAN FULLY DETERMINE THE VALUE OF Z^n. WE WONDER HOW MUCH VARIABILITY THERE CAN BE IN X^n AND STILL RECOVER GAUSSIAN NOISE Z^n. USE OF THE CHANNEL LOOKS LIKE:



ARGUE THAT FOR SOME $R > 0$ THE TRANSMITTER CAN ARBITRARILY SEND ONE OF 2^{nR} DIFFERENT SEQUENCES OF X^n WITHOUT AFFECTING THE RECOVERY OF THE NOISE IN THE SENSE THAT:

$$P_e\{\hat{Z}^n \neq Z^n\} \to 0 \quad \text{AS } n \to \infty$$

FOR WHAT R IS THIS POSSIBLE?

$$nR = H(W) = I(W;\hat{W}) + H(W|\hat{W}) \leq I(X^n;Y^n) + n\epsilon_n =$$

$$= H(Y^n) - H(Y^n|X^n) + n\epsilon_n = H(Y^n) - H(Z^n) + n\epsilon_n \leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} \underbrace{H(Z_i)}_{\text{GAUSS}}$$

$$\leq \sum_{i=1}^{n} \frac{1}{2} \operatorname{ld}(2\pi e)^n (P_i + N_i) - \sum_{i=1}^{n} \frac{1}{2} \operatorname{ld}(2\pi e)^n N_i + n\epsilon_n = \sum_{i=1}^{n} \frac{1}{2} \operatorname{ld}\left(1 + \frac{P_i}{N_i}\right) + n\epsilon_n$$

(1c)

PROBLEM 7.19 CAPACITY OF CARRIER PIGEON CHANNEL

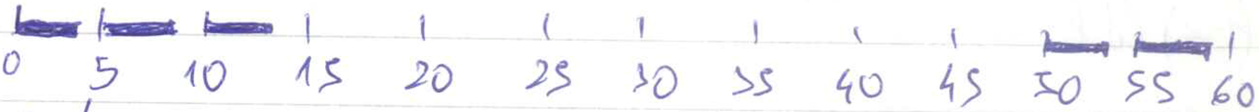
CONSIDER A COMMANDER OF AN ARMY BESIEGED IN A FORT FOR WHOM THE ONLY MEANS OF COMMUNICATION TO HIS ALLIES IS A SET OF CARRIER PIGEONS. ASSUME THAT EACH CARRIER PIGEON CAN CARRY ONE LETTER (8 BITS), THAT PIGEONS ARE RELEASED ONCE EVERY 5 MIN, AND THAT EACH PIGEON TAKES EXACTLY 3 MINUTES TO REACH ITS DESTINATION.

(a) ASSUMING THAT ALL THE PIGEONS REACH SAFELY, WHAT IS THE CAPACITY OF THIS LINK IN BITS/HOUR

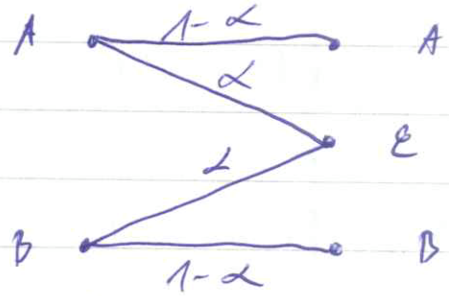
(b) NOW ASSUME THAT THE ENEMIES TRY TO SHOOT DOWN THE PIGEONS AND THAT THEY MANAGE TO HIT FRACTION α OF THEM. SINCE THE PIGEONS ARE SENT AT CONSTANT RATE THE RECEIVER KNOWS WHEN THE PIGEONS ARE MISSING. WHAT IS THE CAPACITY OF THIS LINK?

(c) NOW ASSUME THAT THE ENEMY IS MORE CUNNING AND THAT EVERY TIME THEY SHOT DOWN THE PIGEON, THEY SEND OUT A DUMMY PIGEON CARRYING A RANDOM LETTER (CHOSEN UNIFORMLY FROM ALL 8-BIT LETTERS). WHAT IS THE CAPACITY OF THIS LINK IN BITS/HOUR?

SET UP AN APPROPRIATE MODEL FOR THE CHANNEL IN EACH OF THE ABOVE CASES, AND INDICATE HOW TO GO ABOUT FINDING THE CAPACITY.

(a) 

8 bits / 5 min $\qquad 12 \times 8 \text{ bits} / 1h = 96 \text{ bits}/h$

(b) 

$C = H(Y) - H(Y|X)$

$H(Y|X) = P(X=A)\cdot H(Y|A) + P(X=B)\,H(Y|B)$

$= \pi\, H(\alpha) + (1-\pi)\cdot H(\alpha) = H(\alpha)$

$H(Y) = -\sum_y P(y)\,\mathrm{ld}\,P(y)$

($)

$P(Y) = \{P(Y=A),\ P(Y=E),\ P(Y=B)\} = \{P(X=A)\cdot(1-\alpha),\ P(X=A)\cdot\alpha + P(X=B)\cdot\alpha,\ P(X=B)\cdot(1-\alpha)\} = \{\pi(1-\alpha),\ \pi\cdot\alpha + (1-\pi)\cdot\alpha,\ (1-\pi)(1-\alpha)\} = \{\pi(1-\alpha),\ \alpha,\ (1-\pi)(1-\alpha)\}$

$\pi - \alpha\pi + \alpha + 1 - \alpha - \pi + \alpha\pi = 1$

$H(Y) = -\pi(1-\alpha)\,\mathrm{ld}\,\pi(1-\alpha) - \alpha\,\mathrm{ld}\,\alpha - (1-\pi)(1-\alpha)\,\mathrm{ld}\,(1-\pi)(1-\alpha)$

$= -\pi(1-\alpha)\,\mathrm{ld}\,\pi - \pi(1-\alpha)\,\mathrm{ld}(1-\alpha) - \alpha\,\mathrm{ld}\,\alpha - (1-\pi)(1-\alpha)\,\mathrm{ld}(1-\pi) - (1-\pi)(1-\alpha)\,\mathrm{ld}(1-\alpha)$

$= -\alpha\,\mathrm{ld}\,\alpha - (1-\alpha)\,\mathrm{ld}(1-\alpha) - (1-\alpha)\,\mathrm{ld}(1-\pi) + \pi(1-\alpha)\,\mathrm{ld}(1-\pi) - \pi(1-\alpha)\,\mathrm{ld}\,\pi$

$H(Y) = H(\alpha) - (1-\alpha)\, ld(1-\pi) + \pi(1-\alpha)\, ld \frac{1-\pi}{\pi} =$

$= H(\alpha) - [struck]\, ld(1-\pi) + \alpha\, ld(1-\pi) + \pi\, ld \frac{1-\pi}{\pi} - \pi\alpha\, ld \frac{1-\pi}{\pi}$

$= H(\alpha) - ld(1-\pi) + \alpha\, ld(1-\pi) + \pi\, ld \frac{1}{\pi} + \pi\, ld(1-\pi) - \pi\alpha\, ld(1-\pi) -$

$- \pi\alpha\, ld \frac{1}{\pi}$

$= H(\alpha) - (1-\alpha)\, ld(1-\pi) + \pi(1-\alpha)\, ld(1-\pi) + \pi(1-\alpha)\, ld \frac{1}{\pi} =$

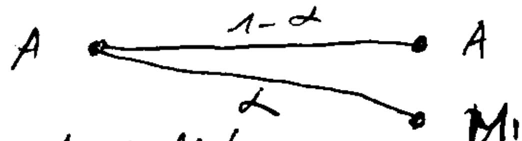
$= H(\alpha) - (1-\alpha)(1-\pi)\, ld(1-\pi) + \pi(1-\alpha)\, ld \frac{1}{\pi} =$

$= H(\alpha) + (1-\alpha)\left[-(1-\pi)\, ld(1-\pi) - \pi\, ld\, \pi\right] = \underline{H(\alpha) + (1-\alpha) H(\pi)}$

$C = \left[H(Y) - H(Y|X)\right]\Big|_{max} = \cancel{H(\alpha)} + (1-\alpha) H(\pi) - \cancel{H(\alpha)}$

$I(X;Y) = (1-\alpha) H(\pi)$ IF: $\pi = \left\{\frac{1}{2}, \frac{1}{2}\right\} \Rightarrow$

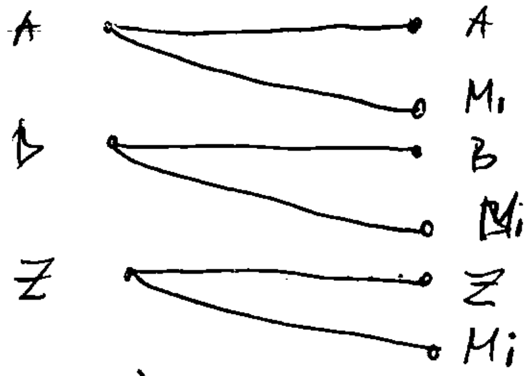
$\boxed{I(X;Y) = 1-\alpha}$



„Mi" − MISSING PIGEON

$I(X;Y) = H(Y) - H(Y|X) =$

$= -(1-\alpha)\, ld(1-\alpha) - \alpha\, ld\, \alpha - H(\alpha) = H(\alpha) - H(\alpha) = 0$



$x \in \{A, B, \ldots Z\}$

$H(X) = \left(\frac{1}{26}\, ld\, 26\right) \cdot 26 = \underline{ld\, 26}$

$H(Y|X) = P(A) \cdot H(Y=A|A) +$

$+ P(X=B) \cdot H(Y=A|B) + \ldots + P(X=Z) \cdot H =$

$= \left[\sum_x P(x)\right] \cdot H(\alpha) = H(\alpha)$

$I(X;Y) = H(Y) - H(\alpha)$

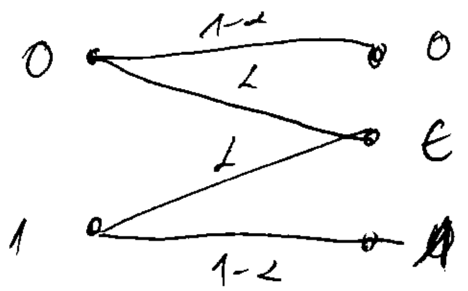
$H(Y) = H(A, B, \ldots Z, Mi) = -\left[(1-\alpha) \sum p(x) \cdot [struck]\, p(x) + \alpha\, ld\, \alpha\right]$

$= \left[(1-\alpha) \cdot [struck]\, H(X)\; [struck]\; \alpha\, ld\, \alpha\right]$

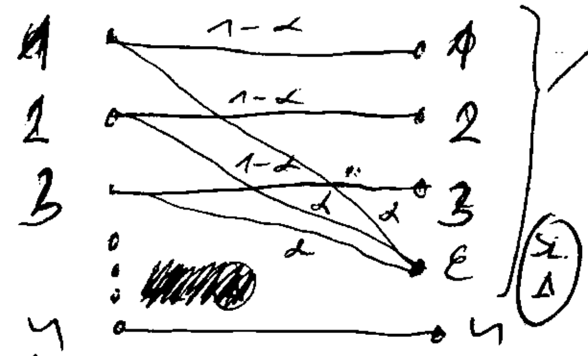
$H(Y) = (1-\alpha)\, [struck]\, H(X) - [struck]\, \alpha\, ld\, \alpha$

$\max\left[H(Y)\right] = (1-\alpha)\, [struck]\, ld\, m - [struck]\, \alpha\, ld\, \alpha$

$C = (1-\alpha)\, [struck]\, ld\, m - [struck]\, \alpha\, ld\, \alpha - H(\alpha)$

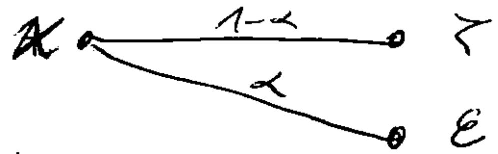


$$P(Z) = \{\pi(1-\alpha),\ \pi\alpha + (1-\pi)\alpha,\ (1-\pi)(1-\alpha)\} =$$
$$= \{\pi(1-\alpha),\ \alpha,\ (1-\pi)(1-\alpha)\}$$



НЕ Е ВАКА ЗОШТО „α" Е ПРОЦЕНТ ОД ВКУПНО ИСПРАТЕНИОТ БЛОК НА ГУЛАБИ!!! МОЖЕ СЕКАК Е!!

ДЕФИНИТИВНО КАНАЛОТ Е:



α Е ПРОЦЕНТ ОД ВКУПНИОТ БЛОК НА ГУЛАБИ. СЕКАК АКО ГИ СОРТИРАШ НА ГУЛАБИ ШТО НОСАТ ПОРАКИ БУКВИ 1 ЌЕ ИМАШ

$N_1\, G(1)$; $N_2\, G(2)$; $N_3 \cdot G(3)$ - - - ; $N_n\, G(n)$;

$N = N_1 + N_2 + \dots N_n$

$\alpha \cdot N = \alpha N_1 + \alpha N_2 + \dots + \alpha N_n$

ОВОЛКУ ГУЛАБИ НОСАТ НА ПОРАКА „1" ЌЕ СТИГНАТ НА ДЕСТИНАЦИЈАТА.

БРОЈ НА ГУЛАБИ НОСАТ НА ПОРАКАТА „2" ШТО ЌЕ СТИГНАТ НА ДЕСТИНАЦИЈАТА.

ЗНАЧИ СЕКАК КАНАЛОТ Е КАКО НА СЛ. Δ

$$P(Z) = \{p_1(1-\alpha),\ p_2(1-\alpha),\ \dots (1-\alpha)p_n,\ p_1\alpha + p_2\alpha + \dots p_n\alpha\} =$$
$$= \{p_1(1-\alpha),\ p_2(1-\alpha),\ \dots (1-\alpha)\cdot p_n,\ \alpha\}$$
$$H(Z) = -\left[\sum_{i=1}^{n} p_i(1-\alpha)\, ld\, p_i(1-\alpha) + \alpha\, ld\, \alpha\right]$$

НА ИСТ НАЧИН МОЖЕВ ДА ГО ДОКАЖАМ (1) НА НАБС. 1

$$H(Z) = \sum_{i=1}^{n} p_i(1-\alpha)\, ld\, \frac{1}{p_i(1-\alpha)} + \alpha\, ld\, \frac{1}{\alpha} =$$
$$= \sum_{i=1}^{n} p_i(1-\alpha)\left[ld\, \frac{1}{p_i} + ld\, \frac{1}{(1-\alpha)}\right] + \alpha\, ld\, \frac{1}{\alpha} =$$
$$= (1-\alpha)\sum_{i=1}^{n} p_i\, ld\, \frac{1}{p_i} + (1-\alpha)\sum_{1} p_i\, ld\, \frac{1}{(1-\alpha)} + \alpha\, ld\, \frac{1}{\alpha}$$

(4c)

$H(Y) = (1-\alpha) \cdot H(X) + H(\alpha)$

$I(X;Y) = H(Y) - H(Y|X) = (1-\alpha) H(X) + H(\alpha) - H(\alpha)$

$I(X;Y) = (1-\alpha) H(X)$

$C_B = \max_{p(x)} I(X;Y) = (1-\alpha) \, ld\, m$ BYTES

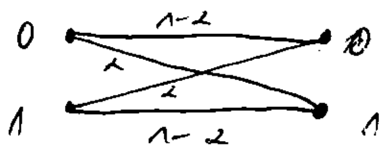
IF $m = 2^8 = 256 \Rightarrow$ $C_B = (1-\alpha) \cdot 8$

KAPACITETOT VO BITE:

$C_b = C_B / 8 = \frac{(1-\alpha)\, ld\, m}{8}$

$C_b = (1-\alpha)$

(c)

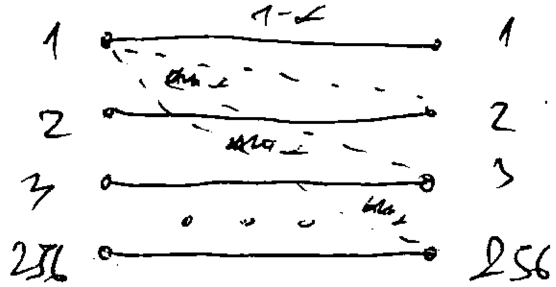


$I(X;Y) = H(Y) - H(Y|X)$

$H(Y|X) = P(X=0) \cdot H(\alpha) + P(X=1) \cdot H(\alpha) = H(\alpha)$

$I(X;Y) = H(Y) - H(\alpha)$

$C = \max_{p(x)} I(X;Y) = 1 - H(\alpha)$



$I(X;Y) = H(Y) - H(Y|X)$

$H(Y|X) = \sum_x p(x) \, H(Y|X=x)$

$H(Y|X=1) = -[P(Y=1) \cdot ld\, P(Y=1) + P(Y=e) \cdot ld\, P(Y=e)] = -[(1-\alpha) \cdot ld(1-\alpha) + \alpha\, ld\, \alpha] = H(\alpha)$

$H(Y|X) = \sum_x p(x) \cdot H(Y|X=x) = p(X=x_1) \cdot H(\alpha) + \dots + p(X=x_m) H_\alpha$

$H(Y|X) = H(\alpha) \cdot \sum_{i=1} p(x_i) = H(\alpha)$

$C = \max I(X;Y)$

$C = ld\, m - H(\alpha)$

$C = ld\, 256 - H(\alpha) = ld\, 2^8 - H(\alpha) = 8 - H(\alpha)$

JOHN HOPKINGS SOLUTIONS

(b) $\max[(1-\alpha) H(X)] = \frac{(1-\alpha)\, ld\, 256}{8} = (1-\alpha)$

8 USES OF THE CHANNEL

$C = (1-\alpha) \cdot 96 \frac{bits}{h}$

ZNAČI PRVO GO SVEDUVAŠ VO BITI I POTOA ŠO MNOŽIŠ SO 96 bits/hour !!!

$$H(Y|x_i) = H\left(1-\alpha+\frac{\alpha}{256}, \frac{\alpha}{256}, \frac{\alpha}{256} \cdots \frac{\alpha}{256}\right)$$

ИЛИ ЌЕ СТИГНЕ БЕЗ ДА ГО ПОГОДАТ ИЛИ ЌЕ СТИГНЕ СО ПОГОДОК НО ПОКАЖУВА КОЈА СЛУЧАЈНО СЕ ДА ИЗБЕРАТ ЌЕ СЕ ПОГОДИ ДА БИДЕ ИСТА СО ИСПРАТЕНАТА ТОЧКА.

$$H(Y|x_i) = -\left(1-\alpha+\frac{\alpha}{256}\right) ld\left(1-\alpha+\frac{\alpha}{256}\right) - 255\frac{\alpha}{256} ld\frac{\alpha}{256}$$

$$H(Y|x_i) = -\left(1-\frac{255\alpha}{256}\right) ld\left(1-\frac{255\alpha}{256}\right) - \frac{255\alpha}{256} ld\frac{\alpha}{256} =$$

$$= -\left(1-\frac{255\alpha}{256}\right) ld\left(1-\frac{255\alpha}{256}\right) - \frac{255\alpha}{256} ld\frac{255\alpha}{256} + \frac{255\alpha}{256} ld\, 255 =$$

$$= H\left(\frac{255\alpha}{256}\right) + \frac{255\alpha}{256} ld\, 255$$

$$\max_{g(x)} I(X;Y) = ld\,256 - H\left(\frac{255\alpha}{256}\right) - \frac{255\alpha}{256} ld\,255 =$$

$$= 8\left(1-\frac{1}{8}H\left(\frac{255\alpha}{256}\right) - \frac{255\alpha}{2048} ld\,255\right)$$ BITES

ВО БИТИ ГО ДЕЛИШ КАРАКОТ СО 8 ЗАТОА ШТО ИМАШ 8 ОДОТЕНИ НА ИСТИОТ ~~ЗАТОА~~

(*) $$C = \left(1-\frac{1}{8}H\left(\frac{255\alpha}{256}\right) - \frac{255\alpha}{2048} ld\,255\right) \text{ bits}$$

• КАПАЦИТЕТОТ ВО bits/hour ГО ДОБИВАШ АКО (*) ГО ПОМНОЖИШ СО $96 \frac{bits}{hou}$ Т.Е.:

$$C = 96\left(1-\frac{1}{8}H\left(\frac{255\alpha}{256}\right) - \frac{255\alpha}{2048} ld\,255\right)$$

IF $\alpha = 1$

$$C = 96\left(1-\frac{1}{8}H\left(\frac{255}{256}\right) - \frac{255}{2048} ld\,255\right)$$

$$-H = \frac{255}{256} ld\frac{255}{256} + \frac{1}{256} ld\frac{1}{256} = \frac{255}{256} ld\frac{1}{256} + \frac{255}{256} ld\,255$$

$$C = 96\left(1+\frac{1}{8} ld\frac{1}{256} + \frac{255}{256} ld\,255 - \frac{255}{2048} ld\,255\right) = \emptyset$$

$$C(\alpha=0) = 96\frac{bits}{h}$$

PROBLEM 7.25 [CONTINUE FROM (9)] BOTTLENECK CHANNEL

$X \to V \to Y$

$X = \{1,2,\dots,m\}$ $\quad Y = \{1,2,\dots,m\}$

$V = \{1,2,\dots,k\}$

$p(v|x)$ $\quad p(y|v)$ ARE ARBITRARY

$p(y|x) = \sum_v p(v|x)\, p(y|v)$ $\qquad$ $C \le \mathrm{ld}(k)$

$I(X;V) = I(V) - H(V|X)$

$I(X,V;Y) = I(X;Y) + I(V;Y|X)$

DATAPROCESSING: $X \to Y \to Z$

$I(X;Y) \ge I(X;Z)$

$I(X;Y,Z) = I(X;Z) + I(Y;Z|X) = I(Y;Z) + \underbrace{I(X;Z|Y)}_{0}$

$I(Y;Z|X) = H(Z|X) - H(Z|XY) = H(Z|X) - H(Z|Y)$

$I(X;Z|Y) = H(Z|Y) - \underbrace{H(Z|YX)}_{=H(Z|Y)} = H(Z|Y) - H(Z|Y) = 0$

$\to$ $I(Y;Z) \ge I(X;Z)$

$I(X;Y,Z) = I(X;Y) + \underbrace{I(X;Z|Y)}_{0} = I(X,Z) + I(X;Y|Z)$

$I(X;Y) \ge I(X;Z)$

MMV

$C = \max\left[I(X;Y)\right]$ $\qquad I(X;Y) \le I(V;Y)$ $\qquad I(X;Y) \le I(X;V)$

$C = \max[I(X;Y)] \le \max[I(V;Y)] = \max[H(Y) - H(Y|V)]$

$C = \max[I(X;Y)] \le \max[I(X;V)] = \max[H(X) - H(X|V)] =$

$= \max[H(V) - H(V|X)] \le \max[H(V)] = \mathrm{ld}\,k$

$\Rightarrow$ $C \le \mathrm{ld}\,k$ $\to$ ZA NEJEDNAKOST VAŽI DA JE KAPACITET ZA MARKOV TIP

$p(XY|Z) = \dfrac{P(X,Y,Z)}{P(Z)} = \dfrac{P(XY)\cdot P(Z|XY)}{P(Z)} = \dfrac{P(X)P(Y|X)P(Z|Y)}{P(Z)}$

$p(XZ|Y) \stackrel{?}{=} p(X|Y)\cdot P(Z|Y) = \dfrac{P(XYZ)}{P(Y)} = \dfrac{P(XY)\cdot P(Z|XY)}{P(Y)} =$

$= \dfrac{P(Y)\cdot P(X|Y)\cdot P(Z|Y)}{P(Y)} = P(X|Y)\cdot P(Z|Y)$

$\to$ OVA 6 EDINSTVENO REŠENJE!!! 26.1.2010 ...

VC GENERALNI SLUČAJ (DOKAZ NEJEDNAKOSTI ZA MARKOV. CH.)

$$C = \max_{p(x)} [I(X;Z)] \; ; \; I(X;Z) = H(Z) - H(Z|X)$$

$$H(Z|X) = \sum_x p(x) \cdot H(Z|X=x)$$

$$\begin{bmatrix} p(z_1) \\ p(z_2) \end{bmatrix} = \begin{bmatrix} p(z_1|x_1) ; p(z_1|x_2) \\ p(z_2|x_1) ; p(z_2|x_2) \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \end{bmatrix}$$

$$\begin{bmatrix} p(z_1) \\ p(z_2) \\ p(z_3) \end{bmatrix} = \begin{bmatrix} p'_{11} & p'_{21} \\ p'_{12} & p'_{22} \\ p'_{13} & p'_{23} \end{bmatrix} \begin{bmatrix} p(v_1) \\ p(v_2) \end{bmatrix} \qquad \begin{bmatrix} p(v_1) \\ p(v_2) \end{bmatrix} = \begin{bmatrix} p''_{11} & p''_{21} & p''_{31} \\ p''_{12} & p''_{22} & p''_{32} \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \end{bmatrix}$$

$$\begin{bmatrix} p(z_1) \\ p(z_2) \\ p(z_3) \end{bmatrix} = \begin{bmatrix} p'_{11} & p'_{21} \\ p'_{12} & p'_{22} \\ p'_{13} & p'_{23} \end{bmatrix} \begin{bmatrix} p''_{11} & p''_{21} & p''_{31} \\ p''_{12} & p''_{22} & p''_{32} \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \end{bmatrix}$$

VIDI 9c

$$\begin{bmatrix} p(z_1) \\ p(z_2) \\ p(z_3) \end{bmatrix} = \begin{bmatrix} \overbrace{p'_{11} p''_{11} + p'_{21} p''_{12}}^{=p(z_1|x_1)} ; & \overbrace{p'_{11} \cdot p''_{21} + p'_{21} \cdot p''_{22}}^{p(z_1|x_2)} ; & \overbrace{p'_{11} \cdot p''_{31} + p'_{21} \cdot p''_{32}}^{p(z_1|x_3)} \\ p'_{12} p''_{11} + p'_{22} p''_{12} ; & p'_{12} p''_{21} + p'_{22} p''_{22} ; & p'_{12} p''_{31} + p'_{22} p''_{32} \\ p'_{13} p''_{11} + p'_{23} p''_{12} ; & p'_{13} p''_{21} + p'_{23} p''_{22} ; & p'_{13} p''_{31} + p'_{23} p''_{32} \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \\ p(x_3) \end{bmatrix}$$

$$I(X;Z) = H(Z) - H(Z|X) \leq H(Z)$$

$$H(Z) = -\sum_z p(z) \, ld \, p(z)$$

$$p(z_1) = \overbrace{\Big[\sum_v p(z_1|v) \cdot p(v|x_1)\Big]}^{①} \cdot \overbrace{p(x_1)}^{②} + \Big[\sum_v p(z_1|v) p(v|x_2)\Big] p(x_2) +$$

$$+ \Big[\sum_v p(z_1|v) \cdot p(v|x_3)\Big] p(x_3) = \sum_{x=x_1}^{x_3} p(x) \sum_v p(z_1|v) \cdot p(v|x)$$

① $p'_{11} p''_{11} + p'_{21} p''_{12} = p(z_1|v_1) \cdot p(v_1|x_1) + p(z_1|v_2) \cdot p(v_2|x_1)$

DALJE NA ... $= p(z_1|x_2)$ DA!!

② $p'_{11} p''_{21} + p'_{21} \cdot p''_{22} = p(z_1|v_1) \cdot p(v_1|x_2) + p(z_1|v_2) \cdot p(v_2|x_2)$

③ $p'_{11} p''_{31} + p'_{21} \cdot p''_{32} = p(z_1|v_1) \cdot p(v_1|x_3) + p(z_1|v_2) \cdot p(v_2|x_3)$ $= p(z_1|x_3)$

$$p(z_1) = \sum_{x=x_1}^{x_3} p(x) \cdot \sum_v p(z_1|v) \cdot p(v|x) = \sum_x \sum_v p(x) p(z_1|v) p(v|x)$$

$$p(y_1) = \sum_{v} p(y_1|v) \sum_{x} p(x)\cdot p(v|x) = \sum_{v} p(y_1|v)\left[\sum_{x} p(x,v)\right]$$

$$①\cdot(*)\ \left[p(y_1|v_1)\cdot p(v_1|x_1) + p(y_1|v_2)\cdot p(v_2|x_1)\right]\cdot p(x_1) =$$

$$= p(y_1|v_1)\cdot p(x_1 v_1) + p(y_1|v_2)\cdot p(x_1, v_2)$$

$$p(y_1) = \sum_{x,v} p(x,v)\cdot p(y_1|v) \qquad p(y_i) = \sum_{x,v} p(x,v)\cdot p(y_i|v)$$

$$H(Y) = \sum_{y}\left[\sum_{xv} p(x,v)\cdot p(y,v)\right] \operatorname{ld}\left[\sum_{xv} p(x,v)\cdot p(y|v)\right]$$

$$①(**) = p(y_1|v_1)\cdot p(v_1 x_2) + p(y_1|v_2)\cdot p(v_2 x_2)$$

$$①(***) = p(y_1|v_1)\cdot p(v_1 x_3) + p(y_1|v_2)\cdot p(v_2 x_3)$$

$$p(y_1) = p(y_1|v_1)\underbrace{\sum_{x_i} p(v_1, x_i)}_{p(v_1)} + p(y_1|v_2)\underbrace{\sum_{x_i} p(v_2, x_i)}_{p(v_2)}$$

$$p(x) = \int_{-\infty}^{\infty} p(x,y)\,dy$$

$$(□) \Rightarrow p(y_1) = \sum_{v} p(y_1|v)\cdot\underbrace{\sum_{x} p(x,v)}_{=p(v)} = \sum_{v} p(y_1|v)\cdot p(v) = p(y_1)$$

$$H(Y|X) = \sum_{x} p(x)\, H(Y|X=x) = ?$$

$$-H(Y|x_1) = (p'_{11} p''_{11} + p'_{21} p''_{12}) \operatorname{ld}(p'_{11} p''_{11} + p'_{21} p''_{12}) +$$

$$+ (p'_{12} p''_{11} + p'_{22} p''_{12}) \operatorname{ld}(p'_{12} p''_{11} + p'_{22} p''_{12}) + (p'_{13} p''_{11} + p'_{23} p''_{12})\cdot$$

$$\cdot \operatorname{ld}(p'_{13} p''_{11} + p'_{23} p''_{12}) = \left[p(y_1|v_1)\cdot p(v_1|x_1) + p(y_1|v_2)\cdot p(v_2|x_1)\right]\cdot$$

$$\cdot \operatorname{ld}\left[p(y_1|v_1)\cdot p(v_1|x_1) + p(y_1|v_2)\cdot p(v_2|x_1)\right] + \left(p(y_2|v_1)\cdot p(v_1|x_1) + p(y_2|v_2)\cdot p(v_2|x_1)\right)\cdot$$

$$\operatorname{ld}\left[p(y_2|v_1)\cdot p(v_1|x_1) + p(y_2|v_2)\cdot p(v_2|x_1)\right] + \left[p(y_3|v_1)\cdot p(v_1|x_1) + p(y_3|v_2)\cdot\right.$$

$$\left.\cdot\, p(v_2|x_1)\right]\cdot \operatorname{ld}\left[p(y_3|v_1)\cdot p(v_1|x_1) + p(y_3|v_2)\cdot p(v_2|x_1)\right]$$

$-H(\tau|x_1) = p(v_1|x_1)\left[p(\tau_1|v_1)\,ld(A) + p(\tau_2|v_1)\,ld(B) + p(\tau_3|v_1)\,ld(C)\right]$

IF: $p(\tau_1|v_1)\cdot p(v_1|x_1) = p(\tau_1|x_1)$ $\quad p(\tau_1|v_2)\cdot p(v_2|x_2) = p(\tau_1|x_2)$

$p(v_1)$

$$\boxed{p(\tau_1|x_1) = p(\tau_1|v_1)\cdot p(v_1|x_1) + p(\tau_1|v_2)\cdot p(v_2|x_1)} \quad \text{MMV}$$

SVE KONSTANTE NA MATRICI JE OSNOVNA LOGIKA !!!

(OVA E USLOVNO IZRAŽENO ZA 6c)

$p(\tau_1|x_1)$

$-H(\tau|x_1) = \left[p(\tau_1|v_1)p(v_1|x_1) + p(\tau_1|v_2)p(v_2|x_1)\right] ld\, p(\tau_1|x_1) +$
$+ \left[p(\tau_2|v_1)\cdot p(v_1|x_1) + p(\tau_2|v_2)\cdot p(v_2|x_1)\right]\cdot ld\, p(\tau_2|x_1) +$
$+ \left[p(\tau_3|v_1)\cdot p(v_1|x_1) + p(\tau_3|v_2)\cdot p(v_2|x_1)\right] ld\left[p(\tau_3|x_1)\right]$

$„—“ = p(v_1|x_1)\left[p(\tau_1|v_1)\cdot ld\, p(\tau_1|x_1) + p(\tau_2|v_1)\cdot ld\, p(\tau_2|x_1) +\right.$
$\left.+ p(\tau_3|v_1)\cdot ld\, p(\tau_3|x_1)\right] \geq p(v_1|x_1)\left\{p(\tau_1|v_1)\cdot ld\left[p(\tau_1|v_1)\cdot p(v_1|x_1)\right] + \dots\right.$

$p(\tau_1|x_1) = p(\tau_1|v_1)\cdot p(v_1|x_1) + p(\tau_1|v_2)\cdot p(v_2|x_1)$

$$\boxed{p(\tau_1|x_1) \geq p(\tau_1|v_1)\cdot p(v_1|x_1)} \quad \text{KLJUČNO VEZANO}$$

$\left.\dots + p(\tau_2|v_1)\cdot ld\left[p(\tau_2|v_1)\cdot p(v_1|x_1)\right] + p(\tau_3|v_1)\, ld\left[p(\tau_3|v_1)\cdot p(v_1|x_1)\right]\right\}$
$= p(v_1|x_1)\left[-H(\tau|V=v_1)\right] - p(v_1|x_1)\left[p(\tau_1|v_1)\, ld\frac{1}{p(v_1|x_1)} +\right.$
$\left.+ p(\tau_2|v_1)\cdot ld\frac{1}{p(v_1|x_1)} + p(\tau_3|v_1)\cdot ld\frac{1}{p(v_1|x_1)}\right] =$
$= -p(v_1|x_1)H(\tau|V=v_1) - p(v_1|x_1)\cdot\left[ld\frac{1}{p(v_1|x_1)}\right]\cdot\sum_{\gamma} p(\tau=\gamma|v_1)$

$p(\tau|V=v_1)$

~~$H(\tau|x_1) \leq p(v_1|x_1)\cdot H(\tau|V=v_1) + p(\tau|v_2)\cdot p(v_2|x_1)\, ld\frac{1}{p(v_2|x_1)}$~~

$„\sim\sim“ = p(v_2|x_1)\left[p(\tau_1|v_2)\, ld\, p(\tau_1|x_1) + p(\tau_2|v_2)\cdot ld\, p(\tau_2|x_1) +\right.$
$\left.+ p(\tau_3|v_2)\, ld\left(p(\tau_3|x_1)\right)\right] \geq p(v_2|x_1)\left[p(\tau_1|v_2)\cdot ld\, p(\tau_1|v_2) +\right.$
$\left.+ p(\tau_2|v_2)\cdot ld\, p(\tau_2|v_2) + p(\tau_3|v_2)\cdot ld\, p(\tau_3|v_2)\right] +$
$+ p(v_2|x_1)\left[p(\tau_1|v_2)\cdot ld\, p(v_2|x_1) + p(\tau_2|v_2)\cdot ld\, p(v_2|x_1) + p(\tau_3|v_2)\, ld\, p(v_2|x_1)\right]$
$= p(v_2|x_1)\left[-H(\tau|V=v_2)\right] + p(v_2|x_1)\cdot\left[ld\, p(v_2|x_1)\right]\cdot P(\tau|V=v_2)$

$$H(Z|x_1) \leq p(v_1|x_1)\cdot H(Z|V=v_1) + p(v_1|x_1)\cdot P(Z|V=v_1)\cdot \operatorname{ld}\frac{1}{p(v_1|x_1)}$$
$$+ p(v_2|x_1)\cdot H(Z|V=v_2) + p(v_2|x_1)\cdot P(Z|V=v_2)\cdot \operatorname{ld}\frac{1}{p(v_2|x_1)}$$

$$\underline{H(Z|X)} = \sum_x H(Z|x)\cdot p(x) \leq \sum_x p(x)\cdot\Big[p(v_1|x)\cdot H(Z|V=v_1)$$
$$+ p(v_2|x)\cdot H(Z|V=v_2)\Big] + \sum_{x_i} p(x_i)\Big[p(v_1|x_i)\cdot P(Z|v_1)\cdot$$
$$\cdot \operatorname{ld}\frac{1}{p(v_1|x_i)} + p(v_2|x_i)\cdot P(Z|V=v_2)\cdot \operatorname{ld}\frac{1}{p(v_2|x_i)}\Big]$$

$$\leq \underbrace{\sum_x p(x, v_1)\, H(Z|V=v_1)}_{p(v_1)\cdot H(Z|V=v_1)} + \underbrace{\sum_x p(x, v_2) H(Z|V=v_2)}_{+\; p(v_2) H(Z|v_2)} +$$

$$+ \underbrace{\sum_x p(x, v_1)\, P(Z|v_1)\cdot \operatorname{ld}\Big(\frac{1}{v_1|x}\Big) + \sum_x p(v_2|x) P(Z|v_2) \operatorname{ld}\frac{1}{p(v_2|x)}}_{P(Z|v_1)\cdot H(v_1|X) \;+\; P(Z|v_2)\cdot H(v_2|X)}$$

$$= p(v_1)\cdot H(Z|v_1) + p(v_2)\cdot H(Z|v_2) + P(Z|v_1)\cdot H(v_1|X) +$$
$$+ P(Z|v_2)\cdot H(v_2|X) = \underline{H(Z|V)} + \underbrace{\sum_v p(Z|v)\cdot H(v|X)}_{\epsilon_1}$$

$$I(X;Z) = H(Z) - H(Z|X) \geq H(Z) - H(Z|V) - \epsilon_1$$

$$C = \max_{p(x)} I(X;Z) \geq \max\big(\underbrace{H(Z) - H(Z|V)}_{H(V;Z) = H(V) - H(V|Z)} - \epsilon_1\big)$$

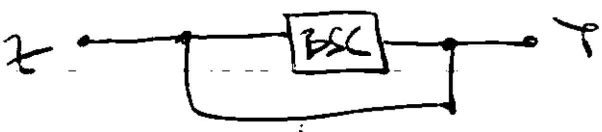
PROBLEM 7.33 BSC WITH FEEDBACK. SUPPOSE THAT FEEDBACK IS USED ON BINARY SYMMETRIC CHANNEL WITH PARAMETER "p". EACH TIME Y IS RECEIVED, IT BECOMES THE NEXT TRANSMISSION. THUS X_1 IS BERN$\left(\frac{1}{2}\right)$, $X_2 = Y_1$, $X_3 = Y_2, \ldots, X_n = Y_{n-1}$.

(a) FIND $\lim_{n\to\infty} \frac{1}{n} I(X^n; Y^n)$

(b) SHOW THAT FOR SOME VALUES OF "p" THIS CAN BE HIGHER THAN CAPACITY

(c) USING THIS FEEDBACK TRANSMISSION SCHEME, $X^n(W, Y^n) = (X_1(W), Y_1, Y_2, \ldots, Y_{n-1})$; WHAT IS ASYMPTOTIC COMMUNICATION RATE ACHIEVED; THAT IS, WHAT IS $\lim_{n\to\infty} \frac{1}{n} I(W; Y^n)$

(11c)



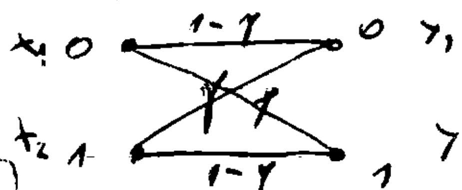
$X_1 = \text{Bern}\left(\frac{1}{2}\right)$

$X_2 = Z_1$; $X_3 = Z_2$ . . .

$X_n = Z_{n-1}$.

(a) $\lim_{n\to\infty} \frac{I(X^n;Z^n)}{n} = \lim_{n\to\infty}\left[\frac{H(X^n) - H(X^n|Z^n)}{n}\right]$

$H(X^n) = H(X_1\, Z_1\, Z_2 \ldots Z_{n-1})$ (with $Z_1 = X_2$, $Z_2 = X_3$, $Z_{n-1} = X_n$)



$$\begin{bmatrix} p(y_1) \\ p(y_2) \end{bmatrix} = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \begin{bmatrix} p(x_1) \\ p(x_2) \end{bmatrix}$$

$p(x_1) = \left\{ \frac{1}{2} \;\; \frac{1}{2} \right\}$ $\qquad X_1 = \begin{cases} 0 & \text{WITH } p = \frac{1}{2} \\ 1 & \text{WITH } p = \frac{1}{2} \end{cases}$

$$H(X^n) = \underbrace{H(X_1)}_{H(\frac{1}{2}) = 1} + \underbrace{H(Z_1|X_1)}_{H(p)} + \underbrace{H(Z_2|X_1 Z_1)}_{= H(Z_2|Z_1) = H(p)} + \underbrace{H(Z_3|X_1 Z_1 Z_2)}_{= H(Z_3|Z_2) = H(p)}$$

$$+ \cdots + \underbrace{H(Z_{n-1}|Z_{n-2})}_{H(p)} = 1 + (n-1)H(p)$$

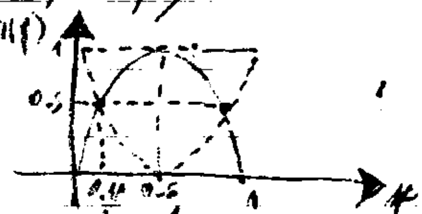
$$H(X^n|Z^n) = H(X_1|Z^n) + H(X_2|X_1 Z^n) + H(X_3|X_1^2 Z^n) + \cdots + H(X_n|X_1^{n-1} Z^n) = H(X_1) + H(X_2|X_1 Z_1^n) + H(X_3|X_1^2 Z_1^n)$$

$$+ \cdots + H(X_n|X_1^{n-1} Z_1^n) = H(X_1) + \underbrace{H(X_2|Z_1)}_{0} + \underbrace{H(X_3|Z_2)}_{0}$$

$$+ \cdots + \underbrace{H(X_n|Z_{n-1})}_{0} = H(X_1) = 1$$

$$I(X^n;Z^n) = 1 + (n-1)H(p) - 1 = (n-1)H(p)$$

$$\lim_{n\to\infty}\left[\frac{n-1}{n} \cdot H(p)\right] = H(p)$$



(b) $C_{BSC} = 1 - H(p)$

$H(p) = 1 - H(p) \qquad 2H(p) = 1 \qquad H(p) = \frac{1}{2}$

$-p \log p - (1-p)\log(1-p) = \frac{1}{2} \rightarrow$ $\boxed{p = 0.11}$

(c) $X^n(W, Y^n) = (X_1(W), Y_1, Y_2, \ldots Y_{n-1})$ (12c)

$\lim_{n\to\infty} \frac{1}{n} I(W; Y^n)$ $\qquad W \in \{1, 2, \ldots 2^{nR}\}$

$W \to X^n \to Y^n \to \hat{W}$

$n\cdot R = H(W) = H(W|\hat{W}) + I(W;\hat{W}) \leq$

$\leq 1 + P_e^{(n)} \cdot n \cdot R + I(X^n; Y^n)$

$I(X^n;Y^n) = H(X^n) - H(Y^n|X^n) =$

$= \sum_{i=1}^{n} H(X_i|X_1^{i-1}) - \sum_{i=1}^{n} H(Y_i|X^n Y_1^{i-1}) \leq$

$\leq \sum_{i=1}^{n} H(X_i) - \sum_{i=1}^{n} H(Y_i|X_i) = \sum_{i=1}^{n} I(X_i;Y_i) \leq n\cdot C$

$I(W;Y^n) = H(W) - H(W|Y^n)$

$I(W,X^n;Y^n) = I(X^n;Y^n) + \underbrace{I(W;Y^n|X^n)}_{0} =$

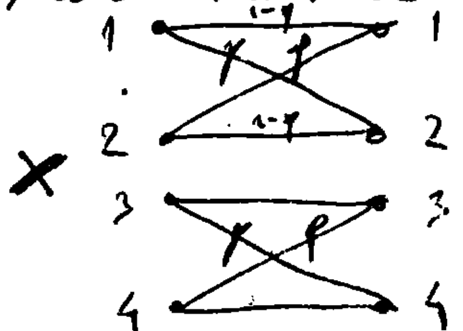
$= I(W;Y^n) + I(X^n;Y^n|W)$

$I(X^n;Y^n) \geq I(W;Y^n)$

$\lim_{n\to\infty} \frac{I(W;Y^n)}{n} \leq \lim_{n\to\infty} \frac{I(X^n;Y^n)}{n}$ $\qquad \lim_{n\to\infty} \frac{n-1}{n} H(p) = H(p)$

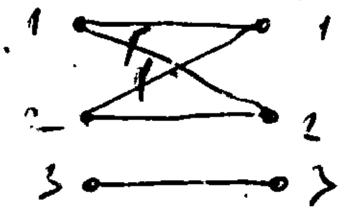
Problem 7.34 Find the capacity of:

(a) Two parallel BSC:

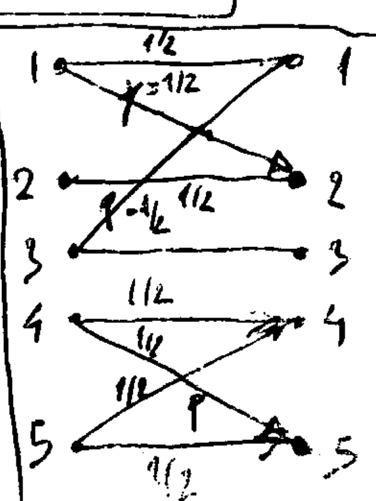


$2^C = 2^{C_1} + 2^{C_2}$

(b) BSC and single symbol



(c) BSC and ternary channel



(d) TOLNAR CHANNEL

$$p(y|x) = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

(a) $2^C = 2^{1-H(q)} + 2^{1-H(q)} = 2\cdot 2^{1-H(q)} = 2^{2-H(q)}$

$\boxed{C = 2-H(q)}$ ⊕

RECALL: $H(Z_1|X_1) = H(q)$ $\quad H(Z_2|X_2) = H(q)$

$I(X;Z) = I(X_1,X_2;Z_1Z_2) = \underline{H(Z_1Z_2)} - H(Z_1Z_2|X_1X_2)$

$H(Z_1Z_2) = H(Z_1) + H(Z_2)$

$H(Z_1Z_2|X_1X_2) = H(Z_1|X_1X_2) + H(Z_2|Z_1X_1X_2) =$

$= H(Z_1|X_1) + H(Z_1|X_2)$

$I(X;Z) = H_1(Z_1) - H(Z_1|X_1) + H(Z_2) - H(Z_2|X_2)$

$I(X;Z) = I(X_1;Z_1) + I(X_2;Z_2)$

$H(Z_1|X_1) =$ [illegible] $p(X_1=1)\cdot H(Z|X_1=1) + p(X_1=2)\cdot H(Z|X_1=2)$

$= H(q)\left[p(X_1=1) + p(X_2=2)\right] = \alpha\cdot H(q)$

$H(Z_2|X_2) = p(X_2=3)\cdot H(Z|X_2=3) + p(X_2=4)\cdot H(Z|X_2=4)$

$\boxed{H(Z_2|X_2) = (1-\alpha)\cdot H(q)}$

$p(Z_1=1) = p(X_1=1)\cdot(1-q) + p(X_1=2)\cdot q = \frac{\alpha}{2}(1-q) + \frac{\alpha q}{2} = \frac{\alpha}{2}$

$H(Z_1) = 2\cdot\frac{\alpha}{2}\,ld\,\frac{2}{\alpha} = \alpha\,ld\,\frac{2}{\alpha}$

$I(X;Z) = \alpha\,ld\,\frac{2}{\alpha} - \alpha H(q) + (1-\alpha)\,ld\,\frac{2}{1-\alpha} - (1-\alpha)H(q)$

$I(X;Z) = H(\alpha) + \alpha\,ld\,2 + (1-\alpha)\,ld\,2 + H(q)$

$I(X;Z) = H(\alpha) + ld\,2 - H(q) = \underline{H(\alpha) + 1 - H(q)}$

$\frac{dI(X;Z)}{d\alpha} = 0 \qquad \frac{d}{d\alpha}\left[\alpha\,ld\,\alpha + (1-\alpha)\,ld(1-\alpha)\right] = ld\,\alpha + \alpha\frac{1}{\alpha} +$

$- ld(1-\alpha) + (1-\alpha)\cdot\frac{1}{(1-\alpha)}\cdot(-1) = ld\,\alpha + 1 - ld(1-\alpha) - 1$

$= ld\,\frac{\alpha}{1-\alpha}$

$I(X;Z) = -ld\,\frac{\alpha}{(1-\alpha)} + (1-H(q))' = -ld\left(\frac{\alpha}{1-\alpha}\right) = 0$

$\frac{\alpha}{1-\alpha} = 1 \qquad \alpha = 1-\alpha \qquad 2\alpha = 1 \qquad \boxed{\alpha = \frac{1}{2}}$

$\boxed{C = H\left(\tfrac{1}{2}\right) + 1 - H(q) = 2 - H(q)}$ 1ST RESULT FOUND ⊕

• Ако двата BSC се спојуваат во еден:

$$I(X;Z) = \alpha \, ld \frac{2}{\alpha} + (1-\alpha)\, ld \frac{2}{1-\alpha} - \alpha H(p) - (1-\alpha) H(q)$$

(H_1 над $H(p)$, H_2 над $H(q)$)

$$= H(\alpha) + \alpha \, ld 2 + (1-\alpha)\, ld 2 - \alpha H_1 - (1-\alpha) H_2$$

$$\frac{dI(X;Z)}{d\alpha} = ld\left(\frac{1-\alpha}{\alpha}\right) - H_1 + H_2 = 0$$

$$ld\left(\frac{1-\alpha}{\alpha}\right) = H_1 - H_2 \qquad \frac{1-\alpha}{\alpha} = 2^{H_1 - H_2}$$

$$1-\alpha = \alpha 2^{H_1 - H_2} \qquad 1 = \alpha\left(1 + 2^{H_1-H_2}\right) \qquad \alpha = \frac{1}{1+2^{H_1-H_2}}$$

$$\boxed{\alpha_m = \frac{2^{H_2}}{2^{H_1} + 2^{H_2}}} \qquad C = I(X;Z)\Big|_{\alpha=\alpha_m} = ld\left(2^{1-H_1} + 2^{1-H_2}\right)$$

$$2^C = 2^{C_1} + 2^{C_2}$$

(b) $2^C = 2^{1-H(p)} + 2^0 = 1 + 2^{1-H(p)}$

$$\boxed{C = ld\left(1 + 2^{1-H(p)}\right)}$$

(d)

$$p(Z|X) = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

$$p(z_1) = \frac{2}{3} p(x_1) + \frac{1}{3} p(x_2)$$

$$p(z_2) = \frac{1}{3} p(x_2) + \frac{2}{3} p(x_3)$$

$$I(X;Z) = H(Z) - H(Z|X)$$

$$H(Z|X) = p(x_1) \cdot H(Z|x_1) + p(x_2) \cdot H(Z|x_2) + p(x_3) \cdot H(Z|x_3) = ~~p(x_1)\cdot \frac{2}{3} ld \frac{3}{2}~~ + p(x_2) \cdot \left[\frac{1}{3} ld 3 + \frac{1}{3} ld 3\right] + ~~p(x_3)\frac{2}{3} ld \frac{3}{2}~~$$

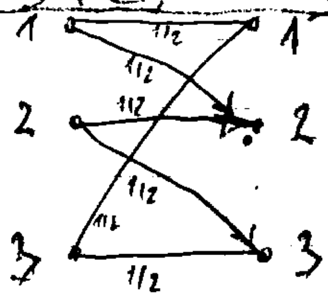
$$= p(X=x_2) \cdot \frac{2}{3} ld 3$$

$$H(Z|X) = p(X=x_2) \cdot \left[\frac{1}{3} ld 3 + \frac{1}{3} ld 3\right]$$

$$p(x_3) = 1 - p(x_1) - p(x_2)$$

$$\Rightarrow p(z_2) = \frac{1}{3} p(x_2) + \frac{2}{3}\left(1 - p(x_1) - p(x_2)\right) = \;?$$

$$= \frac{1}{3} p(x_2) + \frac{2}{3} - \frac{2}{3} p(x_1) - \frac{2}{3} p(x_2) = \frac{2}{3} - \frac{2}{3} p(x_1) - \frac{1}{3} p(x_2)$$



Ова ми личи на NOISY TYPEWRITER CHANNEL!!!

$$I(X;Z) = H(Y) - H(Z|X) = H(Z) - H\left(\tfrac{1}{2}\right)$$

$$I(X;Z) = H(Z) - 1 \qquad C = \max_{p(x)} I(X;Z) =$$

$$= ld 3 - 1 = ld \frac{3}{2}$$

$I(X;Z) = 1 + H(\alpha) - \alpha H(p) - (1-\alpha) H(Z_2|X_2)$

$H(Z_2|X_2) = p(X=x_1)\cdot H(Z|X=x_1) + p(X=x_2) H(Z|X=x_2) + p(X=x_3)\cdot H(Z|X=x_3) = \left(\sum_i p(X=x_i)\right) H\left(\frac{1}{2}\right) = 1$

$I(X;Z) = 1 + H(\alpha) - \alpha H(p) - (1-\alpha) H_2$

$2^C = 2^{C_1} + 2^{C_2} = 2^{1-H(p)} + 2^{ld\frac{3}{2}}$

$C = ld\left[2^{1-H(p)} + 2^{ld\,3/2}\right]$ $\quad p = \frac{1}{2} \quad H(p) = 1$

$C = ld\left[1 + 2^{ld\,3/2}\right] = ld\left(1 + \frac{3}{2}\right) = ld\,\frac{5}{2}$

(d) RENCITER: $C = \max_{p(x)}\left[H(Z) - H(Z|X)\right] =$

$= \max_{p(x)}\left[H(Z) - p(x_2)\frac{2}{3}ld\,3\right] =$

$p(x) = \left[\frac{1}{3}; \frac{1}{3}, \frac{1}{3}\right] \Rightarrow p(z_1) = \frac{2}{3}\frac{1}{3} + \frac{1}{3}\frac{1}{3} = \frac{2}{9} + \frac{1}{9} = \frac{3}{9}$

$p(z_2) = \frac{1}{3}\cdot\frac{1}{3} + \frac{2}{3}\cdot\frac{1}{3} = \frac{3}{9}$

MAKSIMUM SE DOBIVA ZA: $p(Z) = \left\{\frac{1}{2}, \frac{1}{2}\right\}$

$\frac{1}{2} = \frac{2}{3}p(x_1) + \frac{1}{3}p(x_2) \Rightarrow p(x_2) = \left[\frac{1}{2} - \frac{2}{3}p(x_1)\right]\cdot 3$

$\frac{1}{2} = \frac{1}{3}p(x_2) + \frac{2}{3}\left(1 - p(x_1) - p(x_2)\right) = \frac{1}{3}p(x_2) + \frac{2}{3} - \frac{2}{3}p(x_1) - \frac{2}{3}p(x_2)$

$= \frac{2}{3} - \frac{2}{3}p(x_1) - \frac{1}{3}p(x_2) \qquad \frac{2}{3}p(x_1) + \frac{1}{3}p(x_2) = \frac{2}{3} - \frac{1}{2}$

$2p(x_1) + p(x_2) = 2 - \frac{3}{2} = \frac{1}{2} \qquad p(x_1) = \left[\frac{1}{2} - p(x_2)\right]\cdot\frac{1}{2}$

$p(x_1) = \frac{1}{4} - \frac{p(x_2)}{2} \quad ; \quad p(x_2) = \frac{1}{2} - \frac{2}{3}\frac{1}{4} + \frac{2}{3}\frac{p(x_2)}{2} = \frac{1}{2} - \frac{1}{6} + \frac{p(x_2)}{3}$

$p(x_2) = \frac{2}{6} + \frac{p(x_2)}{3} = \frac{1}{3} + \frac{p(x_2)}{3} \qquad 3p(x_2) - p(x_2) = 1$

$2p(x_2) = 1 \qquad \boxed{p(x_2) = \frac{1}{2}} \qquad \boxed{p(x_1) = \frac{1}{4} - \frac{1}{4} = 0}$

$\boxed{p(x_3) = 1 - p(x_1) - p(x_2) = 1/2}$

THE MAXIMIZING DISTRIBUTION OF INPUT SYMBOLS IS: $p(x) = \left[0, \frac{1}{2}, \frac{1}{2}\right]$ FOR THIS DISTRIB THE CAPACITY OF THE THERNACE CHANNEL IS:

$C = ld\,2 - \frac{1}{2}\cdot\frac{2}{3}ld\,3 = ld\,2 - \frac{1}{3}ld\,3 = 1 - \frac{1}{3}ld\,3 =$
$= 0.472$

PROBLEM 7.25. SUPPOSE THAT CHANNEL P HAS CAPACITY "C", WHERE P IS AN $m \times n$ CHANNEL MATRIX

(a) WHAT IS THE CAPACITY OF

$$\tilde{P} = \begin{bmatrix} P & 0 \\ 0 & 1 \end{bmatrix}$$

(b) WHAT ABOUT THE CAPACITY OF

$$\hat{P} = \begin{bmatrix} P & 0 \\ 0 & I_k \end{bmatrix}$$

WHERE I_k IS THE $k \times k$ IDENTITY MATRIX.

(a) $2^C = 2^{C_1} + 2^{C_2}$ $\quad C_1 = ld\, n - H(Y|X) = C$

$C_2 = 2^{ld\, n - H(Y|X)} + 2^0 = 2^{\overbrace{ld\, n - H(Y|X)}^{C}} + 1 = 2^C + 1$

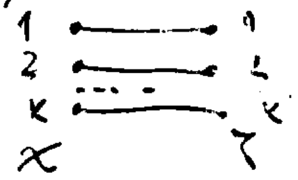
(b) $2^C = 2^{C_1} + \sum^{k} 2^{C_i} = \left| C_i = 0 \right| = 2^{C_1} + k \cdot 1$

$2^C = 2^{\overbrace{ld\, n - H(Y|X)}^{C}} + k$

$$\boxed{C = ld\left[2^C + k\right]}$$

$$\boxed{C = ld\left[\frac{n}{2^{H(Y|X)}} + k\right]}$$

→ OVA VO solns4.pdf OD TORONTO UNIVERSITY SOLUTIONS GO TRETIRAAT KAKO NOISELESS K-AR CHANNEL.



$I(X,Y) = H(Y) - \underbrace{H(Y|X)}_{0} = H(Y)$

$C = \max_{p(x)} H(Y) = ld\, k$

$2^{C_{ef}} = 2^C + 2^{ld\, k}$

$$\boxed{C_{ef} = ld\left[2^C + k\right]}$$

• CHICAGO MIDTERM 1 - 2009 SOLUTIONS (DOPOLNUVANJE NA PROBLEM 7.25) CHANNELS WITH MEMORY

(a) A DISCRETE CHANNEL IS MEMORYLESS IF

$$\boxed{p(y_n | x_1 x_2 \ldots x_n, y_1 y_2, \ldots y_{n-1}) = p(y_n | x_n)}$$

THAT IS, THE OUTPUT AT CHANNEL USE "n" DEPENDS ONLY ON THE INPUT AT CHANNEL USE "n" AND NOT ON ANY OTHER INPUTS AND OUTPUTS.

(b) DO YOU EXPECT THE CAPACITY OF A CHANNEL WITH MEMORY TO BE LARGER OR SMALLER THAN MEMORYLESS VERSION?

$$I(X^n;Y^n) = H(Y^n) - H(Y^n|X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|X^n Y_1^{i-1})$$

$$= H(Y^n) - \underbrace{\sum_{i=1}^{n} H(Y_i|X_i)}_{\text{MEMORYLESS CHANNEL}} \leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) = \sum_{i=1}^{n} I(X_i;Y_i)$$

$$H(Y^n) = \sum_{i=1}^{n} H(Y_i|Y_1^{i-1}) = H(Y_1) + H(Y_2|Y_1) + \dots + H(Y_n|Y_1^{n-1})$$

$$\leq \sum_{i=1}^{n} H(X_i) \rightarrow$$ CONDITIONING REDUCES ENTROPY.

FOR MEMORYLESS: $H(Y^n) = \sum_{i=1} H(Y_i)$

FOR CHANNEL WITH MEMORY: $H(Y^n) \leq \sum_{i=1} H(Y_i)$

FOR MEMORYLESS: $H(Y^n|X^n) = \sum_{i=1}^{n} H(Y_i|X_i)$

FOR CHAN. WITH MEMO. $H(Y^n|X^n) \leq \sum_{i=1}^{n} H(Y_i|X_i)$

IF: $H(Y^n) = \sum_{i=1}^{n} H(Y_i)$ } FOR BOTH MEMORYLESS AND CHAN. WITH MEMORY.

$$I_M(X^n;Y^n) = \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X^n Y_1^{i-1}) \geq \underbrace{\sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i)}_{I_{ML}(X^n;Y^n)}$$

$$\boxed{I_M(X^n;Y^n) \geq I_{ML}(X^n;Y^n)}$$

(c) „Z" VARIES FOR EACH CHANNEL USE

(d) $Y_i = (X_i + Z_i) \bmod 2 \qquad X_i, Y_i, Z_i \in \{0,1\}$

$P(Z_i = 1) = p = 1 - P(Z_i = 0)$

X_i Z_i \ Y	0	1	P(X Y)
0 0	(1-p)/2	0	(1-p)/2
0 1	0	p/2	p/2
1 0	0	(1-p)/2	(1-p)/2
1 1	p/2	0	p/2
$P(Y_i)$	1/2	1/2	

$$I(X;Y) = H(Y) - H(Y|X) = 1 - H(Y|X)$$

$$H(Y|X) = P(X=0)\left[\frac{p}{2} ld\frac{2}{p} + \frac{1-p}{2} ld\frac{2}{1-p}\right] + P(X=1)\left[\frac{p}{2} ld\frac{2}{p} + \frac{1-p}{2} ld\frac{2}{1-p}\right] =$$

$$= \frac{p}{2} ld\frac{2}{p} + \frac{(1-p)}{2} ld\frac{2}{1-p} = \frac{1}{2}\left[p\, ld\frac{2}{p} + (1-p)\, ld\frac{2}{1-p}\right] = \frac{1}{2}\left[p\, ld\frac{1}{p} + p + (1-p) + (1-p)\, ld\frac{1}{1-p}\right] =$$

$$= \frac{1}{2}\left[1 + H(p)\right]$$

$$\boxed{C = \max_{p(x)} I(X;Y) = 1 - \frac{1}{2} - \frac{1}{2}H(p) = \frac{1}{2} - \frac{1}{2}H(p)}$$

(e) SUPPOSE NOW THAT CHANNEL HAS MEMORY OF LENGTH "K" WHICH WE DEFINE TO MEAN THAT "K" CONSECUTIVE INSTANCES OF NOISE Z_i ARE EQUAL, I.E. $Z_1 = Z_2 = \dots = Z_K$ THEN THE NEXT "K" ARE INDEPENDENT OF THE FIRST "K" AND ARE AGAIN EQUAL:

$Z_{K+1} = Z_{K+2} = \dots = Z_{2K}$ AND SO FORTH

WE STILL HAVE THAT $P(Z_1=1) = p = 1 - P(Z_1=0)$

$$C^{(K)} = \max_{p(x_1 \dots x_K)} \boxed{\frac{1}{K} I(X_1^K ; Y_1^K)}$$

$$I(X^n, Y^n) = H(Y^n) - H(Y^n | X^n)$$

$Y = (Z + X) \bmod 2$ → CONST IN K RANGE

$H(Y$

$$I(X^n, Y^n) = H(Y^n) - \sum_{i=1}^{K} H(Y_i | X_1^n Y_1^{i-1}) - \sum_{i=K+1}^{2K} H(Y_i | X_1^n Y_1^{i-1})$$

$$C^{(n)} = \max_{p(x_1^n)} I(X_1^n ; Y_1^n) = \max_{p(x_1^K)} I(X_1^K ; Y_1^K) + \max_{p(x_{K+1}^{2K})} I(X_{K+1}^{2K} ; Y_{K+1}^{2K}) + \dots$$

NEMA NESIGURNOSTI, Z OSTANE KONSTANTNO

$$= H(Y^n) - \underbrace{H(Y_1 | X_1)}_{\frac{1}{2}(1+H(p))} - \underbrace{H(Y_2 | X_2)}_{=0} - 0 \dots 0 -$$

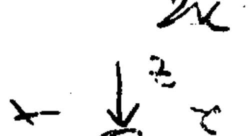
$$- \underbrace{H(Y_{K+1} | X_{K+1})}_{\frac{1}{2}(1+H(p))} + 0 + 0 \dots + H(Y_{2K+1} | X_{2K+1}) + 0 \dots$$

$$+ H(Y_{n-K+1} | X_{n-K+1}) \dots 0$$

$$= n - \frac{n}{K} \cdot \frac{1}{2}(1 + H(p))$$

$$I(X^n ; Y^n) \le n C \qquad C \ge \frac{I(X^n ; Y^n)}{n} =$$

$$= 1 - \frac{1}{2K}(1 + H(p))$$



$$I(X;Y) = H(Y) - H(Y|X$$

$$= (Y = X + Z) = H(Y) - \underbrace{\sum_x p(x) \cdot H(Z+x | X=x)}_{H(Z)}$$

$$= 1 - H(p)$$

DIFFERENCIAL ENTROPY

DEFINITION: LET X BE A RANDOM VARIABLE WITH CUMULATIVE DISTRIBUTION FUNCTION. $F(x) = Pr(X \leq x)$. IF $F(X)$ IS CONTINUOUS THE RANDOM VARIABLE IS SAID TO BE CONTINUOUS. LET $f(x) = F'(x)$ WHEN THE DERIVATIVE IS DEFINED. IF $\int_{-\infty}^{\infty} f(x)\,dx = 1$, $f(x)$ IS CALLED PROBABILITY DENSITY FUNCTION FOR X. THE SET WHERE $f(x) > 0$ IS CALLED SUPPORT SET OF X.

DEFINITION THE DIFFERENTIAL ENTROPY $h(X)$ OF A CONTINUOUS RANDOM VARIABLE X WITH DENSITY $f(x)$ IS DEFINED AS:

$$h(X) = -\int_S f(x) \cdot \operatorname{ld} f(x)\,dx$$

WHERE S IS SUPPORT SET OF RANDOM VARIABLE

$$h(f) = h(X)$$

EXAMPLE 8.1.1. (UNIFORM DISTRIBUTION). CONSIDER A RANDOM VARIABLE DISTRIBUTED UNIFORMLY FROM 0 TO a SO THAT ITS DENSITY IS $1/a$ FROM 0 TO a AND 0 ELSEWHERE. THEN:

$$h(x) = +\int_0^a \frac{1}{a} \operatorname{ld} a\,x = \frac{1}{a} \operatorname{ld} a \cdot x\Big|_0^a = \operatorname{ld} a$$

FOR $a < 1$ $\operatorname{ld} a < 0$, AND THE DIFFERENTIAL ENTROPY IS NEGATIVE! HOWEVER $2^{h(X)} = 2^{\operatorname{ld} a} = a$ IS VOLUME OF THE SUPPORT SET WHICH IS ALWAYS NEGATIVE.

EXAMPLE 8.1.2 (NORMAL DISTRIBUTION) $X \sim \phi(x) = \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$

$$h(X) = +\int_{-\infty}^{\infty} \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \left[\left(\ln(\sigma\sqrt{2\pi})\right) + \left(\frac{x^2}{2\sigma^2}\right)\right] dx$$

$$= \left| \int_{-\infty}^{\infty} e^{-x^2} = \sqrt{\pi} \right| = \underbrace{+\frac{1}{2\sigma\sqrt{2\pi}} \ln(\sigma\sqrt{2\pi})}_{K} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx$$

$$= \left| \frac{x^2}{2\sigma^2} = u \qquad \frac{2x\,dx}{2\sigma^2} = du \qquad x^2 = 2\sigma^2 u \right| =$$

$$= K \cdot \int_{-\infty}^{\infty} 2\sigma^2 u\, e^{-u} \frac{2\sigma^2 du}{2\sigma\sqrt{2u}} = \frac{K \cdot 2\sigma^3}{\sqrt{2}} \int_{-\infty}^{\infty} \sqrt{u}\, e^{-x}\,dx =$$

$$= \left| u = \sqrt{x} \qquad du = \frac{1}{2} x^{-\frac{1}{2}} dx \qquad v = \int e^{-x} dx = -e^{-x} \right| =$$

$$= -\sqrt{x} \cdot e^{-x} + \int_{-\infty}^{\infty} \frac{e^{-x}}{\sqrt{x}}\,dx$$

$$\int_{-\infty}^{\infty} t^2 e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$$

$u = t^2 \quad du = 2t\,dt$

$v = \int e^{-t^2} dt$

$$\int_{-\infty}^{\infty} t \cdot t e^{-t^2} dt = \left| \begin{array}{l} u = t \\ du = dt \\ v = \frac{1}{2}\int e^{-t^2} dt^2 = -\frac{1}{2} e^{-t^2} \end{array} \right. \quad v = \int t e^{-t^2} dt = \int \frac{e^{-t^2} d(t^2)}{2}$$

$$= -\frac{t}{2} e^{-t^2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{e^{-t^2}}{2} dt = -\lim_{t \to \infty} \frac{t}{e^{t^2}} + \lim_{t \to -\infty} \frac{t}{e^{t^2}} + \frac{1}{2}\sqrt{\pi} = -0 + 0 + \frac{\sqrt{\pi}}{2}$$

PROOF

$$h(x) = \frac{1}{2\sigma^2\sqrt{2\pi}} \ln(\sigma\sqrt{2\pi}) \int_{-\infty}^{\infty} t^2 e^{-\frac{t^2}{2\sigma^2}} dt =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \ln(\sigma\sqrt{2\pi})\, \sigma\sqrt{2} \int_{-\infty}^{\infty} \frac{t^2}{2\sigma^2} e^{-\frac{t^2}{2\sigma^2}} d\left(\frac{t}{\sigma\sqrt{2}}\right) =$$

$$= \frac{1}{\sqrt{\pi}} \ln(\sigma\sqrt{2\pi}) \cdot \frac{\sqrt{\pi}}{2} = \frac{1}{2} \ln \sigma\sqrt{2\pi} =$$

$$h(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2\sigma^2}} \left[\ln(\sigma\sqrt{2\pi}) + \frac{t^2}{2\sigma^2}\right] dt$$

$$= \frac{\sigma\sqrt{2}}{\sigma\sqrt{2\pi}} \Bigg[\ln(\sigma\sqrt{2\pi}) \underbrace{\int_{-\infty}^{\infty} e^{-\frac{t^2}{2\sigma^2}} d\left(\frac{t}{\sigma\sqrt{2}}\right)}_{\sqrt{\pi}} - \underbrace{\int_{-\infty}^{\infty} \frac{t^2}{2\sigma^2} e^{-\frac{t^2}{2\sigma^2}} d\left(\frac{t}{\sigma\sqrt{2}}\right)}_{\frac{\sqrt{\pi}}{2}}\Bigg]$$

$$= \frac{1}{\sqrt{\pi}} \left[\ln(\sigma\sqrt{2\pi}) \cdot \sqrt{\pi} - \frac{\sqrt{\pi}}{2}\right] = \left[\ln \sigma\sqrt{2\pi} - \frac{1}{2}\right]$$

$$= \ln \sigma\sqrt{2\pi} - \frac{1}{2} \ln e = \ln \sigma\sqrt{2\pi e} \text{ nats}$$

$$h(x) = \mathrm{ld}(\sigma\sqrt{2\pi e}) \text{ bits}$$

$$h(x) = \frac{1}{2} \mathrm{ld}(2\pi e\sigma^2) \text{ bits}$$

8.2 AEP FOR CONTINUOUS RANDOM VARIABLES

- FOR THE SEQUENCE OF I.I.D RANDOM VARIABLES $p(x_1, x_2 \ldots x_n)$ IS CLOSE TO $2^{-H(x)}$ WITH HIGH PROBABILITY !!! (FOR DISCRETE RANDOM VARIABLES)

THEOREM 8.2.1 LET $X_1, X_2, \ldots X_n$ BE A SEQUENCE OF RANDOM VARIABLES DRAWN I.I.D ACCORDING TO DENSITY $f(x)$. THEN

$$-\frac{1}{n} \operatorname{ld} f(X_1, X_2, \ldots X_n) \to E[-\operatorname{ld} f(X)] = h(X)$$ IN PROBABILITY

DEFINITION: FOR $\varepsilon > 0$ AND ANY n, WE DEFINE THE TYPICAL SET $A_\varepsilon^{(n)}$ WITH RESPECT TO $f(x)$

$$A_\varepsilon^{n} = \left\{ (x_1, x_2, \ldots x_n) \in S^n : \left| -\frac{1}{n} \operatorname{ld} f(x_1, x_2, \ldots x_n) - h(X) \right| \le \varepsilon \right\}$$

WHERE $f(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} f(x_i)$

THE ANALOG OF CARDINALITY OF THE TYPICAL SET FOR THE DISCRETE CASE IS THE VOLUME OF THE TYPICAL SET FOR CONTINUOUS RANDOM VARIABLES.

DEFINITION THE $\operatorname{Vol}(A)$ OF A SET $A \subset \mathbb{R}^n$ IS DEFINED AS:

$$\operatorname{Vol}(A) = \int_A dx_1\, dx_2 \ldots dx_n$$

THEOREM 8.2.2 THE TYPICAL SET $A_\varepsilon^{(n)}$ HAS THE FOLLOWING PROPERTIES:

1. $\Pr(A_\varepsilon^{(n)}) > 1-\varepsilon$ FOR n SUFFICIENTLY LARGE
2. $\operatorname{Vol}(A_\varepsilon^{n}) \le 2^{n(h(X)+\varepsilon)}$ FOR n SUFFICIENTLY LARGE
3. $\operatorname{Vol}(A_\varepsilon^{(n)}) \ge (1-\varepsilon)\, 2^{n(h(X)-\varepsilon)}$ FOR n SUFFICIENTLY LARGE

$$1 = \int_{S^n} f(x_1, x_2, \ldots x_n)\, dx_1\, dx_2 \ldots dx_n \ge \int_{A_\varepsilon^{(n)}} f(x_1, x_2 \ldots x_n)\, dx_1 \ldots dx_n$$

$$= \int_{A_\varepsilon^{(n)}} f(x_1^n)\, dx_1^n \ge \int_{A_\varepsilon^n} 2^{-n(h(X)+\varepsilon)}\, dx_1^n = 2^{-n(h(X)+\varepsilon)} \operatorname{Vol}(A)$$

(T.8.2.1)

$$\operatorname{Vol}(A) \le 2^{n(h(X)+\varepsilon)}$$

$$(1-\varepsilon) \leq \int_{A_\varepsilon^n} f(x_1^n)\, dx_1 dx_2 \cdots dx_n \leq \int 2^{-n(h(X)-\varepsilon)} dx_1 dx_2 \cdots dx_n$$

$$2^{-n(h(X)-\varepsilon)} \cdot Vol(A)$$

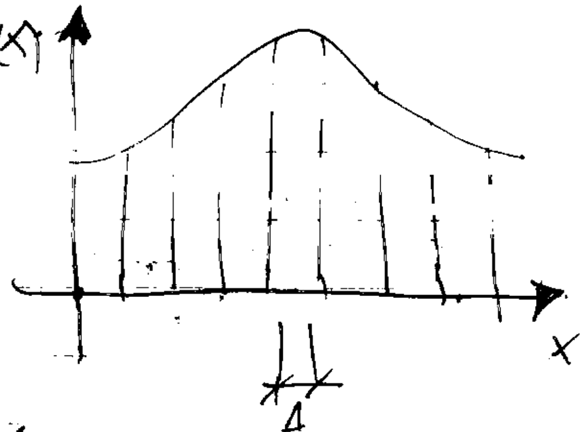
$$Vol(A) \geq (1-\varepsilon)\, 2^{+n(h(X)-\varepsilon)}$$

$$(1-\varepsilon)\, 2^{+n(h(X)-\varepsilon)} \leq Vol(A) \leq 2^{n(h(X)+\varepsilon)}$$

THEOREM 8.2.3 THE SET $A_\varepsilon^{(n)}$ IS THE SMALLEST VOLUME SET WITH PROBABILITY $\geq 1-\varepsilon$, TO FIRST ORDER IN EXPONENT.

SMALL ENTROPY IMPLIES THAT RANDOM VARIABLE IS CONFINED TO A SMALL EFFECTIVE VOLUME AND HIGH ENTROPY INDICATES THAT THE RANDOM VARIABLE IS WIDELY DISPERSED. MMV

8.3 RELATION OF DIFFERENTIAL ENTROPY TO DISCRETE ENTROPY



$$f(x_i)\cdot\Delta = \int_{i\Delta}^{(i+1)\Delta} f(x)\,dx$$

BY MEAN VALUE THEOREM

CONSIDER A QUANTIZED RANDOM VARIABLE X^Δ

$X^\Delta = x_i$ IF $i\cdot\Delta \leq X < (i+1)\Delta$

THE PROBABILITY THAT $X^\Delta = x_i$ IS:

$$p_i = \int_{i\Delta}^{(i+1)\Delta} f(x)\,dx = f(x_i)\Delta$$

$$H(X^\Delta) = -\sum f(x_i)\cdot\Delta\, ld\, f(x_i)\Delta = -\sum f(x_i)\Delta \cdot ld\, f(x_i)$$

$$-\underbrace{\sum f(x_i)\cdot\Delta}_{p_i} \cdot ld\,\Delta = -\sum f(x_i)\Delta \cdot ld\, f(x_i) - ld\,\Delta =$$

$$\sum_\Delta \int_{i\Delta}^{i+1\Delta} f(x)\,dx = 1$$

$$H(X^\Delta) + ld\,\Delta \to h(f) = h(X) \quad \text{as } \Delta \to 0$$

THEOREM 8.3.1 IF THE DENSITY $f(x)$ OF THE RANDOM VARIABLE X IS RIEMANN INTEGRABLE

$$H(X^\Delta) + \operatorname{ld}\Delta \to h(f) = h(X), \quad \text{AS } \Delta \to 0$$

THUS THE ENTROPY OF AN n-BIT QUANTIZATION OF A CONTINUOUS RANDOM VARIABLE X IS APPROXIMATELY $h(X) + n$.

EXAMPLE 8.3.1 IF X HAS UNIFORM DISTRIBUTION ON $[0,1]$ AND WE LET $\Delta = 2^{-n}$

$$h(X) = -\int_0^1 1 \underbrace{\operatorname{ld} 1}_{0} \, dx = 0 \qquad H(X^\Delta) + \operatorname{ld}\Delta \to 0$$

$$H(X^\Delta) + \operatorname{ld}\frac{1}{2^n} \to 0 \qquad H(X^\Delta) \to \operatorname{ld} 2^n = n$$

$H(X^\Delta) \to n$ — n-BITS SUFFICE TO DESCRIBE X TO n-BIT ACCURACY

2. $X \sim [0, \frac{1}{8}]$ $\qquad h(X) = -\int_0^{1/8} (8) \operatorname{ld} 8 \cdot dx =$

$$= -8 \cdot 3 \left[\frac{1}{8} - 0\right] = -8 \cdot 3 \cdot \frac{1}{8} = -3$$

$H(X^\Delta) \to n-3$ TO DESCRIBE X WITH n bit ACCURACY REQUIRE $n-3$ BITS

3. IF $X \sim N(0, \sigma^2)$ WITH $\sigma^2 = 100$, DESCRIBING X TO n bit ACCURACY WOULD REQUIRE:

$$n + h(X) = n + \frac{1}{2}\operatorname{ld} 2\pi e \sigma^2 = n + \frac{1}{2}\operatorname{ld} 2\pi e \cdot 100$$

$n + h(X) = n + 5.37$ bits

IN GENERAL, $h(X) + n$ IS THE NUMBER OF BITS ON AVERAGE REQUIRED TO DESCRIBE X TO n-bit ACCURACY.

THE DIFFERENTIAL ENTROPY OF DISCRETE RANDOM VARIABLE CAN BE CONSIDERED TO BE $-\infty$. NOTE THAT $2^{-\infty} = 0$ AGREEING WITH THE IDEA THAT SUPPORT SET OF DISCRETE RANDOM VARIABLE IS 0.

8.4 JOINT AND DIFFERENTIAL ENTROPY

DEFINITION: THE DIFFERENTIAL ENTROPY OF A SET $X_1, X_2, \ldots X_n$ OF RANDOM VARIABLES WITH DENSITY $f(x_1, x_2, \ldots x_n)$ IS DEFINED AS:

$$h(x_1, x_2, \ldots x_n) = -\int f(x^n) \log f(x^n)\, dx^n$$

DEFINITION: IF X, Y HAVE JOINT DENSITY FUNCTION $f(x,y)$ WE CAN DEFINE THE CONDITIONAL DIFFERENTIAL ENTROPY $h(X|Y)$ AS:

$$h(X|Y) = -\int f(x,y) \log f(x|y)\, dx\, dy$$

$$f(x|y) = f(x,y) / f(y)$$

$$h(X|Y) = h(X,Y) - h(Y)$$

THEOREM 8.4.1 (ENTROPY OF MULTIVARIATE NORMAL DISTRIBUTION) LET $X_1, X_2, \ldots X_n$ HAVE A MULTIVARIATE NORMAL DISTRIBUTION WITH MEAN μ AND COVARIANCE MATRIX K. THEN

$$h(X_1, X_2, \ldots X_n) = h(\mathcal{N}_n(\mu, K)) = \frac{1}{2} \log (2\pi e)^n |K| \text{ bits}$$

$-|K|$ DETERMINANT OF K.

PROOF: THE PROBABILITY DENSITY FUNCTION OF $(X_1, X_2, \ldots X_n)$ IS:

$$f(\mathbf{x}) = \frac{1}{(\sqrt{2\pi})^n |K|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T K^{-1} (\mathbf{x}-\mu)}$$

$$h(f) = -\int f(x) \left[-\frac{1}{2}(\mathbf{x}-\mu)^T K^{-1} (\mathbf{x}-\mu) - \ln (\sqrt{2\pi})^n |K|^{1/2} \right] d\mathbf{x} =$$

$$= \frac{1}{2} E\left[\sum_{i,j} (X_i - \mu_i)(K^{-1})_{ij} (X_j - \mu_j) \right] + \frac{1}{2} \ln (2\pi)^n |K| =$$

$$= \frac{1}{2} E\left[\sum_{i,j} (X_i - \mu_i)(X_j - \mu_j)(K^{-1})_{ij} \right] + \frac{1}{2} \ln (2\pi)^n |K| =$$

$$= \frac{1}{2} \sum_{i,j} E\left[(X_j - \mu_j)(X_i - \mu_i) \right] (K^{-1})_{ij} + \frac{1}{2} \ln (2\pi)^n |K| =$$

$$= \frac{1}{2} \sum_j \sum_i K_{ji} (K^{-1})_{ij} + \frac{1}{2} \ln (2\pi)^n |K| =$$

$$= \frac{1}{2} \sum_j (K K^{-1})_{jj} + \frac{1}{2} \ln (2\pi)^n |K| = \frac{1}{2} \sum_j I_{jj} + \frac{1}{2} \ln (2\pi)^n |K|$$

$$= \frac{n}{2} + \ln (2\pi)^n |K| = \frac{1}{2} \ln e^n + \frac{1}{2} \ln (2\pi)^n |K| = \frac{1}{2} \ln (2\pi e)^n |K| \text{ nats}$$

MMU $h(f) = h(X) = \frac{1}{2} \log (2\pi e)^n |K| \text{ bits}$

8.5 RELATIVE ENTROPY AND MUTUAL INFORMATION

DEFINITION: THE RELATIVE ENTROPY (KULLBACK-LEIBLER DISTANCE) $D(f\|g)$ BETWEEN TWO DENSITIES f & g IS DEFINED AS:

$$\boxed{D(f\|g) = \int f \log \frac{f}{g}}$$

$D(f\|g)$ IS FINITE ONLY IF SUPPORT SET OF f IS CONTAINED IN THE SUPPORT SET OF g.

DEFINITION: THE MUTUAL INFORMATION $I(X;Y)$ BETWEEN TWO RANDOM VARIABLES WITH JOINT DENSITY $f(x,y)$ IS DEFINED AS:

$$\boxed{I(X;Y) = \int f(x,y) \log \frac{f(x,y)}{f(x)f(y)}\, dx\, dy}$$

$$I(X;Y) = h(X) - h(X|Y) = h(Y) - h(Y|X) = h(X) - h(X,Y) + h(Y)$$

$$\boxed{I(X;Y) = D\left[f(x,y)\,\|\,f(x)f(y)\right]}$$

$$I(X^\Delta;Y^\Delta) = H(X^\Delta) - H(X^\Delta|Y^\Delta) \approx h(X) - \log\Delta - h(X|Y) + \log\Delta = h(X) - h(X|Y) = \underline{I(X;Y)}$$

MORE GENERALLY WE CAN DEFINE MUTUAL INFORMATION IN TERMS OF FINITE PARTITIONS OF THE RANGE OF RANDOM VARIABLES. LET $\mathcal{X}$ BE RANGE OF A RANDOM VARIABLE X. A PARTITION $\mathcal{P}$ OF $\mathcal{X}$ IS FINITE COLLECTION OF DISJOINT SETS P_i SUCH AS $\bigcup_i P_i = \mathcal{X}$. THE QUANTIZATION OF $\mathcal{X}$ BY $\mathcal{P}$ (DENOTED $[X]_\mathcal{P}$) IS DISCRETE RANDOM VARIABLE DEFINED AS:

$$P_r([X]_\mathcal{P} = i) = P_r(X \in P_i) = \int_{P_i} dF(x)$$

$$\boxed{I(X;Y) = \sup_{\mathcal{P},\mathcal{Q}} I([X]_\mathcal{P};[Y]_\mathcal{Q})}$$

WHERE SUPREMUM IS OVER ALL FINITE PARTITIONS $\mathcal{P}$ AND $\mathcal{Q}$.

THIS IS MASTER DEFINITION OF MUTUAL INFORMATION THAT ALWAYS APPLIES.

EXAMPLE 8.5.1 (MUTUAL INFORMATION BETWEEN CORRELATED GAUSSIAN RANDOM VARIABLES WITH CORRELATION ρ. LET $(X,Y) \sim N(0,K)$

$$K = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \qquad |K| = \sigma^4 - \rho^2\sigma^4 = \sigma^4(1-\rho^2)$$

$$h(X) = \frac{1}{2} ld(2\pi e)\sigma^2 = h(Y) \qquad h(X,Y) = \frac{1}{2} ld(2\pi e)^2|K|$$

$$= \frac{1}{2} ld(2\pi e)^2(\sigma^2\sqrt{1-\rho^2})^2 = \frac{1}{2} ld(2\pi e)^2\sigma^4(1-\rho^2)$$

$$I(X;Y) = h(X) + h(Y) - H(X,Y) = ld(2\pi e)\sigma^2 -$$

$$- \frac{1}{2} ld(2\pi e)^2\sigma^4(1-\rho^2) = ld(2\pi e)\sigma^2 - ld(2\pi e)\sigma^2 -$$

$$- \frac{1}{2} ld(1-\rho^2) \qquad \boxed{h(X) = -\frac{1}{2} ld(1-\rho^2)} \quad \text{MMV}$$

BIVARIATE GAUSSIAN DISTRIBUTION (DENSITY FUNCTION) — WOLFRAM MATHWORLD

$$f(x_1 x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{z}{2(1-\rho^2)}\right]$$

$$z = \frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}$$

$$f(x_1) = \int_{-\infty}^{\infty} f(x_1 x_2)\, dx_2 = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}}$$

$$f(x_2) = \int_{-\infty}^{\infty} f(x_1 x_2)\, dx_2 = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}}$$

$$\mu_1 = \mu_2 = 0 \qquad z = \frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}$$

$$\boxed{f(x,y) = \frac{1}{(\sqrt{2\pi})^2|K|^{1/2}} e^{-\frac{1}{2}\left(\begin{bmatrix}x\\y\end{bmatrix}-\mu\right)^T \cdot K^{-1}\left(\begin{bmatrix}x\\y\end{bmatrix}-\mu\right)}}$$

$\boxed{\mu=0}$

$$f(x,y) = \frac{1}{2\pi|K|^{1/2}} e^{-\frac{1}{2}\begin{bmatrix}x\\y\end{bmatrix}^T K^{-1}\cdot\begin{bmatrix}x\\y\end{bmatrix}}$$

$$[x\ y]\begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = [x\ y]\begin{bmatrix} \sigma^2 x + \rho\sigma^2 y \\ \rho\sigma^2 x + \sigma^2 y \end{bmatrix} =$$

$$= \sigma^2x^2 + \rho\sigma^2 xy + \rho\sigma^2 xy + \sigma^2y^2 = \sigma^2x^2 + 2\rho\sigma^2xy + \sigma^2y^2$$

$$f(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho}} \exp\left[-\frac{\sigma^2x^2 + 2\rho\sigma^2xy + \sigma^2y^2}{2}\right]$$

$$K^{-1} = \frac{1}{|K|}\begin{bmatrix} \sigma^2 & -\rho\sigma^2 \\ -\rho\sigma^2 & \sigma^2 \end{bmatrix} = \frac{1}{\sigma^4(1-\rho^2)}\begin{bmatrix} \sigma^2 & -\rho\sigma^2 \\ -\rho\sigma^2 & \sigma^2 \end{bmatrix} =$$

$$= \frac{1}{\sigma^2(1-\rho^2)}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$

$$\frac{1}{2}\frac{[x\ y]}{\sigma^2(1-\rho^2)}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{2\cdot\sigma^2(1-\rho^2)}[x\ y]\begin{bmatrix} x-\rho y \\ -\rho x+y \end{bmatrix} =$$

$$= \frac{1}{2\sigma^2(1-\rho^2)}\left(x^2-\rho xy - \rho xy + y^2\right) = \frac{x^2-2\rho xy+y^2}{2\cdot\sigma^2(1-\rho^2)}$$

$$z = \frac{x^2-2\rho xy+y^2}{\sigma^2}$$

$$f(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\, e^{-\frac{z}{2(1-\rho^2)}}$$

ISTO KAKO NA WOLFRAM MATHWORLD

$$f(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\, e^{-\frac{x^2-2\rho xy+y^2}{2\sigma^2(1-\rho^2)}}$$

$$\int f(x,y)\,dx = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{y^2}{2\sigma^2}}$$ } GAUSSIAN $\mathcal{N}(0,\sigma^2)$

(!) ⇒ IF $\rho=0$, X AND Y ARE INDEPENDENT AND THE MUTUAL INFORMATION IS $\emptyset$. IF $\rho=\pm 1$, X AND Y ARE PERFECTLY CORRELATED AND THE MUTUAL INFORMATION IS INFINITE !!! [MMI]

8.6 PROPERTIES OF DIFFERENTIAL ENTROPY, RELATIVE ENTROPY, AND MUTUAL INFORMATION

THEOREM 8.6.1 $D(f\|g) \geq \emptyset$
WITH EQUALITY IF $f=g$ ALMOST EVERYWHERE (A.E.)

PROOF: S – SUPPORT SET OF f

$$-D(f\|g) = \int_S f \log\frac{g}{f} \leq \log\int_S f\frac{g}{f} = \log\int_S g \leq \log 1 = \emptyset$$

(↑ JENSEN INEQUALITY)

$E[f(x)] \leq f(E[x])$ ⇒ CONCAVE FUNCTION $\quad \boxed{D(f\|g)\geq\emptyset}$

EQUALITY IS FOR: $\boxed{f=g}$

- COROLLARY $I(X;Y) \geq 0$ WITH EQUALITY IF $(\tilde{X}, \tilde{Y})$

$$I(X;Y) = \int_{x,y} f(xy) \, ld \frac{f(xy)}{f(x) \cdot f(y)} \, dx\,dy = D(f(xy) \| f(x)f(y))$$

$$I(X;Y) = \sum_{xy} p(xy) \, ld \frac{p(xy)}{p(x) \cdot p(y)} = -\sum_{xy} p(xy) \, ld \frac{1}{p(xy)} +$$

$$- \sum_{xy} p(xy) \, ld \, p(x) - \sum_{xy} p(x,y) \, ld \, p(y) =$$

$$= -H(X,Y) - \sum_{x} ld \, p(x) \sum_{y} p(xy) - \sum_{y} ld \, p(y) \underbrace{\sum_{x} p(x,y)}_{p(y)}$$

$$= -H(X,Y) + H(X) + H(Y) = H(X) + H(Y) -$$

$$- H(X) - H(Y|X) = H(Y) - H(Y|X)$$

- COROLLARY $h(X|Y) \leq h(X)$ WITH EQUALITY $(\tilde{X}, \tilde{Y})$
- THEOREM 8.6.2 CHAIN RULE FOR DIFFERENTIAL ENTROPY.

$$h(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} h(X_i | X_1^{i-1})$$

- COROLLARY:

$$h(X_1, X_2, \ldots X_n) \leq \sum_{i=1}^{n} h(X_i)$$

WITH EQUALITY IF $(X_1, X_2, \ldots X_n)$ ARE INDEPENDENT.

- APPLICATION (HADAMARD'S INEQUALITY) IF WE LET
- $X \sim N(0, K)$ BE A MULTIVARIATE NORMAL RANDOM VARIABLE, CALCULATING THE ENTROPY IN THE ABOVE INEQUALITY GIVES:

$$\boxed{|K| \leq \prod_{i=1}^{n} K_i}$$

THEOREM 8.6.3 $h(X+c) = h(X)$ TRANSLATION DOESN'T CHANGE THE DIFFERENTIAL ENTROPY.

THEOREM 8.6.4 $h(aX) = h(X) + ld|a|$

PROOF: LET $Y = aX$ $\quad p(y) = \left. \frac{p(x)}{\frac{dy}{dx}} \right|_{x = f(y)} =$

$$= \frac{p(\frac{y}{a})}{|a|} \qquad \boxed{f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right)}$$

$h(aX) = -\int f_Y(y) \, ld \, f_Y(y) \, dy = \int \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \cdot ld \frac{1}{|a|} f_X\left(\frac{y}{a}\right) dy.$

$= \int \frac{1}{|a|} f_X(x) \cdot ld \, f_X(x) \, |a| \, dx + \int f_Y(y) \, ld|a| \, dy =$

$= -\int f_X(x) \, ld \, f_X(x) \, dx + ld|a| = h(X) + ld|a|$

COROLLARY: $\boxed{h(AX) = h(X) + ld|\det(A)|}$

→ THEOREM 8.6.4 FOR VECTOR-VALUED RANDOM VAR.

- WE NOW SHOW THAT MULTIVARIATE NORMAL DISTRIBUTION MAXIMIZES THE ENTROPY OVER ALL DISTRIBUTIONS WITH SAME COVARIANCE.

THEOREM 8.6.5 LET THE RANDOM VECTOR $X \in \mathbb{R}^n$ HAVE ZERO MEAN AND COVARIANCE $K = E[XX^T]$ (i.e. $K_{ij} = E[X_i X_j]$, $1 \le i, j \le n$). THEN $h(X) \le \frac{1}{2} ld (2\pi e)^n |K|$ WITH EQUALITY IF: $X \sim N(0, K)$

PROOF: LET $g(x)$ BE ANY DENSITY SATISFYING $\int g(x) x_i x_j dx = K_{ij}$ FOR ALL i, j. LET ϕ_K BE DENSITY OF $N(0,K)$ VECTOR AS GIVEN IN (8.35) WHERE WE SET $\mu = 0$. NOTE THAT $ld \, \phi_K(X)$ IS QUADRATIC FORM AND $\int x_i x_j \phi_K(X) dX = K_{ij}$ THEN

$0 \le D(g \| \phi_K) = \int g \, ld \frac{g}{\phi_K} = -h(g) - \int g \, ld \, \phi_K =$

$-h(g) - \int \phi_K \, ld \, \phi_K = -h(g) + h(\phi_K)$ $\boxed{h(\phi_K) \ge h(g)}$

WHERE THE SUBSTITUTION $\int g \, ld \, \phi_K = \int \phi_K \, ld \, \phi_K$ FOLLOWS FROM THE FACT THAT g AND ϕ_K YIELD THE SAME MOMENTS OF QUADRATIC FORM $ld \, \phi_K(X)$

$\boxed{h(g) \le h(\phi_K) = \frac{1}{2} ld (2\pi e)^n |K|}$

PROBLEMS (CONTINUE FROM NAbg)

8.1 Differential entropy. EVALUATE DIFFERENTIAL ENTROPY: $h(X) = -\int f \ln f$ FOR THE FOLLOWING:

(a) THE EXPONENTIAL DENSITY $\boxed{f(x) = \lambda e^{-\lambda x}}$ $x \ge 0$

(b) LAPLACE DENSITY $f(x) = \frac{1}{2} \lambda e^{-\lambda |x|}$.

(c) THE SUM OF X_1 AND X_2 WHERE X_1 AND X_2 ARE INDEPENDENT RANDOM VARIABLES (NORMAL

WITH MEANS μ_i AND VARIANCES σ_i^2, $i=1,2$.

(a) $f(x) = \lambda e^{-\lambda x}$ $\qquad h(X) = -\int_0^\infty f(t)\ln f(t)\,dt =$

$$= -\int_0^\infty \lambda e^{-\lambda t}\ln(\lambda\cdot e^{-\lambda t})\,dt = -\lambda\left[\int_0^\infty e^{-\lambda t}\ln\lambda\,dt + \int_0^\infty e^{-\lambda t}\ln e^{-\lambda t}\,dt\right]$$

$$= -\lambda\left[\ln(\lambda)\int_0^\infty e^{-\lambda t}\,dt - \int_0^\infty \lambda t e^{-\lambda t}\,dt\right] =$$

$$= -\lambda\left[\ln(\lambda)\frac{1}{\lambda}\int_0^\infty e^{-\lambda t}\,d(\lambda t) - \lambda\int_0^\infty \underset{u}{t}\,\underset{v'}{e^{-\lambda t}}\,dt\right] =$$

$$= -\lambda\left[\ln(\lambda)\frac{(-1)}{\lambda}\cdot e^{-\lambda t}\Big|_0^\infty - \lambda\left(t\cdot\frac{-e^{-\lambda t}}{\lambda}\Big|_0^\infty + \int_0^\infty \frac{e^{-\lambda t}}{\lambda}\,dt\right)\right]$$

$$\left| v = \int e^{-\lambda t}\,dt = \frac{1}{\lambda}\int e^{-\lambda t}\,d(\lambda t) = \frac{(-1)}{\lambda}e^{-\lambda t}\right|$$

$$= -\lambda\left\{-\frac{\ln(\lambda)}{\lambda}\left[\overset{0}{e^{-\infty}} - \overset{1}{e^{0}}\right] - \lambda\left(\frac{t}{\lambda e^{\lambda t}}\Big|_0^\infty + \frac{-e^{-\lambda t}}{\lambda^2}\Big|_0^\infty\right)\right\} =$$

$$= -\lambda\left\{+\frac{\ln(\lambda)}{\lambda} - \lambda\left(0 - \frac{0}{\lambda} - \frac{1}{\lambda^2}(0-1)\right)\right\}$$

$$= -\ln(\lambda) + \lambda^2\left(+\frac{1}{\lambda^2}\right) = \underline{1 - \ln(\lambda)}$$

(b) $f(x) = \frac{1}{2}\lambda e^{-\lambda|x|}$ $\qquad -\int_{-\infty}^{\infty} f(t)\ln f(t)\,dt = 1 + \ln 2 - \ln(\lambda)$

$$\boxed{h(X) = 1 + \ln\frac{2}{\lambda}} \qquad h(x) = -\frac{\lambda}{2}\int_{-\infty}^{\infty} e^{-\lambda|x|}\left[\ln e^{-\lambda|x|} + \ln\frac{\lambda}{2}\right]dt =$$

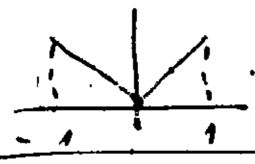
$$= -\frac{\lambda}{2}\left[\int_{-\infty}^{\infty} -\lambda|x|\, e^{-\lambda|x|}\,dx + \int_{-\infty}^{\infty} e^{-\lambda|x|}\ln\frac{\lambda}{2}\,dt\right] =$$

$$= -\frac{\lambda}{2}\left[-\lambda\int_0^\infty x e^{-\lambda x}\,dx - \lambda\int_{-\infty}^0 -x e^{\lambda x}\,dx + \ln\frac{\lambda}{2}\int_0^\infty e^{-\lambda x}\,dx + \ln\frac{\lambda}{2}\int_{-\infty}^0 e^{+\lambda x}\,dx\right] = (*)$$

$$|x| = \begin{cases} x & x \ge 0 \\ -x & < 0 \end{cases}$$

$$\int_{-1}^{1} |x|\,dx = \int_{0}^{1} x\,dx - \int_{-1}^{0} x\,dx = \frac{x^2}{2}\Big|_0^1 - \frac{x^2}{2}\Big|_{-1}^{0} =$$

$$= \frac{1}{2} - \left(0 - \frac{1}{2}\right) = 1$$



$$\frac{1\cdot 1}{2} + \frac{1\cdot 1}{2} = 1$$

$$(*) = -\frac{\lambda}{2}\Big[-\lambda\cdot 2 \int_0^{\infty} x e^{-\lambda x} dx + 2\ln\frac{\lambda}{2} \int_0^{\infty} e^{-\lambda x} dx \Big] =$$

ГАММА ФУНКЦИЯ ... ГАММА Ф.

$$= -\frac{\lambda}{2}\Big[-2\cdot\lambda\cdot\frac{1}{\lambda^2} + 2\ln\frac{\lambda}{2}\cdot\frac{(-1)}{\lambda}\, e^{-\lambda x}\Big|_0^{\infty}\Big] =$$

$$= -\frac{\lambda}{2}\Big[-\frac{2}{\lambda} - \frac{2}{\lambda}\left(\ln\frac{\lambda}{2}\right)\cdot(0-1)\Big] = 1 + \left(\ln\frac{\lambda}{2}\right)\cdot(-1)$$

$$= 1 - \ln\frac{\lambda}{2} = 1 + \ln\frac{2}{\lambda}$$

(e) $x_1 \sim N(\mu_1, \sigma_1)$ $\quad x_2 \sim N(\mu_2, \sigma_2)$

$x = x_1 + x_2$ $\quad f(x) = ?$

$$M_1(-s) = \int f(x_1)\cdot e^{-s x_1} dx_1 \qquad M(s) = \int f(x)\cdot e^{sx} dx$$

$$M(-s) = M_1(-s)\cdot M_2(-s) \qquad W(s) = W_1(s)\cdot W_2(s)$$

MGF. $\quad$ CHARACTERISTIC F.

$$W(s) = \int e^{-sx} \frac{e^{-\frac{(x-\mu)^2}{\sigma^2}}}{\sigma\sqrt{2\pi}} dx \qquad W(s) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{\sigma^2} - sx} dx$$

$$W_1(s) = \sigma\sqrt{2\pi}\; e^{-\frac{s(4\mu-s)}{8\sigma^2}}$$

$$W = (W_1(s))^2 = 2\pi\sigma^2\, e^{-\frac{s(4\mu-s)}{4\sigma^2}}\, e^{+sx}$$

$$f(x) = \int_{-\infty}^{\infty} 2\pi\sigma^2 e^{-\frac{s(4\mu-s)}{4\sigma^2}} ds =$$

13

CONTINUE FROM N16a | CONCAVITY OF DETERMINANTS.

$$|\lambda K_1 + \bar\lambda K_2| \geq |K_1|^\lambda |K_2|^{\bar\lambda} \qquad 0 \leq \lambda \leq 1, \quad \bar\lambda = 1-\lambda$$

$z = x_\theta \quad x_1 \sim N(0, K_1) \quad x_2 \sim N(0, K_2) \quad \theta = \text{Bernoulli}(\lambda)$

$h(z|\theta) \leq h(z)$

$h(z|\theta) = p(\theta=1)\cdot h(z|\theta=1) + p(\theta=2)\, h(z|\theta=2) =$

$= \lambda h(x_1) + \bar\lambda h(x_2)$

$h(x_1) = \frac{1}{2} \operatorname{ld} (2\pi e)^n |K_1| \qquad h(x_2) = \frac{1}{2} \operatorname{ld} (2\pi e)^n |K_2|$

$h(z|\theta) = \frac{\lambda}{2} \operatorname{ld} (2\pi e)^n |K_1| + \frac{\bar\lambda}{2} \operatorname{ld} (2\pi e)^n |K_2| =$

$= \frac{\lambda}{2} \operatorname{ld} (2\pi e)^n + \frac{1}{2} \operatorname{ld} |K_1|^\lambda + \frac{1}{2} \operatorname{ld} (2\pi e)^n + \frac{1}{2} \operatorname{ld} |K_2| -$

$- \frac{\lambda}{2} \operatorname{ld} (2\pi e)^n - \frac{\lambda}{2} \operatorname{ld} |K_2| = \frac{1}{2} \operatorname{ld} (2\pi e)^n + \frac{1}{2} \operatorname{ld} |K_1|^\lambda +$

$+ \frac{\bar\lambda}{2} \operatorname{ld} |K_2| = \frac{1}{2} \operatorname{ld} (2\pi e)^n + \frac{1}{2} \operatorname{ld} |K_1|^\lambda \cdot |K_2|^{\bar\lambda} =$

$= \frac{1}{2} \operatorname{ld} (2\pi e)^n \cdot |K_1|^\lambda \cdot |K_2|^{\bar\lambda}$

$h(z, \theta) = h(z) + h(\theta|z) = h(z) + \lambda h(\theta|z=x_1) + \bar\lambda h(\theta|z=x_2) = h(z) = H(\theta) + h(z|\theta)$

(under $h(\theta|z=x_1)$ … $h(\theta|z=x_2)$: θ; under $H(\theta)$: $H(\lambda)$; under $h(z|\theta)$: $\neq \theta$)

$$h(z) \geq h(z|\theta)$$

$\frac{1+2+3+4+5+6}{6} = \frac{21}{6} = \frac{7}{2}$

$E[zz^T] = E[E[zz^T|\theta]]$

ITERATION FOR TOTAL EXPECTATION

$$E[z] = p(\theta=1)\cdot E[z|\theta=1] + p(\theta=2)\, E[z|\theta=2] =$$

$= \lambda E[x_1] + \bar\lambda E[x_2]$

$E[zz^T] = p(\theta=1)\cdot E[zz^T|\theta=1] + p(\theta=2)\, E[zz^T|\theta=2] =$

$= \lambda \underbrace{E[x_1 x_1^T]}_{|K_1|} + \bar\lambda \underbrace{E[x_2 x_2^T]}_{|K_2|}$

$\sigma^2 = \overline{(x-\bar x)^2} \quad \bar x = 0 \quad \sigma^2 = \overline{x^2} = \int x^2 p(x)\,dx = |K|$

$E[zz^T] = E[\lambda E[x_1] + \bar\lambda E[x^2]] = \lambda|K_1| + \bar\lambda |K_2|$

AMONG JOINT DENSITY WITH GIVEN COVARIANCE MATRIX A MULTIVARIATE GAUSSIAN DENSITY WITH THAT COVARIANCE MATRIX IS MAXIMIZING THE ENTROPY

$$h(Z) \le \frac{1}{2} \operatorname{ld} (2\pi e)^n \left| \lambda K_1 + \bar{\lambda} K_2 \right|$$

$$h(Z) \ge h(Z|\theta) \Rightarrow \frac{1}{2} \operatorname{ld} (2\pi e)^n \left| \lambda K_1 + \bar{\lambda} K_2 \right| \ge \frac{1}{2} \operatorname{ld} (2\pi e)^n |K_1|^\lambda \cdot |K_2|^{\bar{\lambda}}$$

$$\left| \lambda K_1 + \bar{\lambda} K_2 \right| \ge |K_1|^\lambda |K_2|^{\bar{\lambda}}$$

POKAZANO !!!

PROBLEM 8.3 UNIFORMLY DISTRIBUTED NOISE. LET THE INPUT RANDOM VARIABLE X TO A CHANNEL BE UNIFORMLY DISTRIBUTED OVER THE INTERVAL $-1/2 \le x \le 1/2$. LET THE OUTPUT OF THE CHANNEL BE $Y = X + Z$ WHERE THE NOISE RANDOM VARIABLE IS UNIFORMLY DISTRIBUTED OVER THE INTERVAL $-a/2 \le Z \le a/2$

(a) FIND $I(X;Y)$ AS A FUNCTION OF a.

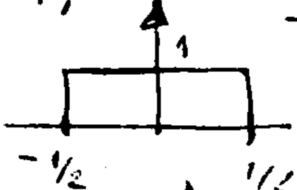
(b) FOR $a=1$ FIND THE CAPACITY OF THE CHANNEL WHEN THE INPUT X IS PEAK-LIMITED; THAT IS THE RANGE OF X IS LIMITED TO $-\frac{1}{2} \le x \le \frac{1}{2}$. WHAT PROBABILITY DISTRIBUTION ON X MAXIMIZES THE MUTUAL INFORMATION $I(X;Y)$?

(c) (OPTIONAL) FIND CAPACITY OF THE CHANNEL FOR ALL VALUES OF a, AGAIN ASSUMING THAT RANGE OF X IS LIMITED TO $-\frac{1}{2} \le x \le \frac{1}{2}$.

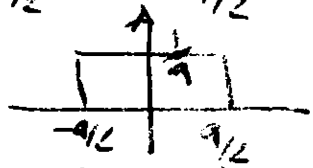
(a) $I(X;Y) = h(X) - h(X|Y) = h(X) - h(Z) = h(Y) - h(Z)$

$h(X|Y) = \int p(Y=y)\, h(y - Z | Y=y)\, dy = h(Z)$

$-h(X) = \int_{-1/2}^{1/2} p(x) \operatorname{ld} p(x)\, dx$



$-h(X) = \int 1 \operatorname{ld} 1\, dx = 0$



$a \cdot \frac{1}{a} = 1$

VAŽNA VO SOLUTIONS UNIVERSITY OF ILLINOIS OBAVEZNO

$+h(Z) = \int_{-a/2}^{a/2} \frac{1}{a} \operatorname{ld} a\, dx = \frac{1}{a} \operatorname{ld} a \left(\frac{a}{2} + \frac{a}{2} \right) = \operatorname{ld} a$

$$\boxed{I(X;Y) = \operatorname{ld} a}$$

(b) $C = \max_{p(x)} I(X;Y) = h(X) - h(Z) = h(X) - \operatorname{ld} a$

$= |a=1| = h(X)$

GAUSSIAN DENSITY FUNCTION MAXIMIZES THE ENTROPY.

$g(x) = \int_{-\infty}^{x} f(t)\,dt$ $\quad \boxed{\frac{dg(x)}{dx} = f(x)}$ FTC 1.

$h(x) \le \frac{1}{2} \operatorname{ld} 2\pi e \sigma^2$

$-h(x) = \int_{-1/2}^{1/2} \gamma(x) \operatorname{ld} \gamma(x)\, dx = \int_{-\infty}^{1/2} \gamma(t) \operatorname{ld} \gamma(t)\, dt$

$\int_{-\infty}^{-1/2} \gamma(x) \operatorname{ld} \gamma(x)\, dx = \int_{-\infty}^{1/2} \gamma(t) \operatorname{ld} \gamma(x)\, dx + \int_{+\infty}^{1/2} \gamma(t) \operatorname{ld} \gamma(t)\, dt$ IF $\gamma(x) = \gamma(-x)$

$\frac{d\,(-h(x))}{dx} = \gamma\left(\frac12\right) \operatorname{ld} \gamma\left(\frac12\right) - \gamma\left(-\frac12\right) \operatorname{ld} \gamma\left(-\frac12\right) = 0$

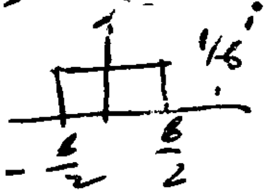
$\gamma\left(\frac12\right) \operatorname{ld} \gamma\left(\frac12\right) = \gamma\left(-\frac12\right) \operatorname{ld} \gamma\left(-\frac12\right) \Rightarrow$

$\boxed{\gamma\left(\frac12\right) = \gamma\left(-\frac12\right)}$

$h(x) = \int_{-1/2}^{1/2} \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \cdot \operatorname{ld} \sigma\sqrt{2\pi}\, e^{\frac{(x-\mu)^2}{2\sigma^2}}\, dx$

$\mu = 0, \quad \sigma = 1 \qquad \boxed{h(X) = 1.32832}$

$\frac12 \operatorname{ld}(2\pi e \sigma^2) = \frac12 \operatorname{ld}(2\pi e) = 4.09419$



$C = h(x) = \int_{-b/2}^{b/2} \frac1b \operatorname{ld} b\, dx = \operatorname{ld} b$

$\frac{dC}{db} = \frac{d}{db}(\operatorname{ld} b) = \frac{1}{\ln 2} \frac{d}{db} \ln b = \frac{1}{b \ln 2}$

$b \to 0 \Rightarrow \boxed{C \to \infty}$

(c) $C = 1.32832 - \operatorname{ld} a \qquad \frac{dC}{da} = 0 \cdot 0 + \frac{1}{a \cdot \ln 2} = 0$

$a \to \infty$

SOLUTION (UNIVERSITY OF ILLINOIS) HW7s.pdf (25)

$I(Y;Z) = h(Y) - h(Y|X) = h(Y) - h(Z)$

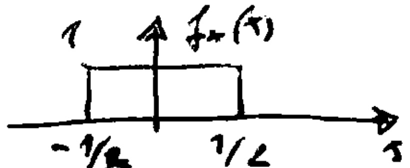
$Y = X + Z$ $\quad h(Z) = \int_{-a/2}^{a/2} \frac{1}{a} \log a \, dt = \log a$

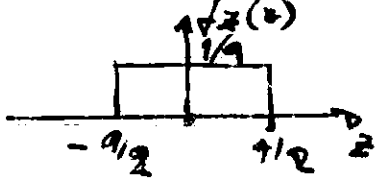
• DISTRIBUTION OF THE SUM OF TWO RANDOM VARIABLES IS CONVOLUTION OF THEIR PDFs

$M(s) = \int_{-\infty}^{\infty} p(x)\, e^{+sx} dx \qquad M(-s) = \int_{-\infty}^{\infty} p(t)\, e^{-st} dt$

$\varphi(s) = \int_{-\infty}^{\infty} p(x)^{+jsx} dx$ — CHARACTERISTIC FUNCTION

• FOR $a < 1$





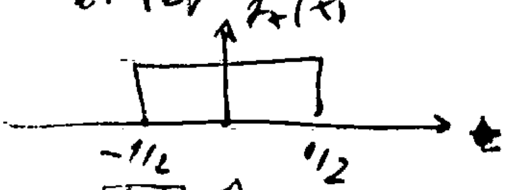
$f_Y(y) = \int_{-\infty}^{\infty} f_X(t)\, f_Z(y-t)\, dt$

$f_Z(z) = \begin{cases} \frac{1}{a} & t = \frac{a}{2} \\ \frac{1}{a} & t = -\frac{a}{2} \end{cases}$

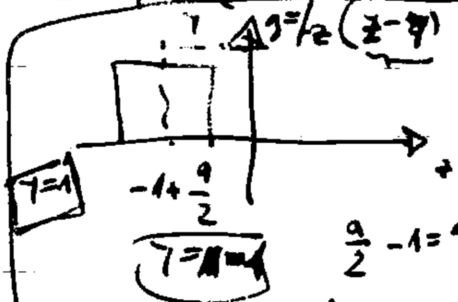
$f_X(t) = \begin{cases} 1 & -\frac{1}{2} \le t \le \frac{1}{2} \\ 0 & t > \frac{1}{2} \quad t < -\frac{1}{2} \end{cases}$ $\qquad f_Z(-t) = \begin{cases} \frac{1}{a} & -\frac{1}{2} < t < \frac{1}{2} \\ 0 & \text{other} \end{cases}$

$f_Z(-t) = \begin{cases} \frac{1}{a} & -\frac{1}{2} < t \le \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$

$Y > \frac{1}{2}$



$\int_{-1/2}^{y} 1 \cdot \frac{1}{a}\, dt = \frac{1}{a}\left(y + \frac{1}{2}\right) = \frac{2y+1}{2a}$

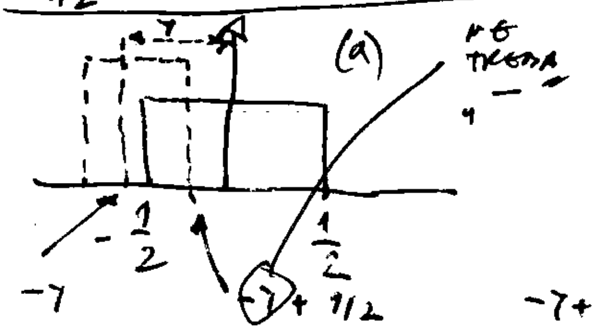


$\frac{a}{2} - 1 = \frac{a-2}{2}$

$g\left(1+\frac{a}{2}\right) = f_Z\left(\frac{a}{2} + 1 - 1\right) = f_Z\left(\frac{a}{2}\right)$

$g(z) = f_Z(z - y)$

$g\left(-1-\frac{a}{2}\right) = f_Z\left(-1-\frac{a}{2}+1\right) = f_Z\left(-\frac{a}{2}\right)$



$\int_{-1/2}^{-y + a/2} \frac{1}{a} \cdot 1\, dt = \frac{1}{a}\left(-y + \frac{a}{2} + \frac{1}{2}\right) = -\frac{1}{a}\left(y - \frac{a+1}{2}\right)$

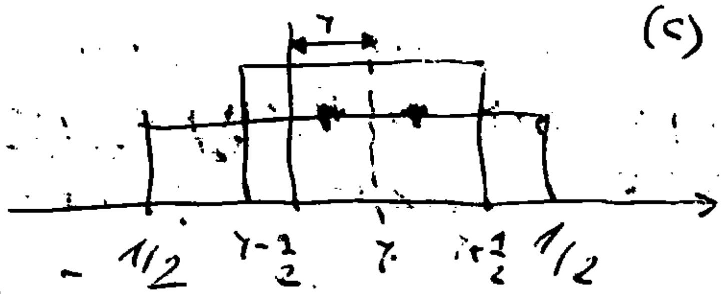
$-y + \frac{1}{2} \ge -\frac{1}{2} \qquad -y \ge -\frac{a}{2} - \frac{1}{2} \qquad y \le \frac{a+1}{2}$

$\gamma + \frac{a}{2} \le \frac{1}{2}$ $\quad$ $\gamma \le \frac{1}{2} - \frac{a}{2}$

$\gamma \le \frac{1-a}{2}$

$\gamma - \frac{a}{2} \ge -\frac{1}{2}$ $\quad$ $\gamma \ge \frac{a-1}{2}$

(c)



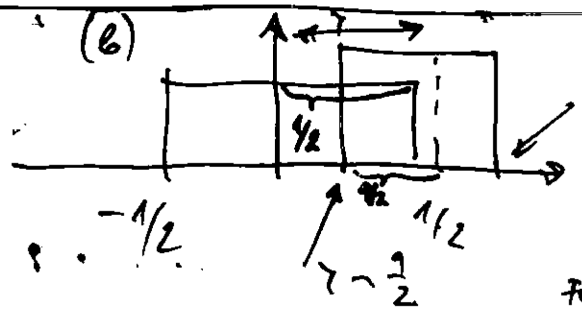
$$\varphi(\gamma) = \int_{\gamma-\frac{a}{2}}^{\gamma+\frac{a}{2}} \frac{1}{a}\, 1\, dt = \frac{1}{a}\left(\gamma + \frac{a}{2} - \gamma + \frac{a}{2}\right) = 1$$

$\varphi(\gamma) = 1$ $\quad$ $\frac{a-1}{2} \le \gamma \le \frac{1-a}{2}$

(a) CONTINUO $\quad$ $-\frac{1}{2} \ge -\gamma - \frac{a}{2}$ $\quad$ $-\gamma \le \frac{a}{2} - \frac{1}{2}$ $\quad$ $\gamma \ge -\frac{a-1}{2}$

FOR: $-\frac{a-1}{2} \le \gamma \le \frac{a+1}{2}$ $\quad$ $\varphi(\gamma) = -\frac{1}{a}\left(\gamma - \frac{a+1}{2}\right)$

(b)



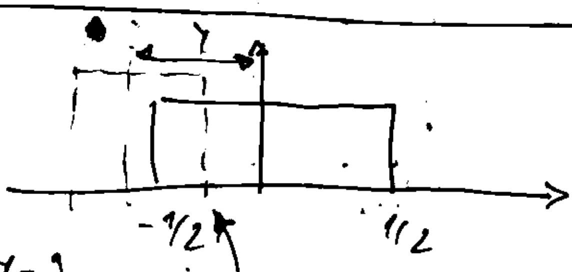
$$\int_{\gamma-a/2}^{1/2} \frac{1}{a} \cdot 1 \cdot dt = \frac{1}{a}\left(\frac{1}{2} - \gamma + \frac{a}{2}\right)$$

FOR $\gamma - \frac{a}{2} \le \frac{1}{2}$ $\quad$ $\gamma \le \frac{1}{2} + \frac{a}{2}$

$\gamma + \frac{a}{2} \ge \frac{1}{2}$ $\quad$ $\gamma \ge \frac{1}{2} - \frac{a}{2}$

$\frac{1-a}{2} \le \gamma \le \frac{a+1}{2}$

(a) REVISITED



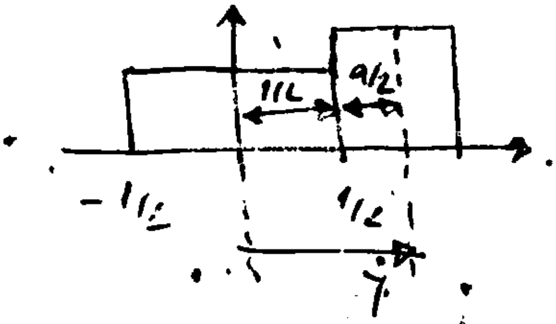
$\gamma - \frac{a}{2} \le -\frac{1}{2}$ $\quad$ $\gamma \le \frac{a-1}{2}$

$\gamma + \frac{a}{2} \ge -\frac{1}{2}$ $\quad$ $\gamma \ge -\frac{a+1}{2}$

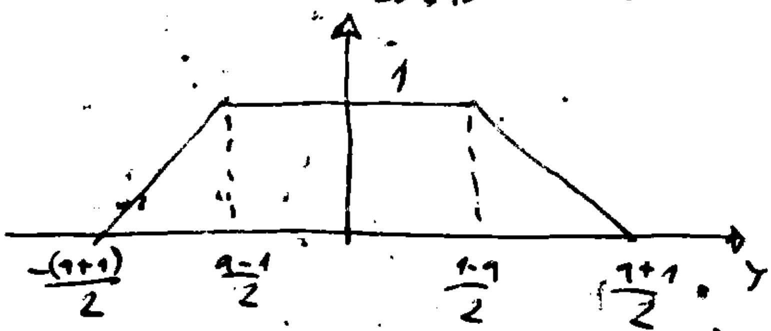
$-\frac{a+1}{2} \le \gamma \le \frac{a-1}{2}$ $\Rightarrow$

$$\int_{-1/2}^{\gamma+\frac{a}{2}} \frac{1}{a} \cdot 1\, dt = \frac{1}{a}\left(\gamma + \frac{a}{2} + \frac{1}{2}\right) = \frac{1}{a}\left(\gamma + \frac{a+1}{2}\right)$$

$$\varphi(\tau) = \begin{cases} \frac{1}{a}\left(\tau + \frac{a+1}{2}\right) & -\frac{a+1}{2} \le \tau \le \frac{a-1}{2} \quad (a) \\ 1 & \frac{a-1}{2} \le \tau \le \frac{1-a}{2} \quad (c) \\ -\frac{1}{a}\left(\tau - \frac{a+1}{2}\right) & \frac{1-a}{2} \le \tau \le \frac{a+1}{2} \end{cases}$$



VO OVOJ EKSTREMEN SLUČAJ IMAM:



WE CAN SEE τ AS TWO DISJOINT VARIABLES τ_1 AND τ_2 WHICH HAPPEN TO BE τ DEPENDING ON CERTAIN PROBABILITY. τ_1 CAN BE ASSIGNED TO THE UNIFORM PART AND τ_2 TO THE TRIANGULAR PART OF THE TOTAL DISTRIBUTION.

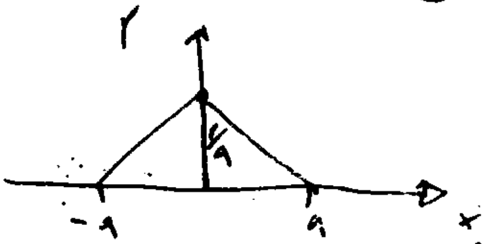
$$\tau = \begin{cases} \tau_1 & \text{WITH PROBABILITY } \lambda \\ \tau_2 & \text{WITH PROBABILITY } 1-\lambda \end{cases}$$

$$\theta = f(\tau) = \begin{cases} 1 & \tau = \tau_1 \\ 2 & \tau = \tau_2 \end{cases}$$

$$H(\tau,\theta) = H(\tau) + \underbrace{H(\theta|\tau)}_{0} = H(\theta) + H(\tau|\theta) = H(\lambda) + \lambda \cdot H(\tau_1) + (1-\lambda)\cdot H(\tau_2)$$

$$\lambda = \int_{\frac{a-1}{2}}^{\frac{1-a}{2}} 1 \cdot d\tau = \left[\frac{1-a}{2} - \frac{a-1}{2}\right] = \frac{1-a-a+1}{2} = 1-a$$

THIS MEANS THAT τ ADOPTS UNIFORM DISTRIBUTION WITH PROBABILITY $1-a$



$y = \frac{1}{a}\cdot\frac{1}{a}(x+a) = \frac{1}{a^2}(x+a)$

$x=-a \quad y = \frac{1}{a^2}(-a+a) = 0$

$x=0 \quad y = \frac{1}{a^2}(a) = \frac{1}{a}$

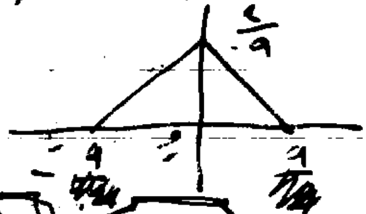
$$h(x) = \int_{-a}^{a} p(x)\, ld \frac{1}{p(x)}\, dx$$

$$= \int_{-a}^{0} \frac{1}{a^2}(x+a)\, ld \frac{a^2}{(x+a)}\, dx + \int_{0}^{a} \frac{-x+a}{a^2}\, ld \frac{a^2}{(-x+a)}\, dx = \frac{2\ln a + 1}{4\ln(2)}$$

$$+ \frac{1}{4}\,\frac{2\ln a + 1}{\ln 2} = \frac{2\ln a + 1}{2\ln 2} = ld\, a + \frac{ld\, e}{2}$$

$$I(X;Z) = H(Z) - H(Z) = H(\lambda) + \lambda \cdot h(Z_1) + (1-\lambda)\, h(Z_2) - h(Z)$$

- $h(Z_2) = ?$



$$\frac{a+1}{2} - \frac{1-a}{2} = \frac{a+1-1+a}{2} = \frac{2a}{2} = a$$

$$\boxed{h(Z_2) = \ell d\left(\tfrac{a}{\ }\right) + \frac{\ell d\, e}{2}}$$

- $h(Z_1) = ?$

$$\frac{1-a}{2} - \frac{a-1}{2} = \frac{1-a-a+1}{2} = 1-a$$

$$h(Z_1) = \int_{\frac{a-1}{2}}^{\frac{1-a}{2}} \frac{1}{(1-a)}\, \ell d \ldots \, dy = -K\left[\frac{1-a}{2} - \frac{a-1}{2}\right] = \ldots$$

$$K = \frac{1}{1-a}\, \ell d\,(1-a)$$

$$h_1(Z_1) = \left[(1-a)^{-1} \cdot \ell d\,(1-a)\right] \cdot (1-a) = \ell d\,(1-a)$$

$$I(X;Z) = \lambda\, \ell d\, \frac{1}{\lambda} + (1-\lambda)\,\ell d\,(1-\lambda) + \lambda\, \ell d\,(1-a) + (1-\lambda)\left[\ell d\,(a) + \frac{\ell d\, e}{2}\right] - h(Z)$$

SVEDONO ... $\ell_n(x)$:

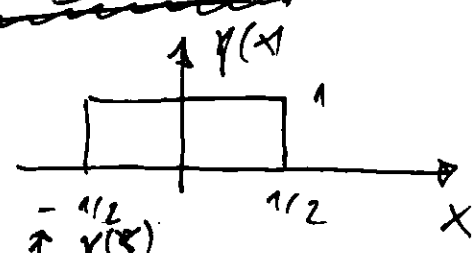
$$I(X;Z) = \lambda\, \ell_n \frac{1}{\lambda} + (1-\lambda)\,\ell_n(1-\lambda)^{-1} + \lambda\, \ell_n(1-a) + (1-\lambda)\left[\ell_n(a) + \frac{1}{2}\right] - h(Z)$$

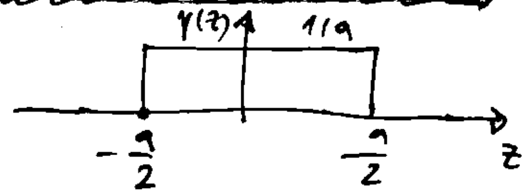
$$I(X;Z) = (1-a)\, \ell_n \frac{1}{1-a} + a\, \ell_n a^{-1} + (1-a)\,\ell_n(1-a) + \ldots + \frac{a}{2} - \ell_n a$$

$$\boxed{I(X;Z) = \frac{a}{2} - \ell_n a} \text{ nats}$$

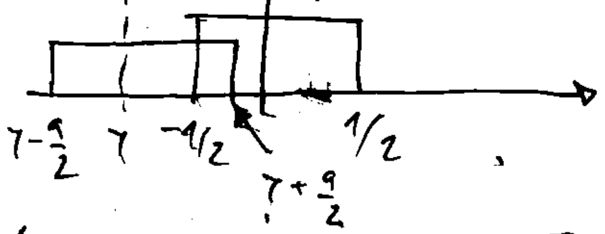
$$\boxed{I(X;Z) = \frac{a\, \ell d\, e}{2} - \ell d\, a}$$

$\boxed{a > 1}$





(a)



$$y + \frac{a}{2} \geq -\frac{1}{2} \qquad y \geq -\frac{a+1}{2}$$

$$y + \frac{a}{2} \leq \frac{1}{2} \qquad y \leq \frac{1-a}{2}$$

$$p(z) = \int^{y+1/2} \frac{1}{a} \cdot 1 \cdot dy = \frac{1}{a}\left(y + \frac{a}{2} + \frac{1}{2}\right)$$

(b) $y + \frac{a}{2} \geq \frac{1}{2}$ $\boxed{y \geq \frac{1-a}{2}}$ $\quad y - \frac{a}{2} \leq -\frac{1}{2}$ $\boxed{y \leq \frac{a-1}{2}}$

$$p(z) = \int_{-1/2}^{1/2} \frac{1}{a}\, 1 \cdot dy = 1/a \qquad \boxed{\frac{1-a}{2} \leq y \leq \frac{a-1}{2}}$$

(98)

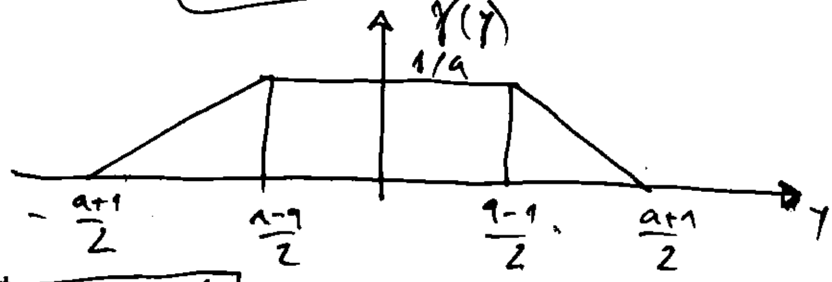
(c) $\gamma - \frac{a}{2} \geq -\frac{1}{2}$ $\quad \gamma \geq \frac{a-1}{2}$

$\gamma - \frac{a}{2} \leq \frac{1}{2}$ $\quad \gamma \leq \frac{1+a}{2}$

$$\boxed{\frac{a-1}{2} \leq \gamma \leq \frac{a+1}{2}}$$

$$\int_{\gamma-\frac{a}{2}}^{1/2} \frac{1}{a}\cdot 1\, d\gamma = \frac{1}{a}\left(-\gamma + \frac{a}{2} + \frac{1}{2}\right) = \frac{1}{a}\left(-\gamma + \frac{a+1}{2}\right)$$

$$\boxed{\varphi(\gamma) = \begin{cases} \frac{1}{a}\left(\gamma + \frac{a+1}{2}\right) & -\frac{(a+1)}{2} \leq \gamma \leq \frac{1-a}{2} \quad (a) \\ 1/a & \frac{1-a}{2} \leq \gamma \leq \frac{a-1}{2} \quad (b) \\ \frac{1}{a}\left(-\gamma + \frac{a+1}{2}\right) & \frac{a-1}{2} \leq \gamma \leq \frac{a+1}{2} \quad (c) \end{cases}}$$



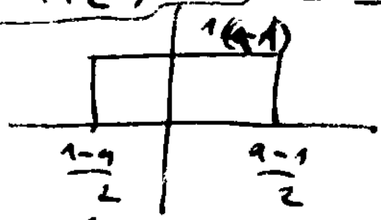
$$\lambda = \int_{\frac{1-a}{2}}^{\frac{a-1}{2}} \frac{1}{a}\, d\gamma = \frac{1}{a}\left(\frac{a-1}{2} - \frac{1-a}{2}\right) = \frac{a-1-1+a}{2a} = \frac{2a-2}{2a} = \frac{a-1}{a}$$

$\boxed{\lambda = 1 - \frac{1}{a}}$ $\quad 1 - \lambda = 1 - 1 + \frac{1}{a} = \frac{1}{a}$

$$\boxed{H(\zeta) = H(\lambda) + \lambda H(\zeta_1) + (1-\lambda)\cdot H(\zeta_2) - H(z)}$$

$\frac{1-a}{2} + \frac{a-1}{2} = \frac{-1+a+a-1}{2} = \frac{2a}{a+1}$

$$H(\zeta_1) = \int_{1-a/2}^{\frac{a-1}{2}} \frac{1}{1-a}\, ld(1-a)\, d\gamma$$

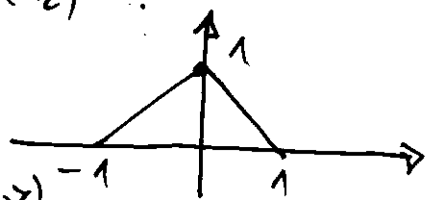


$$H(\gamma_1) = \frac{1}{a-1}\, ld(a-1)\left[\frac{a-1}{2} - \frac{1-a}{2}\right] = \underline{ld(a-1)}$$

$H(\zeta_2) = ?$

$$\frac{a+1}{2} - \frac{a-1}{2} = \frac{a+1-a+1}{2} = 1$$

$$\frac{1-a}{2} + \frac{a+1}{2} = 1$$



$$\boxed{H(\zeta_2) = \ln a + \frac{1}{2} = ld\, 1 + \frac{1}{2} = \frac{1}{2}}$$

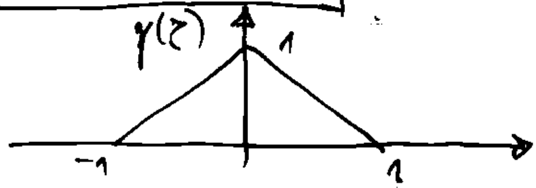
$I(x;\zeta) = \left(1 - \frac{1}{a}\right) ld \frac{1}{1-\frac{1}{a}} + \frac{1}{a}\, ld\, a + \left(1 - \frac{1}{a}\right) ld(a-1) + \frac{1}{a}\frac{1}{2} - \ln a =$

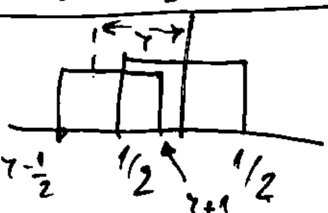
$= \frac{a-1}{a}\, ld \frac{a}{a-1} + \frac{1}{a}\ln a + \frac{a-1}{a}\ln(a-1) + \frac{1}{2a} - \ln a =$

$= \left(1 - \frac{1}{a}\right)\ln a + \frac{1}{a}\ln a + \frac{1}{2a} - \ln a = \frac{1}{2a}$

$\boxed{I(x;\zeta) = 1/2a}$

$\boxed{a = 1}$





$$\varphi(z) = \begin{cases} \gamma + 1 & -1 \leq \gamma \leq 0 \\ -\gamma + 1 & 0 \leq \gamma \leq 1 \end{cases}$$

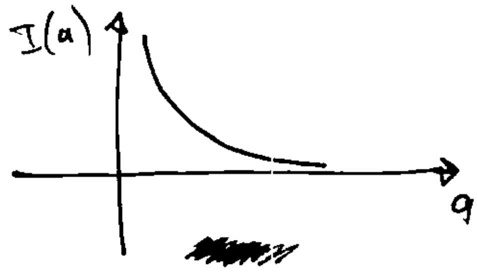
$H(\gamma_1) = 0 \quad H(\lambda) = 0$

$H(\zeta_2) = 1/2$

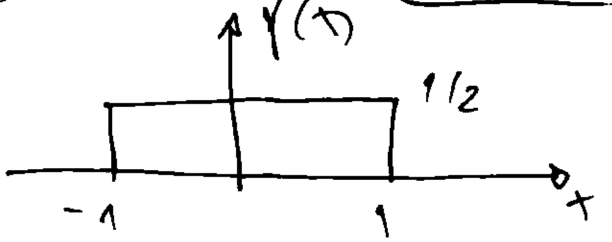
$I(x;\zeta) = 0 + 0 + 1\cdot 1/2 - \ln a = \frac{1}{2}$

(39)

$$I(x;y) = \begin{cases} \frac{a}{2} - \ln a & a < 1 \\ 1/2a & a > 1 \\ 1/2 & a = 1 \end{cases}$$



(b) $a = 1$ $\quad I(x;y) = \frac{1}{2}$ ⟹ FOR ANY $y(x)$



$$h(x) = \int_{-1}^{1} \frac{1}{2}\, ld\, 2 \; dx = \frac{1}{2}(1+1) = \frac{1}{2}\cdot 2 = 1$$

PROBLEM 8.4 QUANTIZED RANDOM VARIABLES.

ROUGHLY HOW MANY BITS ARE REQUIRED ON AVERAGE TO DESCRIBE TO THREE DIGIT ACCURACY THE DECAY TIME (IN YEARS) OF RADIUM ATOM IF THE HALF-LIFE OF RADIUM IS 80 YEARS? NOTE THAT HALF TIME IS MEDIAN OF THE DISTRIBUTION.

$$f(x_i)\cdot\Delta = \int_{i\Delta}^{(i+1)\Delta} f(x)\,dx \qquad x^{\Delta} = x_i \qquad i\Delta \le x_i \le (i+1)\Delta$$

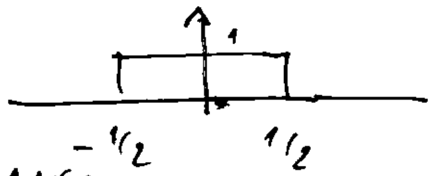
$$H(X^{\Delta}) + ld\,\Delta = h(X) \qquad \Delta = 2^{-n}$$

$$H(X^{\Delta}) \to h(X) + n$$ ⟹ NUMBER OF BITS ON THE AVERAGE REQUIRED TO DESCRIBE X TO n-BIT ACCURACY

(n: BIT ACCURACY)

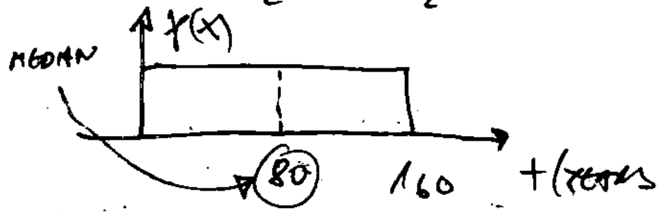
$\mu = 80$ years } MEDIAN OF DISTRIBUTION

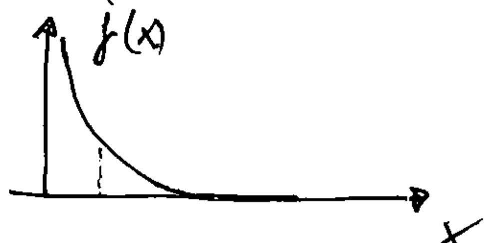
$$h(x) = \int_0^{160} \frac{1}{160}\, dt = \frac{1}{160}(160-0) = 1$$



$$h(x) = \int_{-1/2}^{1/2} 1 \cdot dx = 1\cdot\left(\frac{1}{2}+\frac{1}{2}\right) = 1$$

$$H(X^{\Delta}) = 1 + 3 = 4 \text{ bits}$$





$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

$$\bar{x} = \int_0^\infty x \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

MEAN LIFETIME

$$\frac{1}{\bar{x}} = \lambda \Rightarrow$$ INVERSE OF THE MEAN LIFETIME

WIKIPEDIA FORMULA

$$t_{1/2} = \frac{\ln(2)}{\lambda} = 80 \text{ years} \quad (*)$$

$$\lambda = \frac{\ln(2)}{80} = 0.00866$$

$$h(x) = -\int_0^\infty f(x)\, ld\, f(x) = -\int_0^\infty \lambda \cdot e^{-\lambda x} ld(\lambda^{-1} e^{+\lambda x}) dx$$

$$= -\int_0^\infty \lambda e^{-\lambda x}(+\lambda x)\, ld(\lambda^{-1} e)\, dx = -ld(\lambda e) \cdot \lambda \int_0^\infty \lambda x e^{-\lambda x} dx, \quad \bar{x} = \frac{1}{\lambda}$$

$$h(x) = ld\left(\frac{e}{\lambda}\right)$$

$$h(x) = -\lambda \cdot ld(\lambda^{-1} e) \cdot \frac{1}{\lambda} = ld(\lambda^{-1} e) = \frac{\ln(\lambda^{-1} e)}{\ln 2} = \frac{\ln(\lambda^{-1})}{\ln 2} + \frac{\ln e}{\ln 2}$$

$$h(x) = ld\left(\left[\frac{\ln(2)}{80}\right]^{-1} \cdot e\right) = 8.29$$

$$H(x^\Delta) = h(x) + n = 8.29 + 3 = 11.29 \sim \underline{12 \text{ bits}}$$

- CONCORDIA UNIVERSITY SOLUTIONS

$$\int_0^{80} \lambda e^{-\lambda x} dx = \frac{1}{2}$$

[illegible]

$$\int_0^{80} e^{-\lambda x} d(\lambda x) = -e^{-\lambda x}\Big|_0^{80} = \frac{1}{2}$$

$$-\left[e^{-\lambda 80} - e^0\right] = \frac{1}{2} \qquad e^{-\lambda 80} = 1 - \frac{1}{2} = \frac{1}{2}$$

$$e^{\lambda 80} = 2 \qquad \lambda 80 = \ln 2 \qquad \boxed{\lambda = \frac{\ln(2)}{80}}$$

NAMESTO SO GOTOVATA FORMULA OD WIKIPEDIA MOZE VAZI!!

TREBA DA SE ZEMAT DEKA VAZI ZA ISKORISTUVA MEDIJANA

MMV

$$\int_0^{\text{MEDIAN}} f(x)\, dx = \frac{1}{2}$$

MMV

3 DIGITS ~ 10 BITS $\quad 2^{10} = 1024$

$$H(x^\Delta) = 8.29 + 10 \doteq 18.3 \text{ bits}$$

PROBLEM 8.5 **SCALING.** Let $h(\mathbf{x}) = -\int f(\mathbf{x})\, ld\, f(\mathbf{x})\, d\mathbf{x}$

SHOW $h(A\mathbf{x}) = ld\,|\det(A)| + h(\mathbf{x})$.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad Ax = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

$$h(\underbrace{Ax}_{Y}) = \int f(Ax)\, ld\, f(Ax)\, dx$$

$$f(Ax) = \frac{f(x)}{|J|}\Bigg|_{x = \frac{Y}{A}} \qquad J = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{vmatrix}$$

$$Y = A \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix} \quad \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 \\ y_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned} \qquad dY = J\, dx = A^2 dx$$

$$|J| = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \underbrace{a_{11}a_{22} - a_{12}a_{21}}_{|\det(A)|} \qquad f(Y) = \frac{f(x)}{|\det[A]|} = \frac{f(x)}{|\det[A]|}$$

$$h(Y) = -\int f(Y)\, ld\, f(Y)\, dY = -\int \frac{f(x)}{\det[A]}\, ld\left(\frac{f(x)}{\det[A]}\right) \det[A]\, dx \qquad dY = |J|\, dx$$

$$= -\int f(x)\, ld\left(\frac{f(x)}{\det[A]}\right) dx = -\int f(x)\, ld\, f(x)\, dx + \int f(x)\, ld\, \det[A]\, dx$$

$$= + h(x) + ld\,|\det A| \int f(x)\, dx = h(x) + ld\,|\det(A)| =$$

$$= h(x) + ld\,|\det A|$$

- ISTI DOKAZ ĆE VAŽI I ZA: $x = [x_1\, x_2 \dots x_n]$

~~ZAKLJUČAK ... DIJAGONALNA MATRICA~~

~~$|J| = |A| = \det$~~

$$Y = AX \qquad \boxed{X = A^{-1} \cdot Y} \qquad g(Y) = \frac{f(A^{-1}Y)}{|\det A|}$$

$$h(Y) = -\int g(Y)\, ld\, g(Y)\, dY = -\int \frac{f(A^{-1}Y)}{|\det A|} \cdot ld\left(\frac{f(A^{-1}Y)}{\det A}\right) |\det A|\, dx$$

$$= -\int f(x) \cdot ld\, f(x)\, dx + \int f(x)\, ld\,|\det(A)|\, dx =$$

$$= h(x) + ld\,|\det(A)|$$

8.7 DIFFERENTIAL ENTROPY BOUND ON DISCRETE ENTROPY (CONTINUE FROM №166)

THEOREM 8.6.6. ESTIMATION ERROR AND DIFFERENTIAL ENTROPY

$$E(X-\hat{X})^2 \geq \frac{1}{2\pi e} e^{2h(X)}$$

WITH EQUALITY IF $X \sim$ GAUSSIAN AND $\hat{X}$ IS THE MEAN OF X.

PROOF $E(X-\hat{X})^2 \geq \min_{\hat{X}} (X-\hat{X})^2 = E(X-E(X))^2$

$= \mathrm{VAR}(X) \geq e^{2h}/2\pi e$

$h(X) \leq \frac{1}{2}\ln 2\pi e \sigma^2 \qquad \sigma^2 \geq \frac{e^{2h(X)}}{2\pi e}$

$E(X)$ IS BEST ESTIMATOR FOR X

$$E(X-\hat{X}(Y))^2 \geq \frac{e^{2h(X)}}{2\pi e} \geq \frac{e^{2h(X|Y)}}{2\pi e} \quad ; h(X) \geq h(X|Y)$$

$X = \{a_1, a_2, \ldots\} \qquad \Pr(X = a_i) = p_i$

$$H(p_1, p_2, \ldots) \leq \frac{1}{2}\ln(2\pi e)\left(\sum_{i=1}^{\infty} p_i i^2 - \left(\sum_{i=1}^{\infty} i p_i\right)^2 + \frac{1}{12}\right)$$

$H(X^\Delta) = h(X) - \ln \Delta = |\Delta = 2^n| = h(X) - n$

$X' \sim \Pr(X' = i) = p_i \qquad U \sim [0,1]$

$Y = X' + U$

$I(X';Y) = H(Y) - H(Y|X') = H(Y) - \sum p(x) H(x' + U/x')$

$= H(Y) - H(U) \qquad H(U) = 0$

$h(U) = \int_0^1 1 \cdot \ln 1 \, dx = 0 \qquad I(X';Y) = H(Y)$

$E(X') = \sum_{i=1}^{\infty} i p_i \quad \left|p_i = \frac{1}{2^i}\right| = \sum_{i=1}^{\infty} i q^i = \sum_{i=0}^{\infty} i q^i = \frac{q}{(1-q)^2}$

$= \frac{\frac{1}{2}}{\frac{1}{4}} = 2 \qquad E(X'^2) = \sum_{i=1}^{\infty} i^2 p_i = \sum_{i=1}^{\infty} i^2 \frac{1}{2^i} = 6$

$\sigma^2 = E[(X'-\bar{X})^2] = E[X'^2] - 2E[X']\cdot\bar{X} + \bar{X}^2 = E[X'^2] - \bar{X}^2$

$X \in \{a_1, a_2, \dots\}$ $\quad$ $\boxed{H(X) = H(X')}$

$X' \in \{1, 2, \dots\}$ $\quad$ $Y = X' + V$

recte:

$X_1 \in \{1, 2, \dots, m\}$ $\quad$ $\theta = \begin{cases} 1 & X = X_1 \quad p \\ 2 & X = X_2 \quad 1-p \end{cases}$

$X_2 \in \{m, \dots, n\}$

$H(X, \theta) = H(X) + \underbrace{H(\theta | X)}_{0} = H(X) = \underbrace{H(\theta)}_{H(p)} + H(X|\theta)$

$H(X|\theta) = \underset{p}{H(\theta=1)} \cdot \underbrace{H(X|\theta=1)}_{H(X_1)} + \underset{(1-p)}{H(\theta=2)} \underbrace{H(X|\theta=2)}_{H(X_2)}$

$= p\, H(X_1) + (1-p) H(X_2)$

$H(X) = H(p) + p \cdot [H(X_1)] + (1-p) H(X_2)$

$Y \in \{Y_1, Y_2, \dots\}$ $\quad$ $Y_i \in \{i, i+1\}$

$I(X'; Y) = \underbrace{H(X')}_{= H(X)} - \underline{H(X'|Y)} = H(Y) - \underbrace{H(\theta)}_{H(Y|X')}$

$H(Y) = H(X) - H(X'|Y) = H(Y) - H(X|Y) = \underline{I(X;Y)}$

$H(X) = H(Y) + H(X|Y) = \underline{H(X, Y)}$

$H(X) \leq \frac{1}{2} \operatorname{ld} 2\pi e \sigma^2$

$H(Y) = \lim_{n \to \infty} \frac{1}{n} H(Y_1, Y_2, \dots, Y_n) =$

$= \frac{\Sigma}{n} \underline{H(Y_i)} = \left/ \begin{matrix} \text{UNIFORM ECC} \\ 0 \div 0.5 \Rightarrow 0 \\ 0.5 \div 1 \Rightarrow 1 \end{matrix} \right/ = H\left(\frac{1}{2}\right) = 1$

$H(X) = 1 + H(X|Y) = 1 + P(Y=i)\, H(X|Y=i)$

$= 1 + \sum P(Y=i) \left[P(X=i|Y=i) \operatorname{ld} P(X=i|Y=i) + P(X=i-1|Y=i) \right.$

$\left. \cdot \operatorname{ld} P(X=i-1|Y=i) \right] = 1 + \sum_i P(Y=i) \cdot 1$

PROBLEM 8.9 (CONTINUE FROM 164)

$$Y(x,Y) = \frac{1}{(2\pi)\sqrt{|K_1|}} e^{-[x\,Y] K_1^{-1} \begin{bmatrix} x \\ Y \end{bmatrix}} \qquad K_1 = \begin{bmatrix} \sigma_x^2 & \rho_1 \\ \rho_1 & \sigma_Y^2 \end{bmatrix}$$

$$K_1 = \begin{bmatrix} \sigma^2 & \rho_1 \sigma^2 \\ \rho_1 \sigma^2 & \sigma^2 \end{bmatrix} \qquad |K_1| = \sigma_x^2 \sigma_Y^2 - \rho_1^2$$

MARGINAL

$$Y(x) = \frac{\sqrt{\pi}}{2\pi \sqrt{\sigma_x^2 \sigma_Y^2 - \rho_1^2}} \cdot e^{-\frac{x^2(\sigma_Y^2 - 1)}{(\sigma_x^2 \sigma_Y^2 - \rho_1^2)\sigma_x^2}} \cdot \sqrt{\frac{\sigma_x^2 \sigma_Y^2 - \rho_1^2}{\sigma_x^2}}$$

$$Y(x) = \frac{1}{2\sqrt{\pi}\,\sigma} e^{-\frac{x^2(\sigma^2 - 1)}{(\sigma^4 - \rho_1^2)\sigma^2}} = \frac{1}{2\sigma\sqrt{\pi}} e^{-\frac{x^2(\sigma^2-1)}{(\sigma^4 - \rho_1^2)\sigma^2}}$$

$$h(x) = h(Y) = \frac{1}{2} \mathrm{ld}\, (2\pi e) \sigma^2$$

$$h(x,Y) = \frac{1}{2} \mathrm{ld}\, (2\pi e)^2 |K| = \frac{1}{2} \mathrm{ld}\, (2\pi e)^2 (\sigma^4 - \rho_1^2 \sigma^4)$$

$$I(x;Y) = h(x) - h(x|Y) = h(x) - h(x,Y) + h(Y)$$

$$h(x,Y) = h(x) + h(x|Y) \quad \cdot \quad h(x|Y) = h(x,Y) - h(Y)$$

$$I(x;Y) = \text{[crossed out]}\ \mathrm{ld}\,(2\pi e)\sigma^2 - \frac{1}{2}\mathrm{ld}(2\pi e)^2 \sigma^4 (\text{[crossed out]} - \rho_1^2)$$

$$= \mathrm{ld}\,(2\pi e)\sigma^2 - \mathrm{ld}\,(2\pi e)\sigma^2 - \frac{1}{2}\mathrm{ld}\,\text{[illegible]}(1 - \rho_1^2)$$

$$\boxed{I(x;Y) = -\frac{1}{2}\mathrm{ld}\,\text{[illegible]}(1 - \rho_1^2)}$$

$$\boxed{I(Y;Z) = -\frac{1}{2}\mathrm{ld}\,(1 - \rho_2^2)}$$

$$I(xY;Z) = I(x;Z) + I(Y;Z|x) = I(Y;Z) + \overbrace{I(x;Z|Y)}^{0}$$

$$I(x;Z|Y) = H(x|Y) - H(x|YZ) = H(Z|Y) - H(Z|x,Y) = 0$$

$$I(xY;Z) = I(x;Z) + I(Y;Z|x) = I(Y,Z)$$

$$I(x,Z) = I(Y;Z) - I(Y;Z|x) = -\frac{1}{2}\mathrm{ld}(1-\rho_2^2) - I(Y;Z|x)$$

$$I(Y;Z|x) = H(Y|x) - H(Y|Zx) = H(Z|x) - H(Z|xY) =$$

$$= H(Z|x) - H(Z|Y) = H(xZ) - H(x) - H(YZ) + H(Y)$$

$$H(xZ) = H(xYZ) - H(Y|XZ) = H(xY,Z) - H(Z|x)$$

$$K = \begin{bmatrix} \sigma^2 & \rho_1\sigma^2 & 0 \\ \rho_1\sigma^2 & \sigma^2 & \rho_2\sigma^2 \\ 0 & \rho_2\sigma^2 & \sigma^2 \end{bmatrix} \qquad H(X,Y,Z) = \frac{1}{2} ld (2\pi e)^3 |K|$$

$$K = \sigma^2 \begin{vmatrix} \sigma^2 & \rho_2\sigma^2 \\ \rho_2\sigma^2 & \sigma^2 \end{vmatrix} - \rho_1\sigma^2 \begin{vmatrix} \rho_1\sigma^2 & \rho_2\sigma^2 \\ 0 & \sigma^2 \end{vmatrix} = \sigma^2(\sigma^4 - \rho_2^2\sigma^4) - \rho_1\sigma^2 \cdot \rho_1\sigma^4$$

$$K = \sigma^6(1-\rho_2^2) - \rho_1^2\sigma^6 = \underline{\sigma^6(1-\rho_1^2-\rho_2^2)}$$

$$\boxed{H(X,Y,Z) = \frac{1}{2} ld (2\pi e)^3 \sigma^6 (1-\rho_1^2-\rho_2^2)}$$

$$I(Y;Z|X) = H(XZ) - \cancel{H(X)} - H(YZ) + \cancel{H(X)}$$

$$= H(XZ) - H(YZ) = H(XYZ) - H(Y|XZ) - H(YZ) =$$

$$= H(XYZ) - H(Y|X) - H(YZ) = H(XYZ) - (H(XY) - H(X)) -$$

$$- H(YZ) = H(XYZ) - \underline{H(XY)} + H(X) - \underline{H(YZ)}$$

$$= \frac{1}{2} ld (2\pi e)^3 \sigma^6 (1-\rho_1^2-\rho_2^2) - \frac{1}{2} ld (2\pi e)^2 \sigma^4 (1-\rho_1^2) +$$

$$+ \frac{1}{2} ld (2\pi e) \sigma^2 - \frac{1}{2} ld (2\pi e)^2 \sigma^4 (1-\rho_2^2)$$

$$I(X;Z) = -\frac{1}{2} ld (1-\rho_2^2) - \frac{1}{2} ld (2\pi e)^3 \sigma^6 (1-\rho_1^2-\rho_2^2) -$$

$$- \frac{1}{2} ld (2\pi e)^2 \sigma^4 (1-\rho_1^2) + \underline{\frac{1}{2} ld (2\pi e) \sigma^2} - \underline{ld (2\pi e) \sigma^2} - \underline{\frac{1}{2} ld (1-\rho_2^2)}$$

$$= \underline{ld (1-\rho_2^2) - \frac{1}{2} ld (2\pi e)^3 \sigma^6 (1-\rho_1^2-\rho_2^2) - \frac{1}{2} ld (2\pi e)^2 \sigma^4 (1-\rho_1^2)}$$

$$\underline{- \frac{1}{2} ld (2\pi e) \sigma^2}$$

CONTINUE FROM N.16R :

$$H(X|Y) = \sum_x \sum_y p(x,y) \cdot ld \frac{1}{p(x|y)} = \underline{\sum_y p(y) H(X|y)}$$

$$E(X|Y) \qquad E(X) = \sum_y p(y) E(X|y)$$

$$\boxed{E(X|y) = \sum_x x \; Pr(x|y)}$$

$$E[X|Y] = \int x \, p(x|y) \, dx \qquad E[X|Y] = \int_{-\infty}^{\infty} p(y) \left[\int_{-\infty}^{\infty} x p(x|y) dx \right] dy$$

$$E[x|y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, p(xy)\, dx\, dy = \int_{-\infty}^{\infty} x\, p(x|y)\, dx$$

$$E[xz|y] = ?$$

$$p(xz|y) = \frac{p(xyz)}{p(y)} = \frac{p(xy)\cdot p(z|xy)}{p(y)} = \frac{p(xy)\,p(z|y)}{p(y)} = p(x|y)\cdot p(z|y)$$

MARKOVITY

$$E[xz|y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xz \cdot p(xz|y)\, dx\, dz = \int_{-\infty}^{\infty} x\, p(x|y)\, dx \int_{-\infty}^{\infty} z\, p(z|y)\, dz$$

$$= E(x|y)\cdot E(z|y)$$

FIX y

$E[xz] =$ FOR ALL y

$$= E[xz|y] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} xz\, p(xz|y)\, dx\, dz\right] p(y)\, dy = E_y\left[E_x(x|y)\, E_z(z|y)\right]$$

THEOREM FOR TOTAL EXPECTATION; VIDI PREV. [illegible]

MMV

$$p(x,y) = \frac{1}{2\pi\sqrt{|K|}}\, e^{-[x\ y]\, K^{-1} \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]} \qquad - \frac{x^2 - 2xy\rho_1 + y^2}{2\sigma^2(1-\rho_1^2)}$$

$$p(x|y) = \frac{p(xy)}{p(y)} = \frac{\frac{1}{2\pi\sqrt{\sigma^4\cdot(1-\rho_1^2)}}\cdot e}{\frac{1}{\sigma\sqrt{2\pi}}\, e^{-x^2/2\sigma^2}}$$

MAPLE

$$E[x|y] = \int_{-\infty}^{\infty} x\, p(x|y)\, dx = y\cdot\rho_1$$

$$E[z|y] = \int_{-\infty}^{\infty} z\, p(z|y)\, dz = y\cdot\rho_2$$

$$E_y\left[\frac{y^2}{2\rho_1^2\rho_2^2}\right] = \rho_1\rho_2\int_{-\infty}^{\infty} y^2 p_y(y)\, dy = \sigma^2\rho_1\rho_2$$

~~[illegible]~~

MMV

$$\rho_{xz} = \frac{E[xz]}{\sigma_x\sigma_z} = \frac{E_z[xz|y]}{\sigma_x\sigma_z} = \frac{E_y\left[E_x[x|y]\cdot E_z[z|y]\right]}{\sigma_x\sigma_y} = \frac{\rho_1\rho_2\, E_y[y^2]}{\sigma_x\sigma_y}$$

$$\rho_{xz} = \rho_1\rho_2$$

$$\frac{y\rho_1\sigma}{\sqrt{\sigma^4(1-\rho_1^2)}\sqrt{\frac{1}{\sigma^2(1-\rho_1^2)}}} = \frac{y\rho_1\sigma}{\sqrt{\frac{\sigma^4}{\sigma^2}}} = \frac{y\rho_1\sigma}{\sigma}$$

TOTAL EXPECTATION

$$E[XZ] = \int p(y)\, E[XZ|Y]\, dy$$

→ VIDI WIKIPEDIA ARTICLE: Conditional Expectation and Total Expectation.pdf

PROBLEM 8.11 NON-ERGODIC GAUSSIAN PROCESS. CONSIDER A CONSTANT SIGNAL V IN THE PRESENCE OF IID OBSERVATIONAL NOISE $\{Z_i\}$. THUS $X_i = V + Z_i$ WHERE: $V \sim \mathcal{N}(0,S)$ AND Z_i ARE IID $\sim \mathcal{N}(0,N)$. ASSUME THAT V AND $\{Z_i\}$ ARE INDEPENDENT.

(a) Is $\{X_i\}$ STATIONARY?

(b) FIND $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i$. IS THE LIMIT RANDOM?

(c) WHAT IS THE ENTROPY RATE h OF $\{X_i\}$?

(d) FIND THE LEAST-MEAN-SQUARED ERROR PREDICTOR $\hat{X}_{n+1}(X^n)$ AND FIND $\sigma_\infty^2 = \lim_{n\to\infty} E[\hat{X}_n - X_n]^2$.

(e) DOES $\{X_i\}$ HAVE AEP? THAT IS DOES $-\frac{1}{n}\,ld\, f(X^n) \to h$?

(a) MISLAM DEKA E STACIONAREN $\{X_i\}$ ZATOA ŠTO SUMA OD DVE GAUSOVI SLUČAJNI PROMENLIVI E POVTORNO GAUSOVA SO SREDNA VREDNOST KOJA E SUMA OD DVETE SREDNI VREDNOSTI I VARIJANSA KOJA E SUMA OD DVETE VARIJANSI, T.E.

$$X_i \sim \mathcal{N}(0, S+N)$$

ZA BILO KOE i, X_i E ISTO DISTRIBUIRANA PA ZATOA PROCESOT E STACIONAREN.

(b) SO OGLED NA TOA ŠTO PROCESOT E STACIONAREN

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i = E[X_i] = 0$$

} STATISTIČKATA SREDNA VREDNOST E EDNAKVA NA SREDNATA VREDNOST VO VREME.

$$\frac{1}{n}\, ld \frac{1}{p(x^n)} \xrightarrow[n\to\infty]{\text{STRONG LAW}} E\left[ld \frac{1}{p(X^n)}\right] = h(X)$$

(c) $\lim_{n\to\infty} \frac{h(X_1^n)}{n} = h(X) = \underline{\frac{1}{2}\, ld\, 2\pi e (S+N)}$

(d) MEAN SQUARE PREDICTOR

$$E[(X-\hat{X})^2] \geq \min_{\hat{X}} E(X-\hat{X})^2 = E(X-\bar{X})^2 = var(X)$$

$$h(X) \leq \frac{1}{2}\, ln(2\pi e \sigma^2)$$

$$\sigma^2 \geq \frac{2^{2h(x)}}{2\pi e} \qquad \boxed{E[(X-\hat{X})^2] \geq \frac{e^{2h(x)}}{2\pi e}}$$

$$E\left[\left(X-\hat{X}(Y)\right)^2\right] \geq \frac{e^{2h(X|Y)}}{2\pi e}$$

$$E\left[\left(X-\hat{X}(X^n)\right)^2\right] \xrightarrow[n\to\infty]{} \frac{e^{2h(X|X^n)}}{2\pi e} = \frac{e^{\ln 2\pi e(S+N)}}{2\pi e}$$

$$h(X_{n+1}|X^n) = h(X_{n+1}) = h(X_i) = \frac{1}{2}\ln 2\pi e\,(S+N)$$

$$E\left[\left(X-\hat{X}(X^n)\right)^2\right] = \frac{2\pi e(S+N)}{2\pi e} = S+N$$

(e) $f(x^n) = f(x_1)\cdot f(x_2)\cdots f(x_n)$

$$\frac{1}{n}\ln f(x^n) = \frac{1}{n}\sum_{i=1}^{n}\ln f(x_i) \xrightarrow[n\to\infty]{} E[\ln f(x_i)] = h(X)$$

$$h(X) = \frac{1}{2}\ln 2\pi e(S+N)$$

(DA)

STANFORD UNIVERSITY SOLUTIONS

(a) YES! $E[X_i] = E[V] + E[Z_i] = 0$

$$E[X_iX_j] = E\left[(V+Z_i)(V+Z_j)\right] = \begin{cases} S & i \neq j \\ S+N & i = j \end{cases}$$

SINCE X_i IS GAUSSIAN DISTRIBUTED IT IS COMPLETELY CHARACTERIZED BY ITS FIRST AND SECOND MOMENTS. SINCE THE MOMENTS ARE STATIONARY, X_i IS WIDE SENSE STATIONARY WHICH FOR GAUSSIAN DISTRIBUTION IMPLIES THAT X_i IS STATIONARY.

$$E\left[(V+Z_i)^2\right] = E\left[V^2 + 2VZ_i + Z_i^2\right] = \underbrace{E[V^2]}_{S} + 2\underbrace{E[VZ_i]}_{0} + \underbrace{E[Z_i^2]}_{N}$$

$$E\left[(V+Z_i)(V+Z_j)\right] = E\left[V^2 + VZ_j + VZ_i + Z_iZ_j\right] = E[V^2] + E[VZ_j] + E[VZ_i] + E[Z_iZ_j] = S + \underbrace{0 + 0 + 0}_{\text{STOCHAST. INDEPENDENT}}$$

(b) $$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}X_i = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(V+Z_i) = V + \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}Z_i$$

$= V + \underbrace{E[Z_i]}_{0} = V$ THE LIMIT IS RANDOM - VARIABLE $\mathcal{N}(0,S)$

↑ STRONG LAW OF LARGE NUMBERS

(c) $x^n \sim N(0, K_{xn})$ WHERE K_{xn} HAS DIAGONAL VALUES S+N AND OFF DIAGONAL VALUES OF S.

$$K = \begin{bmatrix} S+N & S & S \\ S & S+N & S \\ S & S & S+N \end{bmatrix}$$

$$|K| = 3SN^2 + N^3 = \left(\frac{3S}{N} + 1\right) N^3$$

MAKE

$$|K| = \left(\frac{nS}{N} + 1\right) \cdot N^n$$

(n=3) $\quad |K_3| = \left(\frac{3S}{N} + 1\right) N^3$

n=k $\quad |K_k| = \left(\frac{kS}{N} + 1\right) N^k$

n=k+1 $\quad |K_{k+1}| = \left(\frac{(k+1)S}{N} + 1\right) N^{k+1}$?

$$|K_{k+1}| = \begin{vmatrix} K_k & S \\ S & N+S \end{vmatrix} = |K_k| \cdot (N+S) + S^2 = \left(\frac{kS}{N} + 1\right) N^k (N+S) - S^2 =$$

$$= \left(\frac{kS}{N} \cdot N + N + \frac{kS^2}{N} + S\right) N^k - S^2 = kSN^k + N^{k+1} + SN^k - S$$

$$h(X) = \lim_{n \to \infty} \frac{h(x_1^n)}{n} = \lim_{n \to \infty} \frac{1}{2n} \ln (2\pi e)^n \cdot |K| =$$

$$= \lim_{n \to \infty} \frac{1}{2n} \ln \left[(2\pi e)^n \cdot \left(\frac{nS}{N} + 1\right) \cdot N^n \right] =$$

$$= \lim_{n \to \infty} \frac{1}{2} \ln(2\pi e) + \frac{1}{2n} \lim_{n \to \infty} \ln \left(\frac{nS}{N} + 1\right) N^n =$$

~~$= \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \ln \lim_{n \to \infty} \left(\frac{nS}{N} + 1\right) N^n =$~~

~~$= \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \ln \lim_{n \to \infty} \frac{nS + N}{N} \cdot N^{n-1} =$~~

~~$= \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \ln \lim_{n \to \infty} (nSN^{n-1} + N^n)$~~

$$= \lim_{n \to \infty} \frac{1}{2n} \ln (2\pi e N)^n + \lim_{n \to \infty} \frac{1}{2n} \ln \left(\frac{nS}{N} + 1\right) \to 0$$

$$= \frac{1}{2} \ln 2\pi e N$$

(d) $E[(x_{n+1} - \hat{x}_{n+1}(x^n))^2] = E_{x^n}\{E[(x_{n+1}-\hat{x}_{n+1}(x^n))^2 | x^n]\} =$

$= E[\, E[x_{n+1}^2 - 2x_{n+1}\cdot \hat{x}_{n+1}(x^n) + \hat{x}_{n+1}^2(x^n)]]$

$$\frac{\partial E[(x_{n+1}-\hat{x}_{n+1}(x^n))^2 / x^n]}{\partial x_{n+1}} = E[2x_{n+1} - 2\hat{x}_{n+1}(x^n) / x^n + 0] = 0$$

$E[2x_{n+1} / x^n] - 2\hat{x}_{n+1} = 0$ $\quad \boxed{\hat{x}_{n+1}(x^n) = E[x_{n+1} / x^n]}$

$$f(V|x^n) = \frac{f(x^n, V)}{f(x^n)}$$

$$f(x^n, V) = \frac{1}{(2\pi)^{n+1}\cdot |K|} e^{-\frac{[x_1^n V]\cdot K^{-1}\cdot \begin{bmatrix} x_1^n \\ V\end{bmatrix}}{2}}$$

$n - \cancel{n} - 1 + n$

$$f(x^n) = \frac{1}{(2\pi)^n \cdot |K_x|} e^{-\frac{[x_1^n] K^{-1} [x_1^n]}{2}}$$

$$f(V|x^n) \sim N\left(\frac{S}{nS+N}\sum_{i=1}^{n} x_i,\ \frac{SN}{nS+N}\right)$$

$\hat{x}_{n+1}(x^n) = E[x_{n+1}|x^n] = E[V|x^n] + E[Z_{n+1}|x^n]$

$$= \frac{S}{nS+N}\sum_{i=1}^{n} x_i + 0$$

THEREFORE THE LIMITING SQUARE ERROR IS:

$$e^2 = \lim_{n\to\infty} E[(\hat{x}_n - x_n)^2] = \lim_{n\to\infty} E\left(\frac{S}{(n-1)S+N}\sum_{i=1}^{n-1} x_i - x_n\right)^2 = \circledast$$

$$= \lim_{n\to\infty} E\left[\frac{S}{(n-1)S+N}\sum_{i=1}^{n-1}(V+Z_i) - Z_n - V\right]^2$$

$\frac{S}{(n-1)S+N}\cdot n\cancel{V} - V = \frac{S n V - (n-1)SV - NV}{(n-1)S+N} = \frac{\cancel{SnV - SnV + SV - NV}}{(n-1)S+N}$

$= \cancel{\frac{S-N}{(n-1)S+N}\cdot V}$; $\frac{S(n-1)V - \cancel{(n-1)SV} - NV}{(n-1)S+N} = -\frac{NV}{(n-1)S+N}$

$$\circledast = \lim_{n\to\infty} E\left[\frac{S}{(n-1)S+N}\sum_{i=1}^{n-1} Z_i - Z_n - \frac{NV}{(n-1)S+N}\right] =$$

$$= \lim_{n\to\infty} \left(\frac{S}{(n-1)S+N}\right)^2 \sum_{i=1}^{n-1} E[Z_i^2] - E[Z_n^2] - \left(\frac{N}{(n-1)S+N}\right)^2 \cdot E[V^2]$$

$$= \lim_{n\to\infty}\left[\underbrace{\left(\frac{S}{(n-1)S+N}\right)^2}_{0}(n-1)N + N + \underbrace{\left(\frac{N}{(n-1)S+N}\right)^2}_{0}S\right]$$

$$= N \qquad \boxed{e^2 = \lim_{n\to\infty} E[\hat{x}_n - x_n]^2 = N}$$

PROBLEM 9.4 (CONTINUE FROM 16R)

$$M_x(-s) = \frac{M_Y(-s)}{M_Z(-s)} \qquad p_x(x) = \mathcal{L}^{-1}[M_x(-s)]$$

$$M_Y(-s) = e^{s(s\sigma^2 - 2\mu)/2} \qquad M_Z(-s) = \frac{\lambda}{s+\lambda}$$

$$M_X(-s) = \left(1 + \frac{s}{\lambda}\right) e^{s(s\sigma^2 - 2\mu)}$$

$$p_x(x) = p_Y(x) + \frac{1}{\lambda}\mathcal{L}^{-1}\left[s e^{s(s\sigma^2 - 2\mu)}\right] + \frac{1}{\lambda} p_Y(0)$$

$$\mathcal{L}\left[\frac{df(t)}{dt}\right] = sF(s) - f(0) \qquad sF(s) = \mathcal{L}\left[\frac{df(t)}{dt}\right] + f(0)$$

$$\mathcal{L}\left[\frac{1}{\lambda}\frac{df(t)}{dt}\right] = \frac{sF(s)}{\lambda} - \frac{f(0)}{\lambda} \qquad \frac{sF(s)}{\lambda} = \frac{1}{\lambda}\mathcal{L}\left(\frac{df(t)}{dt}\right) + \frac{f(0)}{\lambda}$$

$$p_x(x) = -\frac{1}{\sqrt{2}}\frac{(x-\mu-\lambda\sigma^2)\, e^{\frac{x(2\mu-x)-\mu^2}{2\sigma^2}}}{\sqrt{\pi}\lambda\sigma^3} + \frac{1}{\sigma\sqrt{2\pi}\lambda} e^{-\mu^2/2\sigma^2} =$$

$$= -\frac{x\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\lambda\sigma^3\sqrt{2\pi}} + \frac{\mu+\lambda\sigma^2}{\lambda\sigma^3\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \frac{e^{-\mu^2/2\sigma^2}}{\lambda\sigma\sqrt{2\pi}}$$

PROBLEM 9.5 FADING CHANNEL (CONTINUE FROM N16R)

$$Y = XV + Z$$

KNOWLEDGE OF FADING FACTOR V INCREASES CAPACITY

$$I(X;Y|V) \geq I(X;Y)$$

$$I(X;Y|V) = H(Y|V) - H(Y|XV) = H(Y|V) - H(Z) \leq H(Y) - H(Z)$$

$$I(X;Y) = H(Y) - H(XV+Z|X) = H(Y) - H(V+Z)$$

$$H(f(X)) \leq H(X) \Rightarrow H(V+Z) = H(f(V,Z)) \leq H(V,Z)$$

$$I(X;Y) = H(Y) - H(V+Z) \geq H(Y) - H(V,Z) = H(Y) -$$

$$- [H(Z) + H(V|Z)] = H(Y) - H(Z) - H(V)$$

INDEPENDENT

$I(X;Z|Y) = I(X,Y;Z) - I(Y;Z)$

$I(X,Y;Z) = \underline{I(X;Z)} + I(Y;Z|X) = I(Y;Z) + \underline{I(X;Z|Y)}$

$I(Y;Z|X) = H(Z|X) - H(Z|X,Y) = H(Z|X) - H(Z)$

$H(Z|X) = H(Y+Z) \leq H(Y,Z) = H(Z) + H(Z|Y) = 2H(Z)$

$\boxed{I(Y;Z|X) \leq 2H(Z) - H(Z) = H(Z)}$

$I(Y;Z) = H(Z) - H(Z|Y) = H(Z) - H(X+Z)$

$I(Y;Z) = H(Y) - H(Y|Z)$

$I(X;Z) = I(Y;Z) - I(Y;Z|X) + I(X;Z|Y)$

$I(Y;Z) - I(Y;Z|X) = H(Z) - H(Z|Y) - \left(H(Z|X) - H(Z|X,Y)\right)$

$= \underbrace{H(Z) - H(Z|X)}_{>0} - \left(\underbrace{H(Z|Y) - H(Z|X,Y)}_{>0}\right)$

$= H(Z) - H(X+Z) - H(Y+Z) + H(Z) = (*)$

$= H(Z) + H(Z) - H(X+Z) - H(Y+Z) \geq H(Z) + H(Z) - H(X,Z) - H(Y,Z)$

$|H(X+Z) < H(X,Z)| \geq H(Z) + H(Z) - H(X) - H(Z|X) - H(Y) - H(Z|Y)$

$= H(Z) + \cancel{H(Z)} - H(X) - \cancel{H(Z)} - H(Y) - H(Z)$

$\boxed{H(Z) \leq H(X+Y+Z) \leq H(X,Y,Z) = H(X,Y) + \underbrace{H(Z|X,Y)}_{H(Z)} = H(X) + H(Y|X) + H(Z) = \underline{H(X) + H(Y) + H(Z)}}$

$(*) \leq H(X) + H(Y) + H(Z) - H(X+Z) - H(Y+Z) + H(Z) =$

$= H(X) + H(Y) + H(Z) - H(Z|Y) - H(Z|X) + H(Z) \leq$ | CONDITIONING REDUCES ENTROPY |

$\leq H(X) + H(Y) + H(Z) - \underbrace{H(Z|Y,X)}_{H(Z)} - \underbrace{H(Z|X,Y)}_{H(Z)} + H(Z) = (\#)$

$= H(X) + H(Y) + \cancel{2H(Z)} - \cancel{2H(Z)} = H(X) + H(Y)$

$I(X;Z) = I(Y;Z) - I(Y;Z|X) + I(X;Z|Y) \leq H(X) + H(Y) + I(X;Z|Y)$

$(\#) \leq H(X) + \cancel{H(Y)} + H(Z) - \underbrace{H(Z|Y,X)}_{\cancel{H(Z)}} - \underbrace{H(Z|X,Z)}_{H(Y)} + \cancel{H(Z)} = \underline{H(X) + H(Z)}$

$I(X;Z|Y) = H(Z|Y) - H(Z|X,Y) = H(Z|Y) - H(Z)$

$I(X;Z) = H(Z) - H(Z|X) \leq H(Z) - H(Z|X,Y) = H(Z) - H(Z)$

$I(X;Z|Y) = H(Z|Y) - H(Z|XY) = H(Z|Y) - H(Z)$

$I(X;Z) = H(Z) - H(Z|X) \le H(Z) - H(Z|XY) = H(Z) - H(Z)$

$Z = a \cdot X \qquad p(z) = \left.\frac{p(x)}{\frac{dg(x)}{dx}}\right|_{x=g^{-1}(z)} = \left.\frac{p(x)}{a}\right|_{x=\frac{z}{a}}$

$H(Z) = -\sum_{z} p(z)\cdot ld\, p(z) = -\sum_{z} \frac{p(x)}{a}\, ld \frac{p(x)}{a} =$

$= -a^{-1}\sum_{x} p(x)\left[ld\, p(x) - ld\, a\right] = \frac{1}{a}\sum_{x} p(x)\, ld\, p(x) + \frac{ld\, a}{a}$

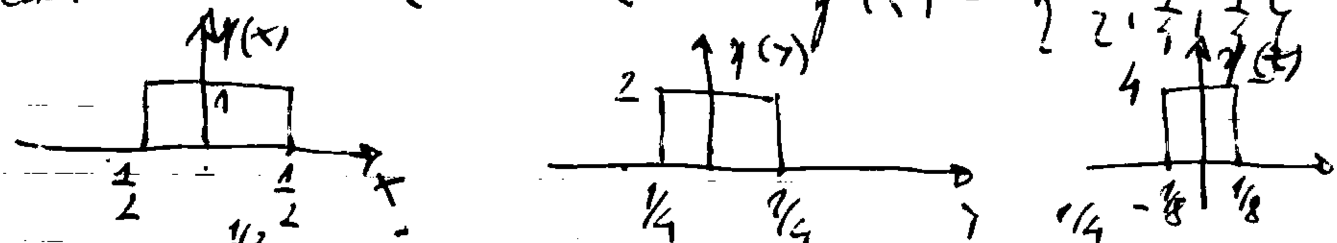
NEMA NIŠTA ZA DISCRETNI PROMENLJIVI

Za DISCRETNI PROMENLJIVI MISLAM DEKA $H(X) = H(Y)$

PRIMER: $x \in \{1, 2, 3\} \qquad p(x) = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\}$

$y \in \{2, 4, 6\} \qquad p(z) = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\}$

• CONT

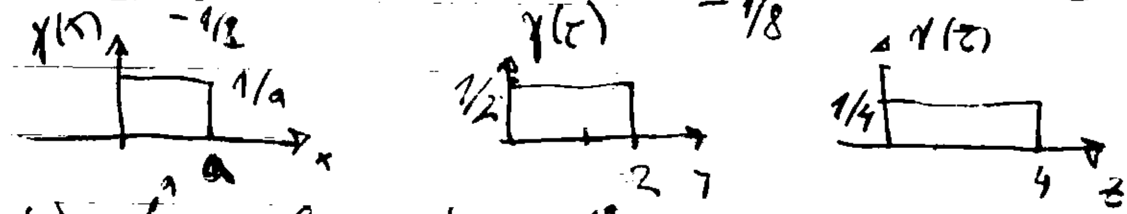


$h(X) = -\int_{-1/2}^{1/2} p(x)\, ld\, p(x)\, dx = 0 \qquad h(Y) = -\int_{-1/4}^{1/4} 2\cdot ld\, 2\, dy$

$h(Y) = -2y\Big|_{-1/4}^{1/4} = -2\left(\frac{1}{4} + \frac{1}{4}\right) = -2\,\frac{1}{2} = -1 \quad ?$

$h(Z) = \int_{-1/8}^{1/8} p(z)\, ld \frac{1}{p(z)}\, dz = \int_{-1/8}^{1/8} 4\, ld \frac{1}{4}\, dz = -4\cdot 2\left(\frac{1}{8} + \frac{1}{8}\right)$

$= -8\cdot\frac{1}{4} = -2$



$h(X) = -\int_0^a p(x)\cdot ld\, p(x)\, dx = +\int_0^a \frac{1}{a}\, ld\, a\, dx = +\frac{1}{a}\, ld\, a\; x\Big|_0^a = ld\, a$

$h(Y) = |a = 2| = 1 \qquad h(Z) = ld\, 4 = 2$

$Z = 2\cdot Y = 6\cdot X \qquad p(z) = \frac{p(x)}{6} = \frac{\frac{1}{2}}{2} = \frac{1}{4}$

$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$ $\quad y = 2x \quad$ $p(y) = \frac{p(x)}{\frac{dy}{dx}}\Big|_{x=\frac{y}{2}}$

$p(y) = \frac{1}{2\cdot\sigma_x\sqrt{2\pi}} e^{-\frac{y^2}{2\cdot 4\sigma_x^2}} = \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma_y^2}}$ $\quad \sigma_y = 2\sigma_x$

$h(x) = \frac{1}{2} ld(2\pi e)\sigma_x^2 \qquad h(y) = \frac{1}{2} ld(2\pi e)\sigma_y^2$

$h(y) = \frac{1}{2} ld(2\pi e) 4\cdot\sigma_x^2 = \frac{1}{2} ld(2\pi e)\sigma_x + \frac{1}{2} ld(2\pi e)\cdot 4$

$\boxed{h(y) \geq h(x)}$

BEST PROOF

$I(x;z|v) = H(z|v) - H(z|xv) = H(z|v) - H(Z)$

$I(x;z) = H(z) - H(z|x) = H(z) - H(z|x)$

$I(x;z) - I(x;z|v) = H(z) - H(z|x) - H(z|v) + H(z|xv)$

$= \underbrace{H(z) - H(z|v)}_{\geq 0} - [\underbrace{H(z|x) - H(z|xv)}_{\geq 0}] = I(v;z) - I(v;z|x)$ NONNEGATIVITY OF MUTUAL ENTROPY

$H(z|v) = H(v\cdot x + Z|v) \leq H(v,x,Z)$

$I(v;z|x) = H(v|x) - H(v|xy) =$ [scribbled out]

$z = vx + Z \qquad \boxed{z - Z = vx}$

VX → GLEDAJ GO KAKO EDNA PROMENLIVA.

(?) — DA PROVERAM DALI E TOCNO!!! $\qquad$ (a) CONDITIONING REDUCES ENTROPY

$H(z|x) - H(z|xv) \geq H(z) - H(z|v)$

$H(z|x) - H(Z) \geq H(z) - H(z|v) \overset{(a)}{\geq} H(z|x) - H(z|v)$

$H(Z) \leq H(z|v) = H(xv + Z|v) \leq H(x,v,Z) = H(x) + H(v) + H(Z)$ (b)

$\boxed{(b)\ H(f(x)) \leq H(x)}$ $\qquad H(Z) \leq H(x) + H(v) + H(Z)$

SO OVA DOKAZOV DEKA:

$H(x) + H(v) \geq 0$ } $\quad I(x;z) - I(x;z|v) \leq 0$ i.e.

$I(x;z) \leq I(x;z|v)$

OVA VAZI POVOLKU $h(x)$, $h(v)$ SE ≥ 0 (MORE DIFFERENTIAL [illegible])

$I(x;y) - I(x;z|v) = H(z) - H(z|x) - H(z|v) + H(z|xv) =$

$= H(z) - H(z|x) - H(z|v) + H(Z) \leq H(z) - H(z|xv) - H(z|v) + H(Z) \leq H(z) -$

$- H(z|v)$ REVERSED: $I(x;z|v) - I(x;z) = H(z|v) - H(z|xv) - H(z) + H(z|x) =$

$= H(z|x) - H(z) + H(z|v) - H(z|xv) \leq H(z|x) - H(z) + H(z|v) - H(Z) \leq H(z|x) - H(x)$

(?) $+ H(z|v) - H(Z) = H(xv+Z|v) - H(Z) \leq H(x,v,Z|v) - H(Z) \leq H(x,v,Z) - H(Z)$

$\leq H(x) + H(z) + H(Z) - H(Z) \Rightarrow \boxed{I(x;z|v) = I(x;z) + H(x) + H(z)}$ PROOF SS

REVISITED (THIRD TRY) PROOF: ⑦

$I(X;Z|V) - I(X;Z) = H(Z|V) - H(Z|XV) - \overbrace{H(Z)}^{H(Z)} + \overbrace{H(Z|X)}^{\geq H(Z|X)} \leq$
$\leq H(Z|V) - H(Z) - H(Z|X) + H(Z|X) = H(VX+Z|V) - H(Z)$
$\leq H(X,V,Z|V) - H(Z) = H(X) + H(Z) - H(Z) = \underline{H(X)}$!

$H(XVZ|V) = H(X|V) + \overbrace{H(V|XV)}^{0} + H(Z|XVV) = \overbrace{H(X|V)}^{H(X)} + \overbrace{H(Z|XV)}^{H(Z)}$

$H(X_1X_2X_3) = \sum_{i=1} H(X_i|X^{i-1}) = H(X_1) + H(X_2|X_1) + H(X_3|X_1X_2)$

$I(X;Z|V) - I(X;Z) = H(Z|V) - H(Z|XV) - H(Z) + H(Z|X)$
$= H(Z|V) - H(Z) + H(Z|X) - H(Z|XV) = \underbrace{H(Z|X) - H(Z|XV)}_{I(V;Z|X)\geq 0} - \underbrace{(H(Z) - H(Z|V))}_{I(Z;V)\geq 0}$

~~LA DA~~ PROVE $I(X;Z|V) \geq I(X;Z)$, I.E.:

$~~H(Z|X)~~ - \underbrace{H(Z|XV)}_{H(Z)} \geq H(Z) - H(Z|V) \geq_{(?)} ~~H(Z|X)~~ - H(Z|V)$

$H(Z) \leq H(Z|V) = H(XV+Z|V) \leq H(XVZ|V) =$
$H(X|V) + H(V|XV) + H(Z|XVV) = \underline{H(X)} + \theta + \underline{H(Z)}$
$~~H(Z)~~ \leq H(X) + ~~H(Z)~~$, V $\boxed{H(X) \geq \theta}$ TRUE !!

HENCE: $I(X;Z|V) - I(X;Z) \geq \theta$ $I(X;Z|V) \geq I(X;Z)$

- BERKLEY SOLUTION (soln_5_229ASpr07.pdf)

$I(X;VZ) = \underbrace{I(X;V)}_{\theta} + I(X;Z|V) = I(X;Z|V)$

$I(X;V) = H(V) - H(V|X) = H(V) - H(V) = \theta$

$I(X;VZ) = I(X;Z) + \underbrace{I(X;V|Z)}_{\geq\theta} \geq I(X;Z)$

$I(X;Z|V) = I(X;Z) + I(X;V|Z) \geq I(X;Z)$

$\boxed{I(X;Z|V) \geq I(X;Z)}$

PROBLEM 9.6 PARALLEL CHANNELS AND WATER-FILLING.

CONSIDER A PAIR OF PARALLEL GAUSSIAN CHANNELS:

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

WHERE: $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right)$

AND THERE IS A POWER CONSTRAINT $E[x_1^2 + x_2^2] \leq 2P$

ASSUME THAT $\sigma_1^2 > \sigma_2^2$. AT WHAT POWER DOES THE CHANNEL STOP BEHAVING LIKE A SINGLE CHANNEL WITH NOISE VARIANCE σ_2^2 AND BEGIN BEHAVING LIKE A PAIR OF CHANNELS?

RECALL

$$I(x_1^n; z_1^n) = h(x_1^n) - h(z_1^n | x_1^n) = h(z_1^n) - h(z_1^n)$$

$$= h(z_1^n) - \sum_{i=1} h(z_i) \leq \sum_i h(z_i) - h(z_i) \leq$$

$$\leq \sum_i \frac{1}{2} \log\left(\frac{\sigma_{x_i}^2 + \sigma_{z_i}^2}{\sigma_{z_i}^2}\right) = \sum_i \frac{1}{2} \log\left(1 + \frac{P_{x_i}}{N_i}\right)$$

$P_i = E[x_i^2] \qquad \sum P_i = P$

$$x_1^n \sim N\left(0, \begin{bmatrix} P_1 & 0 & \cdots & 0 \\ 0 & P_2 & \cdots & 0 \\ 0 & \cdots & \cdots & P_n \end{bmatrix}\right)$$

$$J(P_1^k) = \sum_i \frac{1}{2} \ln\left(1 + \frac{P_i}{N_i}\right) + \lambda \sum_i P_i$$

$$\frac{dJ(P_1^k)}{dP_i} = 0 \qquad \frac{1}{2} \frac{1}{1 + \frac{P_i}{N_i}} \cdot \left(\frac{1}{N_i}\right) + \lambda = 0$$

$$\frac{1}{2} \frac{1}{N_i + P_i} + \lambda = 0 \qquad -\lambda = \frac{1}{2(N_i + P_i)} \qquad N_i + P_i = -\frac{1}{2\lambda}$$

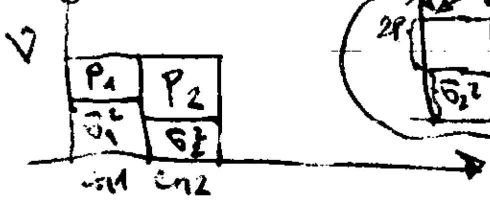
$$\boxed{P_i = -\frac{1}{2\lambda} - N_i = \nu - N_i} \qquad P_i = (\nu - N_i)^+$$

$$\boxed{\sum_i (\nu - N_i)^+ = n \cdot P} \qquad \boxed{(x)^+ = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}}$$

$$(\nu - \sigma_1^2)^+ + (\nu - \sigma_2^2)^+ = 2P \qquad \boxed{\sigma_1^2 > \sigma_2^2}$$

$$2\nu - (\sigma_1^2 + \sigma_2^2) = 2P \qquad 2(\nu - P) = \sigma_1^2 + \sigma_2^2$$

$$2\nu = 2P + \sigma_1^2 + \sigma_2^2 \qquad \nu = \frac{2P + \sigma_1^2 + \sigma_2^2}{2} \qquad \boxed{(\nu - \sigma_1^2)^+ = 0}$$



(★) $\boxed{\sigma_1^2 > \nu}$ ⟹ $(\nu - \sigma_2^2)^+ = 2P$; $\nu \geq \sigma_2^2$

$\boxed{\nu = 2P + \sigma_2^2}$ SINGLE CHANNEL

σ_2^2 WITH NOISE VARIANCE σ_2^2 AND POWER 2P

$$(\nu-\sigma_1^2)^+ + (\nu-\sigma_2^2)^+ = 2P \qquad (\nu-\sigma_1^2)^+ = \nu-\sigma_1^2$$

IF $\sigma_1^2 \le \nu$ AND $\sigma_2^2 \le \nu$ $\Rightarrow$ $2\nu = 2P + \sigma_1^2 + \sigma_2^2$ } (IF $\nu \ge \sigma_2^2$)

$$\nu = \frac{2P+\sigma_1^2+\sigma_2^2}{2}$$

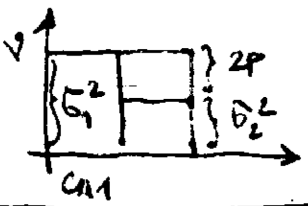
VSŠNOST KOGA: $\sigma_1^2 \le \nu$ TOGAŠ: $(\nu-\sigma_2^2)^+ = 2P - \nu + \sigma_1^2$

– OD (*) SE GLEDA DEKA $\boxed{2P = \sigma_1^2 - \sigma_2^2}$ } OVA NE MI TOČNA ACE DA PROŠTAM SOLUTION.

NA OVAA SNAGA E GRANICATA!!!

• University of Toronto Solution.

We will put all the signal power into channel with less noise until the total power of noise + signal in the channel equals the noise power in other channel. After that, we will split any additional power evenly between the two channels.



So the combined channel begins to behave like a pair of parallel channels when signal power: $2P = \sigma_1^2 - \sigma_2^2$

PROBLEM 9.13 Feedback Capacity (CONTINUE FROM NNbr)

$$(Z_1, Z_2) \sim N(0, K) \qquad K = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

$$\max_{tr(K_x)\le 2P} \frac{1}{2}\log\frac{|K_{x+z}|}{|K_z|} = C_{n,FB}$$

$$C_n = \max_{tr(K_x)\le 2P} \frac{1}{2}\log\frac{|K_x| + |K_z|}{|K_z|} = \frac{1}{2n}\sum_{i=1}^{n}\log\left(1+\frac{(\lambda-\lambda_i^{(z)})^+}{\lambda_i^{(z)}}\right)$$

$$K = V\Lambda V^{-1} \qquad \Lambda = \begin{bmatrix} \rho+1 & 0 \\ 0 & 1-\rho \end{bmatrix} \qquad \lambda_1 = \rho+1,\ \lambda_2 = 1-\rho$$

• WITHOUT FEEDBACK

$$(\lambda-\lambda_1)^+ = P_1 \qquad (\lambda-\lambda_2)^+ = P_2$$

$$P_1 - \lambda_1 = P_2 - \lambda_2 \qquad \therefore\ P_1 + P_2 = 2P$$

$$P_1 = 2P - P_2$$

$$2P - P_2 - \lambda_1 = P_2 - \lambda_2$$

$$2P - 2P_2 - \lambda_1 + \lambda_2 = 0 \qquad 2P_2 = 2P - \lambda_1 + \lambda_2 \qquad \boxed{P_2 = \frac{2P-\lambda_1+\lambda_2}{2}}$$

$$P_1 = 2P - P_2 = 2P - \frac{2P-\lambda_1+\lambda_2}{2} = \frac{4P - 2P + \lambda_1 - \lambda_2}{2}$$

$$\boxed{P_1 = \frac{2P+\lambda_1-\lambda_2}{2}} \qquad P_1 = \frac{2P + 1 + \rho - 1 + \rho}{2} = \frac{2P+2\rho}{2} = P+\rho$$

$$P_2 = \frac{2P - \rho - 1 + 1 - \rho}{2} = \frac{2P-2\rho}{2} = \rho - \rho$$

$P_1 + P_2 = P + \varrho + P - \varrho = 2P$ $\quad P_1 = P + \varrho \quad P_2 = P - \varrho$

$$C_2 = \frac{1}{4}\left\{ ld\left(1 + \frac{P_1}{\lambda_1}\right) + ld\left(1 + \frac{P_2}{\lambda_2}\right)\right\} = \frac{1}{4} ld\left(1 + \frac{P+\varrho}{1+\varrho}\right)$$

$$+ \frac{1}{4} ld\left(1 + \frac{P-\varrho}{1-\varrho}\right) = \frac{1}{4} ld\left(1 + \frac{P-\varrho}{1-\varrho}\right)\left(1 + \frac{P+\varrho}{1+\varrho}\right) =$$

$$= \frac{1}{4} ld \frac{(1 + P - 2\varrho)(1 + P + 2\varrho)}{1-\varrho^2} = \frac{1}{4} ld \frac{(1+P)^2 - (2\varrho)^2}{1-\varrho^2}$$

$$\boxed{C_2 = \frac{1}{4} ld \frac{1 + 2P + P^2 - 4\varrho^2}{1-\varrho^2}}$$

52550
44000
8550

■ WITH FEEDBACK

$$C_{n,FB} = \max_{tr(K_x) \le 2P} \frac{1}{2} ld \frac{|K_x + z|}{|K_z|} \le \max_{tr(K_x)\le 2P} \frac{1}{2} ld \frac{|K_x + K_z|}{|K_z|} + \frac{1}{2} ld\, 2$$

$$\le \frac{1}{2n} ld \frac{|K_x|^{\frac{1}{2}} |K_z|^{\frac{1}{2}} 2^n}{|K_z|} = \frac{1}{2 \cdot 2} ld \frac{|K_x|}{K_z} = \frac{1}{4} ld \frac{|K_x|}{|K_z|} + \frac{1}{2}$$

$$C_{2,FB} \le \frac{1}{4} ld \frac{|K_x|}{(1-\varrho^2)} + \frac{1}{2} = \frac{1}{4} ld \frac{P^2}{(1-\varrho^2)} + \frac{1}{2}$$

MOZEM: $|K_x| = P_1^2 + P_2^2 = (P+\varrho)^2 + (P-\varrho)^2 =$

$$= P^2 + 2P\varrho + \varrho^2 + P^2 - 2P\varrho + \varrho^2 = 2P^2 + 2\varrho^2$$

$$C_{2,FB} = \frac{1}{4} ld \frac{2P^2 + 2\varrho^2}{(1-\varrho^2)} + \frac{1}{2}$$

$$C_{2,FB} - C_2 = \frac{1}{4} ld \frac{2P^2 + 2\varrho^2}{1 + 2P + P^2 - 4\varrho^2} + \frac{1}{2} > 0 \Rightarrow$$

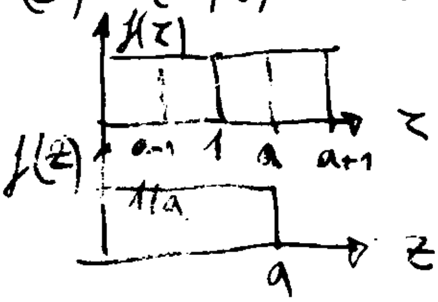
$$\boxed{C_{2,FB} \ge C_2}$$

PROBLEM 9.15 (CONTINUE FROM N16n) DISCRETE INPUT CONTINUOUS OUTPUT

$Pr(X=1) = p \;;\; Pr(X=0) = 1-p \;;\; Y = X + Z \quad Z \sim U[0,a],\ a>1$

(a) $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y - Z | Y) = H(Y) - ld\, a$, $= H(Z)$

(b) $I(X;Y) = H(Y) - H(Y|X) = h(Y) - H(Z) = h(Y) - ld\, a$



$H(Y) = ?$

$$\gamma(0 \le Y \le 1) = \gamma(X=0)\cdot\gamma(Z \le 1) = \frac{p_0}{a}$$

$$\gamma(1 \le Y \le a) = \gamma(X=1)\cdot\gamma(0 \le Z \le a-1) + \gamma(X=0)\cdot\gamma(1 \le Z \le a) = (p_0 + p_1)\cdot\frac{a-1}{a}$$

$$\int_0^{a-1} \frac{1}{a} dz = \frac{1}{a} z \Big|_0^{a-1} = \frac{a-1}{a} \qquad \int_1^a \frac{1}{a} dz = \frac{1}{a}(a-1)$$

$$P(a \le \tau \le a+1) = P(t=1) \cdot P(z \ge a-1) = p_1 \cdot \int_{a-1}^{a} \frac{1}{a} dz = p_1 \frac{a-a+1}{a} = p_1/a$$

$$P(\tau) = \left\{ \underset{P(\tau_1)}{\overset{0 \div 1}{\frac{p_0}{a}}},\ \underset{P(\tau_2)}{\overset{1 \div a}{\frac{a-1}{a}}},\ \underset{P(\tau_3)}{\overset{a \div a+1}{\frac{p_1}{a}}} \right\}$$

$$P(\tau_2) = 1 - \frac{p_0}{a} - \frac{p_1}{a} = 1 - \frac{1}{a} = \frac{a-1}{a} = 1 - \frac{1}{a}$$

$$-h(\tau) = \int_0^1 \frac{p_0}{a} \operatorname{ld} \frac{p_0}{a} d\tau + \int_1^a \frac{a-1}{a} \operatorname{ld} \frac{a-1}{a} d\tau + \int_a^{a+1} \frac{p_1}{a} \operatorname{ld} \frac{p_1}{a} d\tau$$

$$= \frac{p_0}{a} \operatorname{ld} \frac{p_0}{a} + \frac{a-1}{a}\left(\operatorname{ld} \frac{a-1}{a}\right)(a-1) + \frac{p_1}{a} \operatorname{ld} \frac{p_1}{a} \cdot 1 =$$

$$= \frac{1}{a}\left(p_0 \operatorname{ld} p_0 + p_1 \operatorname{ld} p_1\right) + \frac{p_0}{a} \operatorname{ld} \frac{1}{a} + \frac{p_1}{a} \operatorname{ld} \frac{1}{a} + \frac{(a-1)^2}{a} \operatorname{ld} \frac{a-1}{a}$$

$$= -\frac{H(p_0)}{a} + \frac{1}{a} \operatorname{ld} \frac{1}{a} + \frac{(a-1)^2}{a} \operatorname{ld} \frac{a-1}{a}$$

$$\boxed{h(\tau) = +\frac{H(p)}{a} + \frac{1}{a} \operatorname{ld} a + \frac{(a-1)^2}{a} \operatorname{ld}\left(\frac{a}{a-1}\right)}$$

(*) $I(t;\tau) = h(\tau) - \operatorname{ld} a$

$$I(t;\tau) = +\frac{H(p)}{a} + \frac{1}{a} \operatorname{ld} a + \frac{a^2-2a+1}{a}\left(\operatorname{ld} a + \operatorname{ld} \frac{1}{a-1}\right) - \operatorname{ld} a$$

$$= +\frac{H(p_0)}{a} + \frac{1}{a} \operatorname{ld} a + a \operatorname{ld} a - 2 \operatorname{ld} a + \frac{1}{a} \operatorname{ld} a + \frac{(a-1)^2}{a} \operatorname{ld} \frac{1}{a-1} - \operatorname{ld} a$$

$$= +\frac{H(p_0)}{a} + \frac{2}{a} \operatorname{ld} a - 3 \operatorname{ld} a + \frac{(a-1)^2}{a} \operatorname{ld} \frac{1}{a-1} + a \operatorname{ld} a =$$

$$= +\frac{H(p_0)}{a} + \operatorname{ld} a \left(\frac{2}{a} - 3 + a\right) + \frac{(a-1)^2}{a} \operatorname{ld} \frac{1}{a-1} =$$

$$= +\frac{H(p_0)}{a} + \operatorname{ld} a \left(\frac{2-3a+a^2}{a}\right) + \frac{(a-1)^2}{a} \operatorname{ld} \frac{1}{a-1} =$$

$$= +\frac{H(p_0)}{a} + \operatorname{ld} a \left[\frac{2-3a+a^2}{a} - \frac{a^2-2a+1}{a}\right] + \frac{(a-1)^2}{a} \operatorname{ld} \frac{a}{a-1}$$

$$= +\frac{H(p_0)}{a} + \operatorname{ld} a \left[\frac{1-a}{a}\right] + \frac{(a-1)^2}{a} \operatorname{ld} \frac{a}{a-1} =$$

$$= +\frac{H(p_0)}{a} + \left(\frac{1}{a} - 1\right) \operatorname{ld} a + \frac{(a-1)^2}{a} \operatorname{ld} \frac{1}{1-\frac{1}{a}} =$$

$$= +\frac{H(p_0)}{a} + \left(1 - \frac{1}{a}\right) \operatorname{ld} \frac{1}{a} + \frac{a^2-2a+1}{a} \operatorname{ld} \frac{1}{1-\frac{1}{a}} =$$

$$= +\frac{H(p_0)}{a} + \operatorname{ld} \frac{1}{a} - \frac{1}{a} \operatorname{ld} \frac{1}{a} + \left(a - 2 + \frac{1}{a}\right) \operatorname{ld} \frac{1}{1-\frac{1}{a}}$$

$$\boxed{H\left(\frac{1}{a}\right) = \frac{1}{a} \operatorname{ld} a + \left(1 - \frac{1}{a}\right) \operatorname{ld} \frac{1}{1-\frac{1}{a}}}$$

$$\to (a-1) \operatorname{ld} \frac{1}{1-\frac{1}{a}} - \left(1 - \frac{1}{a}\right) \operatorname{ld} \frac{1}{1-\frac{1}{a}}$$

$$\operatorname{ld} \frac{1}{a} - \frac{1}{a} \operatorname{ld} \frac{1}{a} + (1-a) \operatorname{ld} \frac{1}{1-\frac{1}{a}} - 1 + \frac{1}{a} \operatorname{ld} \frac{1}{1-\frac{1}{a}}$$

$$= \operatorname{ld} \frac{1}{a} + \frac{1}{a}\left(\operatorname{ld} \frac{1}{1-\frac{1}{a}}\right) - (1-a) \operatorname{ld} \frac{a}{a-1}$$

$\boxed{\gamma = \frac{1}{a}}$ $\quad I(X;Y) = +H(\gamma_0)\cdot\gamma + \mathrm{ld}\,\gamma - \gamma\,\mathrm{ld}\,\gamma + \left(\frac{1}{\gamma} - 2 + \gamma\right)\mathrm{ld}\frac{1}{1-\gamma}$

$= +H(\gamma_0)\cdot\gamma + \mathrm{ld}\,\gamma + \gamma\,\mathrm{ld}\frac{1}{\gamma} + \left(\frac{1}{\gamma} - 1\right)\mathrm{ld}\frac{1}{1-\gamma} - (1-\gamma)\,\mathrm{ld}\frac{1}{1-\gamma}$

$= +H(\gamma_0)\cdot\gamma - \mathrm{ld}\left(\frac{1}{\gamma}\right) + \gamma\,\mathrm{ld}\frac{1}{\gamma} + \frac{1-\gamma}{\gamma}\,\mathrm{ld}\frac{1}{(1-\gamma)} - (1-\gamma)\,\mathrm{ld}\frac{1}{1-\gamma} =$

$= +H(\gamma_0)\cdot\gamma - (1-\gamma)\,\mathrm{ld}\frac{1}{\gamma} + (1-\gamma)\left(\mathrm{ld}\frac{1}{1-\gamma}\right)\left(\frac{1}{\gamma} - 1\right) =$

$= +H(\gamma_0)\cdot\gamma - \frac{(1-\gamma)}{\gamma}\left[\gamma\,\mathrm{ld}\left(\frac{1}{\gamma}\right) - (1-\gamma)\,\mathrm{ld}\frac{1}{1-\gamma}\right] =$

$= H(\gamma_0)\cdot\gamma + \left(\frac{1-\gamma}{\gamma}\right)\underbrace{\left[+\gamma\,\mathrm{ld}\frac{1}{\gamma} + (1-\gamma)\,\mathrm{ld}\frac{1}{1-\gamma}\right]}_{H(\gamma)} - 2\left(\frac{1-\gamma}{\gamma}\right)\gamma\,\mathrm{ld}\frac{1}{\gamma}$

$= \frac{H(\gamma_0)}{a} + \left(\frac{1}{\gamma} - 1\right)\cdot H\left(\frac{1}{a}\right) - 2(1-\gamma)\,\mathrm{ld}\frac{1}{\gamma}$

$$\boxed{I(X;Y) = \frac{H(\gamma_0)}{a} + (a-1)\,H\left(\frac{1}{a}\right) - 2\left(1 - \frac{1}{a}\right)\mathrm{ld}\,a}$$ → TOČEN IZRAZ ~~...~~ SO (*)

(c) $\boxed{\frac{\gamma_0}{a} = \frac{a-1}{a}\cdot(a-1) = \frac{\gamma_1}{a} = \frac{1}{3}}$ (*) $\quad \frac{a^2-2a+1}{a} = \frac{1}{3}$

$a^2 - 2a + 1 = \frac{a}{3} \qquad a^2 - \frac{7a}{3} + 1 = 0 \Rightarrow \boxed{a = \frac{7}{6} + \frac{\sqrt{13}}{6}}$

$C = \log_2(3) - \mathrm{ld}\,a = \underline{0.7632} \qquad \gamma_0 = (a-1)^2 = \underline{0.5892}$

$\gamma_1 = 1 - \gamma_0 = 1 - (a-1)^2 = \underline{0.4108}$ } KONTRADIKCIJA SO (*)

$h(Y) = \frac{H(\gamma_0)}{a} + \frac{1}{a}\,\mathrm{ld}\,a + \frac{(a-1)^2}{a}\,\mathrm{ld}\frac{a}{a-1}$

→ GI EDNAČIŠ POVRŠINITE; ZA DA $\gamma(c) = \left\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right\}$ TOGAŠ DOBIVAŠ $h(Y)\big|_{max} = \mathrm{ld}\,3$

OD (*) ⇒ DEKA MORA $\gamma_0 = \gamma_1$ NO NE IZLEGUVA TAKA

~~[illegible]~~

$(a-1)^2 = \frac{1}{2} \qquad a^2 - 2a + 1 = \frac{1}{2}$

$a^2 - 2a + \frac{1}{2} = 0 \Rightarrow \boxed{a^* = 1 + \frac{\sqrt{2}}{2}} \qquad a \approx 1.7071$

ZA VAKVO a^* = $\quad \boxed{\gamma_0 = \left(1 + \frac{\sqrt{2}}{2} - 1\right)^2 = \frac{2}{4} = \frac{1}{2}}$

$\boxed{\gamma_1 = 1 - \gamma_0 = \frac{1}{2}}$ ⇒ MAKSIMALNIOT KAPACITET SE DOBIVA ZA SIMETRIČNA RASPREDELBA NA VLEZNITE SIMBOLI.

$$\boxed{C = I(X;Y)\big|_{max} = \frac{1}{a} + (a-1)\,H\left(\frac{1}{a}\right) - 2\left(1 - \frac{1}{a}\right)\mathrm{ld}\,a = 0.6386}$$

SOLUTION

$$\int_0^{a+1} f(y)\,dy = \frac{f_0}{a}\cdot 1 + \frac{a-1}{a}(a-1) + \frac{f_1}{a} = \frac{1}{a} + \frac{(a-1)^2}{a} = 0.8787$$

TROPPO $\neq 1$

NEÒTO NE E VOLED

CONCORDIA SOLUTION

(a)

$$f(y \mid x=0) = \begin{cases} \frac{1}{a} & 0 \le y \le a \\ 0 & \text{otherwise} \end{cases}$$

$$f(y \mid x=1) = \begin{cases} (1-y)\frac{1}{a} & 0 < y \le 1 \\ \frac{1}{a} & 1 \le y \le a \\ \frac{1}{a}y & a \le y \le a+1 \end{cases}$$

$$f(y) = \begin{cases} (1-y)\frac{1}{a} & 0 < y \le 1 \\ \frac{1}{a} & 1 \le y \le a \\ \frac{1}{a}y & \end{cases}$$

(b) $H(X) = H(Y)$ $\quad h(x \mid y=y)$ is nonzero only for $1 \le y \le a$, and by Bayes rule, conditioned on y, the probability that $x=1$ is:

$$p(x=1 \mid y=y) = \frac{P(x=1)\, f(y \mid x=1)}{P(x=1)\, f(y \mid x=1) + P(x=0)\, f(y \mid x=0)} = p$$

BAYES' THEOREM

$$P(W \mid L) = \frac{P(L \mid W)\, P(W)}{P(L)} \qquad P(W, L) = P(W, L)$$

$$P(L)\cdot P(W \mid L) = P(L \mid W)\, P(W) \qquad P(W \mid L) = \frac{P(L \mid W)\, P(W)}{P(L)}$$

$$p(x=1, y) = p(x=1, y)$$

$$p(x=1 \mid y)\cdot p(y) = p(x=1)\cdot P(y \mid x=1)$$

$$p(x=1 \mid y) = \frac{p(x=1)\, p(y \mid x=1)}{p(y)} = \frac{P(x=1)\, P(y \mid x=1)}{P(x=0)\cdot P(y \mid x=0) + P(x=1)\, P(y \mid x=1)}$$

$$p(x=1 \mid y) = \frac{p \cdot \frac{1}{a}}{(1-p)\frac{1}{a} + p\frac{1}{a}} = \frac{p \cdot \frac{1}{a}}{\frac{1}{a}} = p$$