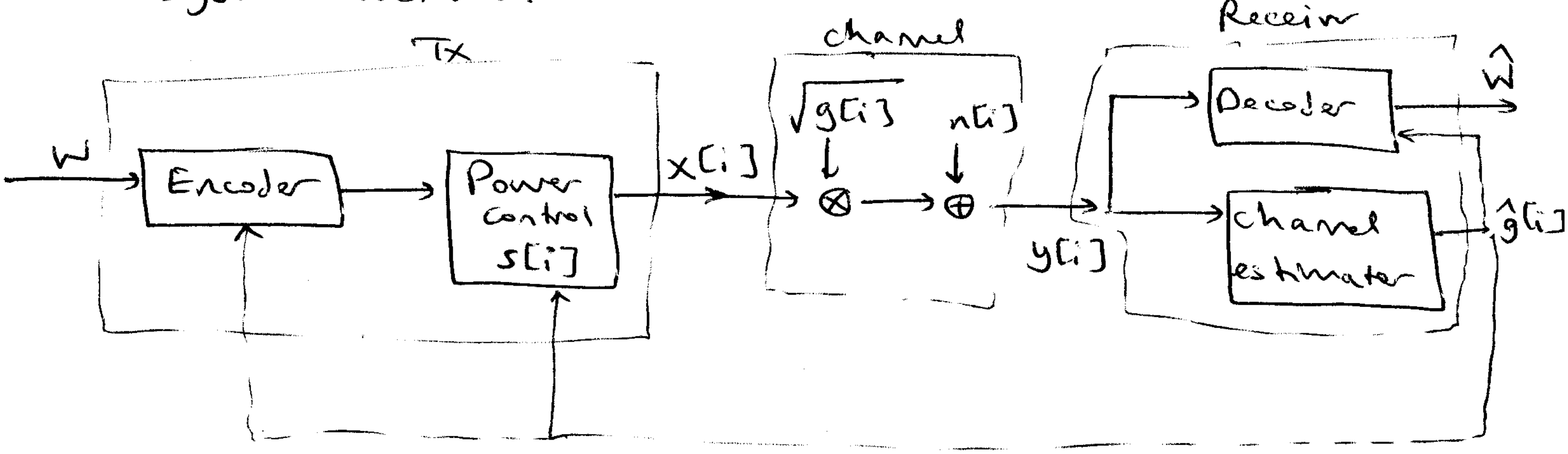
LECTURE #9

— Last time, we examined the capacity of fading channels under different cases of tx/rx side information. Our basic system model was:



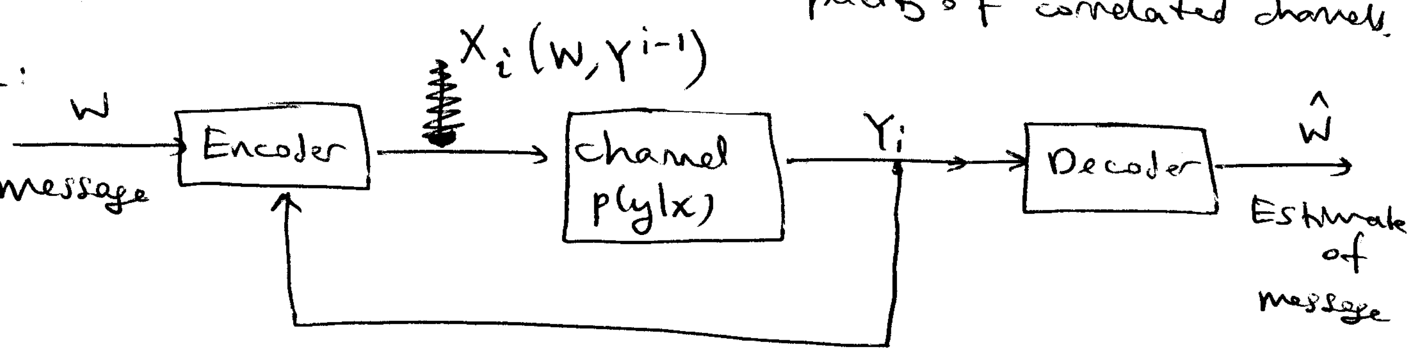
Review of capacity of channels with feedback:

> **Theorem 1.** Feedback does not increase the capacity of discrete memoryless channels, i.e. $C_{FB} = C$.

However, 1) Feedback may lower the complexity of the receiver implementation,

2) Feedback does increase the capacity of correlated channels.

model:



Assumptions:

- Assume that all received symbols are sent back immediately and noiselessly to the Tx.

Pf (Proof of the theorem):

Define a $(2^{nR}, n)$ feedback code as a sequence of

$$(\underbrace{2^{nR}}_{\substack{\text{\# of}\\ \text{codewords}}}, \underbrace{n}_{\substack{\text{length of a}\\ \text{codeword}}})$$

mappings $X_i(W, Y^{i-1})$. [These mappings are performed at the Tx. (Note that each mapping for symbol i uses the message W and the previous $\{Y^{i-1}\}$ $(Y_1, \ldots, Y_{i-1})$ which the Tx received via feedback.] and

a sequence of decoding functions $g: \mathcal{Y}^n \longrightarrow \{1, 2, \ldots, 2^{nR}\}$. Then

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

when $W \sim$ uniform $\{1, \ldots, 2^{nR}\}$

Definition: $C_{FB} = \sup\{\text{achievable rates by feedback codes}\}$

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

Pf (continued):

Note that $C_{FB} \geq C$ (since not using any Y^{i-1} amounts to a code without feedback)

If $W \sim$ Uniform $\{1, \dots, 2^{nR}\}$, then

$$\boxed{nR \overset{\text{since } W \sim \text{uniform}}{=} H(W)} = H(W|Y^n) + I(W;Y^n) \quad \text{(by def'n of mutual information)}$$

$$\leq [1 + P_e^{(n)} \cdot nR] + I(W;Y^n) \quad \text{(by Fano's inequality)}$$

Now,

$$I(W;Y^n) = H(Y^n) - H(Y^n|W) \quad \text{by def'n of } I(\cdot,\cdot).$$

$$= H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1,\dots,Y_{i-1},W) \quad \text{[chain rule]}$$

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

(X_i is a fn of W and $Y_1,\dots,Y_{i-1}$)

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

(since Y_i is independent of W and $Y_1,\dots,Y_{i-1}$ given X_i).

Then:

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

$$\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \quad \left[\text{Since } H(Y^n) \leq \sum_{i=1}^{n} H(Y_i)\right]$$

(= if they ~~pmf~~ marginals)

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

Then: (from p.3)

$$nR \le \left[1 + P_e^{(n)} \cdot nR\right] + nC.$$

$$\Leftrightarrow \quad R \le \frac{1}{n} + P_e^{(n)} \cdot R + C.$$

$$\Rightarrow \quad R \le \frac{1}{n} + C$$

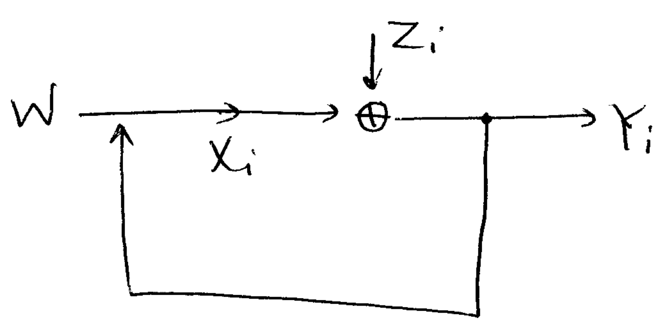
$n \to \infty$: $\quad R \le C.$

Hence: $\quad C_{FB} = C.$

Gaussian Channels with Feedback:

- For memoryless (DMC) Gaussian channels, we know from the above Theorem 1, that feedback does not increase capacity.
- **But**: for correlated channels, feedback **does** increase capacity – we demonstrate this for correlated Gsn channels.

Model:



$$Y_i = X_i + Z_i, \qquad Z_i \sim \mathcal{N}(0, K_Z^{(n)})$$

covariance matrix of the noise process

- A $(2^{nR}, n)$ code for the Gsn channel with feedback is:

  a sequence of mappings $x_i(W, Y^{i-1})$

  (where $W \in \{1, 2, .., 2^{nR}\}$ is the input message)

  Y^{i-1}: sequence of past values of ~~input~~ output

<u>Power constraint</u>:

max power level

$$\underbrace{E}_{\text{over all possible noise sequences}}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2(\underbrace{w}_{\text{a fixed } w}, Y^{i-1})\right] \le P \quad \forall w \in \{1, .., 2^{nR}\}$$

— Because of feedback, X^n and Z^n are <u>not</u> independent.

(X_i may depend on the previous values of Z)

<u>Results</u>:

1. <u>With feedback</u>: $C_{n,FB}$ (bits per transmission of the ~~time-varying~~ correlated Gsn channel)

$$\boxed{C_{n,FB} = \max_{\underbrace{\frac{1}{n}tr(K_X^{(n)})}_{\text{trace}} \le P} \frac{1}{2n}\log_2 \frac{|K_{X+Z}^{(n)}|}{|K_Z^{(n)}|}}$$

where max is taken over <u>all</u> X^n of the form:

$$X_i = \sum_{j=1}^{i-1} b_{ij} Z_j + V_i \qquad i = 1, 2, .., n.$$

and $V^n \perp\!\!\!\perp Z^n$.

[That is, the input encoding is done as a linear combination of the past noise samples, and a linearly additive, independent ~~function~~ sequence V^n.]

[See Cover p. 258 for the argument that this gives no loss of generality since the channel is Gsn.]

Write $X = BZ + V$. ← (i-dimensional)

$Y = X + Z$

$$C_{n,FB} = \max \frac{1}{2n} \log_2 \frac{(B+I) K_Z^{(n)} \cdot (B+I)^T + K_V}{|K_Z^{(n)}|}$$

[no feedback ⇒ $B = 0$]

taken over all ~~≥ 0 def~~ non-negative definite K_V and lower triangular B s.t. $tr(B K_Z^{(n)} + K_V) \leq nP$

2. Without feedback:

$$C_n = \max_{\frac{1}{n} tr(K_X^{(n)}) \leq P} \frac{1}{2n} \log_2 \frac{|K_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|}.$$

THM: $\boxed{C_{n,FB} \leq C_n + \frac{1}{2}}$ (bits per transmission)

∴ Feedback increases the capacity of a correlated Gsn channel by at most $\frac{1}{2}$ bit per transmission.

(Recall. feedback channel assumed noiseless

Now, let us turn to fading channels; in particular

- Side information at the Receiver:

We said last time that:

$$C(S) = \int_0^\infty B \log_2(1+\gamma)\, p(\gamma)\, d\gamma \; .$$

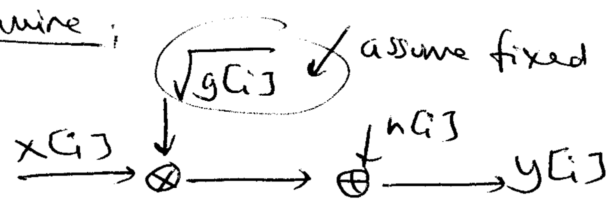
$\Rightarrow$ at Tx
[no power adaptation;
no rate adaptation]

which is the same expression as if tx had the knowledge of $g[i]$, but could not do power adaptation.

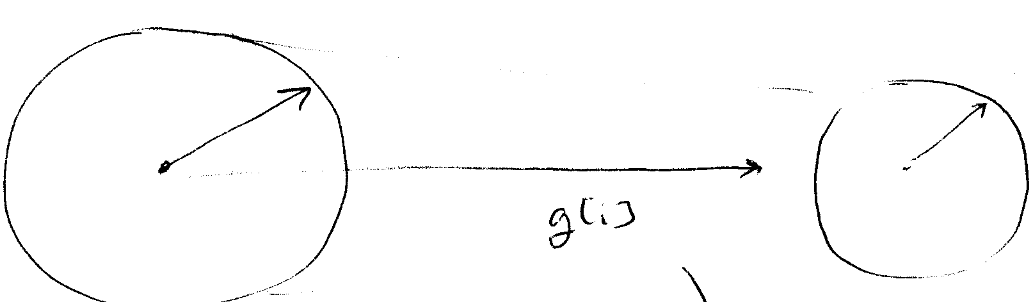
- Why is this the case?

Reason: Note that pf of Shannon capacity not constructive [does not tell you how to code except random codes]

Key Scenario to examine:



Remember, the proof of channel coding thm for the Gsn channel.



generate n-bit codewords randomly with power level

due to channel gain (attenuation), the sphere shrinks: But if receiver knows the $g[i]$, can scale it back.

Note that the receiver knows the codebook $[X^n(1), \ldots, X^n(2^{nR})]$, hence can decode using joint typicality. [Roughly 2^{nI} of the sequences are jointly typical.]

Hence: Capacity of an AWGN channel with gain $g[i]$ known only at the receiver is: $B \log_2(1+\gamma)$. $\left(\gamma = SNR = \frac{g[i] \cdot \mathcal{E}_x}{\sigma^2} \right)$

– using Wolfowitz's result gives: $\int_0^\infty B \log_2(1+\gamma) p(\gamma) \, d\gamma$.

$\triangle$

— Let's go back to Lecture 8, p.13, to continue our development of <u>Adaptive modulation</u> with M-QAM constellations.

— Symbol period: T_s

- M-QAM
- Assume ideal Nyquist pulses

<u>Definitions</u>:

$\overline{S} \triangleq$ avg power constraint at <u>Tx</u>.

$$\int_0^\infty S(\gamma)\, p(\gamma)\, d\gamma \leq \overline{S}$$

B = bandwidth (Hz)

N_0 = noise spectral density

$$\gamma[i] = \frac{\overline{S} \cdot g[i]}{N_0 \cdot B} \quad \text{(SNR of ith symbol)}$$

$$\left[\overline{\gamma} = \frac{\overline{S}}{N_0 \cdot B}\right]$$

$$\left[\frac{E_s}{N_0} = \frac{\overline{S} \cdot T_s}{N_0} = \overline{\gamma}\right]$$

For fixed M:

$$R = \frac{\log_2 M}{T_s}.$$

M-QAM modulation

(with ideal coherent phase detection)

Bounds:

$$BER \le 2\, e^{-1.5\bar{\gamma}/(M-1)} \quad \text{[AWGN channel, M-QAM]}$$

For $M \geq 4$ and $0 \le \bar{\gamma} \le 30$ dB, a tighter bound:

$$BER \le 0.2\, e^{-1.5\bar{\gamma}/(M-1)} \quad \text{[AWGN channel, M-QAM]}$$

Under adaptive modulation:

$$\boxed{BER(\gamma) \le 0.2\, e^{-1.5\left[\gamma \cdot \frac{S(\gamma)}{\bar{S}}\right]/(M-1)}}$$

Idea: Adjust $M(\gamma)$ and $S(\gamma)$ to maintain a fixed BER at the receiver.

(solve for M (in terms of BER))

$$\Leftrightarrow \quad M(\gamma) = 1 + \frac{1.5\gamma}{-\ln(5\,BER)} \cdot \frac{S(\gamma)}{\bar{S}}$$

- The spectral efficiency for M-QAM is $\log_2 M$.

$$\max \underbrace{E[\log_2 M(\gamma)]}_{\text{avg spectral efficiency}} = \int_0^\infty \log_2\left(1 + \frac{1.5\gamma}{-\ln(5\,BER)} \cdot \frac{S(\gamma)}{\bar{S}}\right) p(\gamma)\, d$$

$$\text{s.t.} \quad \int_0^\infty S(\gamma)\, p(\gamma)\, d\gamma \le \bar{S}.$$

Solution: (necessary condition):

$$\frac{S(\gamma)}{\bar{S}} = \begin{cases} \frac{1}{\gamma_{\ast}} - \frac{1}{\gamma} & \gamma \geq \gamma_k = \boxed{\frac{\gamma_0}{K}} \\ 0 & \gamma < \frac{\gamma_0}{K} \; (\equiv \gamma_k) \end{cases}$$

where $K \equiv \frac{-1.5}{\ln(5\,BER)}$ and $\gamma_k \equiv \frac{\gamma_0}{K}$.

Then:

$$\max E[\log_2(M)] = \frac{\max R}{B} = \int_{\gamma_k}^{\infty} \log_2\left[\frac{\gamma}{\gamma_k}\right] p(\gamma)\, d\gamma .$$

$\frac{\gamma_0}{K}$

- This shows that uncoded M-QAM under adaptive modulation has an effective power loss of K relative to the optimal transmission scheme, independent of the fading distribution [i.e. the distribution of the $g(i)$ does not matter in computing the "gap" but of course $g(i)$ came in determining $p(\gamma)$.]

Recall:

$$\max \rho = \int_{\gamma_0}^{\infty} \log_2\left(\frac{\gamma}{\gamma_0}\right) p(\gamma)\, d\gamma$$ (from p. 6 of Lecture 8)

– We also see that K (in dB) is the maximum coding gain.

Channel inversion (using M-QAM).

$$\cancel{\frac{R}{B} = \log_2\left(1 + \frac{-1.5}{\ln(5BER)\, E[1/\gamma]}\right)} \quad \cancel{\text{for } M \geq 4,\ 0 \leq \gamma \leq 30\text{dB}}$$

$$M(\gamma) = 1 + \frac{1.5}{-\ln(5BER)} \cdot \boxed{\frac{\gamma\, S(\gamma)}{\bar{S}}} \rightarrow \sigma$$

σ = constant under channel inversion

$$\boxed{\frac{R}{B}} = \log_2 M(\gamma) = \log_2\left[1 + \frac{1.5}{-\ln(5BER)} \cdot \sigma\right]$$

$$= \boxed{\log_2\left[1 + \frac{-1.5}{\ln(5BER) \cdot E[1/\gamma]}\right]}$$

(since $\sigma = \frac{1}{E[1/\gamma]}$)

Truncated channel inversion (M-QAM)

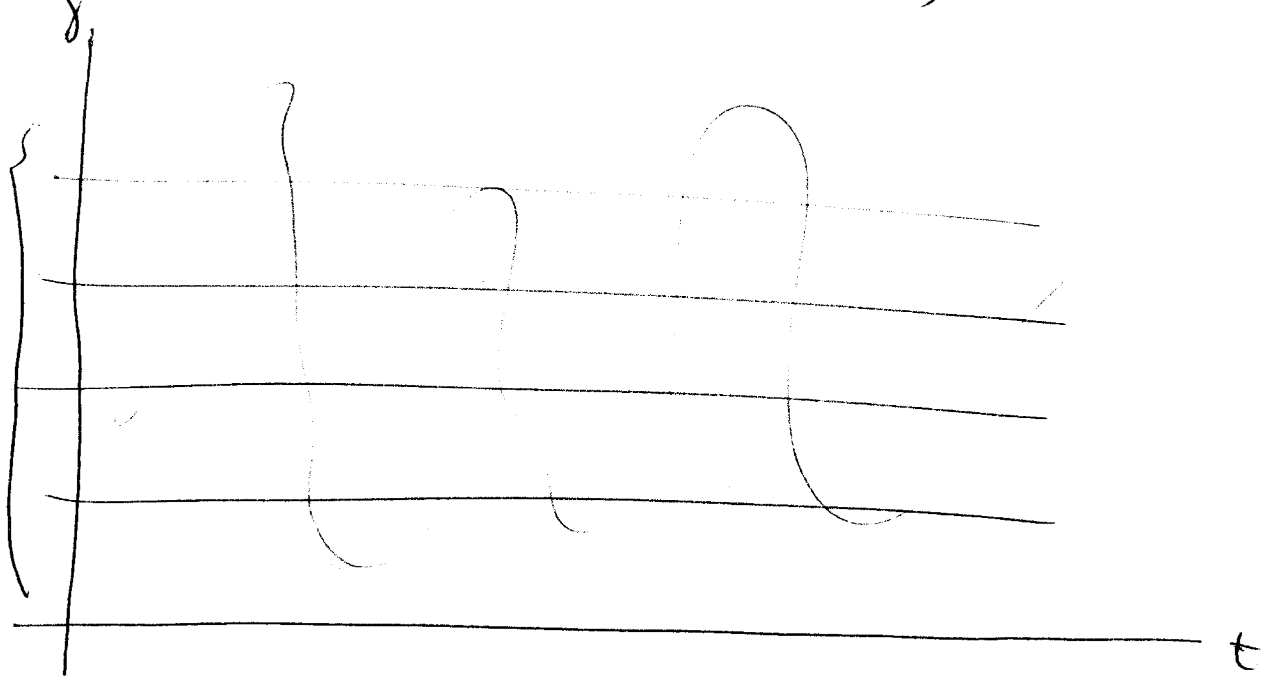
$$\frac{R}{B} = \max_{\gamma_0}\left[\log_2\left[1 + \frac{-1.5}{\ln(5BER) \cdot E_{\gamma_0}[1/\gamma]}\right] \cdot P_r(\gamma > \gamma_0)\right]$$

$$E_{\gamma_0}[1/\gamma] \equiv \int_{\gamma_0}^{\infty} \frac{1}{\gamma}\, p(\gamma)\, d\gamma$$

Constellation restriction

- Now, we restrict $M \in \{0, 2, 4, 16, 64, 256, \ldots\}$
(Square constellations)

discretize to N+1 "fading regions"



- Associate M_j with the jth fading region.

Region boundaries

$$M(\gamma) = \frac{\gamma}{\gamma_k^*} \qquad [\gamma = \gamma_k^* \cdot M_j]$$

where γ_k^* is a parameter to be optimized.

- Define $M_{N+1} = +\infty$. (we will never use this. The M_N constellation will be the largest one used)

For a fixed γ, transmit the largest constellation supportable (to maintain the BER) from the set $M_j \in \{0, 2, 4, 16 \ldots$

Ex) $2 \le \frac{\gamma}{\gamma_k^*} < 4 \Rightarrow$ use 2-QAM.

Region boundaries: $\boxed{\gamma = \gamma_k^* \cdot M_j}$

Power control:

To satisfy a target BER ($\forall M_j > 0$).

$$\frac{S_j(\gamma)}{\bar{S}} = \begin{cases} (M_j - 1)\frac{1}{\gamma K} & M_j \le \frac{\gamma}{\gamma_k^*} \le M_{j+1} \\ 0 & M_j = 0. \end{cases}$$

Fixed BER: $\forall M_j > 0$:

$$\Rightarrow \boxed{\frac{E_s(j)}{N_0} = \frac{\gamma S_j(\gamma)}{\bar{S}} = \frac{(M_j - 1)}{K}}.$$

Region (j)	γ	M_j	$\overbrace{S_j(\gamma)/\bar{S}}^{\left[E_s(j)/N_0\right]\cdot\frac{1}{\gamma}}$
0	$0 \le \frac{\gamma}{\gamma_k^*} < 2$	0	0
1	$2 \le \frac{\gamma}{\gamma_k^*} < 4$	2	$\frac{1}{K\cdot\gamma}$
2	$4 \le \frac{\gamma}{\gamma_k^*} < 16$	4	$\frac{3}{K\cdot\gamma}$
3	$16 \le \frac{\gamma}{\gamma_k^*} < 64$	16	$\frac{15}{K\cdot\gamma}$
4	$64 \le \frac{\gamma}{\gamma_k^*} < +\infty$	64	$\frac{63}{K\cdot\gamma}$

$$\frac{R}{B} = \sum_{j=1}^{N} \log_2(M_j)\cdot \Pr\left\{M_j \le \frac{\gamma}{\gamma_k^*} < M_{j+1}\right\}.$$

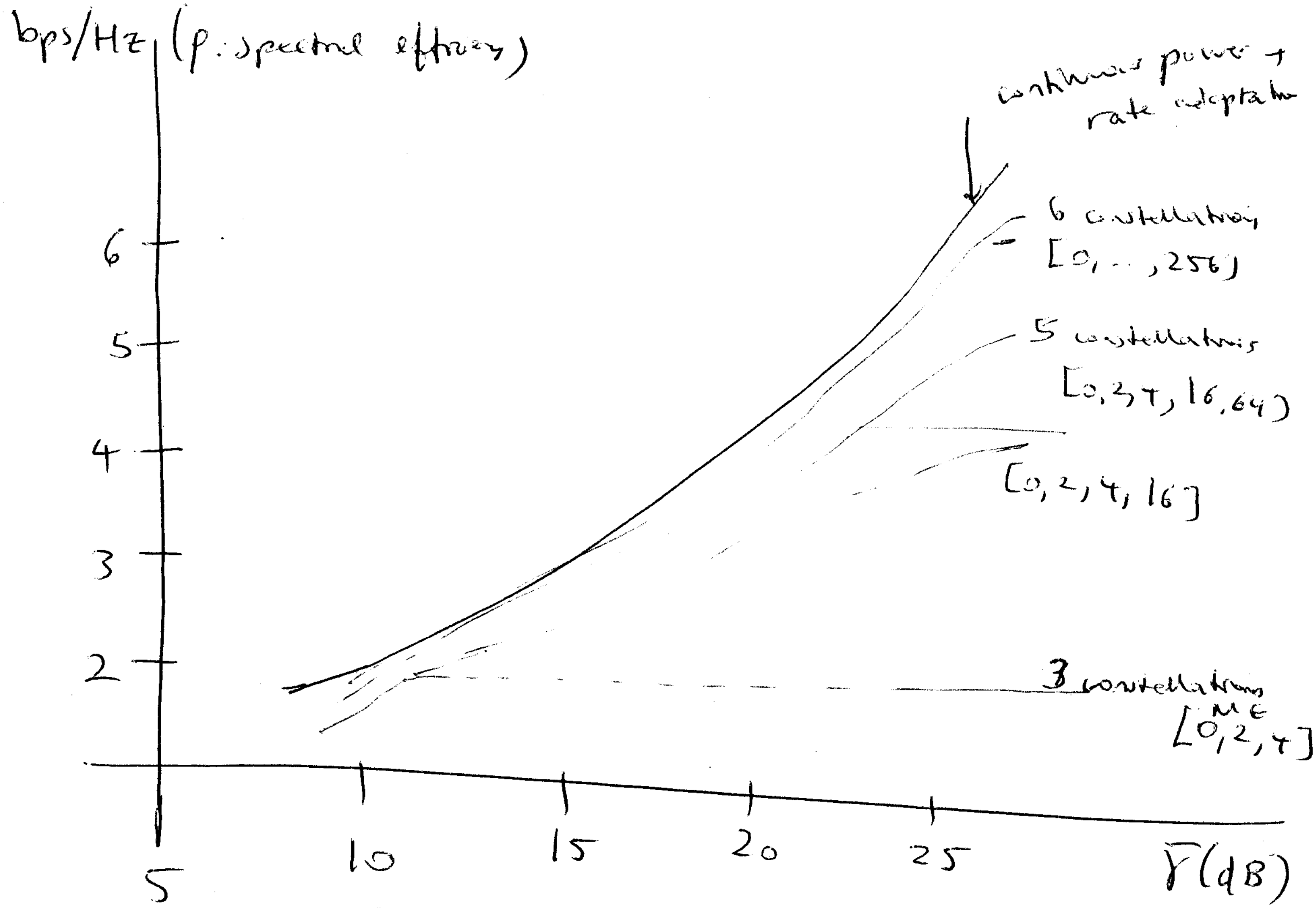
- Maximize $\frac{R}{B}$ w.r.t γ_k^*.

$$\max \sum_{j=1}^{N} \log_2(M_j)\cdot \Pr\left\{M_j \le \frac{\gamma}{\gamma_k^*} < M_{j+1}\right\}$$

$$\text{s.t.} \quad \sum_{j=1}^{N-1} \int_{\gamma_k^*\cdot M_j}^{\gamma_k^*\cdot M_{j+1}} \frac{S_j(\gamma)}{\bar{S}}\, p(\gamma)\, d\gamma = 1.$$

(the total power constraint)

no closed form solution: use numerical search.



(lognormal shadowing: $\sigma = 8$ dB)

– we see that 6 constellations $M \in (0, .., 256)$ gets you very close to what you would achieve by cont. power + rate adaptation.

DIVERSITY:

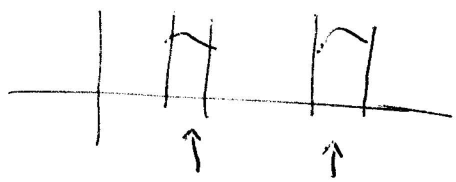
— refers to sending the same signal (info) over several paths (redundancy)

Types of Diversity:

- multiple receive antennas (Space diversity)
- time diversity: continue sending same signal over a period > correlation period of the shadowing ~~(...)~~
  (for mobile applications)

Same symbol then travels over ~~different~~ uncorrelated channels

- frequency diversity:



over far away bands
(~~x~~ → see uncorrelated channels)

- ~~directional~~
  polarization diversity: (send same info signal over vertical & horizontally polarized components)

- Antenna diversity — (MIMO systems)
  — or several receive antennas

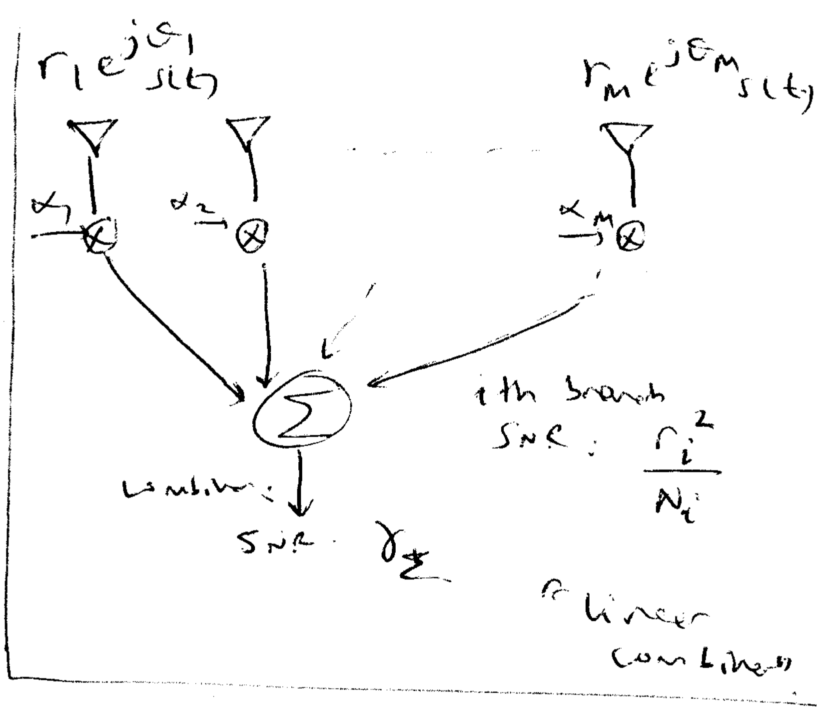
– Combining techniques:

A. Selection

~~B. Threshold~~

B. MRC

C. Equal-gain combining



A. Selection combining:

– Strategy: choose the branch with the highest $SNR = \frac{r_i^2}{N_i}$ (or practically highest S+N, if noise is same on all branches.)

Then: SNR of combiner $\gamma_\Sigma = \max_{1 \le i \le M} (\gamma_i)$

- Assume: M branches
  - uncorrelated Rayleigh fading amplitudes r_i

$$\gamma_i = \frac{r_i^2}{N}$$ (instantaneous SNR on branch i)

- Assume noise power same on all branches (N).

Let $\overline{\gamma_i} \equiv E[\gamma_i]$.

$$\cancel{p(\gamma_i) = \frac{1}{\overline{\gamma_i}} e^{-\gamma_i/\overline{\gamma}}}$$ ~~(pdf of SNR on ith branch)~~

Recall Rayleigh pdf:

$$p(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}, \quad r \geq 0$$

Make a change of variables:

$$p(\gamma_b) \cdot d\gamma_b = p(r)\, dr \qquad \left(p(\gamma_b) = p(r) \cdot \left| \frac{dr}{d\gamma_b} \right| . \right)$$

SNR per bit for a given amplitude r $\equiv \gamma_b = \dfrac{r^2}{2\sigma_n^2}$ ← mean.

σ_n^2: noise var in in-phase & quadrature. component.

Then:

$$p(\gamma_b) = \frac{\sigma_n^2}{\sigma^2} e^{-\gamma_b \cdot \sigma_n^2/\sigma^2} \qquad \left[\bar{\gamma}_b = \frac{\sigma^2}{\sigma_n^2} \right]$$

$$\boxed{p(\gamma_b) = \frac{1}{\bar{\gamma}_b} \cdot e^{-\gamma_b/\bar{\gamma}_b}, \quad \gamma_b \geq 0}$$

Hence, when the envelope r has a Rayleigh distribution, the SNR per bit $\gamma_b \equiv \dfrac{r^2}{2\sigma_n^2}$ is ~ exponential with parameter: $\dfrac{1}{\bar{\gamma}_b}$. (mean $\bar{\gamma}_b$).

Then ith branch

$$p(\gamma_i) = \frac{1}{\bar{\gamma}_i} e^{-\gamma_i/\bar{\gamma}_i} \quad , \gamma_i \geq 0$$

$$\boxed{P_{out}^{(i)}} \triangleq Pr\{\gamma_i < \gamma_0\}$$

$$= \int_0^{\gamma_0} \frac{1}{\bar{\gamma}_i} e^{-\gamma_i/\bar{\gamma}_i} d\gamma_i = \boxed{1 - e^{-\gamma_0/\bar{\gamma}_i}}$$

$$P_{out}(\text{combiner}) = P\{\gamma_\Sigma < \gamma_0\}$$

$$= P\{\max(\gamma_i) \leq \gamma_0\}$$

$$= P\{\gamma_1 \leq \gamma_0, \ldots, \gamma_M < \gamma_0\}$$

$$= \prod_{i=1}^{M} P\{\gamma_i < \gamma_0\}$$ (by independence of the branches)

$$= \prod_{i=1}^{M} \left(1 - e^{-\gamma_0/\bar{\gamma}_i}\right)$$

$$\underset{\substack{\uparrow \\ \text{if } \bar{\gamma}_i = \bar{\gamma} \ \forall i}}{=} \left(1 - e^{-\gamma_0/\bar{\gamma}}\right)^M$$

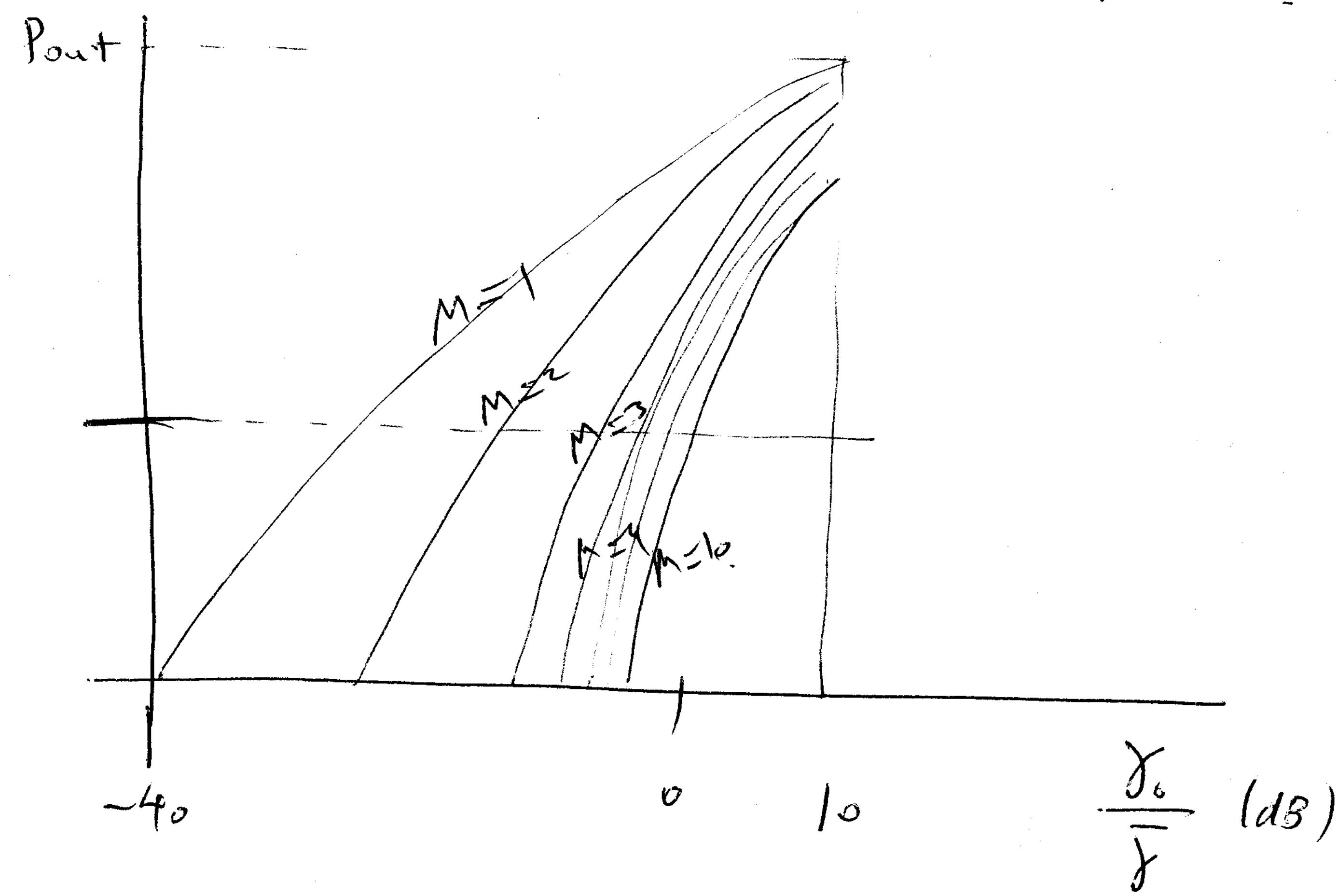
$$p_{\gamma_\Sigma}(\gamma) = \frac{d}{d\gamma} P\{\gamma_\Sigma < \gamma_0\}$$

$$= \frac{M}{\bar{\gamma}} \left(1 - e^{-\gamma/\bar{\gamma}}\right)^{M-1} e^{-\gamma/\bar{\gamma}}$$

$$\bar{\gamma}_\Sigma = \int_0^\infty \gamma \cdot p_{\gamma_\Sigma}(\gamma)\, d\gamma$$

$$\boxed{\bar{\gamma}_\Sigma = \bar{\gamma} \cdot \sum_{i=1}^{M} \frac{1}{i}}$$

– SNR increases with M, but with diminishing returns



- Note that $M \to \infty$ gives $\overline{\gamma}_\Sigma \to +\infty$
  (harmonic series diverges)

  but not physically possible, — the antennas in a given area will become correlated.

Calculation of $\overline{P_b}$:

$$\overline{P_b} = \int_0^\infty \underbrace{P_b(\gamma)}_{\text{the bit error rate of the particular modulation scheme (BPSK, etc.)}} \, P_{\gamma_\Sigma}(\gamma) \, d\gamma$$

after combining

(B) ~~Threshold combining.~~

~~Idea: scan each of the branches in sequential order & output first signal with a power above a threshold.~~

~~- As long as the signal remains above threshold, it outputs that signal otherwise — do another seq. search.~~

~~- If cannot find~~

(B) MRC:

Idea: weight each branch based on its SNR.

- Remove the phase then combine.

$$r = \sum_{i=1}^{M} a_i r_i.$$

(Assume $N_i = N$), at ith branch $\qquad N_{total} = \sum_{i=1}^{M} a_i^2 \cdot N$

$$\gamma_\Sigma = \frac{r^2}{N_{total}} = \frac{\left(\sum_{i=1}^{M} a_i r_i\right)^2}{\left(\sum_{i=1}^{M} a_i^2\right)\cdot N}$$

$\uparrow$ maximize this

– can be shown that optimal $a_i^* = \frac{r_i^2}{N}$.

$$\boxed{\gamma_\Sigma^* = \sum_{i=1}^{M} \frac{r_i^2}{N} = \sum_{i=1}^{M} \gamma_i}$$

$\uparrow$ sum of γ.

- SNR increases linearly with # of diversity branches.

Let $\bar{\gamma}_i = \bar{\gamma}\ \forall i$ $\qquad$ $(\gamma_i \sim \exp(1/\bar{\gamma}_i))$

$\gamma_\Sigma \sim \cancel{\chi^2}$ with ~~2M degree of freedom~~.

~~Chi-squared~~

Gamma $\left(\frac{1}{\bar{\gamma}}, M\right)$

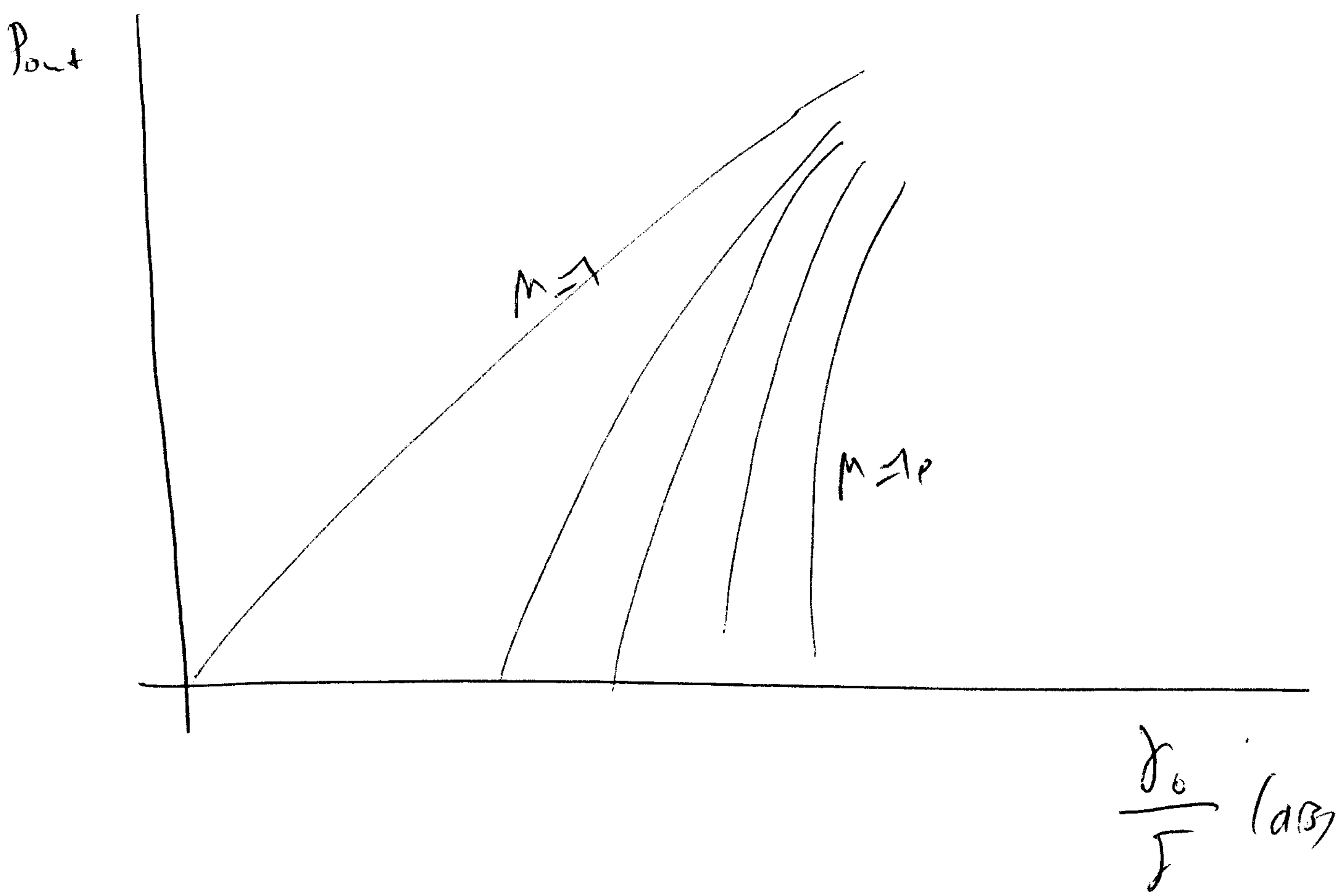
$$\boxed{\overline{\gamma_\Sigma} = M \cdot \overline{\gamma}}$$

$$Var(\gamma_\Sigma) = \frac{\overline{\gamma}}{2}$$

$$p_{\gamma_\Sigma}(\gamma) = \frac{\gamma^{M-1} e^{-\gamma/\overline{\gamma}}}{\overline{\gamma}^M (M-1)!} \quad , \gamma \geq 0$$

$$P_{out} = P_r\{\gamma_\Sigma < \gamma_0\} = \int_0^{\gamma_0} p_{\gamma_\Sigma}(\gamma)\, d\gamma$$

$$= 1 - e^{-\gamma/\overline{\gamma}} \sum_{k=1}^{M} \frac{(\gamma/\overline{\gamma})^{k-1}}{(k-1)!}$$



(C) Equal gain combining

<u>Idea</u>: MRC requires knowledge of the $\bar{\gamma}$.

— EGC: sets $a_i = 1$

$$\gamma_\Sigma = \frac{\left(\sum r_i\right)^2}{N_{total}} = \frac{1}{NM}\left(\sum_{i=1}^{M} r_i\right)^2$$

$N_{total} = N \cdot M$

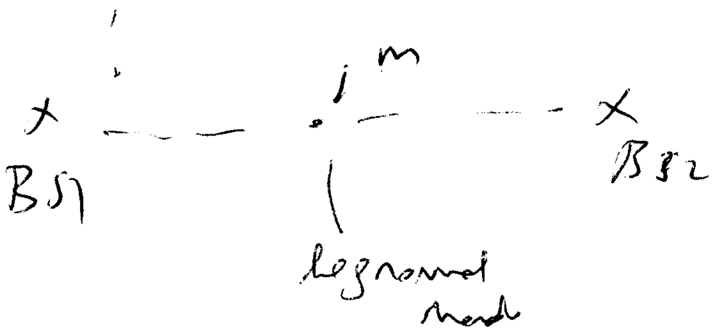
~~Ex~~: <u>M = 2 case</u>

$$P_{out} = \Pr\{\gamma_\Sigma < \gamma_0\}$$

$$= 1 - e^{-2(\gamma_0/\bar\gamma)} - \sqrt{\pi\left(\frac{\gamma_0}{\bar\gamma}\right)}\, e^{-(\gamma_0/\bar\gamma)}\, \mathrm{erf}\left(\sqrt{\gamma_0/\bar\gamma}\right)$$

Performance loss $\simeq$ <u>1.3 dB</u> (small)
with r. to MRC

~~Ex~~ <u>Macroscopic diversity</u> ("BS diversity")



selection diversity
(select BS with best signal or
(~~[illegible]~~))

Fano's Inequality:

For a DMC with codebook $\mathcal{C}$ and W (input messages) uniformly distributed,

$$H(X^n|Y^n) \le 1 + P_e^{(n)} \cdot nR$$

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

$\boxed{\text{Pf}}$ Define $E = \begin{cases} 1 & \text{if } \hat{W} \neq W \\ 0 & \text{otherwise} \end{cases}$ (an error event)

Then,

$$H(E, W|Y^n) = H(W|Y^n) + H(E|W,Y^n)$$ (chain rule; conditioning)

$$= H(E|Y^n) + H(W|E,Y^n)$$ [condition the other way]

- Since E = function of $\{W, g(Y^n)\}$, $H(E|W,Y^n) = 0$.
- Since E is binary-valued, $H(E) \le 1$
- Finally,

$\Leftarrow \hat{W} = W \Rightarrow W = \text{function of } (Y^n)$

$$H(W|E,Y^n) = \overbrace{\Pr(E=0)}^{(1-P_e^{(n)})}\ \overbrace{H(W|E=0,Y^n)}^{0} + \overbrace{P(E=1)}^{P_e^{(n)}} \cdot \underbrace{H(W|Y^n, E=1)}_{\le \log(|W|-1)}$$

nR

$$\le (1-P_e^{(n)}) \cdot 0 + P_e^{(n)} \cdot nR$$

$$= P_e^{(n)} \cdot nR.$$

since $W|Y^n$ can take on at most nR values

Then:

$$\underbrace{H(W|Y^n)}_{} + \overset{0}{H(E|W,Y^n)} = \underbrace{H(E|Y^n)}_{\leq 1} + \underbrace{H(W|E,Y^n)}_{\leq P_e^{(n)} \cdot nR}$$

(note that g is known here g(Y^n) known)

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

– For a fixed code $X^n(W)$, $X^n(W)$ is a function of W.

$$\Rightarrow \boxed{H(X^n(W)|Y^n) \leq H(W|Y^n) \leq 1 + P_e^{(n)} \cdot nR}$$

i.e. $\boxed{H(X^n|Y^n) \leq 1 + P_e^{(n)} \cdot nR}$