COM 5120
Communications Theory

Chapter 6
Information Theory

Prof. Jen-Ming Wu

Inst. of Communications Engineering

Dept. of Electrical Engineering

National Tsing Hua University

Email:jmwu@ee.nthu.edu.tw

Fall, 2021

Outline

- Mathematical Models for Information Source
- Measure of Information
- Channel Models and Channel Capacity
- Achieving Channel Capacity w/ Orthogonal Signals and Channel Reliability

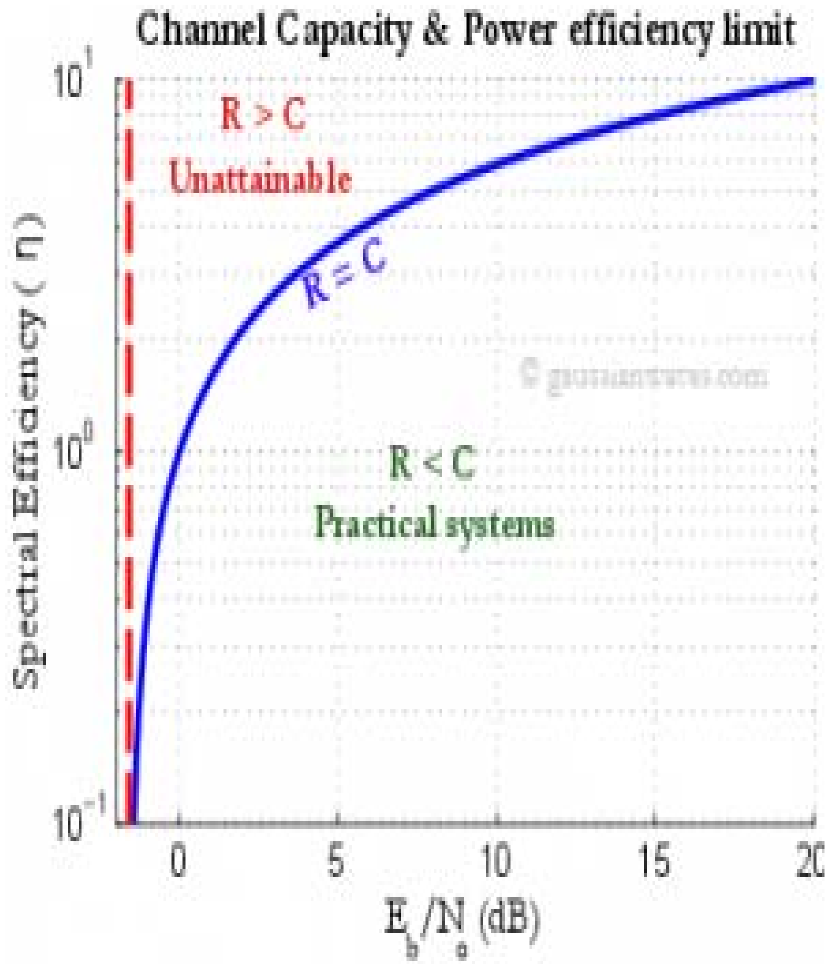
Why Information Theory?

Information theory deals with mathematical modeling and analysis of a communication system. It tries to answer the following questions:

- What is the irreducible complexity that below which a signal source can not be further compressed?

- What is the ultimate transmission rate for reliable communication over a noisy channel?

$$R\uparrow \quad P_e\uparrow \quad \Rightarrow unreliable$$



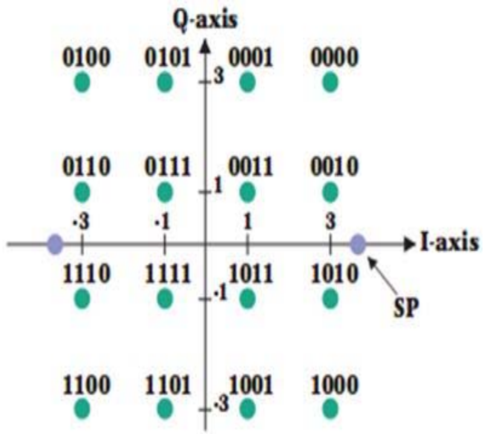
6.1 Mathematical Models for Information Source

- Discrete Memoryless Source (DMS)

  Assume a discrete information source $\{x_1, x_2, \ldots, x_K\}$ each source has a given probability of p_k, $1 \le k \le K$

  where $\sum_{k=1}^{K} p_k = 1$

$\Rightarrow$ A discrete source with statistically independent output sequence is called discrete memoryless source (DMS)



6.2 Measure of Information

- How to measure information?

Given a DMS $X \in \{x_1, x_2, \ldots, x_K\}$,the amount of information for $X = x$ is inverse proportional to its probability $\mathrm{P}(x)$, defined as:

$$I(x) = \log_2 \frac{1}{\mathrm{P}(x)} = -\log_2 \mathrm{P}(x) \quad \text{(bits)}$$

$$= \ln \frac{1}{\mathrm{P}(x)} \quad \text{(nats)}$$

$P(x)\uparrow \quad I(x)\downarrow$

Properties of Information

For the DMS $\{x_1,\ldots,x_K\}$,

(1) Zero informatin event: $I(x_k) = 0$ for $p_k = 1$

$\Rightarrow$ Absolute certainty of the outcome about an event

$\Rightarrow$ No information gains for the message

(2) Non-negativity: $I(x_k) \geq 0$ $\quad \because\ 0 \leq p_k \leq 1$

Given message $X = x_k$, always produce some info or no info.

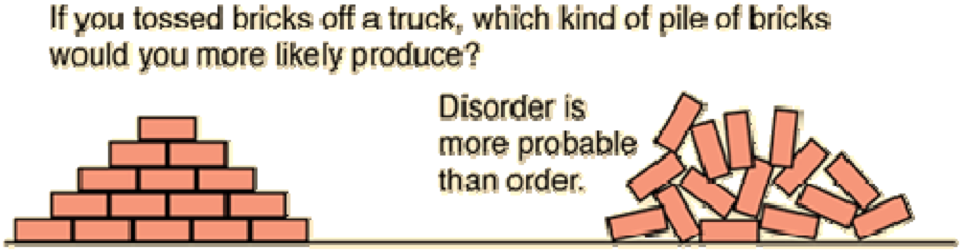
Never bring about a loss of info. (non-negative)

(3) $I(x_k) > I(x_j)$ for $p_k < p_j$

Entropy of Information

- Definition: Entropy represents the mean value of information per source symbol

$$H(X) = \mathrm{E}[I(x_k)] = \sum_{k=1}^{K} p_k I(x_k)$$



$\Rightarrow$ Entropy is a measure of uncertainty about X

➢ Entropy is used to describe the degree of randomness in a system.

- Properties

$$(1) H(X) = 0 \text{ iff } \begin{cases} p_k = 1 & \text{for } k = k^* \\ p_k = 0 & \text{other} \end{cases} \quad \Rightarrow \text{no uncertainty}$$

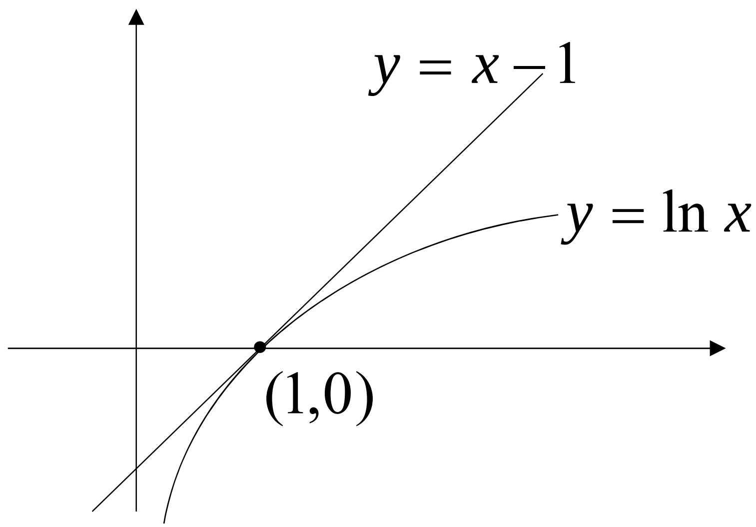
$$(2) H(X) = \log_2 K \text{ iff } p_k = \frac{1}{K} \text{ for all } K$$

$\Rightarrow$ uniform prob. distribution leads to max uncertainty

(3)The entropy is bounded: $0 \le H(X) \le \log_2 K$

Proof :

$$
\begin{aligned}
& H(X) - \log_2 K \\
&= \sum_{k=1}^{K} p_k \log_2 \frac{1}{p_k} - \log_2 K \\
&= \sum_{k=1}^{K} p_k \log_2 \frac{1}{p_k} - \sum_{k=1}^{K} p_k \log_2 K \\
&= \sum_{k=1}^{K} p_k \log_2 \frac{1}{K p_k} = \frac{1}{\ln 2} \sum_{k=1}^{K} p_k \ln \frac{1}{K p_k}
\end{aligned}
$$



By the inequality $\ln x \le x - 1$

$$
\Rightarrow H(X) - \log_2 K \le \frac{1}{\ln 2} \sum_{k=1}^{K} p_k \left(\frac{1}{K p_k} - 1 \right)
$$

$$
= \frac{1}{\ln 2} \sum_{k=1}^{K} \left(\frac{1}{K} - p_k \right) = 0 \quad \therefore H(X) \le \log_2 K
$$

- Joint Entropy:

  The entropy of a pair of random variables X, Y

$$H(X,Y) = -\sum_x \sum_y \mathrm{P}(x,y) \log_2 \mathrm{P}(x,y)$$

  i.e. the mean value of joint information

- Conditional Entropy:

  Given $X = x$, the entropy of Y is

$$H(Y \mid x) = -\sum_y \mathrm{P}(y \mid x) \log_2 \mathrm{P}(y \mid x)$$

  i.e. the mean value of conditional information

- The average conditional entropy over all possible values of X is

$$H(Y \mid X) = \sum_x \mathrm{P}(x) H(Y \mid x) = -\sum_x \sum_y \mathrm{P}(x,y) \log_2 \mathrm{P}(y \mid x)$$

$$Lemma\ 1: H(X,Y) = H(X) + H(Y \mid X)$$

Proof: $$H(Y|X) = -\sum_x \sum_y \mathrm{P}(x,y)\log_2 \mathrm{P}(y \mid x)$$

$\because \mathrm{P}(y \mid x) = \mathrm{P}(x,y)/\mathrm{P}(x)$

$$= -\sum_x \sum_y \mathrm{P}(x,y)\log_2 \mathrm{P}(x,y) + \sum_x \sum_y \mathrm{P}(x,y)\log_2 \mathrm{P}(x)$$

$$= H(X,Y) - H(X)$$

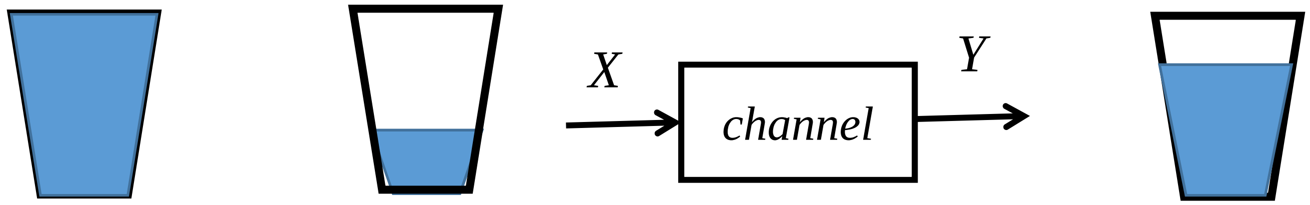
$$\therefore H(X,Y) = H(X) + H(Y|X)$$

- The **conditional entropy** $H(Y|X)$ measures how much entropy a random variable *Y* has remaining if we have already learned the value of another random variable *X*.

$$Collary: H(X,Y) \geq \max\left\{H(X),\quad H(Y)\right\}$$

Mutual Information

The information provided by the occurrence of $Y = y$ about the event $X = x$, is defined by



$\log_2 \frac{1}{P(x)}$ $\log_2 \frac{1}{P(x|y)}$ $I(x,y)$

$$I(x,y) = \log_2 \frac{1}{\mathrm{P}(x)} - \log_2 \frac{1}{\mathrm{P}(x \mid y)} = \log_2 \frac{\mathrm{P}(x \mid y)}{\mathrm{P}(x)} = \log_2 \frac{\mathrm{P}(x, y)}{\mathrm{P}(x)\mathrm{P}(y)}$$

called mutual information between *x* and *y*.

Given Y = y, $-\log_2 \frac{1}{\mathrm{P}(x \mid y)}$ represents the amount of information if X = x is further offered.

Q: If X and Y are independent random variables, $\rightarrow I(x,y) = ?$

Mutual Information

- The mutual information measures how much one random variable tells us about another.

- X, Y are independent:

$$P(x \mid y) = \frac{P(x)P(y)}{P(y)} = P(x) \ \Rightarrow I(x, y) = 0$$

- X, Y are fully dependent:

$$P(x \mid y) = 1, \ \ \Rightarrow I(x, y) = I(x)$$

- The mutual information about random variables X, Y is

$$\begin{aligned} I(X, Y) &= \sum_x \sum_y P(x, y) I(x, y) \\ &= \sum_x \sum_y P(x, y) \log_2 \frac{P(x \mid y)}{P(x)} = \sum_x \sum_y P(x, y) \log_2 \frac{P(y \mid x)}{P(y)} \end{aligned}$$

Mutual Information

- Properties of Mutual Information

(1). $I(x,y) = I(y,x)$ $\quad \because I(x,y) = \log_2 \dfrac{\mathrm{P}(x,y)}{\mathrm{P}(x)\mathrm{P}(y)}$

(2). $I(x,y) \geq 0$, "=" holds when X,Y are independent

(3). $I(X,Y) \leq \min\{K_x, K_y\}$

where $K_x = size(X), K_y = size(Y)$

Mutual Information and Entropy

$$Lemma\ \ 2: I(X,Y) = H(X) - H(X \mid Y) =$$ Reduction of uncertainty about X after observing Y

Uncertainty of source (average information of X)

Uncertainty of X given Observation of Y

= Amount of information about X in Y

Proof:

$$I(X,Y) = \sum_x \sum_y \mathrm{P}(x,y) \log_2 \frac{\mathrm{P}(x \mid y)}{\mathrm{P}(x)}$$

$$= \underbrace{\sum_x \sum_y \mathrm{P}(x,y) \log_2 \mathrm{P}(x \mid y)}_{-H(X \mid Y)} \underbrace{- \sum_x \sum_y \mathrm{P}(x,y) \log_2 \mathrm{P}(x)}_{H(X)}$$

Mutual Information and Entropy

$$Lemma\ 3: \quad I(X,Y) \le \min\{H(X), H(Y)\}$$

$$\Pr oof: \quad \because I(X,Y) = H(X) - H(X \mid Y) \quad \text{(Lemma 2)}$$

$$\because \quad I(X,Y) \le H(X)$$

$$I(X,Y) \le H(Y)$$

$$\rightarrow \quad I(X,Y) \le \min\{H(X), H(Y)\}$$

Furthermore,

$$H(X) \le \log K_x$$

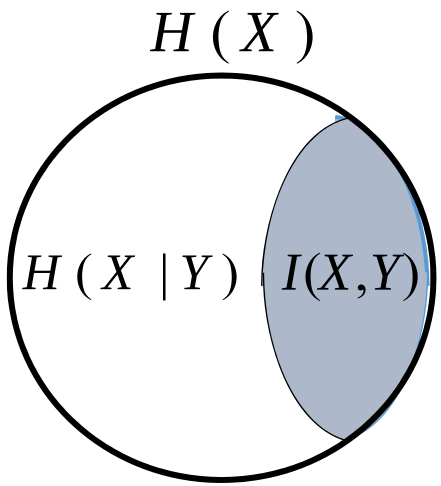
$$H(Y) \le \log K_y$$

$$\therefore I(X,Y) \le \min\{\log K_x, \log K_y\}$$

$$\rightarrow I(X,Y) \le \min\{K_x, K_y\}$$

Remarks

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

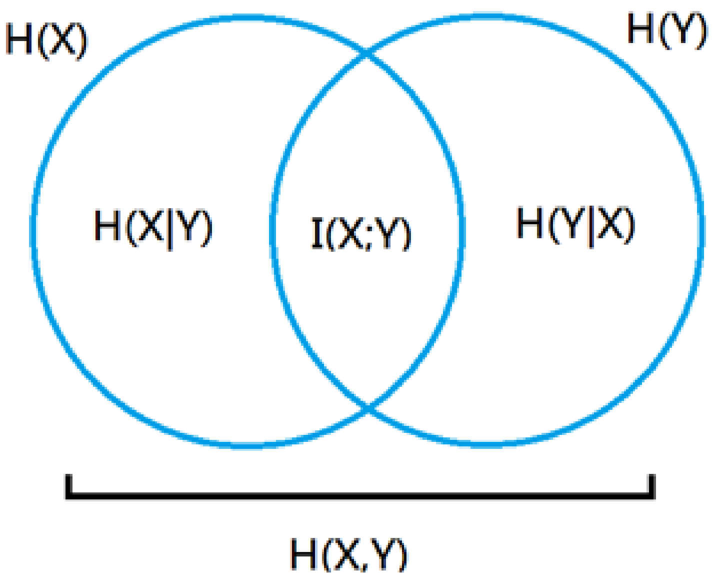


- To decode the transmitted message correctly, the conditional entropy $H(X|Y)$ has to be minimized.
- Consequently, the mutual information is maximized if $H(X|Y)=0$, then $I(X,Y)=H(X)$.

  $\Rightarrow$ The mutual information provides all the entropy about X.

Summary- Venn Diagram Representation

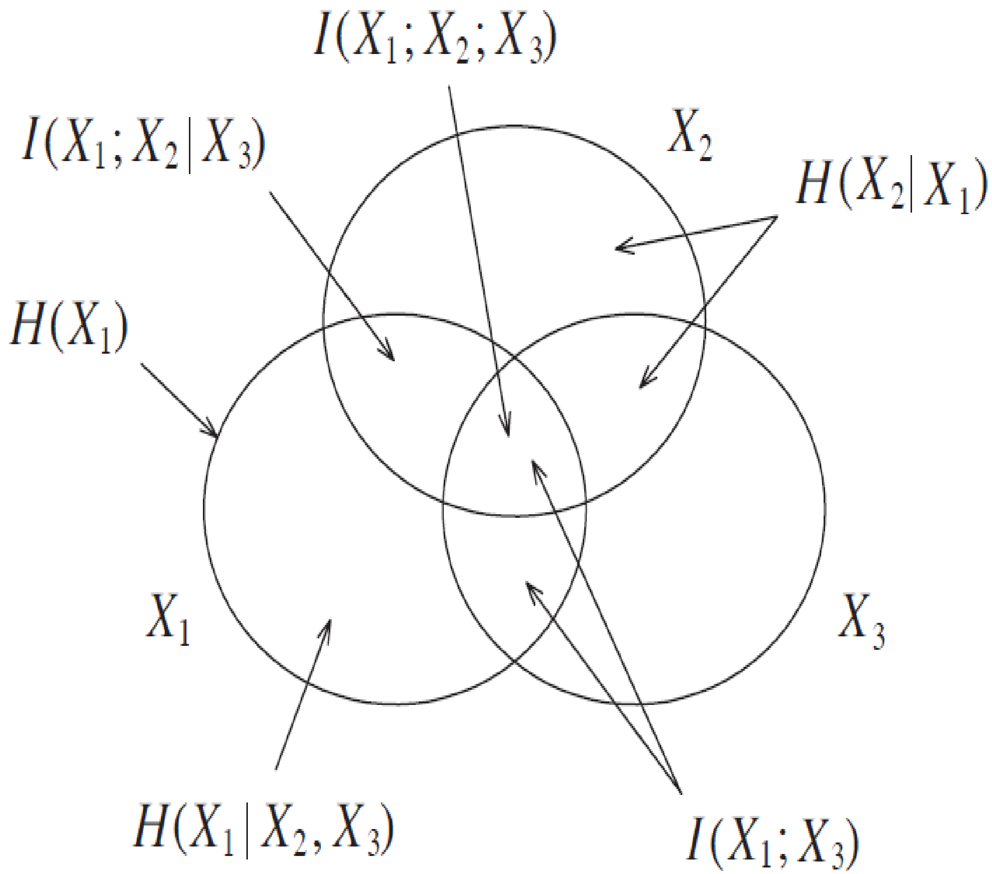
- Given r.v.'s X and Y



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

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

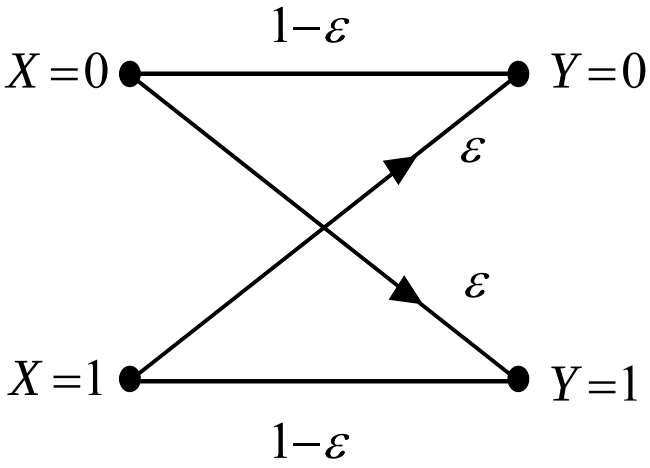
- Given r.v.'s X_1 , X_2 and X_3



Outline

- Mathematical Models for Information Source
- Measure of Information
- **Channel Models and Channel Capacity**
- Achieving Channel Capacity w/ Orthogonal Signals and Channel Reliability

Binary Symmetric Channel (BSC)



$$X \in \{0,1\}$$

$$\mathrm{P}(Y=0 \mid X=1) = \mathrm{P}(Y=1 \mid X=0) = \varepsilon$$

$$\mathrm{P}(Y=1 \mid X=1) = \mathrm{P}(Y=0 \mid X=0) = 1-\varepsilon$$

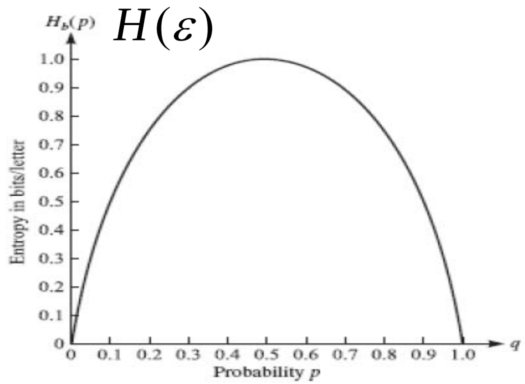
(1) $H(X)$ is maximized when the source prior probability

$$\mathrm{P}(x) = \tfrac{1}{2} \quad \forall x = 0,1 \quad \Rightarrow H(X) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1$$

(2)
$$H(X|Y) = -\sum_x \sum_y \mathrm{P}(x,y)\log_2 \mathrm{P}(x \mid y)$$

$$= -p(0,0)\log_2 p(0|0) - p(1,1)\log_2 p(1|1)$$
$$- p(1,0)\log_2 p(1|0) - p(0,1)\log_2 p(0|1)$$

$$= -(1-\varepsilon)\log_2(1-\varepsilon) - \varepsilon\log_2\varepsilon \;= H(\varepsilon)$$



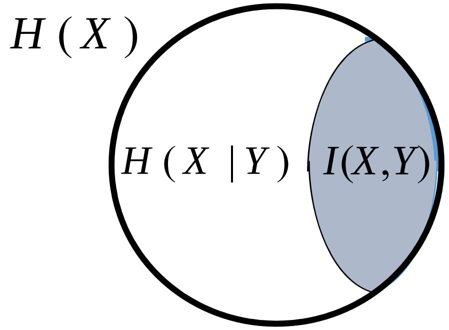
(3) $I(X,Y) = H(X) - H(X \mid Y)$

Capacity of Binary Symmetric Channel (BSC)

- Channel Capacity is defined as the max mutual information being transmitted over the channel.

$$C = \max\{I(X,Y)\}$$

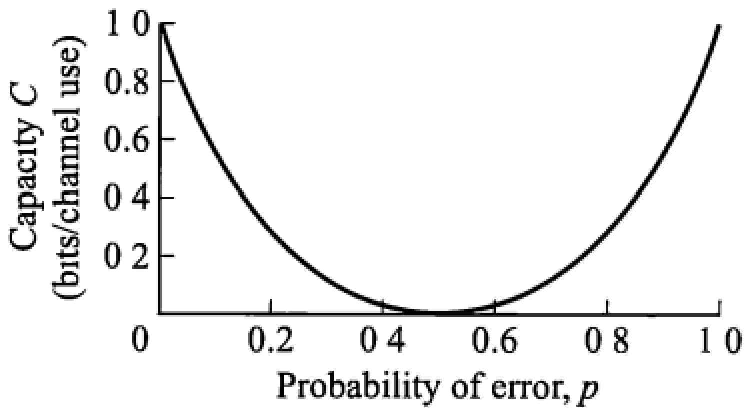
$$i.e. C = \max\{H(X) - H(X \mid Y)\}$$



From (1) and (2)

$$C = 1 + \varepsilon \log \varepsilon + (1-\varepsilon)\log(1-\varepsilon)$$

$$= 1 - H(\varepsilon) \qquad 0 \le \varepsilon \le 1$$



C is maximized when

$$\varepsilon = 0 \text{ or } 1$$

C is minimized when

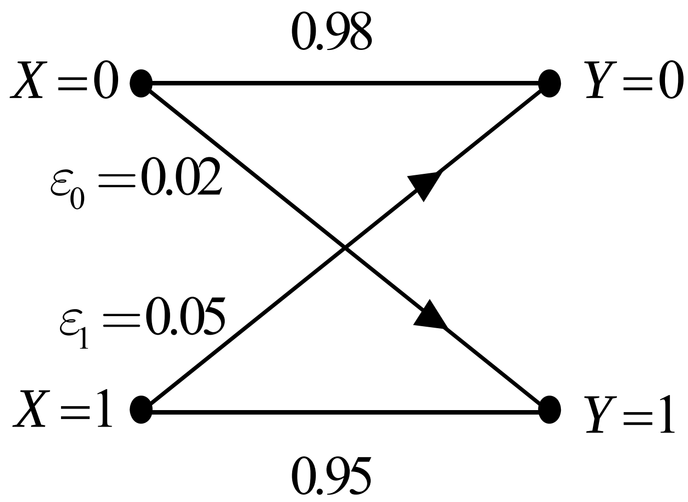
$$\varepsilon = \frac{1}{2}$$

Mutual information of Binary Channel

Example: Let *X* be a binary source which has equal probable symbol {0, 1}. Let *Y* be a binary output {0, 1}. The channel has transition probability matrix

$$P_{ch} = \begin{bmatrix} 0.98 & 0.02 \\ 0.05 & 0.95 \end{bmatrix}$$

Calculate the mutual information of this channel (known *H*(*Y*) = 0.9994 and *H*(*X*) =1.



Solution:

$$H(X) = -\sum_x \mathrm{P}(x)\log_2 \mathrm{P}(x) = 1$$

$$H(X|Y) = -\sum_x \sum_y \mathrm{P}(x,y)\log_2 \mathrm{P}(x\,|\,y)$$

$$= 0.2146$$

$$\rightarrow \; I(X,Y) = H(X) - H(X\,|\,Y) = 0.7854$$

Entropy for Discrete or Continuous Random Variables

- **Entropy**

  Given discrete random variable *X*,

$$H(X) = \sum_{k=1}^{K} p_k \log_2 \frac{1}{p_k}.$$

- **Differential Entropy**

  Given continuous random variable *X*,

$$H(X) = \int_{-\infty}^{\infty} f(x) \log_2 \frac{1}{f(x)} dx,$$

Capacity of DMC over AWGN

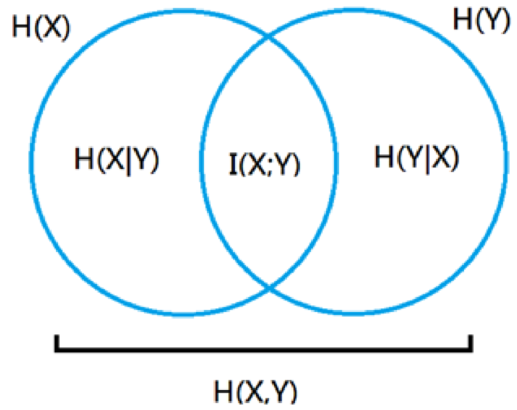
- Consider transmission of random variable $X_k, k = 1, \ldots, K$ over AWGN channel with zero mean and power spectral density $(\mathrm{PSD}) = \dfrac{N_0}{2}$. The received r.v. is $Y_k = X_k + N_k$,

- The transmitted power is limited to $E[X_k^2] = P$, and the noise power is $\frac{N_0}{2} \times 2W = N_0 W$ with transmission bandwidth 2W.

Capacity of DMC over AWGN

$$C = \max \ I(X_k, Y_k)$$
$$= \max\{ \boxed{H(X_k)} - H(X_k \mid Y_k)\}$$

Unknown for receiver



By the duality, $I(X_k, Y_k) = I(Y_k, X_k)$

$$\therefore C = \max\{H(Y_k) - H(Y_k \mid X_k)\}$$

For Rx, only $H(Y_k)$ is available, not $H(X_k)$.

Also, $Y_k = X_k + N_k$

$$\Rightarrow H(Y_k \mid X_k) = H(X_k + N_k \mid X_k)$$
$$= H(X_k \mid X_k) + H(N_k \mid X_k) = H(N_k)$$

$$\Rightarrow C = I(X_k, Y_k) = H(Y_k) - H(N_k)$$

Capacity of DMC over AWGN

$$H(Y_k) = E[I(Y_k)] = -\int_{-\infty}^{\infty} f_{Y_k}(y_k) \log_2 f_{Y_k}(y_k) dy_k$$

$$\text{where } f_{Y_k}(y_k) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(\frac{-y_k^2}{2\sigma_y^2}\right)$$

$$\sigma_Y^2 = \mathrm{E}[X_k^2] + \sigma_N^2 = P + \frac{N_0}{2} \quad \text{Why?} \quad \because Y_k = X_k + N_k, \therefore \sigma_Y^2 = \sigma_X^2 + \sigma_N^2$$

$$\Rightarrow H(Y_k) = -\int_{-\infty}^{\infty} f_{Y_k}(y_k)[-\ln(\sqrt{2\pi\sigma_y^2}) - \frac{y_k^2}{2\sigma_y^2}](\log_2 e) dy_k$$

$$\text{Note } \int_{-\infty}^{\infty} f_{Y_k}(y_k) dy_k = 1$$

$$\int_{-\infty}^{\infty} y_k^2 f_{Y_k}(y_k) dy_k = \sigma_Y^2 = P + \frac{N_0}{2}$$

Capacity of DMC over AWGN

$$\Rightarrow H(Y_k) = -\int_{-\infty}^{\infty} f_{Y_k}(y_k)[-\ln(\sqrt{2\pi\sigma_y^2}) - \frac{y_k^2}{2\sigma_y^2}](\log_2 e)dy_k$$

$$= (\log_2 e)\left\{\ln(\sqrt{2\pi\sigma_y^2})\int_{-\infty}^{\infty} f_{Y_k}(y_k)dy_k + \frac{1}{2\sigma_y^2}\int_{-\infty}^{\infty} y_k^2 f_{Y_k}(y_k)dy_k\right\}$$

$$= \log_2 e[\ln(\sqrt{2\pi\sigma_y^2}) + \frac{1}{2}] = \log_2 e[\frac{1}{2}\ln(2\pi\sigma_y^2) + \frac{1}{2}\ln e]$$

$$= \log_2 e[\frac{1}{2}\ln(2\pi e\sigma_y^2)] = \frac{1}{2}\log_2(2\pi e\sigma_y^2)$$

Similarly,

$$H(N_k) = -\int_{-\infty}^{\infty} f_{N_k}(n_k)\log_2 f_{N_k}(n_k)dn_k = \frac{1}{2}\log_2(2\pi e\sigma_n^2)$$

$$C = H(Y_k) - H(N_k)$$

$$= \frac{1}{2}\log_2 2\pi e(P + \sigma_n^2) - \frac{1}{2}\log_2 2\pi e\sigma_n^2$$

$$= \frac{1}{2}\log_2(1 + \frac{P}{\sigma_n^2}) \qquad \Rightarrow C = \frac{1}{2}\log_2(1 + SNR) \quad \square$$

Geometric Interpretation of Channel Capacity

- The capacity of discrete-time AWGN channel with input power constraint:

Let $Y_i = X_i + N_i, \quad i = 1, \ldots, n,$ where X_i is the transmission symbol at basis $\phi_i(t)$.

$N_i \sim \mathcal{N}(0, \sigma_n^2)$ and the input power of X_i is constrained, $\mathrm{E}[X_i^2] \le P$

- Transmission of $\mathbf{x}$ with n-dimension: $\mathbf{x} = [X_1, \ldots, X_n]^T$

The received signal vector is $\mathbf{y} = \mathbf{x} + \mathbf{n} = [Y_1, \ldots, Y_n]^T$

The average received power on each dimension:

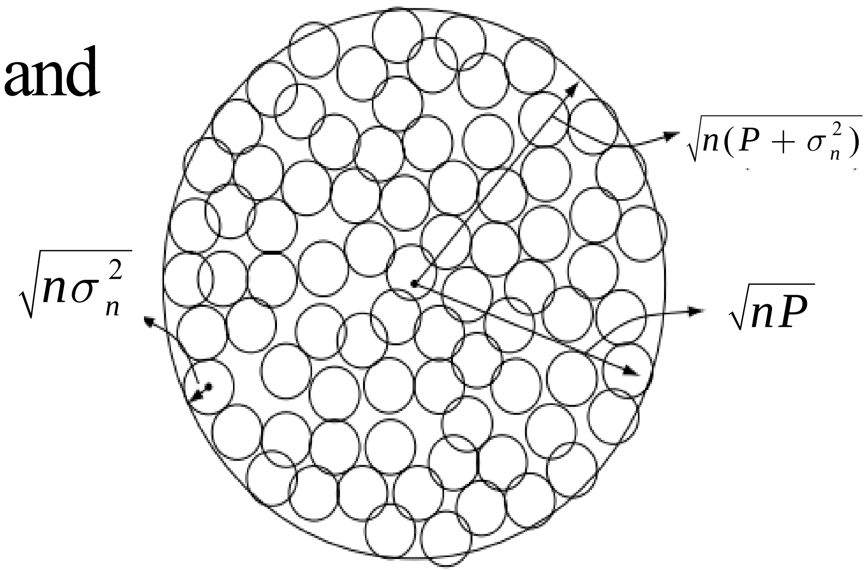
$$\frac{1}{n}\|\mathbf{y}\|^2 = \mathrm{E}[X_i^2] + \mathrm{E}[N_i^2] \le P + \sigma_n^2$$

Geometric Interpretation of Channel Capacity

- $\mathbf{y}$ is inside the n-dimensional sphere of radius $\sqrt{n(P+\sigma_n^2)}$

- If $\mathbf{x}$ is transmitted, $\mathbf{y}$ will be in an n-dimension sphere of radius $\sqrt{n\sigma_n^2}$ and centered at $\mathbf{x}$ with high probability.

- The max number of spheres with radius $\sqrt{n\sigma_n^2}$ that can be packed in a sphere of radius $\sqrt{n(P+\sigma_n^2)}$ is the ratio of the volumes of the sphere.



Geometric Interpretation of Channel Capacity

The volume of a *n*-dimension sphere with radius γ

$$V_n = B_n \gamma^n \propto \gamma^n$$

$$Ex: \; n = 2, V_2 = \pi\gamma^2 \quad n = 3, V_3 = \frac{4}{3}\pi\gamma^3$$

⇒ The maximum number of different messages (or symbols) that can be resolvable at the receiver is

$$M = \frac{B_n \left(\sqrt{n(P+\sigma_n^2)}\right)^n}{B_n \left(\sqrt{n\sigma_n^2}\right)^n} = (1 + \frac{P}{\sigma_n^2})^{\frac{n}{2}}$$

- The resulting transmission bit rate at each dimension:

$$R = \frac{1}{n}(\log_2 M) = \frac{1}{2}\log_2(1 + \frac{P}{\sigma_n^2}) \;\; \text{bits/transmission (symbol)}$$

The Capacity of Band-limited AWGN Channel (Continuous Time)

Given channel bandwidth $2W$, input power constraint P and noise power spectral density $\frac{N_0}{2}$

- The capacity of discrete-time channel in bits/transmission

$$C = \frac{1}{2}\log_2(1+\frac{P}{N_0 W})$$

bits/transmission
(or bits/symbol)

- The capacity of discrete-time channel in bits/sec

$$C = 2W\,\frac{1}{2}\log_2(1+\frac{P}{N_0 W}) = W\,\log_2(1+\frac{P}{N_0 W}) \quad \text{bits/sec}$$

How to Improve the Channel Capacity?

- Primary Communication Resources
  - ✓Transmitted Power (Signal-to-Noise Ratio, SNR)
  - ✓Channel Bandwidth (sampling rate, and noise power)

$$C = W \log_2\left(1 + \frac{P}{N_0 W}\right)$$

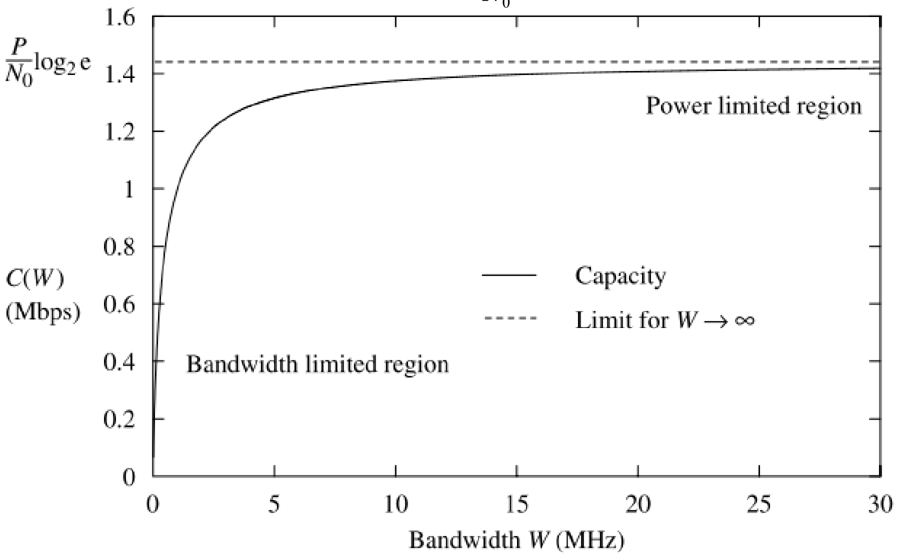
$$\ln(1+x) \cong x, \quad x << 1$$

$$\log_2(1+x) \cong (\log_2 e)x$$

(1) Fixed *P* and increase *W*

$$\text{As } W \uparrow \quad \Rightarrow C_\infty = \lim_{W \to \infty} W \log_2\left(1 + \frac{P}{N_0 W}\right) = \frac{P}{N_0}\log_2 e$$

$$\text{Ex}: \frac{P}{N_0} = 10^6$$



- ➢ With infinite bandwidth, the channel capacity can not increase definitely!

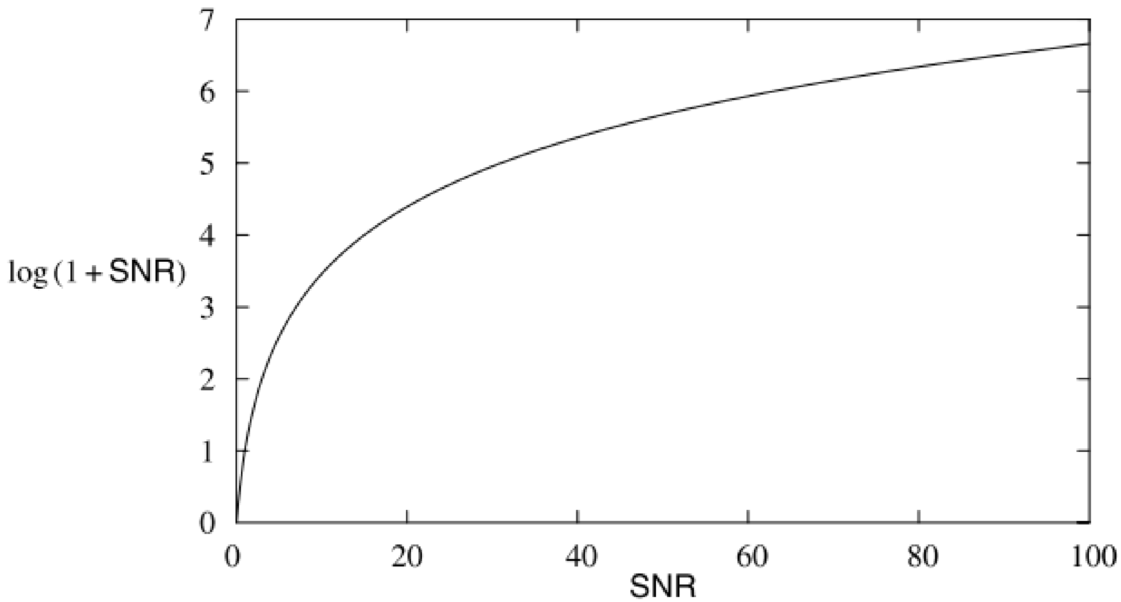
How to Improve the Channel Capacity?

$$C = W \log_2 (1 + \frac{P}{N_0 W})$$

(2) Increase P with fixed W.

$$\text{As } P \uparrow \ \Rightarrow \frac{P}{N_0 W} >> 1 \ \Rightarrow C \cong W \log_2 \frac{P}{N_0 W}$$

- C increases at a logarithmic rate

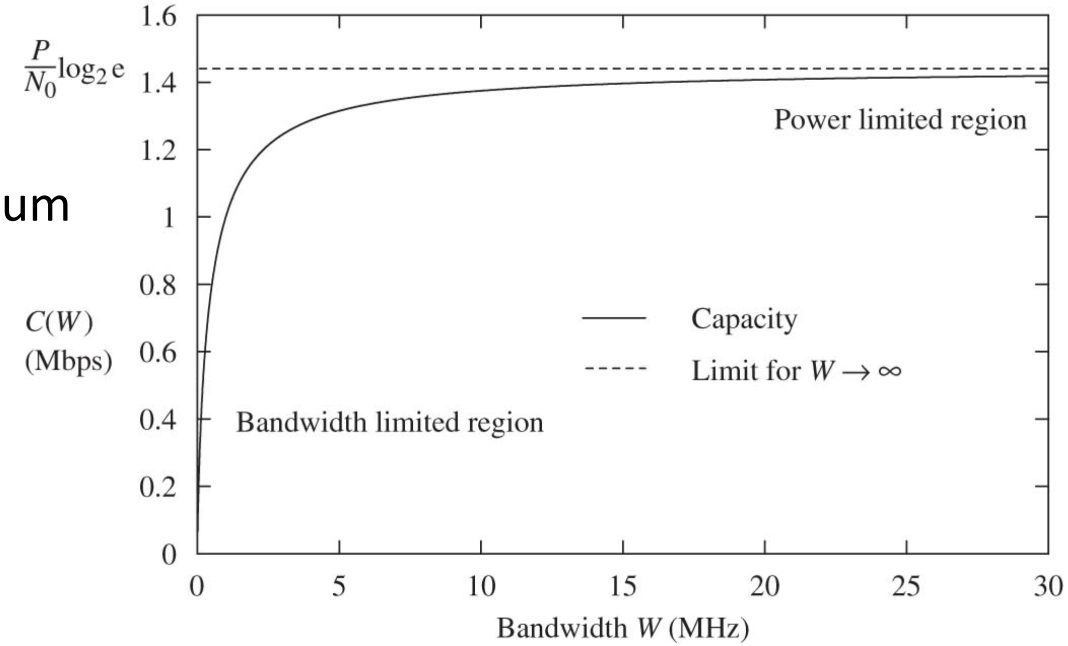


- Power inefficient at high SNR region !

How to Improve the Channel Capacity?

- Types of Communication Channels
  - ✓ Power Limited
    - Noise
    - Finite Power
    - Propagation Loss
  - ✓ Band Limited
    - Freq response of medium
    - Multiple users share the medium

$$C = W \log_2 \left(1 + \frac{\overline{P}}{N_0 W}\right)$$



Fundamental Limit and Relation of Power Efficiency and Bandwidth Efficiency

(1) Power efficiency

Let the power efficiency be defined as $\gamma = \dfrac{P}{C}$ (J/bit)

$$\min \frac{P}{C} = \lim_{W\to\infty} \frac{P}{C} = \lim_{W\to\infty} \frac{P}{W\log_2(1+\frac{P}{N_0W})}$$

$$= \frac{P}{\frac{P}{N_0}\log_2 e} = N_0 \ln 2 \quad \to (\mathrm{A})$$

$$E_b = \frac{E_s}{\log_2 M} = \frac{P T_s}{\log_2 M} = \frac{P}{R} \qquad \Rightarrow R = \frac{P}{E_b} \quad \to (\mathrm{B})$$

$$\frac{T_s}{\log_2 M} = T_b, \qquad R \leftarrow \text{bit rate}$$

For reliable communications $\Rightarrow R \le C$

$$\Rightarrow \frac{P}{E_b} \le C \quad \Rightarrow \frac{P}{C} \le E_b \overset{(A)(B)}{\Rightarrow} N_0 \ln 2 \le E_b \Rightarrow \underbrace{\frac{E_b}{N_0}}_{\text{SNR per bit}} \ge \ln 2 \ (\cong 0.693 = \text{-}1.6\text{dB})$$

What's the minimum energy to transmit one bit at room temperature?

- Shannon-von Neumann-Landauer Limit

$$N_0 \ln 2 \le E_b$$

The noise power $N_0 = kT$ where
k = Boltzmann constant = $1.38 \mathrm{x} 10^{-23}$ (J/K)

- At T = 300K (= 27°C), the min energy required to transmit <u>one bit</u> is

$$E_{b,\min} = kT \ln 2 = 1.38 \times 10^{-23} \times 300 \times \ln 2$$
$$= 2.9 \times 10^{-21} \text{(J/bit)}$$

Communication Energy Efficiency

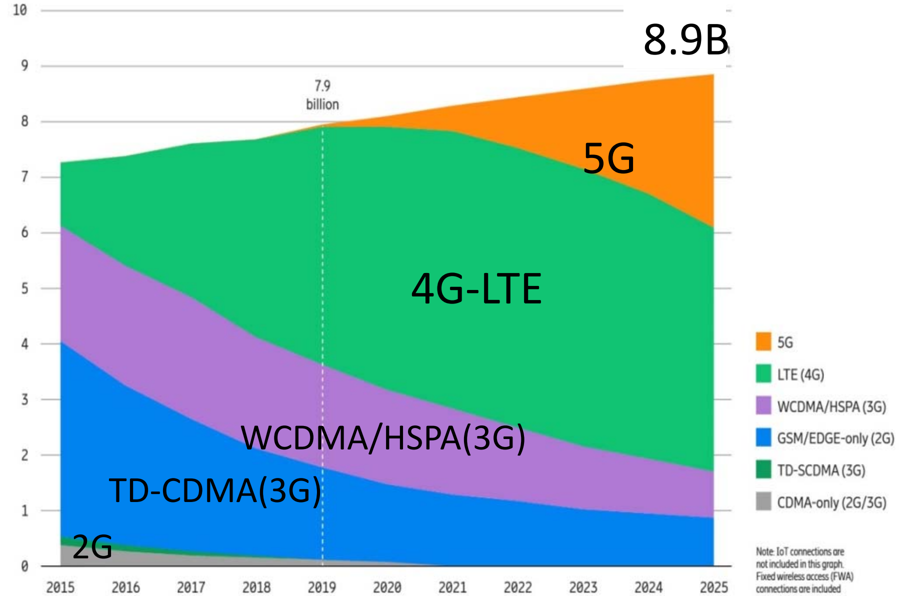
■ Wireless communication performance under energy constraints

- Shannon-Von Neumann-Landauer Bound:

  Minimum energy/bit = $kTln2$ = 2.9×10^{-21} J/bit at 27°C

- Complexity-energy-performance trade-off

	Rate (Mb/s)	P_Rx (mW)	Rx (nJ/bit)	P_Tx (mW)	Tx (nJ/bit)
802.11g	22	140	6.4	450	20.4
802.11n	200	1000	5	1800	9.0
BT 2.0	0.7	45	64.3	62	88.6
BT EDR	2.2	48	21.8	65	29.5



➢ Worldwide mobile subscribers
- Over 8.9-Billions (by 2025)
- Increasing 2M/day

■ Energy efficiency under energy constraints

- Portable device/handset operations are always limited by the battery.
- Every mW saving in wireless transceiver represents MW's of saving for greener communications.

Fundamental Limit and Relation of Power Efficiency and Bandwidth Efficiency

(2) Bandwidth efficiency ($\eta = C/W$)

$$\frac{R}{W} \le \frac{\mathrm{C}}{W} \Rightarrow \log_2\left(1 + \frac{E_b / T_b}{N_0 W}\right) \le \log_2\left(1 + \frac{E_b}{N_0} \cdot \frac{C}{W}\right) \quad ①$$

$$\text{Define } \cancel{C_\infty} = \lim_{W \to \infty} C = \lim_{W \to \infty} W \log_2\left(1 + \frac{P}{N_0 W}\right) \qquad \because P = E_b R \le E_b C$$

$$\le \lim_{W \to \infty} W \log_2\left(1 + \frac{E_b C_\infty}{N_0 W}\right) \cong (\log_2 e) \frac{\cancel{E_b C_\infty}}{N_0} \quad \text{for } R \le C$$

$$(\text{Note that, for } x << 1, \quad \ln(1 + x) \cong x\ ) \Rightarrow \frac{E_b}{N_0} \ge \ln 2 \ = \text{-1.6dB}$$

$$\text{As } W \to \infty \ \Rightarrow C_\infty \to \frac{P}{N_0} \log_2 e \quad \Rightarrow \eta = \frac{C_\infty}{W} \to 0$$

Also from ① $\eta = \dfrac{C}{\mathrm{W}} = \log_2\left(1 + \dfrac{E_b C}{N_0 W}\right) \quad \rightarrow \quad \dfrac{E_b}{N_0} = \dfrac{2^{C/W} - 1}{C/W}$

$$C = W \log_2\left(1 + \frac{P}{N_0 W}\right)$$

$$\eta = \frac{C}{\mathrm{W}} = \log_2\left(1 + \frac{P}{N_0 W}\right)$$

$$\leq \log_2\left(1 + \frac{\mathrm{E}_b C}{N_0 W}\right)$$

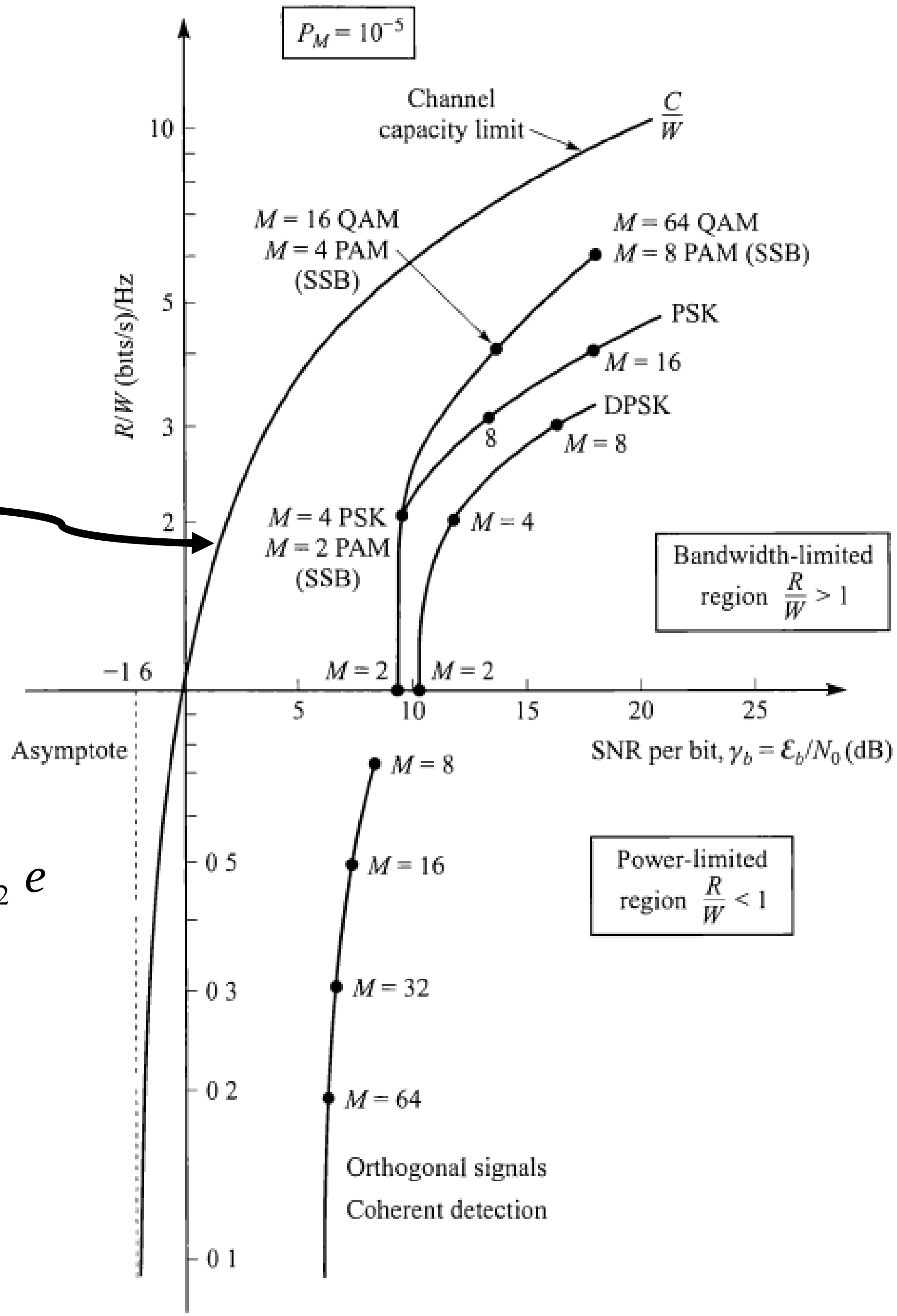
$$\frac{E_b}{N_0} = \frac{2^{C/W} - 1}{C/W}$$

$$\text{As } W \to \infty \quad \Rightarrow C_\infty \to \frac{P}{N_0} \log_2 e$$

$$\Rightarrow \eta \to 0$$

$$\frac{E_b}{N_0} \to -1.6dB$$

$$\Rightarrow \frac{E_b}{N_0} \geq \ln 2 \ (\cong 0.693 = \text{-1.6dB})$$



Outline

- Mathematical Models for Information Source
- Measure of Information
- Channel Models and Channel Capacity
- **Achieving Channel Capacity w/ Orthogonal Signals and Channel Reliability**

Achieving Channel Capacity w/ Orthogonal Signals

- From Ch4, for M-ary orthogonal signals, the error performance is

$$P_c = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (1-Q(x))^{M-1} e^{-\frac{1}{2}\left(x-\sqrt{\frac{2E_s}{N_0}}\right)^2} \, \mathrm{d}\, x$$

$$P_e = 1 - P_c = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [1-(1-Q(x))^{M-1}] e^{-\frac{1}{2}\left(x-\sqrt{\frac{2E_s}{N_0}}\right)^2} \, \mathrm{d}\, x$$

- With the inequality $(1\text{-}x)^n \geq 1\text{-}\, nx$, *for* $0 \leq x \leq 1$

$$1-[1-Q(x)]^{M-1} \leq (M-1)Q(x) < MQ(x) < Me^{-\frac{x^2}{2}}$$

- When x is small, *i.e.* $x < x_0$ for some small x_0, the above union bound is loose. Use the tighter bound $1-[1-Q(x)]^{M-1} \leq 1$

➤ Let the SNR be $\gamma = \frac{E_s}{N_0}$

$$P_e \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_0} e^{-\frac{1}{2}(x-\sqrt{2\gamma})^2} dx + \frac{M}{\sqrt{2\pi}} \int_{x_0}^{\infty} e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}(x-\sqrt{2\gamma})^2} dx \equiv P_e(x_0)$$

- Note that the upper bound of $P_e(x_0)$ is minimized when $\frac{\partial P_e(x_0)}{\partial x_0} = 0$

$$\frac{\partial P_e(x_0)}{\partial x_0} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\sqrt{2\gamma})^2} - \frac{M}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}(x-\sqrt{2\gamma})^2} = 0$$

$$\Rightarrow e^{\frac{x_0^2}{2}} = M, \quad \text{i.e.} \quad x_0 = \sqrt{2\ln M} = \sqrt{2\ln 2\log_2 M} = \sqrt{2k\ln 2}$$

- Finding the upper bound of $P_e(x_0)$. The first term in $P_e(x_0)$:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_0} e^{-\frac{1}{2}(x-\sqrt{2\gamma})^2} dx \overset{Let\ u=\left(x-\sqrt{2\gamma}\right)/\sqrt{2}}{=} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{-\left(\sqrt{2\gamma}-x_0\right)/\sqrt{2}} e^{-u^2} du,$$

$$= Q\left(\sqrt{2\gamma} - x_0\right) < e^{-\frac{1}{2}(\sqrt{2\gamma}-x_0)^2}, \quad x_0 \le \sqrt{2\gamma}$$

- The 2nd term in $P_e(x_0)$:

$$\frac{M}{\sqrt{2\pi}} \int_{x_0}^{\infty} e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}(x-\sqrt{2\gamma})^2} dx \overset{\text{Let } u=(x-\sqrt{\gamma/2})}{=} \frac{M}{\sqrt{2\pi}} e^{-\frac{\gamma}{2}} \int_{x_0-\sqrt{\gamma/2}}^{\infty} e^{-u^2} du$$

$$= \begin{cases} Me^{-\frac{\gamma}{2}}(1-Q(x_0-\sqrt{\gamma/2})), & x_0 \le \sqrt{\gamma/2} \\ Me^{-\frac{\gamma}{2}}Q(x_0-\sqrt{\gamma/2}), & x_0 > \sqrt{\gamma/2} \end{cases} < \begin{cases} Me^{-\frac{\gamma}{2}}, & x_0 \le \sqrt{\gamma/2} \\ Me^{-\frac{\gamma}{2}}e^{-(x_0-\sqrt{\gamma/2})^2}, & x_0 > \sqrt{\gamma/2} \end{cases}$$

- Combining the two terms in $P_e(x_0)$ and $M = e^{x_0^2/2}$,

$$P_e < \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_0} e^{-\frac{1}{2}(x-\sqrt{2\gamma})^2} dx + \frac{M}{\sqrt{2\pi}} \int_{x_0}^{\infty} e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}(x-\sqrt{2\gamma})^2} dx$$

$$< \begin{cases} e^{-\frac{1}{2}(\sqrt{2\gamma}-x_0)^2} + e^{\frac{x_0^2}{2}} e^{-\frac{\gamma}{2}}, & 0 \le x_0 \le \sqrt{\gamma/2} \\ e^{-\frac{1}{2}(\sqrt{2\gamma}-x_0)^2} + e^{\frac{x_0^2-\gamma}{2}} e^{-(x_0-\sqrt{\gamma/2})^2}, & \sqrt{\gamma/2} \le x_0 \le \sqrt{2\gamma} \end{cases}$$

$$= \begin{cases} e^{\frac{x_0^2-\gamma}{2}} (1 + e^{-(x_0-\sqrt{\gamma/2})^2}), & 0 \le x_0 \le \sqrt{\gamma/2} \\ 2e^{-\frac{1}{2}(\sqrt{2\gamma}-x_0)^2}, & \sqrt{\gamma/2} \le x_0 \le \sqrt{2\gamma} \end{cases}$$

$$< \begin{cases} 2e^{\frac{x_0^2-\gamma}{2}}, & 0 \le x_0 \le \sqrt{\gamma/2} \\ 2e^{-\frac{1}{2}(\sqrt{2\gamma}-x_0)^2}, & \sqrt{\gamma/2} \le x_0 \le \sqrt{2\gamma} \end{cases} \qquad \text{(A)}$$

Reliable Communication with $\frac{E_b}{N_0} \geq \ln 2$

- By reliable communication, we mean that $P_e \to 0$ is possible.

$$P_e < \begin{cases} 2e^{-\frac{x_0^2-\gamma}{2}}, & 0 \leq x_0 \leq \sqrt{\gamma/2} \\ 2e^{-\frac{1}{2}(\sqrt{2\gamma}-x_0)^2}, & \sqrt{\gamma/2} \leq x_0 \leq \sqrt{2\gamma} \end{cases} \qquad \text{(A)}$$

- The min P_e occurs at $x_0 = \sqrt{2\ln M} = \sqrt{2k\ln 2}$. Also let

$$\gamma = k\gamma_b = k\frac{E_b}{N_0}$$

$$\Rightarrow P_e < \begin{cases} 2e^{-k\left(\gamma_b - 2\ln 2\right)/2}, & \ln M \leq \gamma/4 \\ 2e^{-k\left(\sqrt{\gamma_b}-\sqrt{\ln 2}\right)^2}, & \gamma/4 \leq \ln M \leq \gamma \end{cases}$$

- As $k \to \infty, P_e \to 0$ is possible, if $\gamma_b > \ln 2$.

Reliable Communication with $R < C_\infty$

- Since $C_\infty = \lim_{W\to\infty} W \log_2\left(1+\frac{P}{N_0 W}\right) = \log_2 e \lim_{W\to\infty} W \ln\left(1+\frac{P}{N_0 W}\right) = (\log_2 e)\left(\frac{P}{N_0}\right)$,

$$x_0 = \sqrt{2k\ln 2} = \sqrt{2RT\ln 2}, \qquad \gamma = \frac{E_s}{N_0} = \frac{TP}{N_0} = TC_\infty \ln 2$$

Consequently, from (A), $P_e < \begin{cases} 2e^{\frac{x_0^2-\gamma}{2}}, & 0 \le x_0 \le \sqrt{\gamma/2} \\ 2e^{-\frac{1}{2}(\sqrt{2\gamma}-x_0)^2}, & \sqrt{\gamma/2} \le x_0 \le \sqrt{2\gamma} \end{cases}$ (A)

$$\Rightarrow P_e < \begin{cases} 2\times 2^{-T\left(\frac{1}{2}C_\infty - R\right)}, & 0 \le R \le \frac{1}{4}C_\infty \\ 2\times 2^{-T\left(\sqrt{C_\infty}-\sqrt{R}\right)^2}, & \frac{1}{4}C_\infty \le R \le C_\infty \end{cases}$$

$\rightarrow$ Show how the gap of C_∞ and R affects P_e.

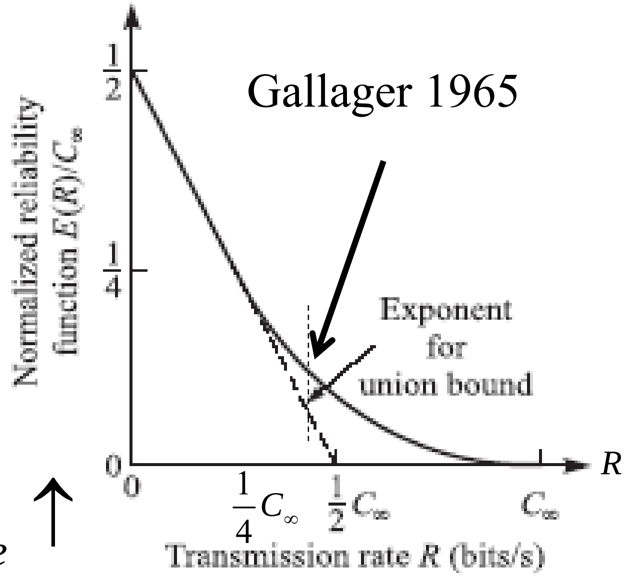
$\Rightarrow$ As $T(=kT_b) \to \infty, P_e \to 0$ is possible, if

$$R < C_\infty = (\log_2 e)\left(\frac{P}{N_0}\right) \qquad (i.e. \frac{E_b}{N_0} \ge \ln 2)$$

The Channel Reliability Function

- Channel reliability function is defined as the exponential factor (Gallager 1965)

$$E(R)=\begin{cases}\frac{1}{2}C_\infty - R, & 0\le R\le \frac{1}{4}C_\infty \\ \left(\sqrt{C_\infty}-\sqrt{R}\right)^2, & \frac{1}{4}C_\infty \le R\le C_\infty\end{cases}$$



$$\Rightarrow P_e < 2\times 2^{-TE(R)}$$ As $R\uparrow, E(R)\downarrow \quad P_e\uparrow$

- Recall union bound on P_{e} of orthogonaling signaling from Ch4,

$$P_e < e^{-\frac{k}{2}\left(\frac{E_b}{N_0}-2\ln 2\right)} \quad \Rightarrow P_e < \frac{1}{2}\times 2^{-T\left(\frac{1}{2}C_\infty - R\right)}, \quad 0\le R\le \frac{1}{2}C_\infty$$

$\Rightarrow$ The exponent $\frac{E_b}{N_0}-2\ln 2 \ge 0$ is not as tight as $E(R), i.e. \frac{E_b}{N_0}\ge \ln 2,$

due to the looseness of the union bound.

Summary

- Measure of information, entropy, and mutual information
- Channel capacity and mutual information
  - ✓ Binary Symmetry Channel
  - ✓ AWGN Channel
- Band-width efficiency and power efficiency
  - ✓ Minimum energy per bit
- Reliable communications and channel reliability function

HW #4

Due: TBD

Midterm II

Time: TBD
Place: Delta 215/217
Coverage: Ch4 and Ch6
Note: Closed book. Open 1 sheet of A4 size note. Calculator allowed.