電子情報通信特論
Advanced Course of Electronics, Information and Communications

Capacity of Wireless Channels

担当:金 ミンソク 准教授
mskim@eng.niigata-u.ac.jp

自然科学研究科　電気情報工学専攻

NIIGATA UNIVERSITY

Outline

- Review of Fundamental Information Theory
- Capacity in AWGN
- Capacity of flat fading channel
- Capacity of frequency selective fading channel

  (To be skipped)

Channel Capacity

- How to calculate how many data we can transmit over the channel ??
- Wireless channel is
  - Time invariant / varying
  - Frequency flat / selective
- Claude Shannon established theory
  - Channel capacity is mutual information maximized over all possible input distribution
  - Shannon’s coding theory

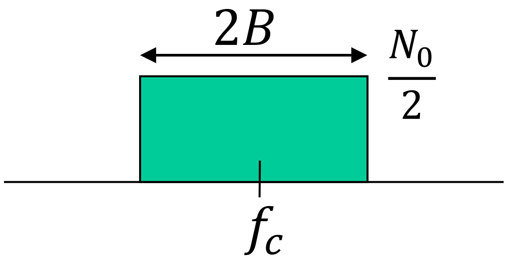
Capacity in AWGN Channel

- Discrete time AWGN (additive white Gaussian noise) channel

$$y[n] = x[n]+w[n]$$

- Channel capacity

$$C = B\log_2(1+\gamma)\ \text{[bps]}$$



$\gamma = \dfrac{P}{N_0 B}$ → SNR (signal to noise ratio)

$P = E\{|x[n]|^2\}$

$\frac{N_0}{2}$: Power spectral density of noise

B: Baseband signal bandwidth

Information and Entropy

- Self information on a discrete system

$$I(x) = -\log_2 p(x)\ [\text{bits}]$$

- Random variable with possible outcomes (alphabet) of

$$x_i \in X$$

- Entropy (Average information) is given by

$$H(X) = \sum_{i=1}^{N} p(x_i) I(x_i) = -\sum_{i=1}^{N} p(x_i) \log p(x_i)$$

Mean of the information over all source alphabets of X

Discrete Memoryless Channel

- Statistical model with an input X and an output Y which is noisy version of X

$$X \longrightarrow \boxed{p(y|x)} \longrightarrow Y$$

$p(x)$

Input or Priori prob.

Transition probability through the channel

$$p(y) = \sum_{x \in X} p(y|x)p(x)$$

Output or Marginal prob.

- Alphabet X and Y are finite set → discrete
- The current output symbol depends only on the current input symbol → memoryless

Mutual Information

- Conditional self information on erroneous channel

$$I\left(x_i|y_j\right) = -\log_2 p(x_i|y_j) \text{ [bits]}$$

- Mutual information

$$I\left(x_i; y_j\right) = I(x_i) - I\left(x_i|y_j\right) \text{ [bits]}$$

  - Error free channel $\left(p\left(x_i|y_j\right) = 1\right)$
    $$I\left(x_i; y_j\right) = I(x_i) \text{ [bits]}$$
  - Independent between x and y $\left(p\left(x_i|y_j\right) = p(x_i)\right)$
    $$I\left(x_i; y_j\right) = 0 \text{ [bits]}$$

Mutual Information (2)

- Entropy

$$H(X|Y) = \sum_{x\in X, y\in Y} p(x,y) I(x|y) = -\sum_{x\in X, y\in Y} p(x,y) \log p(x|y)$$

$$H(X) = -\sum_{x\in X} p(x) \log p(x)$$

- Mutual Information

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= -\sum_{x\in X} p(x) \log p(x) - \left(-\sum_{x\in X, y\in Y} p(x,y) \log p(x|y)\right) \\ &= \sum_{x\in X, y\in Y} p(x,y) \log\left(\frac{p(x|y)}{p(x)}\right) \end{aligned}$$

$$\boxed{p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}}$$

- Commutation Property

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

Mutual Information (3)

- Capacity is defined as the maximum mutual information of channel

$$I(X;Y) = H(X) - H(X|Y)$$
$$= -\sum_{x\in X} p(x)\log p(x) - \left(-\sum_{x\in X, y\in Y} p(x,y)\log p(x|y)\right)$$
$$= \sum_{x\in X, y\in Y} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

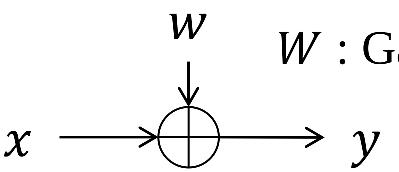
- Capacity

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} \sum_{x\in X, y\in Y} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \qquad (4.3)$$

  - For AWGN, maximizing input distribution is Gaussian

Capacity in AWGN Channel

- Let's consider additive Gaussian noise channel



W : Gaussian

$Y = X + W$

$$E[X^2] = P$$
$$E[W^2] = \sigma^2$$
$$E[Y^2] = E[X^2] + E[W^2] = P + \sigma^2$$

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

$$p(y|x) = p(w)$$

$$p(x,y) = p(y|x)p(x) = p(x)p(w)$$

$$\begin{aligned} H(Y|X) &= -\iint p(x,y)\log p(y|x)\,dxdy \\ &= -\iint p(x)p(w)\log p(w)\,dxdw \\ &= -\int p(w)\log p(w)\,dw = H(W) = \log\sqrt{2\pi e\sigma^2} \end{aligned}$$

$$H(Y) \quad = -\int p(y)\log p(y)\,dy = \log\sqrt{2\pi e(P+\sigma^2)}$$

Capacity in AWGN Channel

- Let's consider additive Gaussian noise channel

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

$$= \log\sqrt{2\pi e(P+\sigma^2)} - \log\sqrt{2\pi e\sigma^2}$$

$$\max_{p(x)} I(X;Y) = \frac{1}{2}\log_2\left(1+\frac{P}{\sigma^2}\right) = \frac{1}{2}\log_2(1+\gamma)$$

$$\Rightarrow \quad C = \frac{I}{T} = \frac{\log_2(1+\gamma)}{2T} = B\log_2(1+\gamma)\ [\text{bps}]$$

Entropy of Gaussian Dist.

The entropy of a one-dimensional Gaussian distribution whose standard deviation is σ is given by

$$H(x) = \log \sqrt{2\pi e}\sigma .$$

This is calculated as follows:

$$
\begin{aligned}
p(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-(x^2/2\sigma^2)} \\
-\log p(x) &= \log \sqrt{2\pi}\sigma + \frac{x^2}{2\sigma^2} \\
H(x) &= -\int p(x) \log p(x)\, dx \\
&= \int p(x) \log \sqrt{2\pi}\sigma\, dx + \int p(x) \frac{x^2}{2\sigma^2}\, dx \\
&= \log \sqrt{2\pi}\sigma + \frac{\sigma^2}{2\sigma^2} \\
&= \log \sqrt{2\pi}\sigma + \log \sqrt{e} \\
&= \log \sqrt{2\pi e}\sigma .
\end{aligned}
$$

C. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, Vol. 27, pp. 379–423, 623–656, July, October, 1948.

Capacity of Flat-Fading Channels

- Channel and System model

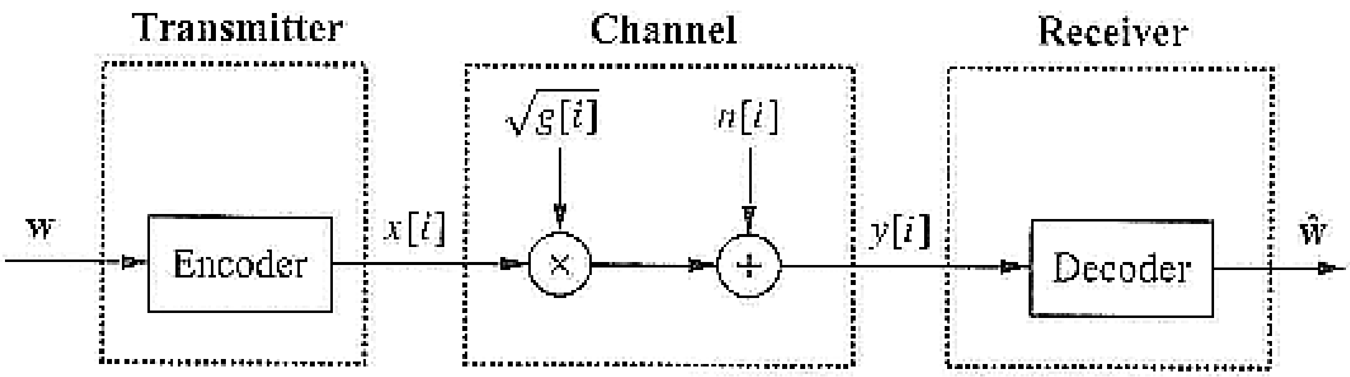


Figure 4.1: Flat fading channel and system model.

- Capacity depends on what is known about channel at Rx/Tx
  - Channel distribution information
  - CSI (channel side information) at Rx
  - CSI at both Rx and Tx

CSI known at Receiver

- Fading Known at Receiver Only
  - via channel estimation
- Shannon Capacity (Ergodic Capacity)
  - Maximum rates that can be sent over the channel with small error probability
- Capacity with outage
  - Maximum rate that can be transmitted over the channel with outage prob. that the transmission cannot be decoded with negligible error prob.
  - Outage prob. characterizes the prob. of data loss or deep fading

CSI at Receiver: Shannon Capacity

- Shannon (ergodic) capacity of fading channel

Jensen's inequality

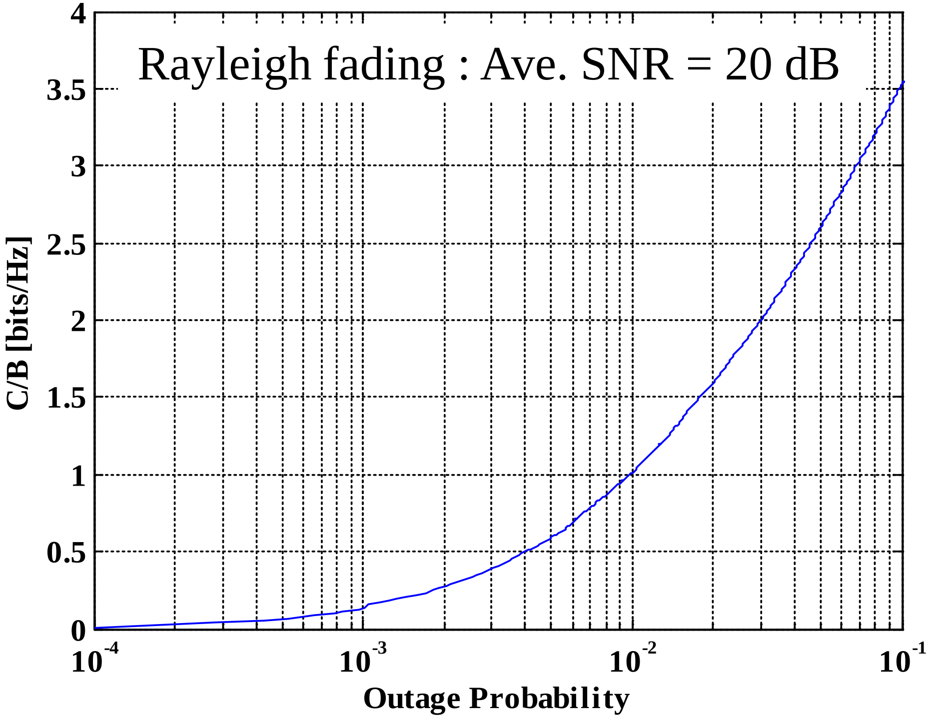
$$C = \int_0^\infty B \log_2(1+\gamma) p(\gamma) d\gamma \leq B \log_2(1+\bar{\gamma}) \qquad (4.4)$$

C @ AWGN channel

- Only the Rx knows the instantaneous SNR, so the data rate transmitted over the channel is constant
- Shannon capacity of a fading channel is less than that of an AWGN channel with the same Ave. SNR

CSI at Receiver: Capacity w/ Outage

- Tx fixes a minimum received SNR and encodes for a data rate as $C = B\log_2(1+\gamma_{\min})$
- The data is correctly received if the instantaneous Rx SNR > min Rx SNR



CSI at Rx/Tx: Shannon Capacity

- For fixed transmit power, same as with only receiver knowledge of fading (variable coding rates)

$$C = \int_0^\infty B \log_2(1+\gamma) p(\gamma) d\gamma \qquad (4.4)$$

- Transmit power $P(\gamma)$ can also be adapted
- Leads to optimization problem

$$C = \max_{P(\gamma):E[P(\gamma)]=\bar{P}} \int_0^\infty B \log_2\left(1 + \frac{\gamma P(\gamma)}{\bar{P}}\right) p(\gamma) d\gamma \qquad (4.9)$$

CSI at Rx/Tx: Shannon Capacity (2)

- It effectively reduces the time-varying channel to a set of time-invariant channels in parallel

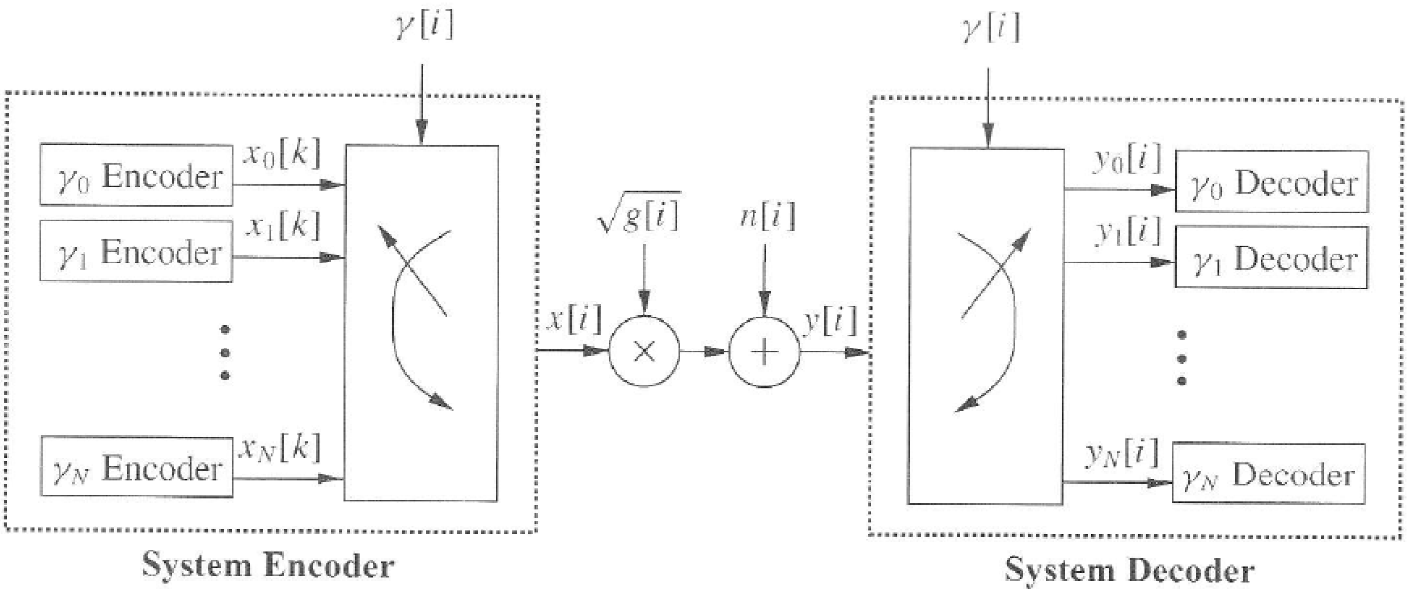


Figure 4.4: Multiplexed coding and decoding.

$$\text{Power constraint : } \int_0^\infty P(\gamma)p(\gamma)d\gamma = \bar{P} \qquad (4.8)$$

Optimal Adaptive Scheme

- Power Adaptation

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} \frac{1}{\gamma_0} - \frac{1}{\gamma} & (\gamma \geq \gamma_0) \\ 0 & (\gamma < \gamma_0) \end{cases} \quad (4.12)$$

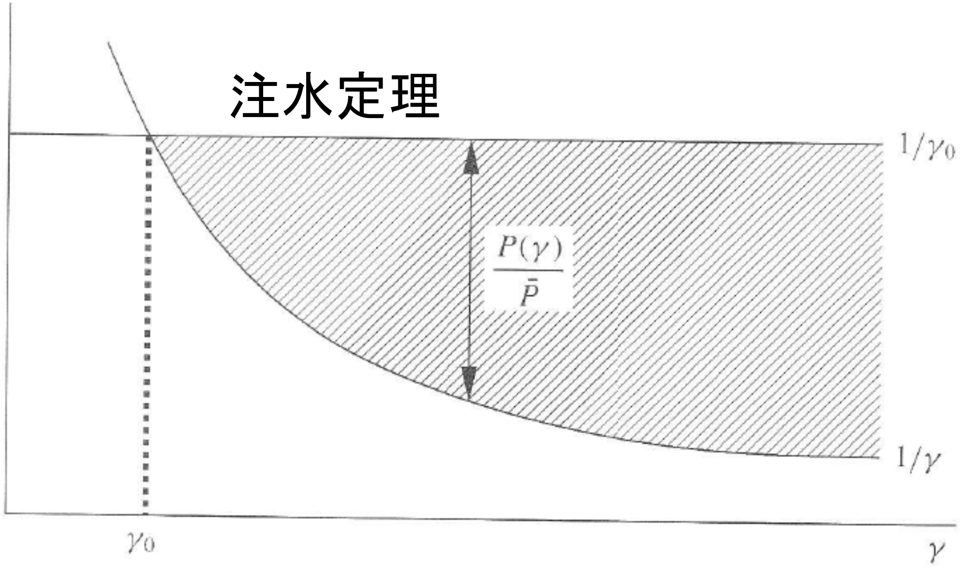


Figure 4.5: Optimal power allocation: water-filling.

- Optimum Capacity

$$\frac{C}{B} = \int_{\gamma_0}^{\infty} \log_2\left(\frac{\gamma}{\gamma_0}\right) p(\gamma) d\gamma \quad (4.13)$$

$$\text{For } \gamma, \int_0^{\infty} \frac{P(\gamma)}{\bar{P}} p(\gamma) d\gamma = \int_{\gamma_0}^{\infty} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma}\right) p(\gamma) d\gamma = 1 \quad (4.14\text{-}15)$$

Zero Outage Capacity

- Fading is inverted to maintain constant SNR (*channel inversion*) by Tx's power control
- Simplifies design (fixed coding rate)
- Capacity

$$\frac{P(\gamma)}{\bar{P}}\gamma = \sigma$$

const

$$C = B\log_2(1+\sigma) = B\log_2\left(1 + 1/E\left[\frac{1}{\gamma}\right]\right) \quad (4.18)$$

$$\because \frac{P(\gamma)}{\bar{P}}\gamma = \sigma$$

$$E\left[\frac{P(\gamma)}{\bar{P}}\right] = E\left[\frac{\sigma}{\gamma}\right] = 1 \qquad \therefore \sigma = \frac{1}{E[1/\gamma]}$$

  - greatly reduced relative to Shannon Capacity
  - becomes zero in Rayleigh fading

Maximum Outage Capacity

- Truncated inversion
  - Invert channel above cutoff fade depth
  - Constant SNR (fixed rate) above cutoff
  - Cutoff greatly increases capacity
    - Close to optimal

$$C = B\log_2\left(1+\frac{1}{E_{\gamma_0}\left[\frac{1}{\gamma}\right]}\right)p(\gamma \geq \gamma_0)$$

$$\frac{P(\gamma)}{\bar{P}} = \begin{cases} \sigma/\gamma & \gamma \geq \gamma_0 \\ 0 & \gamma < \gamma_0 \end{cases}$$

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

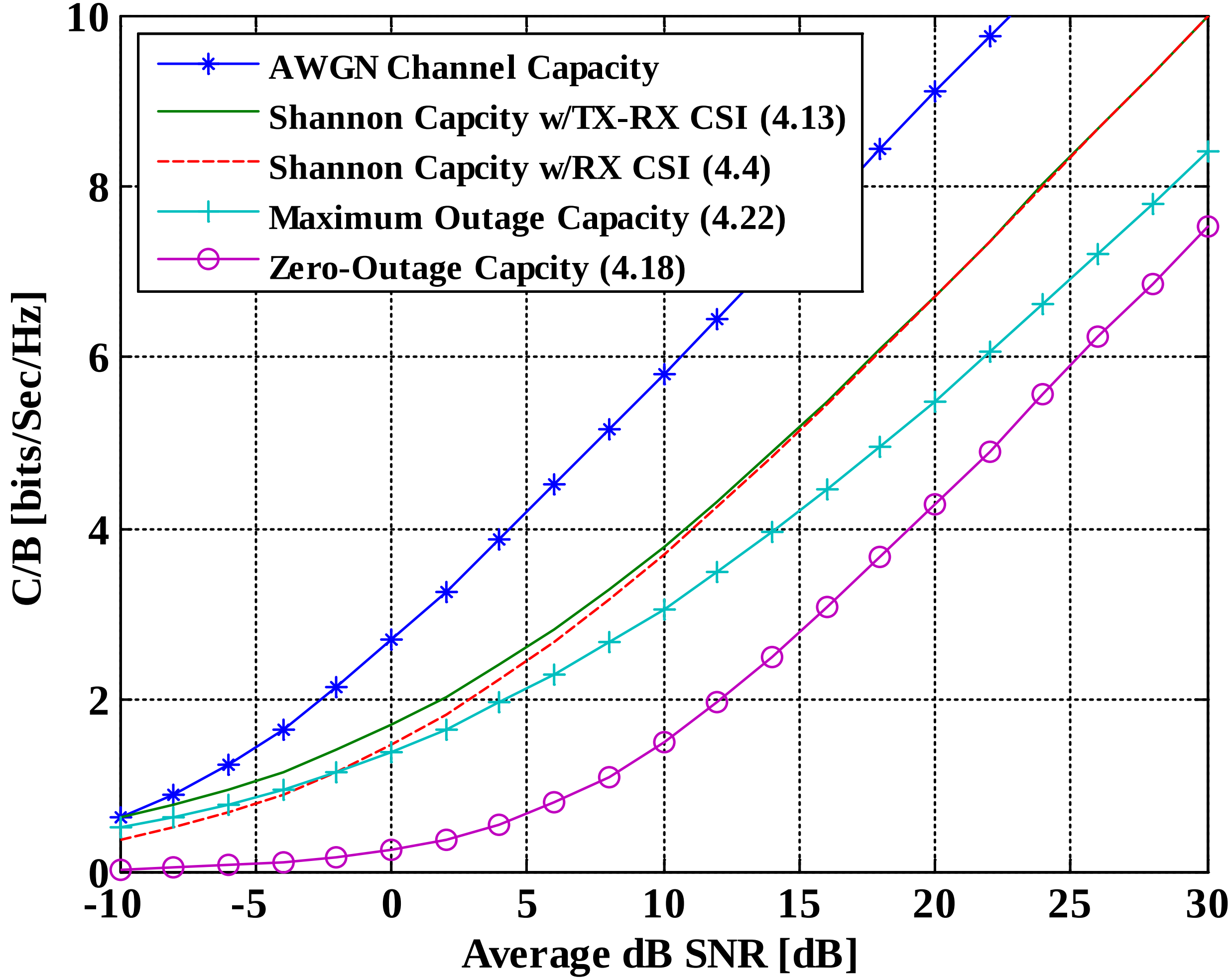
- Find maximum w.r.t. γ_0

$$C = \max_{\gamma_0} B\log_2\left(1+\frac{1}{E_{\gamma_0}\left[\frac{1}{\gamma}\right]}\right)p(\gamma \geq \gamma_0)$$

Capacity comparison in Flat-Fading

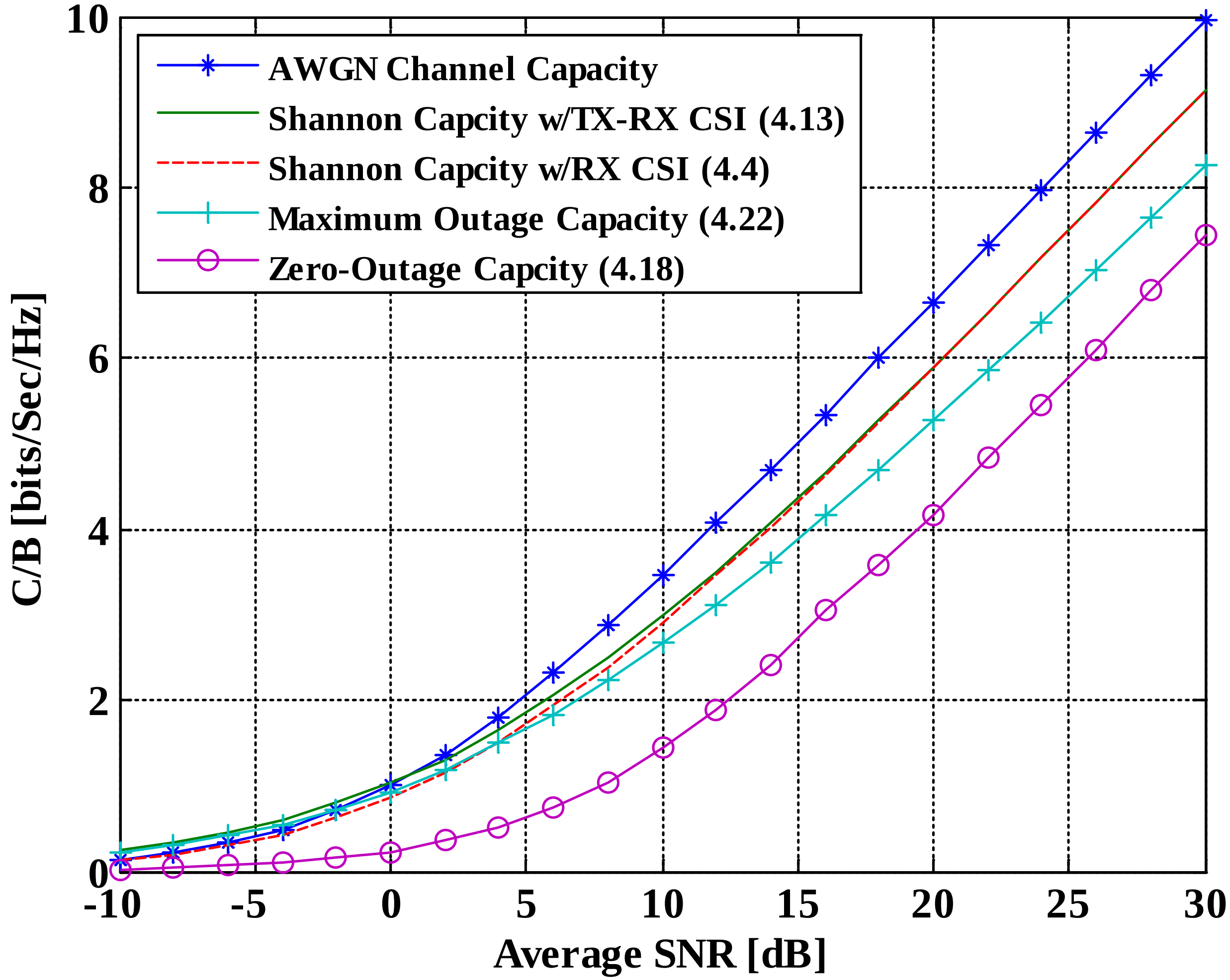
- Log normal distribution

$\sigma_{\text{dB}} = 8$



Capacity comparison in Flat-Fading (2)

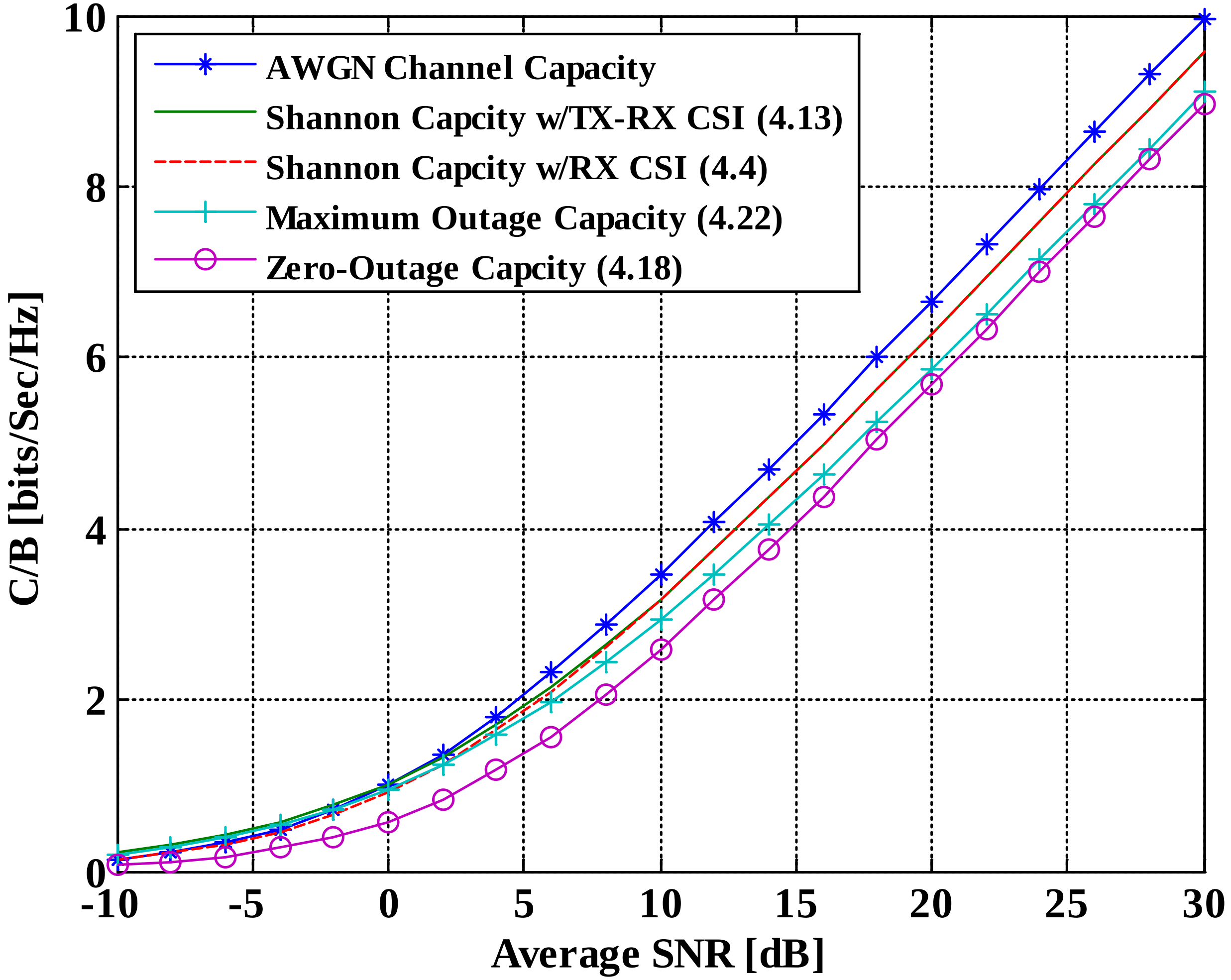
- Rayleigh distribution



Capacity comparison in Flat-Fading (3)

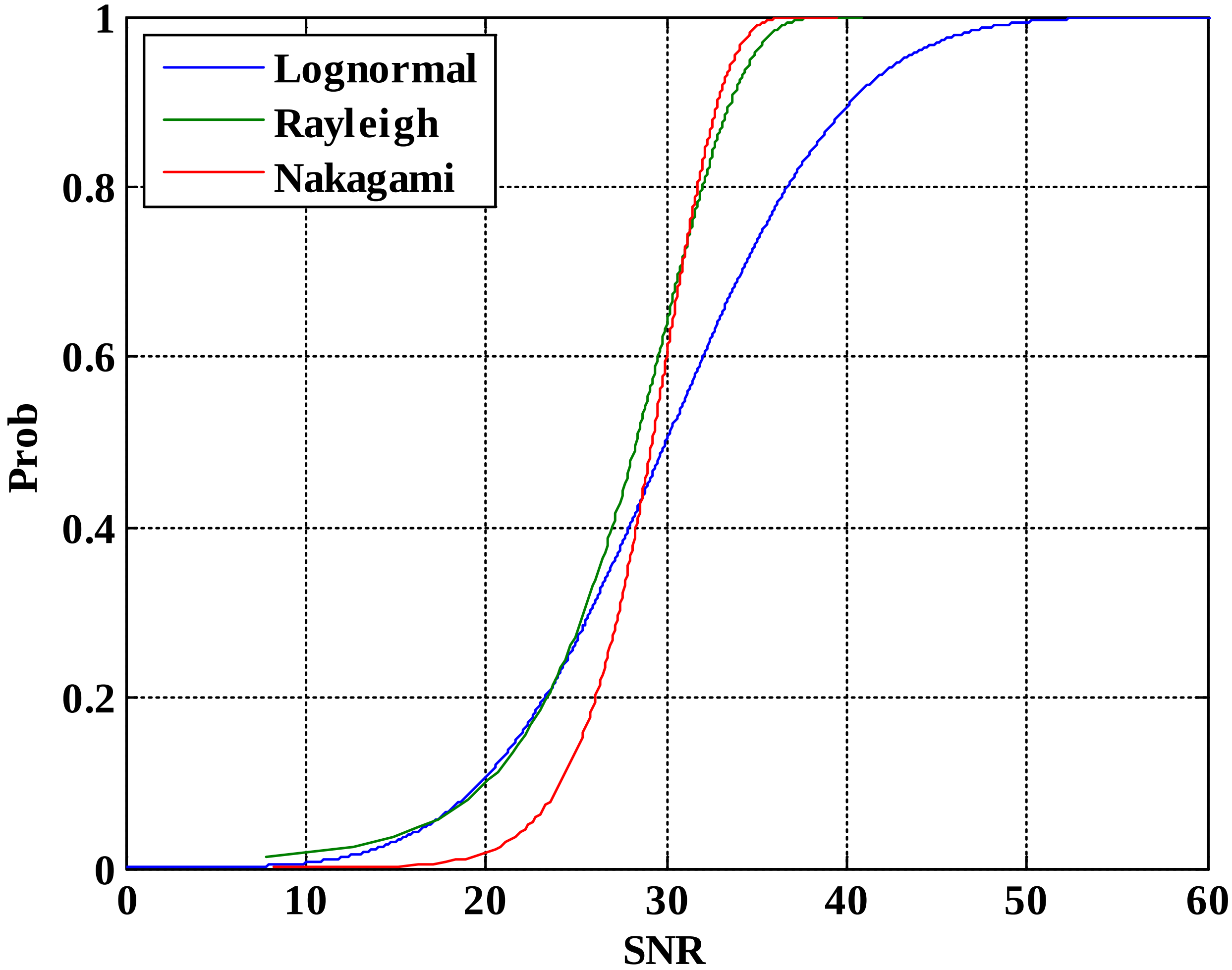
- Nakagami-*m* distribution

$m = 2$



Capacity comparison in Flat-Fading (4)

- SNR distributions (@Ave. SNR=30 dB)



Capacity comparison (1)

- Capacity in AWGN channel is larger than that in fading channel for all cases
- At low SNR, AWGN and fading channel with Tx/Rx CSI have almost the same capacity
- At negative SNR, capacity of the fading channel with Tx/Rx CSI is larger than that in AWGN channel
- For low severity of fading, capacity difference between various adaptive policies decreases and approach that of the AWGN channel

Capacity comparison (2)

- Capacities with Tx/Rx CSI and Rx CSI only are almost same → power adaptation yields a negligible gain since coding rate is adapted relative to the channel
- Trade-off between capacity and complexity
  - Adaptive policy
    - Increase Tx complexity but allows simple Rx decoder
  - Non-adaptive policy
    - Needs relatively simple Tx encoder but its code design must use the channel correlation statistics
    - Decoder complexity is proportional to the channel decorrelation time

Frequency Selective Fading Channels

- For TI channels, capacity achieved by water-filling in frequency
- Capacity of time-varying channel unknown
- Approximate by dividing into subbands
  - Each subband has bandwidth B_c
  - Independent fading in each subband
  - Capacity is the sum of subband capacities