Algorithms for Wireless Communications

Heinrich Meyr, Andreas Burg

Institute for Integrated Signal Processing Systems, RWTH Aachen

Telecommunications Circuits Laboratory, EPFL

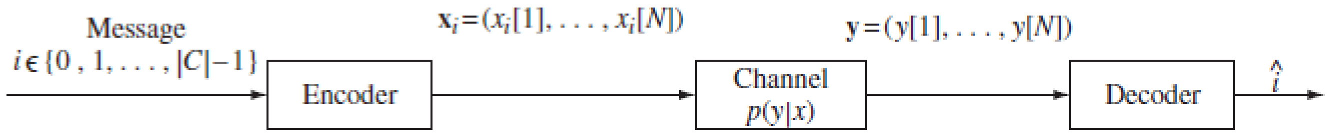
Lecture: Achivable Rate for flat fading channel

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

RWTHAACHEN
UNIVERSITY

Information Theory Fundamentals

- **Abstraction of a Communication system**



- **Reliable communication at rate R**
- **For every δ find a code of rate R and a block of length n such that the error prob p_e < δ**
- **Capacity: Maximum rate for which reliable communication is possible. Shannon proved that C larger than zero exists if a sufficiently large block N is used**

- **Entropy: measure of uncertainty of the r.v. x.**

$$H(x) = \sum_i p_x(i) \log\left(\frac{1}{p_x(i)}\right)$$

$$H(x) \geq 0, \text{ zero if x is deterministic}$$

- **Joint entropy**

$$H(x, y) = \sum_{i,j} p_{x,y}(i, j) \log\left(\frac{1}{p_{x,y}(i, j)}\right)$$

- **Conditional entropy**

$$H(x \mid y = j) = \sum_i p_{x|y}(i \mid j) \log\left(\frac{1}{p_{x|y}(i \mid j)}\right)$$

amount of uncertainty left after observing y=j

- **Averaging over all possible outcomes of y yields**

$$H(x|y) = \sum_j p_y(j) H(x|y=j)$$

- **Chain rule for entropy**

$$H(x,y) = H(x) + H(y|x) = H(y) + H(x|y)$$

The uncertainty of x and y is the uncertainty about x plus the uncertainty of x after one observes y

- **Properties**

$H(x,y) = H(x) + H(y)$ if x, y are statistically independent

Conditioning reduces uncertainty

$0 \le H(x|y) \le H(x)$

- **Mutual information (definition)**

$$I(x;y) = H(x) - H(x|y)$$
$$= H(y) - H(y|x)$$

- **The quantity** $H(x) - H(x|y)$ **is of special significance. Since H(x) is the uncertainty before observing y of x the difference is the reduction of uncertainty after having observed y**
- **Note: The mutual information is symmetric in x and y and therefore the second line in this equation has an analogous interpretation**
- **The mutual information is always non-negative.it is zero only if the r.v. are independent.**
- **The mutual information can also be written as (verify)**

$$I(x;y) = \sum_{i,j} p_{x,y}(i,j)\log\left(\frac{p_{x,y}(i,j)}{p_x(i)\,p_y(j)}\right)$$

- **One can also define a conditional mutual information**

$$I(x;y|z) = H(x|z) - H(y|z)$$

- **Chain rule of information**

$$\begin{aligned} I(x_1, x_2; y) &= I(x_1; y) + I(x_2; y|x_1) \\ &= I(x_2; y) + I(x_1; y|x_2) \end{aligned}$$

- **Information that x1,x2 jointly provide about y equals the information of x1 provides on y plus the information that x2 provides after observation of x1**

- **For reliable communication the conditional entropy of x when observing y must be small. Otherwise there exists to much uncertainty about x**
- **Assume we transmit N symbols**

$$H(\mathbf{x}|\mathbf{y}) = H(\mathbf{x}) - I(\mathbf{x};\mathbf{y})$$

(Note that x,y are vectors of length N)

- **The source generates a number of bits equal to H(x) which are transmitted at rate R, therefore**

$$H(\mathbf{x}) = NR$$

$$R = \frac{1}{N}H(\mathbf{x}) \approx \frac{1}{N}I(\mathbf{x};\mathbf{y})$$

- **Intuitively:The rate of information generated should match the rate which is reliably transmitted**

- **Recall that the capacity is defined as the maximum rate of mutual information optimized over the distribution of the input symbols**

$$C = \max_{P(x)} I(x; y)$$

Capacity of the additive white Gaussian Noise channel

- **The channel is memory less and completely characterized by the conditional density**

$$p(y_k | x_k) = N(0, \sigma_n^2) \qquad p(y_k) = N(0, \sigma_s^2 + \sigma_n^2)$$

- **For a complex additive Gaussian noise channel y(k)=x(k)+w(k) with input power constraint P and noise variance σ^2 the capacity is achieved for a Gaussian input signal distribution (see B.4)**

$$C = \log_2\left(1 + \frac{P}{\sigma^2}\right) \textit{ bits/chann el\_use}$$

- **For a bandlimited (bandwidth W) continuous-time channel we have W complex degrees of freedom**

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

$$\Leftrightarrow \quad C/W = \log_2\left(1 + \frac{P'}{N_0}\right) \textit{ bit/s/Hz} = \log_2(1 + SNR) \textit{ bit/s/Hz} \qquad (P' = P/W)$$

Capacity of the additive white Gaussian Noise channel

- **The channel is memory less and completely characterized by the conditional density**

$$p(y_k | x_k) = N(0, \sigma_n^2)$$

- **For a complex additive Gaussian noise channel y(k)=x(k)+w(k) with input power constraint P and noise variance σ^2 the capacity is achieved for a Gaussian input signal distribution (see B.4)**

$$C = \log_2\left(1 + \frac{P}{\sigma^2}\right) \textit{ bits/channel\_use}$$

- **For a bandlimited (bandwidth W) continuous-time channel we have W complex degrees of freedom**

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

$$\Leftrightarrow \quad C/W = \log_2\left(1 + \frac{P'}{N_0}\right) \textit{ bit/s/Hz} = \log_2(1 + SNR) \textit{ bit/s/Hz} \qquad (P' = P/W)$$

- **C approximately increases by with an SNR increase of 2**
    - “1 bit increase per 3db SNR increase”

- **C as a function of the bandwidth**
  - Two conflicting dependencies
    - The first factor increases linearly with W
    - However, the noise variance also increases linearly with the bandwidth W (noise density N0 watts/Hz is constant) which decreases the SNR
  - Series expansion of C yields

$$C = W \ln\left(1 + \frac{P}{N_0 W}\right) 1/\ln(2)$$

$$C \approx \frac{P}{N_0} 1/\ln 2 \quad \text{for W} \to \infty$$

    - C increases linearly with power

Capacity as a Function of bandwidth

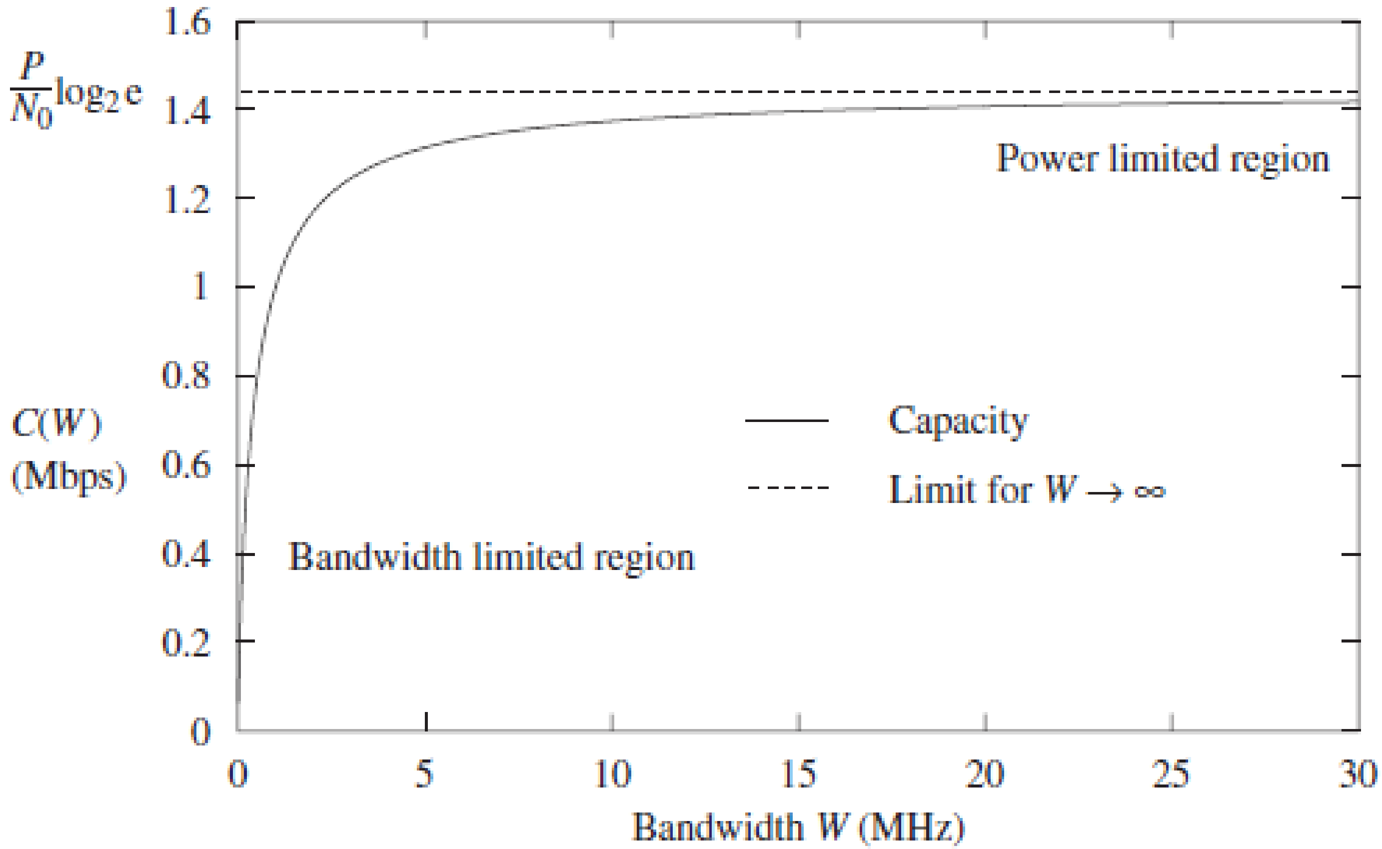


Figure 5.5 Capacity as a function of the bandwidth W. Here $\bar{P}/N_0 = 10^6$.

Capacity in the limit for W

$$P = E_s r = E_b R_b$$

R_b : rate of the information bits/s

r : rate of the coded bits/s

$$\mathrm{R} = \frac{\mathrm{R}_{\mathrm{b}}}{r} : \text{code rate}$$

E_{b} : energy per bit

E_{s} : energy per

$$\mathrm{C} = \mathrm{W}\ln(1 + \frac{E_b R}{N_0})1/\ln 2$$

$$\approx WR\frac{E_b}{N_0}1/\ln 2$$

- **In the limit**

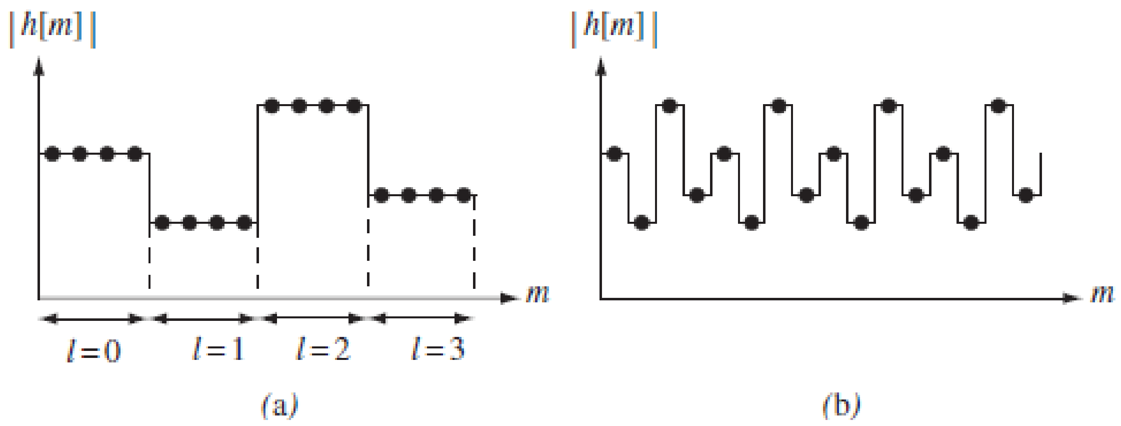
$$C_{\infty} \approx RW$$

we obtain

$$\frac{E_b}{N_0} = \ln 2 \qquad , (-1.59\text{db})$$

Fading Channel with perfect Interleaving

Figure 5.19 (a) Typical trajectory of the channel strength as a function of symbol time under a block fading model. (b) Typical trajectory of the channel strength after interleaving. One can equally think of these plots as rates of flow of information allowed through the channel over time.



- **Ideal interleaving completely breaks the memory of the channel Successive interleaved symbols go trough independent fades**
- **We can write for the de-interleaved channel output**

$$y(m) = h_{\text{int}}(m)\,x(m) + w(m)$$

- **where** $h_{\text{int}}(m)$ **are statistically independent channel values**

  (in the following we omit the index int to simplify notation)

- **Consider a block of length N of transmitted symbols**
- **The mutual information for a *perfectly known* channel (called "coherent communication" in the literature) is**

$$I(\mathbf{x};\mathbf{y},\ \mathbf{h})$$

- **Note that we view the known channel h as part of the channel output**
- **Applying the channel rule of mutual information we obtain**

$$I(\mathbf{x};\mathbf{y},\ \mathbf{h}) = I(\mathbf{x};\ \mathbf{h}) + I(\mathbf{x};\ \mathbf{y}\,|\,\mathbf{h})$$

$$= \mathrm{I}(\ \mathbf{x};\ \mathbf{y}\,|\,\mathbf{h})$$

- **Since the symbols and the channel are independent the first term equals zero**

- **Applying the chain rule of information and using the property that the channel is memoryless we obtain**

$$\frac{1}{N} I(\mathbf{x};\mathbf{y} \mid \mathbf{h}) = \frac{1}{N} \sum_{n=0}^{n=N-1} I(x_n; y_n \mid h_n)$$

- **Due to the ergodicity of the fading process we obtain for large N**

$$\frac{1}{N} I(\mathbf{x};\mathbf{y} \mid \mathbf{h}) = \frac{1}{N} \sum_{n=0}^{n=N-1} I(x_n; y_n \mid h_n) \rightarrow E_h(I(x; y \mid h))$$

- **Conditioned on the coefficient *h* the channel is simply a AWGN one**
- **The optimal input distribution which maximizes the achievable rate is Gaussian for any SNR**

$$C = \max_{p(x)} I(x; y|h) = \log(1 + |h|^2 \mathrm{SNR})$$

- **Therefor, the capacity of the fading channel is obtained by averaging over many possible realization of the fading process**

$$C = E_h(I(x; y|h)) = E_h(\log(1 + |h|^2 SNR))$$

- **Consider a system with a fixed transmission rate R and a “perfect” code**
  - If the rate R is below the capacity of the channel transmission is error free.
  - If the rate is above capacity, error-free transmission is not possible, i.e., the link is down => the link is in **OUTAGE**

- **The channel as Rayleigh fading, i.e., $\boldsymbol{h \sim \mathcal{CN}(0,1)}$**
  - Since the channel is a r.v., the capacity is also a r.v.

- **We are interested in the following question**

“Outage-probability: What is the probability that the channel does not support error-free transmission with rate R“

- **Mathematical formulation:**

$$P_{out} = Pr(C(h) < R)$$

- **Since** $C(h) = \log_2(1 + |h|^2 SNR)$, we can equivalently consider the probability that $|h|^2 SNR$ drops below the SNR required for a capacity of at least R

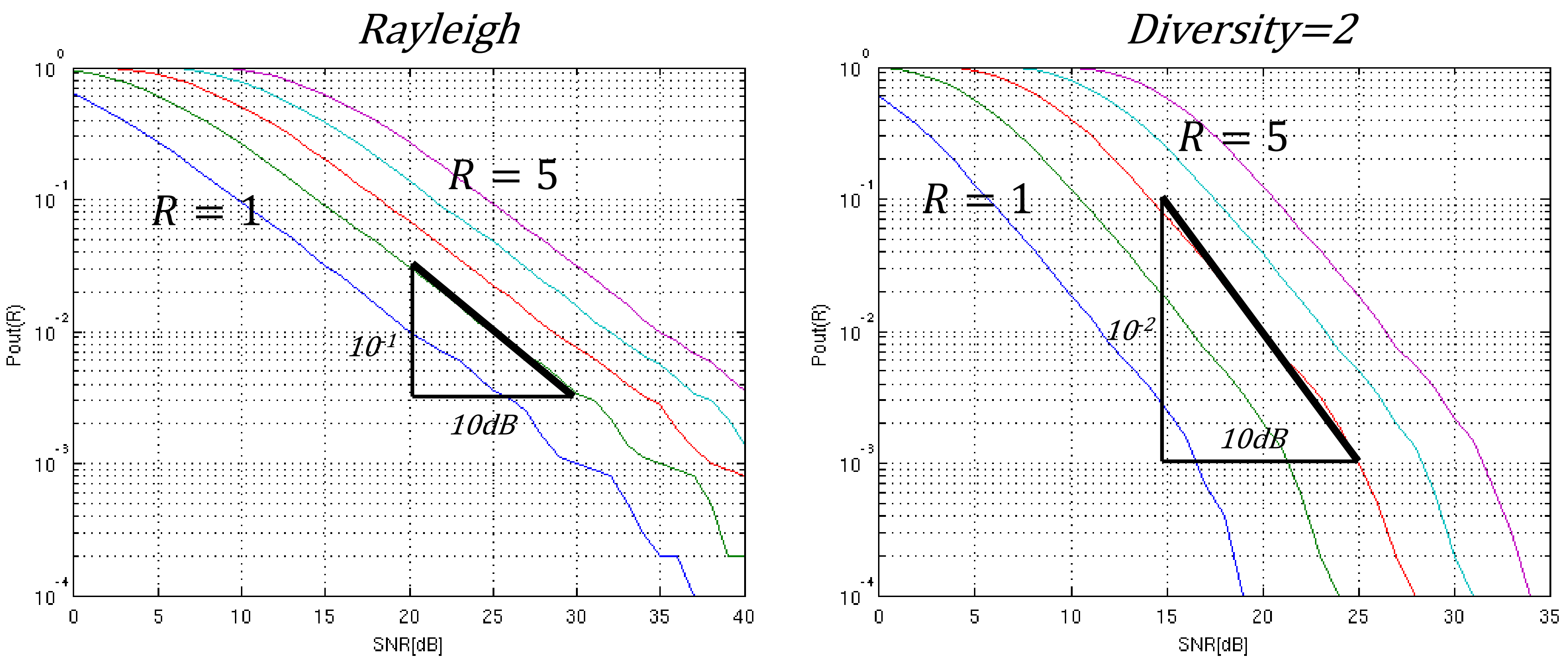
$$Pr(|h|^2 SNR < SNR_R) = Pr\left(|h|^2 < \frac{SNR_R}{SNR}\right)$$

$$SNR_R = 2^R - 1$$

Depends on the distribution of $|h|^2$

- **Note: dependency of outage probability on SNR follows the same criterion as the error-probability**

- **Rayleigh fading channel vs Diversity=2 channel**



- **Outage probability is a lower-bound on block-error-rate (outage-lower-bound)**

- **We can reverse the question as follows:**

"For a given SNR, what is the highest rate that I can choose to transmit such that the outage probability does not exceed X%"

Or in other words

"Outage capacity: For a given SNR, what is the capacity $C_{out}(X\%)$ that is available for 100-X% of the cases"

$$X\% = Pr\left(C(h) < C_{out}(X\%)\right)$$

- **Outage capacity can be derived from CDF**