Outline

- **Capacity of Wireless Channels (Chapter 4 in Goldsmith's Book)**
  - **Capacity in AWGN**
  - **Capacity of Flat-Fading Channels**
    - **Channel and System Model**
    - **Channel Distribution Information (CDI) Known**
    - **Channel Side Information at Receiver**
    - **Channel Side Information at Transmitter and Receiver**
    - **Capacity with Receiver Diversity**
  - **Capacity of Frequency-Selective Fading Channels**

Introduction

- **Channel capacity was pioneered by Claude Shannon in the late 1940s, using a mathematical theory of communication based on the notion of mutual information between the input and output of a channel**
- **C. E. Shannon, "A Mathematical Theory of Communication", Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, July, October, 1948**

1916-2001

- **Proposed Entropy as a measure of information**
- **Showed that Digital Communication is possible.**
- **Derived the Channel Capacity Formula**
- **Showed that reliable information transmission is always possible if transmitting rate does not exceed the channel capacity**

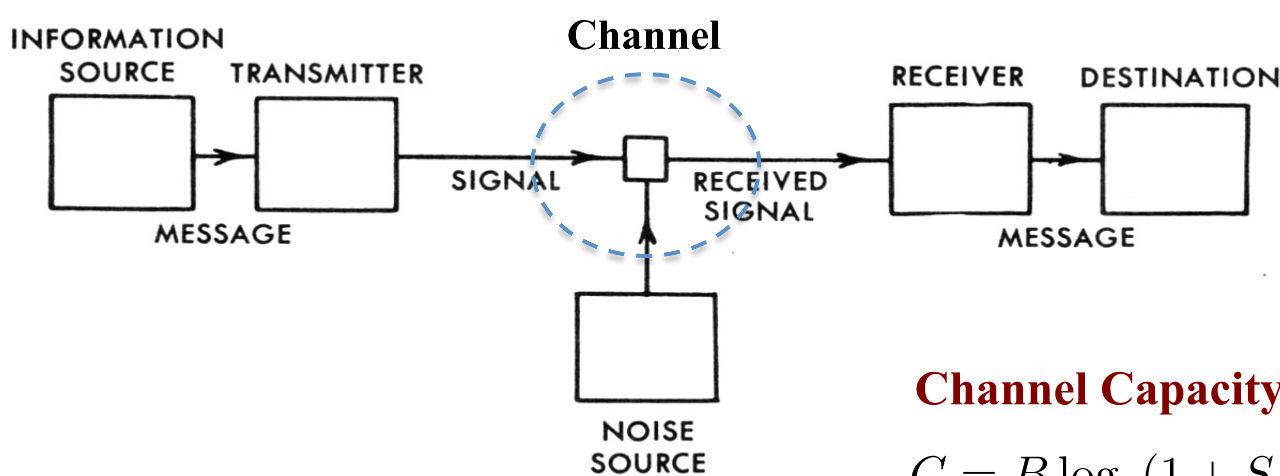
Claude Shannon in 1948 (32 years old)

Introduction

The fundamental problem of communication is that of reproducing at one point, either exactly or approximately, a message selected at another point.

~ Claude Shannon

34 The Mathematical Theory of Communication



Channel Capacity :

$$C = B \log_2(1 + SNR)$$

Fig. 1. — Schematic diagram of a general communication system.

Capacity in AWGN

- **Consider a discrete-time additive white Gaussian noise (AWGN) channel with channel input/output relationship $y[i] = x[i] + n[i]$, where *x*[*i*] is the channel input at time *i*, *y*[*i*] is the corresponding channel output, and *n*[*i*] is a white Gaussian noise random process.**
- **Assume a channel bandwidth *B* and transmit power *P*.**
- **The channel SNR, the power in *x*[*i*] divided by the power in *n*[*i*], is constant and given by $\gamma = P/(N_0 B)$, where N_0 is the power spectral density of the noise.**
- **The capacity of this channel is given by Shannon's well-known formula**

$$C = B \log_2(1 + \gamma),$$

where the capacity units are bits/second (bps).

- **Shannon's coding theorem proves that a code exists that achieves data rates arbitrarily close to capacity with arbitrarily small probability of bit error.**

Capacity in AWGN

- **The converse theorem shows that any code with rate *R>C* has a probability of error bounded away from zero.**
- **For a memoryless time-invariant channel with random input *x* and random output *y*, the channel's mutual information is defined as**

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

where the sum is taken over all possible input and output pairs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ for $\mathcal{X}$ and $\mathcal{Y}$ the input and output alphabets.

- **Mutual information can also be written in terms of the entropy in the channel output *y* and conditional output *y*|*x* as**

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

where

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

Capacity in AWGN

- **Shannon proved that channel capacity equals the mutual information of the channel maximized over all possible input distributions:**

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

For the AWGN channel, the maximizing input distribution is Gaussian, which results in the channel capacity $C = B \log_2(1+\gamma)$ **.**

- **At the time that Shannon developed his theory of information, data rates over standard telephone lines were on the order of 100 bps. Thus, it was believed that Shannon capacity, which predicted speeds of roughly 30 Kbps over the same telephone lines, was not a very useful bound for real systems.**
- **However, breakthroughs in hardware, modulation, and coding techniques have brought commercial modems of today very close to the speeds predicted by Shannon in the 1950s.**

Capacity in AWGN

- **Example 4.1: Consider a wireless channel where power falloff with distance follows the formula $P_r(d) = P_t(d_0/d)^3$ for d_0 = 10 m. Assume the channel has bandwidth B = 30 KHz and AWGN with noise power spectral density of $N_0 = 10^{-9}$ W/Hz. For a transmit power of 1W, find the capacity of this channel for a transmit-receive distance of 100 m and 1 Km.**

Solution: The received SNR is $\gamma = P_r(d)/(N_0B) = .1^3/(10^{-9} \times 30 \times 10^3) = 33 = 15$ dB for $d = 100$ m and $\gamma = .01^3/(10^{-9} \times 30 \times 10^3) = .033 = -15$ dB for $d = 1000$ m. The corresponding capacities are $C = B\log_2(1+\gamma) = 30000\log_2(1+33) = 152.6$ Kbps for $d = 100$ m and $C = 30000\log_2(1+.033) = 1.4$ Kbps for $d = 1000$ m. Note the significant decrease in capacity at farther distances, due to the path loss exponent of 3, which greatly reduces received power as distance increases.

Capacity of Flat-Fading Channels

- **Assume a discrete-time channel with stationary and ergodic time-varying gain $\sqrt{g[i]}, 0 \leq g[i]$, and AWGN *n*[*i*], as shown in Figure 4.1.**

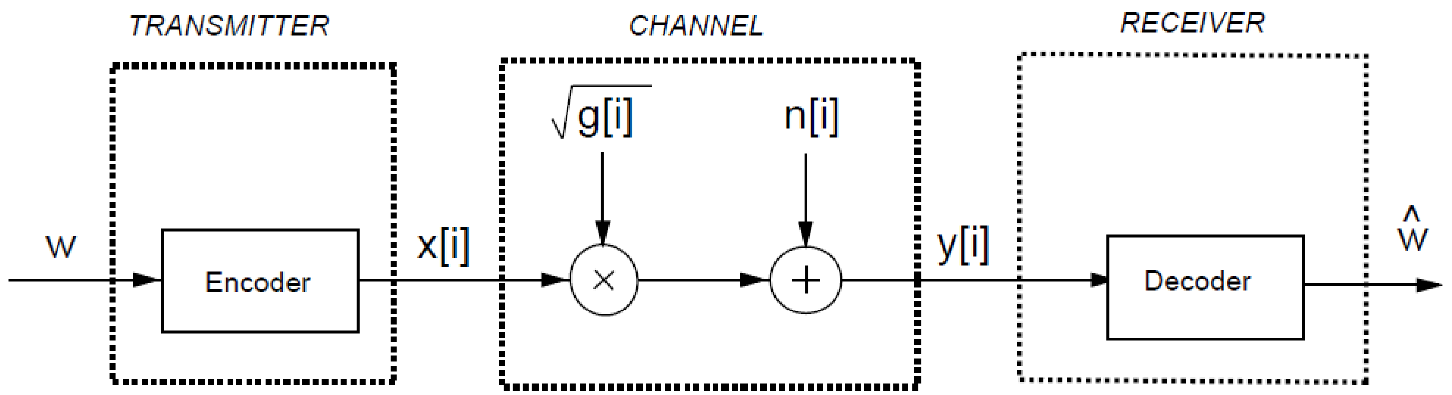


Figure 4.1: Flat-Fading Channel and System Model.

- **The channel power gain *g*[*i*] follows a given distribution *p*(*g*), e.g. for Rayleigh fading *p*(*g*) is exponential. The channel gain *g*[*i*] can change at each time *i*, either as an i.i.d. process or with some correlation over time.**

Capacity of Flat-Fading Channels

- **In a block fading channel $g[i]$ is constant over some block length T after which time $g[i]$ changes to a new independent value based on the distribution $p(g)$.**
- **Let $\overline{P}$ denote the average transmit signal power, $N_0/2$ denote the noise power spectral density of $n[i]$, and B denote the received signal bandwidth. The instantaneous received signal-to-noise ratio (SNR) is then $\gamma[i] = \overline{P}g[i]/(N_0B), 0 \leq \gamma[i] < \infty$ and its expected value over all time is $\overline{\gamma} = \overline{P}\overline{g}/(N_0B)$.**
- **The channel gain $g[i]$, also called the channel side information (CSI), changes during the transmission of the codeword.**
- **The capacity of this channel depends on what is known about $g[i]$ at the transmitter and receiver. We will consider three different scenarios regarding this knowledge: Channel Distribution Information (CDI), Receiver CSI and Transmitter and Receiver CSI**

Capacity of Flat-Fading Channels

- First consider the case where the channel gain distribution $p(g)$ or, equivalently, the distribution of SNR $p(\gamma)$ is known to the transmitter and receiver.
- For i.i.d. fading the capacity is given by (4.3), but solving for the capacity-achieving input distribution, i.e. the distribution achieving the maximum in (4.3), can be quite complicated depending on the fading distribution.
- For these reasons, finding the capacity-achieving input distribution and corresponding capacity of fading channels under CDI remains an open problem for almost all channel distributions.

- Now consider the case where the CSI $g[i]$ is known at the receiver at time i. Equivalently, $\gamma[i]$ is known at the receiver at time i.

- Also assume that both the transmitter and receiver know the distribution of $g[i]$. In this case, there are two channel capacity definitions that are relevant to system design: Shannon capacity, also called ergodic capacity, and capacity with outage.

Capacity of Flat-Fading Channels

- **For the AWGN channel, Shannon capacity defines the maximum data rate that can be sent over the channel with asymptotically small error probability.**
- **Note that for Shannon capacity the rate transmitted over the channel is constant: the transmitter cannot adapt its transmission strategy relative to the CSI.**
- **Capacity with outage is defined as the maximum rate that can be transmitted over a channel with some outage probability corresponding to the probability that the transmission cannot be decoded with negligible error probability.**
- **The basic premise of capacity with outage is that a high data rate can be sent over the channel and decoded correctly except when the channel is in deep fading.**
- **By allowing the system to lose some data in the event of deep fades, a higher data rate can be maintained if all data must be received correctly regardless of the fading state.**

Capacity of Flat-Fading Channels

- **Shannon capacity of a fading channel with receiver CSI for an average power constraint *P* can be obtained as**

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

- **By Jensen's inequality,**

$$\mathbf{E}[B\log_2(1+\gamma)] = \int B\log_2(1+\gamma)p(\gamma)d\gamma \leq B\log_2(1+\mathbf{E}[\gamma]) = B\log_2(1+\overline{\gamma}),$$

where $\overline{\gamma}$ is the average SNR on the channel.

- **Here we see that the Shannon capacity of a fading channel with receiver CSI only is less than the Shannon capacity of an AWGN channel with the same average SNR.**
- **In other words, fading reduces Shannon capacity when only the receiver has CSI.**

Capacity of Flat-Fading Channels

- **Example 4.2: Consider a flat-fading channel with i.i.d. channel gain *g*[*i*] which can take on three possible values: $g_1 = .05$ with probability $p_1 = .1$, $g_2 = .5$ with probability $p_2 = .5$, and $g_3 = 1$ with probability $p_3 = .4$. The transmit power is 10 mW, the noise spectral density is $N_0 = 10^{-9}$ W/Hz,
and the channel bandwidth is 30 KHz. Assume the receiver has knowledge of the instantaneous value of *g*[*i*] but the transmitter does not. Find the Shannon capacity of this channel and compare with the capacity of an**

Solution: The channel has 3 possible received SNRs, $\gamma_1 = P_t g_1/(N_0 B) = .01*(.05^2)/(30000*10^{-9}) = .8333 = -.79$ dB, $\gamma_2 = P_t g_2/(N_0 B) = .01 \times (.5^2)/(30000*10^{-9}) = 83.333 = 19.2$ dB, and $\gamma_3 = P_t g_3/(N_0 B) = .01/(30000*10^{-9}) = 333.33 = 25$ dB. The probabilities associated with each of these SNR values is $p(\gamma_1) = .1$, $p(\gamma_2) = .5$, and $p(\gamma_3) = .4$. Thus, the Shannon capacity is given by

$$C = \sum_i B\log_2(1+\gamma_i)p(\gamma_i) = 30000(.1\log_2(1.8333) + .5\log_2(84.333) + .4\log_2(334.33)) = 199.26 \text{ Kbps}.$$

Capacity of Flat-Fading Channels

The average SNR for this channel is $\bar{\gamma} = .1(.8333) + .5(83.33) + .4(333.33) = 175.08 = 22.43$ dB. The capacity of an AWGN channel with this SNR is $C = B\log_2(1 + 175.08) = 223.8$ Kbps. Note that this rate is about 25 Kbps larger than that of the flat-fading channel with receiver CSI and the same average SNR.

- **Capacity with outage applies to slowly-varying channels, where the instantaneous SNR γ is constant over a large number of transmissions (a transmission burst) and then changes to a new value based on the fading distribution.**
- **If the channel has received SNR γ during a burst, then data can be sent over the channel at rate $B\log_2(1+\gamma)$ with negligible probability of error**
- **Capacity with outage allows bits sent over a given transmission burst to be decoded at the end of the burst with some probability that these bits will be decoded incorrectly.**
- **Specifically, the transmitter fixes a minimum received SNR γ_{min} and encodes for a data rate $C = B\log_2(1+\gamma_{min})$.**
- **The data is correctly received if the instantaneous received SNR is greater than or equal to γ_{min}.**

Capacity of Flat-Fading Channels

- **If the received SNR is below γ_{min} then the receiver declares an outage. The probability of outage is thus $p_{out} = p(\gamma < \gamma_{min})$.**
- **The average rate correctly received over many transmission bursts is $C_o = (1 - p_{out}) B \log_2(1 + \gamma_{min})$ since data is only correctly received on $1 - p_{out}$ transmissions.**

- **Capacity with outage is typically characterized by a plot of capacity versus outage, as shown in Figure 4.2.**
- **In Figure 4.2. we plot the normalized capacity $C/B = \log_2(1 + \gamma_{min})$ as a function of outage prob. $p_{out} = p(\gamma < \gamma_{min})$ for a Rayleigh fading channel with $\overline{\gamma} = 20$ dB.**

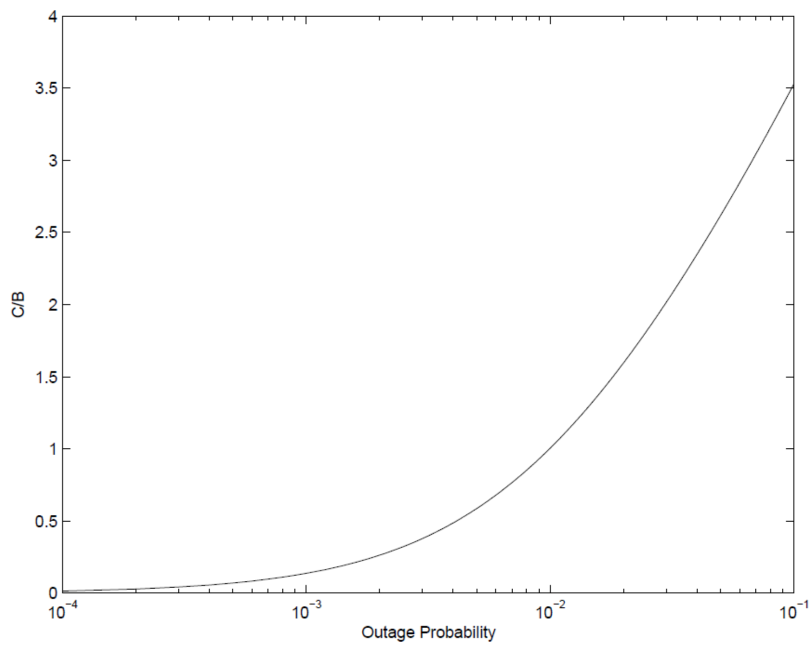


Figure 4.2: Normalized Capacity (C/B) versus Outage Probability.

Capacity of Flat-Fading Channels

- **Example 4.3**: **Assume the same channel as in the previous example, with a bandwidth of 30 KHz and three possible received SNRs** $\gamma_1 = .8333$ **with** $p(\gamma_1) = .1$, $\gamma_2 = 83.33$ **with** $p(\gamma_2) = .5$ **and** $\gamma_3 = 333.33$ **with** $p(\gamma_3) = .4$ **Find the capacity versus outage for this channel, and find the average rate correctly received for outage prob.** $p_{out} < .1$, $p_{out} = .1$ and $p_{out} = .6$.

Solution: For time-varying channels with discrete SNR values the capacity versus outage is a staircase function. Specifically, for $p_{out} < .1$ we must decode correctly in all channel states. The minimum received SNR for p_{out} in this range of values is that of the weakest channel: $\gamma_{min} = \gamma_1$, and the corresponding capacity is $C = B\log_2(1+\gamma_{min}) = 30000\log_2(1.833) = 26.23$ Kbps. For $.1 \leq p_{out} < .6$ we can decode incorrectly when the channel is in the weakest state only. Then $\gamma_{min} = \gamma_2$ and the corresponding capacity is $C = B\log_2(1+\gamma_{min}) = 30000\log_2(84.33) = 191.94$ Kbps. For $.6 \leq p_{out} < 1$ we can decode incorrectly if the channel has received SNR γ_1 or γ_2. Then $\gamma_{min} = \gamma_3$ and the corresponding capacity is $C = B\log_2(1+\gamma_{min}) = 30000\log_2(334.33) = 251.55$ Kbps. Thus, capacity versus outage has $C = 26.23$ Kbps for $p_{out} < .1$, $C = 191.94$ Kbps for $.1 \leq p_{out} < .6$, and $C = 251.55$ Kbps for $.6 \leq p_{out} < 1$.

Capacity of Flat-Fading Channels

For $p_{out} < .1$ data transmitted at rates close to capacity C = 26.23 Kbps are always correctly received since the channel can always support this data rate. For $p_{out} = .1$ we transmit at rates close to C = 191.94 Kbps, but we can only correctly decode these data when the channel SNR is γ_2 or γ_3 so the rate correctly received is (1-0.1)191940 = 172.75 Kbps. For $p_{out} = .6$ we transmit at rates close to C = 251.55 Kbps but we can only correctly decode these data when the channel SNR is γ_3, so the rate correctly received is (1-0.6)251550 = 125.78 Kbps. It is likely that a good engineering design for this channel would send data at a rate close to 191.94 Kbps, since it would only be received incorrectly at most 10% of this time and the data rate would be almost an order of magnitude higher than sending at a rate commensurate with the worst-case channel capacity. However, 10% retransmission probability is too high for some applications, in which case the system would be designed for the 26.23 Kbps data rate with no retransmissions.

Capacity of Flat-Fading Channels

- **When both the transmitter and receiver have CSI, the transmitter can adapt its transmission strategy relative to this CSI, as shown in Figure 4.3.**

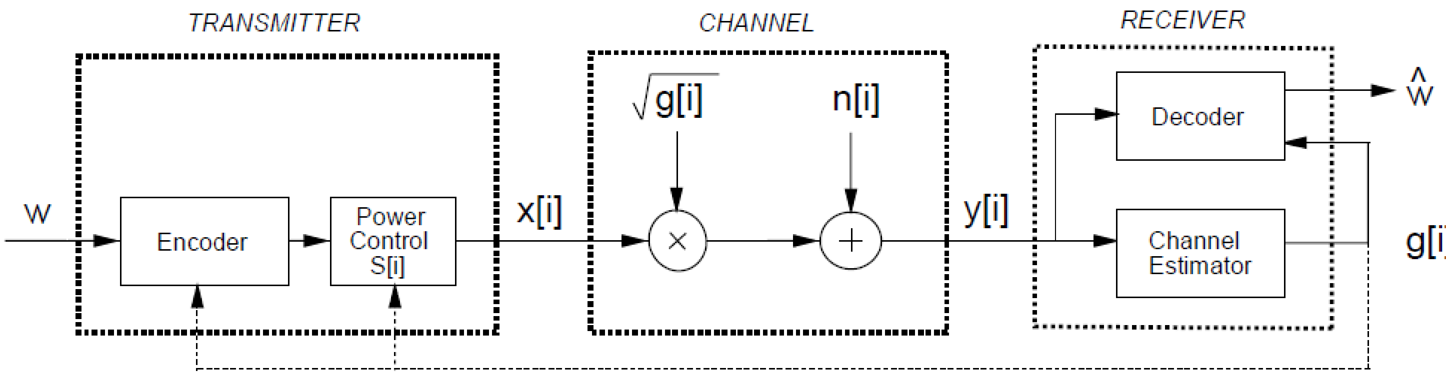


Figure 4.3: System Model with Transmitter and Receiver CSI.

- **In this case, there is no notion of capacity versus outage where the transmitter sends bits that cannot be decoded, since the transmitter knows the channel and thus will not send bits unless they can be decoded correctly.**

Capacity of Flat-Fading Channels

- **Now consider the Shannon capacity when the channel power gain *g*[*i*] is known to both the transmitter and receiver at time *i*.**

- **Let *s*[*i*] be a stationary and ergodic stochastic process representing the channel state, which takes values on a finite set *S* of discrete memoryless channels.**

- **Let C_s denote the capacity of a particular channel $s \in \mathcal{S}$, and *p*(*s*) denote the probability, or fraction of time, that the channel is in state *s*. The capacity of this time-varying channel is then given by**

$$C = \sum_{s \in \mathcal{S}} C_s p(s). \qquad (4.6)$$

- **The capacity of an AWGN channel with average received SNR *γ* is**

$$C_\gamma = B \log_2(1+\gamma)$$

Capacity of Flat-Fading Channels

- **From (4.6) the capacity of the fading channel with transmitter and receiver side information is**

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

- **Let us now allow the transmit power *P*(*γ*) to vary with *γ*, subject to an average power constraint $\overline{P}$:**

$$\int_0^\infty P(\gamma)p(\gamma)d\gamma \le \overline{P}.$$

- **Define the fading channel capacity with average power constraint as**

$$C = \max_{P(\gamma):\int P(\gamma)p(\gamma)d\gamma = \overline{P}} \int_0^\infty B\log_2\left(1+\frac{P(\gamma)\gamma}{\overline{P}}\right)p(\gamma)d\gamma. \qquad \textbf{(4.9)}$$

(Can this capacity be achieved?)

Capacity of Flat-Fading Channels

- **Figure 4.4 shows the main idea of how to achieve the capacity in (4.9)**

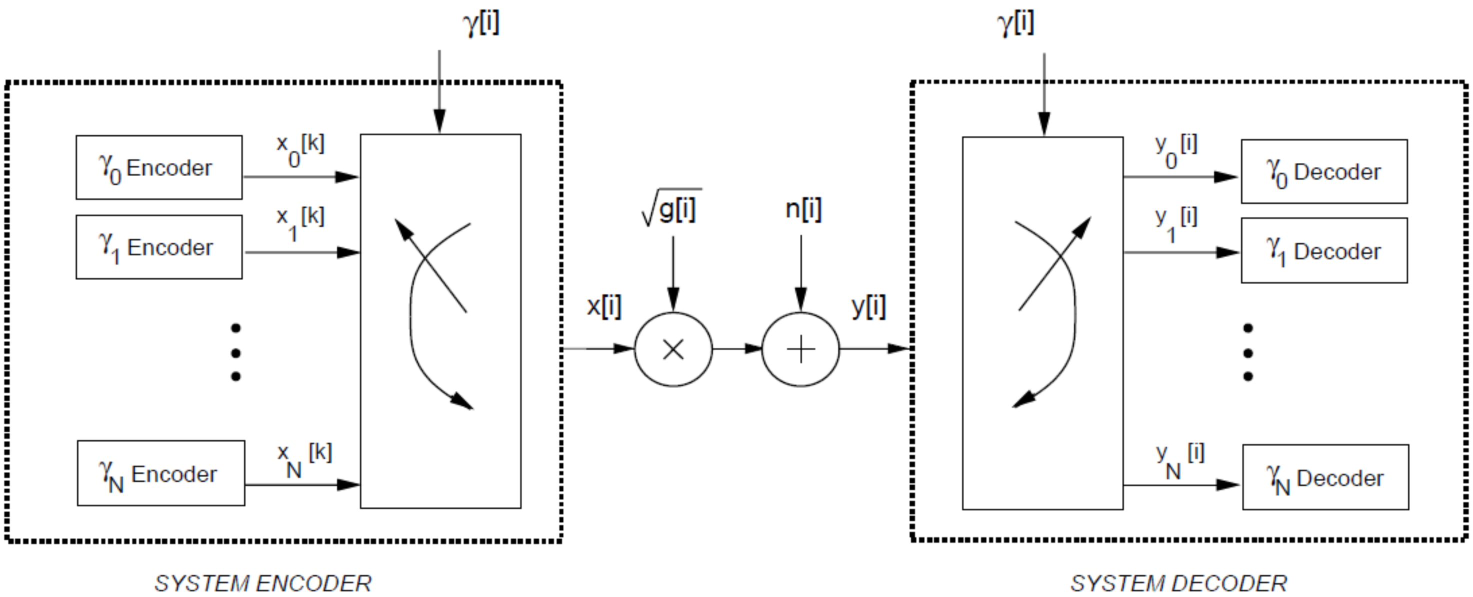


Figure 4.4: Multiplexed Coding and Decoding.

- **To find the optimal power allocation *P*(*γ*), we form the Lagrangian**

$$J(P(\gamma)) = \int_0^\infty B \log_2\left(1 + \frac{\gamma P(\gamma)}{\overline{P}}\right) p(\gamma) d\gamma - \lambda \int_0^\infty P(\gamma) p(\gamma) d\gamma.$$

Capacity of Flat-Fading Channels

- **Next we differentiate the Lagrangian and set the derivative equal to zero:**

$$\frac{\partial J(P(\gamma))}{\partial P(\gamma)} = \left[\left(\frac{B/\ln(2)}{1+\gamma P(\gamma)/\overline{P}}\right)\frac{\gamma}{\overline{P}} - \lambda\right] p(\gamma) = 0.$$

- **Solving for *P*(*γ*) with the constraint that *P*(*γ*) > 0 yields the optimal power adaptation that maximizes (4.9) as**

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

for some "cutoff" value γ_0 **:**

- **If *γ*[*i*] is below this cutoff then no data is transmitted over the *i*th time interval, so the channel is only used at time *i* if** $\gamma_0 \leq \gamma[i] < \infty$**.**
- **Substituting (4.12) into (4.9) then yields the capacity formula:**

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

Capacity of Flat-Fading Channels

- **The multiplexing nature of the capacity-achieving coding strategy indicates that (4.13) is achieved with a time varying data rate, where the rate corresponding to instantaneous SNR γ is** $B\log_2(\gamma/\gamma_0)$ **.**
- **Note that the optimal power allocation policy (4.12) only depends on the fading distribution $p(\gamma)$ through the cutoff value** γ_0 **. This cutoff value is found from the power constraint.**
- **By rearranging the power constraint and replacing the inequality with equality (since using the maximum available power will always be optimal) yields the power constraint**

$$\int_0^\infty \frac{P(\gamma)}{\overline{P}} p(\gamma) d\gamma = 1.$$

- **By using the optimal power allocation (4.12), we have**

$$\int_{\gamma_0}^\infty \left(\frac{1}{\gamma_0} - \frac{1}{\gamma}\right) p(\gamma) d\gamma = 1.$$

Capacity of Flat-Fading Channels

- **Note that this expression only depends on the distribution $p(\gamma)$. The value for γ_0 cannot be solved for in closed form for typical continuous pdfs $p(\gamma)$ and thus must be found numerically.**
- **Since γ is time-varying, the maximizing power adaptation policy of (4.12) is a "water-filling" formula in time, as illustrated in Figure 4.5.**

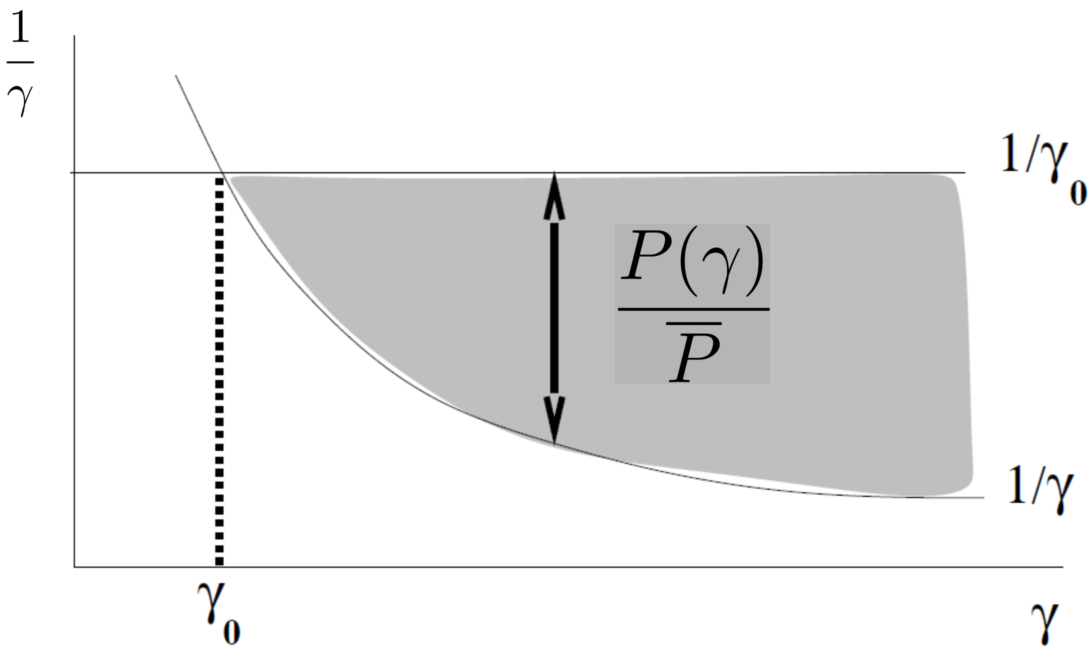


Figure 4.5: Optimal Power Allocation: Water-Filling.

- **The water-filling terminology refers to the fact that the line $1/\gamma$ sketches out the bottom of a bowl, and power is poured into the bowl to a constant water level of $1/\gamma_0$.**

Capacity of Flat-Fading Channels

- **Example 4.4: Assume the same channel as in the previous example, with a bandwidth of 30 KHz and three possible received SNRs: $\gamma_1 = .8333$ with $p(\gamma_1) = .1$, $\gamma_2 = 83.33$ with $p(\gamma_2) = .5$, and $\gamma_3 = 333.33$ with $p(\gamma_3) = .4$. Find the ergodic capacity of this channel assuming both transmitter and receiver have instantaneous CSI.**

Solution: **We know the optimal power allocation is water-filling, and we need to find the cutoff value γ_0 that satisfies the discrete version of (4.15) given by**

$$\sum_{\gamma_i \geq \gamma_0} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma_i} \right) p(\gamma_i) = 1.$$

We first assume that all channel states are used to obtain γ_0, i.e. assume $\gamma_0 \leq \min_i \gamma_i$, and see if the resulting cutoff value is below that of the weakest channel. If not then we have an inconsistency, and must redo the calculation assuming at least one of the channel states is not used. Applying (4.17) to our channel model yields

Capacity of Flat-Fading Channels

$$\sum_{i=1}^{3}\frac{p(\gamma_i)}{\gamma_0}-\sum_{i=1}^{3}\frac{p(\gamma_i)}{\gamma_i}=1 \Rightarrow \frac{1}{\gamma_0}=1+\sum_{i=1}^{3}\frac{p(\gamma_i)}{\gamma_i}=1+\left(\frac{.1}{.8333}+\frac{.5}{83.33}+\frac{.4}{333.33}\right)=1.13$$

Solving for γ_0 yields $\gamma_0 = 1/1.13 = .89 > .8333 = \gamma_1$. Since this value of γ_0 is greater than the SNR in the weakest channel, it is inconsistent as the channel should only be used for SNRs above the cutoff value. Therefore, we now redo the calculation assuming that the weakest state is not used. Then (4.17) becomes

$$\sum_{i=2}^{3}\frac{p(\gamma_i)}{\gamma_0}-\sum_{i=2}^{3}\frac{p(\gamma_i)}{\gamma_i}=1 \Rightarrow \frac{.9}{\gamma_0}=1+\sum_{i=2}^{3}\frac{p(\gamma_i)}{\gamma_i}=1+\left(\frac{.5}{83.33}+\frac{.4}{333.33}\right)=1.0072$$

Solving for γ_0 yields $\gamma_0 = .89$. So by assuming the weakest channel with SNR γ_1 is not used, we obtain a consistent value for γ_0 with $\gamma_1 < \gamma_0 \le \gamma_2$. The capacity of the channel then becomes

$$C=\sum_{i=2}^{3} B\log_2(\gamma_i/\gamma_0)p(\gamma_i)=30000(.5\log_2(83.33/.89)+.4\log_2(333.33/.89))=200.82\ \text{Kbps}.$$

Capacity of Flat-Fading Channels

Comparing with the results of the previous example we see that this rate is only slightly higher than for the case of receiver CSI only, and is still significantly below that of an AWGN channel with the same average SNR. That is because the average SNR for this channel is relatively high: for low SNR channels capacity in flat-fading can exceed that of the AWGN channel with the same SNR by taking advantage of the rare times when the channel is in a very good state.

- **Zero-Outage Capacity and Channel Inversion: now consider a suboptimal transmitter adaptation scheme where the transmitter uses the CSI to maintain a constant received power, i.e., it inverts the channel fading.**

- **The channel then appears to the encoder and decoder as a time-invariant AWGN channel. This power adaptation, called channel inversion, is given by** $P(\gamma)/\overline{P} = \sigma/\gamma,$ **where σ equals the constant received SNR that can be maintained with the transmit power constraint (4.8).**

Capacity of Flat-Fading Channels

- **The constant σ thus satisfies** $\int \frac{\sigma}{\gamma} p(\gamma) d\gamma = 1$, so $\sigma = 1/\mathbf{E}[1/\gamma]$.

- **Fading channel capacity with channel inversion is just the capacity of an AWGN channel with SNR σ:**

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

- **The capacity-achieving transmission strategy for this capacity uses a fixed-rate encoder and decoder designed for an AWGN channel with SNR σ.**
- **This has the advantage of maintaining a fixed data rate over the channel regardless of channel conditions.**
- **For this reason the channel capacity given in (4.18) is called zero-outage capacity.**
- **Zero-outage capacity can exhibit a large data rate reduction relative to Shannon capacity in extreme fading environments. For example, in Rayleigh fading E[1/γ] is infinite, and thus the zero-outage capacity given by (4.18) is zero.**

Capacity of Flat-Fading Channels

- **Example 4.5: Assume the same channel as in the previous example, with a bandwidth of 30 KHz and three possible received SNRs:** $\gamma_1 = .8333$ **with** $p(\gamma_1) = .1$, $\gamma_2 = 83.33$ **with** $p(\gamma_2) = .5$, and $\gamma_3 = 333.33$ **with** $p(\gamma_3) = .4.$**. Assuming transmitter and receiver CSI, find the zero-outage capacity of this channel.**

Solution: The zero-outage capacity is $C = B\log_2[1+\sigma]$, where $\sigma = 1/\mathbf{E}[1/\gamma]$. Since

$$\mathbf{E}[1/\gamma] = \frac{.1}{.8333} + \frac{.5}{83.33} + \frac{.4}{333.33} = .1272,$$

we have $C = 30000\log_2(1+1/.1272) = 9443$ Kbps. Note that this is less than half of the Shannon capacity with optimal water-filling adaptation.

- **The outage capacity is defined as the maximum data rate that can be maintained in all nonoutage channel states times the probability of nonoutage.**

Capacity of Flat-Fading Channels

- **Outage capacity is achieved with a truncated channel inversion policy for power adaptation that only compensates for fading above a certain cutoff fade depth** γ_0**:**

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

- **Since the channel is only used when** $\gamma \geq \gamma_0$**, the power constraint (4.8) yields** $\sigma = 1/\mathbf{E}_{\gamma_0}[1/\gamma]$**, where**

$$\mathbf{E}_{\gamma_0}[1/\gamma] \triangleq \int_{\gamma_0}^{\infty} \frac{1}{\gamma} p(\gamma) d\gamma.$$

- **The outage capacity associated with a given outage probability** p_{out} **and corresponding cutoff** γ_0 **is given by**

$$C(p_{out}) = B \log_2 \left(1 + \frac{1}{\mathbf{E}_{\gamma_0}[1/\gamma]} \right) p(\gamma \geq \gamma_0).$$

Capacity of Flat-Fading Channels

- **We can also obtain the maximum outage capacity by maximizing outage capacity over all possible γ_0:**

$$C = \max_{\gamma_0} B \log_2\left(1 + \frac{1}{\mathbf{E}_{\gamma_0}[1/\gamma]}\right) p(\gamma \geq \gamma_0).$$

- **This maximum outage capacity will still be less than Shannon capacity (4.13) since truncated channel inversion is a suboptimal transmission strategy.**

- **Example 4.6: Assume the same channel as in the previous example, with a bandwidth of 30 KHz and three possible received SNRs: $\gamma_1 = .8333$ with $p(\gamma_1) = .1$, $\gamma_2 = 83.33$ with $p(\gamma_2) = .5$, and $\gamma_3 = 333.33$ with $p(\gamma_3) = .4$. Find the outage capacity of this channel and associated outage probabilities for cutoff values $\gamma_0 = .84$ and $\gamma_0 = 83.4$. Which of these cutoff values yields a larger outage capacity?**

Capacity of Frequency-Selective Fading Channels

Solution: For $\gamma_0 = .84$ we use the channel when the SNR is γ_2 or γ_3, so $\mathbf{E}_{\gamma_0}[1/\gamma] = \sum_{i=2}^{3} p(\gamma_i)/\gamma_i = .5/83.33 + .4/333.33 = .0072$. The outage capacity is $C = B\log_2(1 + 1/\mathbf{E}_{\gamma_0}[1/\gamma])p(\gamma \geq \gamma_0) = 30000\log_2(1 + 138.88) * .9 = 192.457$. For $\gamma_0 = 83.34$ we use the channel when the SNR is γ_3 only, so $\mathbf{E}_{\gamma_0}[1/\gamma] = p(\gamma_3)/\gamma_3 = .4/333.33 = .0012$. The capacity is $C = B\log_2(1 + 1/\mathbf{E}_{\gamma_0}[1/\gamma])p(\gamma \geq \gamma_0) = 30000\log_2(1 + 833.33) * .4 = 116.45$ Kbps. The outage capacity is larger when the channel is used for SNRs γ_2 and γ_3. Even though the SNR γ_3 is significantly larger than γ_2, the fact that this SNR only occurs 40% of the time makes it inefficient to only use the channel in this best state.

- **Consider a time-invariant channel with frequency response *H*(*f*), as shown in Figure 4.9. Assume a total transmit power constraint *P*.**

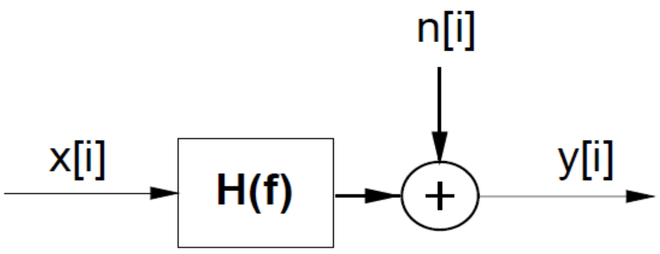


Figure 4.9: Time-Invariant Frequency-Selective Fading Channel.

Capacity of Frequency-Selective Fading Channels

- **First assume that *H*(*f*) is block-fading, so that frequency is divided into subchannels of bandwidth *B*, where $H(f) = H_j$ is constant over each block, as shown in Figure 4.10.**

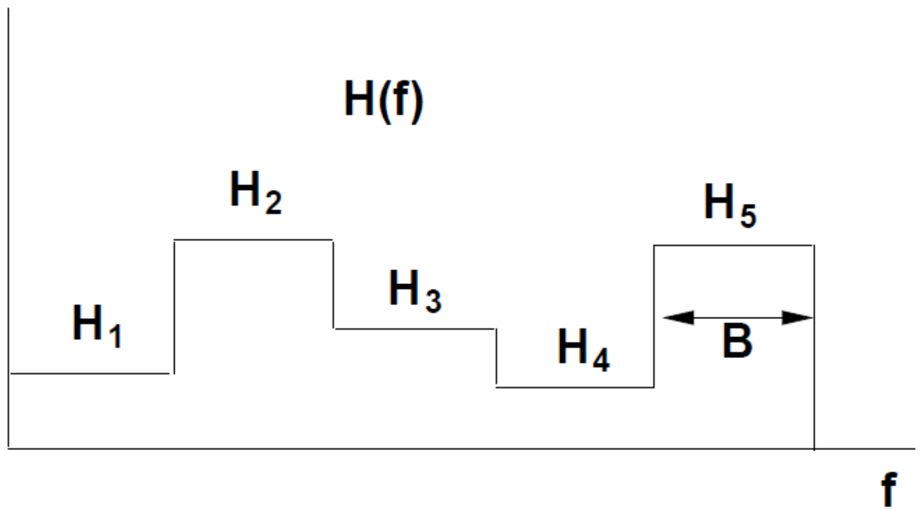


Figure 4.10: Block Frequency-Selective Fading

- **The frequency-selective fading channel thus consists of a set of AWGN channels in parallel with SNR $|H_j|^2 P_j/(N_0 B)$ on the *j*th channel, where P_j is the power allocated to the *j*th channel in this parallel set, subject to the power constraint $\sum_j P_j \leq P$.**

Capacity of Frequency-Selective Fading Channels

- **The capacity of this parallel set of channels is the sum of rates associated with each channel with power optimally allocated over all channels**

$$C = \sum_{\max P_j : \sum_j P_j \le P} B \log_2 \left(1 + \frac{|H_j|^2 P_j}{N_0 B} \right).$$

- **This is similar to the capacity and optimal power allocation for a flat-fading channel, with power and rate changing over frequency in a deterministic way rather than over time in a probabilistic way.**

- **The optimal power allocation is found via the same Lagrangian technique used in the flat-fading case, which leads to the water-filling power allocation**

$$\frac{P_j}{P} = \begin{cases} \frac{1}{\gamma_0} - \frac{1}{\gamma_j} & \gamma_j \ge \gamma_0 \\ 0 & \gamma_j < \gamma_0 \end{cases}$$

for some cutoff value γ_0, where $\gamma_j = |H_j|^2 P/(N_0 B)$ is the SNR

Capacity of Frequency-Selective Fading Channels

associated with the *j*th channel assuming it is allocated the entire power budget.

- This optimal power allocation is illustrated in Figure 4.11.

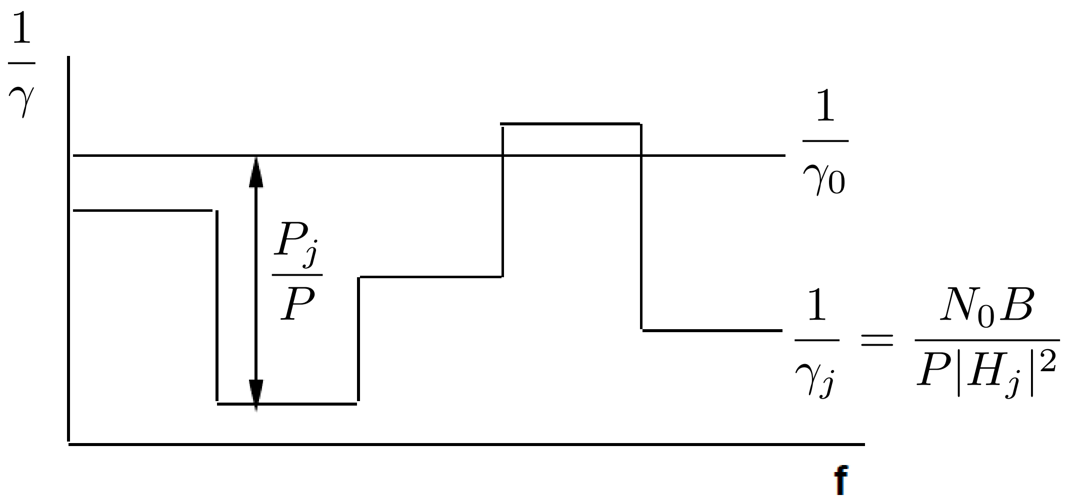


Figure 4.11: Water-Filling in Block Frequency-Selective Fading

- The cutoff value is obtained by substituting the power adaptation formula into the power constraint, so γ_0 must satisfy

Capacity of Frequency-Selective Fading Channels

$$\sum_j \left(\frac{1}{\gamma_0} - \frac{1}{\gamma_j}\right) = 1.$$

- **The capacity then becomes**

$$C = \sum_{j:\gamma_j \geq \gamma_0} B \log_2(\gamma_j/\gamma_0).$$

- **This capacity is achieved by sending at different rates and powers over each subchannel.**

- **When *H*(*f*) is continuous the capacity under power constraint *P* is similar to the case of the block-fading channel, with some mathematical intricacies needed to show that the channel capacity is given by**

$$C = \max_{P(f):\int P(f)df \leq P} \int \log_2\left(1 + \frac{|H(f)|^2 P(f)}{N_0}\right) df. \quad \textbf{(4.27)}$$

Capacity of Frequency-Selective Fading Channels

- The power allocation over frequency, *P*(*f*), that maximizes (4.27) is found via the Lagrangian technique. The resulting optimal power allocation is water-filling over frequency:

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

- This results in channel capacity

$$C = \int_{f:\gamma(f)\geq\gamma_0} \log_2(\gamma(f)/\gamma_0)df.$$

- **Example 4.7**: Consider a time-invariant frequency-selective block fading channel consisting of three subchannels of bandwidth *B* = 1 MHz. The frequency response associated with each channel is $H_1 = 1$, $H_2 = 2$ and $H_3 = 3$. The transmit power constraint is *P* = 10 mW and the noise PSD is $N_0 = 10^{-9}$ W/Hz. Find the Shannon capacity of this channel and the optimal power allocation that achieves this capacity.

Capacity of Frequency-Selective Fading Channels

Solution: We first first find $\gamma_j = |H_j|^2 P/(N_b)$ for each subchannel, yielding $\gamma_1 = 10$, $\gamma_2 = 40$ and $\gamma_3 = 90$. The cutoff γ_0 must satisfy (4.25). Assuming all subchannels are allocated power, this yields

$$\frac{3}{\gamma_0} = 1 + \sum_j \frac{1}{\gamma_j} = 1.14 \Rightarrow \gamma_0 = 2.64 < \gamma_j \ \forall j.$$

Since the cutoff γ_0 is less than γ_j for all j, our assumption that all subchannels are allocated power is consistent, so this is the correct cutoff value. The corresponding capacity is $C = \sum_{j=1}^{3} B \log_2(\gamma_j/\gamma_0) = 1000000(\log_2(10/2.64) + \log_2(40/2.64) + \log_2(90/2.64)) = 10.93$ Mbps.

Appendix:
Channel Capacity for a Gaussian Channel

Introduction to Gaussian Channel

- **The most important continuous alphabet channel is the Gaussian channel depicted in the following figure. This is a time-discrete channel with output Y_i at time *i*, where Y_i is the sum of the input X_i and the noise Z_i .**
- **The noise Z_i is drawn i.i.d. from a Gaussian distribution with variance N. Thus,**

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

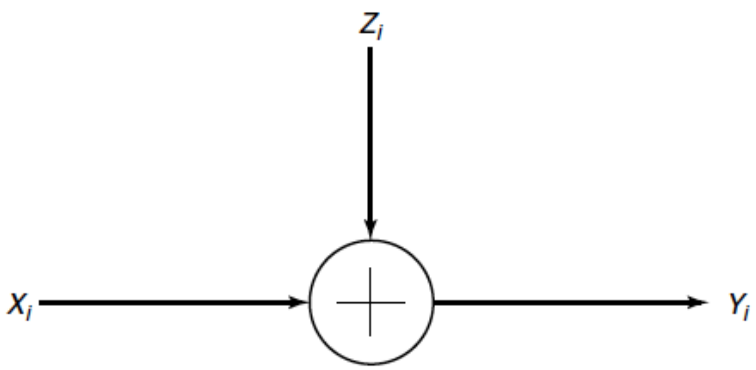


Figure: A Gaussian Channel

- **The noise Z_i is assumed to be independent of the signal X_i .**
- **If the noise variance is zero or the input is unconstrained, the capacity of the channel is infinite.**

Introduction to Gaussian Channel

- Assume an average power constraint. For any codeword $(x_1, x_2, \ldots, x_n)$ transmitted over the channel, we require that

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \le P.$$

- Assume that we want to send 1 bit over the channel in one use of the channel.

- Given the power constraint, the best that we can do is to send one of two levels, $-\sqrt{P}$ or $\sqrt{P}$. The receiver looks at the corresponding Y received and tries to decide which of the two levels was sent.

- Assuming that both levels are equally likely (this would be the case if we wish to send exactly 1 bit of information), the optimum decoding rule is to decide that $\sqrt{P}$ was sent if $Y>0$ and decide $-\sqrt{P}$ was sent if $Y<0$.

Introduction to Gaussian Channel

- **The probability of error with such a decoding scheme is**

$$
\begin{aligned}
P_e &= \frac{1}{2}\Pr(Y < 0|X = +\sqrt{P}) + \frac{1}{2}\Pr(Y > 0|X = -\sqrt{P}) \\
&= \frac{1}{2}\Pr(Z < -\sqrt{P}|X = +\sqrt{P}) + \frac{1}{2}\Pr(Z > \sqrt{P}|X = -\sqrt{P}) \\
&= \Pr(Z > \sqrt{P}) \\
&= 1 - \Phi\left(\sqrt{P/N}\right),
\end{aligned}
$$

where $\Phi(x)$ is the cumulative normal function

$$
\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}\, dt.
$$

- **Using such a scheme, we have converted the Gaussian channel into a discrete binary symmetric channel with crossover probability P_e .**

Differential Entropy

- **Let X now be a continuous r.v. with cumulative distribution**

$$F(x) = \Pr(X \leq x)$$

and $f(x) = \frac{d}{dx}F(x)$ **is the density function.**

- **Let** $S = \{x : f(x) > 0\}$ **be the support set. Then**

- **Definition of Differential Entropy (*h*(*X*)) :**

$$h(X) = -\int_S f(x) \log f(x) \, dx$$

Since we integrate over only the support set, no worries about log 0.

Differential Entropy of Gaussian Distribution

- **If we have:**

$$X \sim N(0,\sigma^2) \qquad \Leftrightarrow \qquad f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-\frac{1}{2}x^2/\sigma^2}$$

- **Let's compute this in nats.**

$$h(X) = -\int f \ln f = -\int f(x)\left[-\frac{x^2}{2\sigma^2} - \ln\sqrt{2\pi\sigma^2}\right]dx =$$

$$\frac{EX^2}{2\sigma^2} + \frac{1}{2}\ln(2\pi\sigma^2) = \frac{1}{2} + \frac{1}{2}\ln(2\pi\sigma^2)$$

$$= \frac{1}{2}\ln e + \frac{1}{2}\ln(2\pi\sigma^2) = \frac{1}{2}\ln(2\pi e\sigma^2)\text{nats}$$

$$= \frac{1}{2}\log_2 2\pi e\sigma^2, \ \text{bits}$$

- **Note: only a function of the variance σ^2, not the mean. Why?**
- **So entropy of a Gaussian is monotonically related to the variance.**

Gaussian Distribution Maximizes Differential Entropy

- **Definition : The relative entropy (or Kullback–Leibler distance) $D(f||g)$ between two densities *f* and *g* is defined by**

$$D(g||\phi) = \int g(x) \log \left(\frac{g(x)}{\phi(x)} \right) dx \geq 0$$

- **Theorem : Let random variable *X* have zero mean and variance σ^2. Then $h(X) \leq \frac{1}{2} \log_2(2\pi e \sigma^2)$, with equality iff $X \sim \mathcal{N}(0, \sigma^2)$.**

Proof: Let $g(x)$ be any density satisfying $\int x^2 g(x) dx = \sigma^2$. Let $\phi(x)$ be the density of a Gaussian random variable with zero mean and variable σ^2. Note that $\log \phi(x)$ is a quadratic form and it is $-\frac{x^2}{2\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)$. Then

$$\begin{aligned} D(g||\phi) &= \int g(x) \log \left(\frac{g(x)}{\phi(x)} \right) dx \\ &= -h(g) - \int g \log(\phi(x)) dx = -h(g) - \int \phi(x) \log(\phi(x)) dx \quad \textbf{(Why?)} \\ &= -h(g) + h(\phi) \geq 0 \end{aligned}$$

- **Therefore, the Gaussian distribution maximizes the entropy overall distributions with the same variance.**

Definitions of Gaussian Channels

- **Definition : The information capacity of the Gaussian channel with power constraint *P* is**

$$C = \max_{f(x): E X^2 \le P} I(X;Y).$$

- **We can calculate the information capacity as follows: Expanding $I(X;Y)$, we have**

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

since *Z* is independent of *X*. Now $h(Z) = \frac{1}{2}\log 2\pi e N$. Also,

$$EY^2 = E(X+Z)^2 = EX^2 + 2EXEZ + EZ^2 = P + N,$$

since *X* and *Z* are independent and *E[Z]*=0.

Definitions of Gaussian Channels

- **Given $EY^2 = P + N$, the entropy of Y is bounded by $\frac{1}{2}\log 2\pi e(P+N)$ (the Gaussian distribution maximizes the entropy for a given variance).**
- **Applying this result to bound the mutual information, we obtain**

$$\begin{aligned} I(X;Y) &= h(Y) - h(Z) \\ &\le \frac{1}{2}\log 2\pi e(P+N) - \frac{1}{2}\log 2\pi e N \\ &= \frac{1}{2}\log\left(1+\frac{P}{N}\right). \end{aligned}$$

- **Hence, the information capacity of the Gaussian channel is**

$$C = \max_{EX^2 \le P} I(X;Y) = \frac{1}{2}\log\left(1+\frac{P}{N}\right),$$

and the maximum is attained when $X \sim \mathcal{N}(0, P)$.