UNIT I

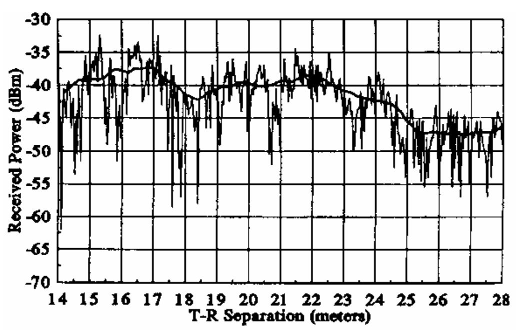
UNIT I WIRELESS CHANNELS

Large scale path loss – Path loss models: Free Space and Two-Ray models -Link Budget design – Small scale fading- Parameters of mobile multipath channels – Time dispersion parameters- Coherence bandwidth – Doppler spread & Coherence time, Fading due to Multipath time delay spread – flat fading – frequency selective fading – Fading due to Doppler spread – fast fading – slow fading.

Large Scale & Small Scale Propagation Models

Propagation models that predict the mean signal strength for an arbitrary transmitter-receiver (T-R) separation distance are useful in estimating the radio coverage area of a transmitter and are called **large-scale propagation models.**

On the other hand, propagation models that characterize the rapid fluctuations of the received signal strength over very short travel distances (a few wavelengths) or short time durations (on the order of seconds) are called **small-scale or fading models.**

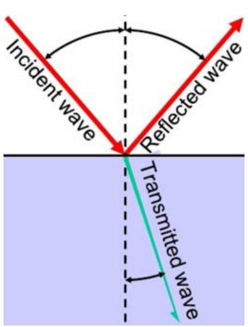


The physical mechanisms that govern radio propagation are complex and diverse, but generally attributed to the following three factors

1. Reflection
2. Diffraction
3. Scattering

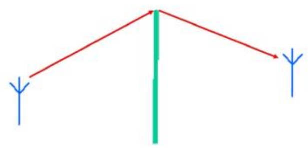
Reflection

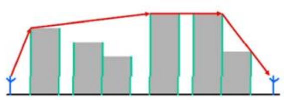
Reflection occurs when an electromagnetic wave falls on an object, which has very large dimensions as compared to the wavelength of the propagating wave. Example: reflections from earth and buildings. These reflections may interfere with the original signal constructively or destructively.



Diffraction

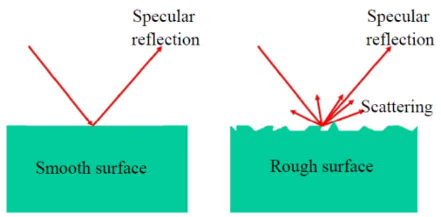
Occurs when the radio path between sender and receiver is obstructed by an impenetrable body and by a surface with sharp irregularities (edges). Explains how radio signals can travel urban and rural environments without a line-of-sight path. Single or multiple edges makes it possible to go behind corners. And it is less pronounced when the wavelength is small compared to objects.





Scattering

Occurs when the radio channel contains objects whose sizes are on the order of the wavelength or less of the propagating wave and also when the number of obstacles are quite large. They are produced by small objects, rough surfaces and other irregularities on the channel. Follows same principles with diffraction. Causes the transmitter energy to be radiated in many directions. Lamp posts and street signs may cause scattering.



Brewster's Angle

When no reflection occurs in the medium of origin, the incident angle would be such that the reflection coefficient is equal to zero. This angle is the Brewster's angle.

$$\sin(\theta_B) = \sqrt{\frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2}} \qquad \sin(\theta_B) = \frac{\sqrt{\varepsilon_r - 1}}{\sqrt{\varepsilon_r^2 - 1}}$$

<u>Free Space Model (Friis Free Space Equation)</u>

The free space model is used to predict the received signal strength when the transmitter and receiver have a clear, unobstructed line of sight between them.

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L}$$

Example 4.1

Find the Fraunhofer distance for an antenna with maximum dimension of 1 m and operating frequency of 900 MHz. If antennas have unity gain, calculate the path loss.

Solution

Operating frequency, $f = 900$ MHz

$\lambda = c/f = 3 \times 10^8$ m/s$/900 \times 10^6$ Hz $= 0.33$ m

Fraunhofer distance, $d_f = 2D^2 / \lambda = 2(1)^2 / 0.33 = 6$ m

Path loss $P_L(\text{dB}) = -10 \log [(\lambda^2)/(4\pi)^2 d^2] = -10 \log [(0.33)^2 / (4 \times 3.14)^2 \times 36] = 47\text{dB}$

Example 4.2

If a transmitter produces 50 W of power, express the transmit power in units of (a) dBm, and (b) dBW. If 50 W is applied to a unity gain antenna with a 900 MHz carrier frequency, find the received power in dBm at a free space distance of 100 m from the antenna. What is $P_r(10\text{ km})$? Assume unity gain for the receiver antenna.

Solution

Given:

Transmitter power, $P_t = 50$ W

Carrier frequency, $f_c = 900$ MHz

Using Equation (4.9),

(a) Transmitter power,

$$P_t(\text{dBm}) = 10\log[P_t(\text{mW})/(1\text{ mW})]$$
$$= 10\log[50 \times 10^3] = 47.0\text{ dBm}.$$

(b) Transmitter power,

$$P_t(\text{dBW}) = 10\log[P_t(\text{W})/(1\text{ W})]$$
$$= 10\log[50] = 17.0\text{ dBW}.$$

The received power can be determined using Equation (4.1)

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} = \frac{50(1)(1)(1/3)^2}{(4\pi)^2 (100)^2 (1)} = (3.5 \times 10^{-6})\text{ W} = 3.5 \times 10^{-3}\text{ mW}$$

$$P_r(\text{dBm}) = 10\log P_r(\text{mW}) = 10\log(3.5 \times 10^{-3}\text{ mW}) = -24.5\text{ dBm}.$$

The received power at 10 km can be expressed in terms of dBm using Equation (4.9), where $d_0 = 100$ m and $d = 10$ km

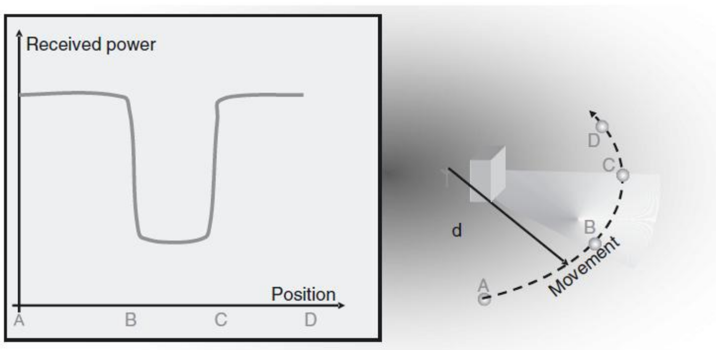
$$P_r(10\text{ km}) = P_r(100) + 20\log\left[\frac{100}{10000}\right] = -24.5\text{ dBm} - 40\text{ dB}$$
$$= -64.5\text{ dBm}.$$

Wired Vs Wireless Communication

Wired communication	Wireless communication
The communication takes place over a more or Less stable medium like copper wires or optical fibers. The properties of the medium are well defined and time-invariant.	Due to user mobility as well as multipath propagation, the transmission medium varies strongly with time.
The range over which communications can be performed without repeater stations is mostly limited by attenuation by the medium (and thus noise); for optical fibers, the distortion of transmitted pulses can also limit the speed of data transmission.	The range that can be covered is limited both by The transmission medium (attenuation, fading, and signal distortion) and by the requirements of spectral efficiency (cell size).
Increasing the transmission capacity can be achieved by using a different frequency on an existing cable, and/or by stringing new cables.	Increasing the transmit capacity must be achieved by more sophisticated transceiver concepts and smaller cell sizes (in cellular systems), as the amount of available spectrum is limited.
Interference and crosstalk from other users either do not happen or the properties of the interference are stationary.	Interference and crosstalk from other users are inherent in the principle of cellular communications. Due to the mobility of the users, they also are time-variant.
The delay in the transmission process is also constant, determined by the length of the cable and the group delay of possible repeater amplifiers.	The delay of the transmission depends partly on The distance between base station and Mobile Station (MS), and is thus time-variant.
The *Bit Error Rate* (BER) decreases strongly (approximately exponentially) with increasing *Signal-to-Noise Ratio* (SNR). This means that a relatively small increase in transmit power can greatly decrease the error rate.	For simple systems, the average BER decreases only slowly (linearly) with increasing average SNR. Increasing the transmit power usually does not lead to a significant reduction in BER. However, more sophisticated signal processing helps.
Due to the well-behaved transmission medium, the quality of wired transmission is generally high.	Due to the difficult medium, transmission quality is generally low unless special measures are used.
Jamming and interception of dedicated links With wired transmission is almost impossible without transmitted signal.	Jamming a wireless link is straightforward, unless special measures are taken. Interception of the air signal is possible. Encryption is therefore necessary to prevent unauthorized use of the channel

Fading

A simple RX cannot distinguish between the different *Multi Path Components* (MPCs); it just adds them up, so that they interfere with each other. The interference between them can be constructive or destructive, depending on the phases of the MPCs. The phases, in turn, depend mostly on the run length of the MPC, and thus on the position of the Mobile Station (MS) and the IOs. For this reason, the interference, and thus the amplitude of the total signal, changes with time if either TX, RX, or IOs is moving. This effect – namely, the changing of the total signal amplitude due to interference of the different MPCs – is called *small-scale fading*.



The principle of shadowing.

Imagine, e.g., the MS in Figure that at first (at position A) has LOS to the Base Station (BS). As the MS moves behind the high-rise building (at position B), the amplitude of the component that propagates along the direct connection (LOS) between BS and MS greatly decreases. This is due to the fact that the MS is now in the radio shadow of the high-rise building, and any wave going through or around that building is greatly attenuated – an effect called *shadowing*. Of course, shadowing can occur not only for an LOS component but also for *any* MPC. Note also that obstacles do not throw "sharp" shadows: the transition from the "light" (i.e., LOS) zone to the "dark" (shadowed) zone is gradual. The MS has to move over large distances (from a few meters up to several hundreds of meters) to move from the light to the dark zone. For this reason, shadowing gives rise to *large-scale fading*.

Link Budget

A link budget is the clearest and most intuitive way of computing the required TX power. It tabulates all equations that connect the TX power to the received SNR. As most factors influencing the SNR enter in a multiplicative way, it is convenient to write all the equations in a logarithmic form – specifically, in dB. It has to be noted, however, that the link budget gives only an approximation (often a worst case estimate) for the total SNR, because some interactions between different effects are not taken into account.

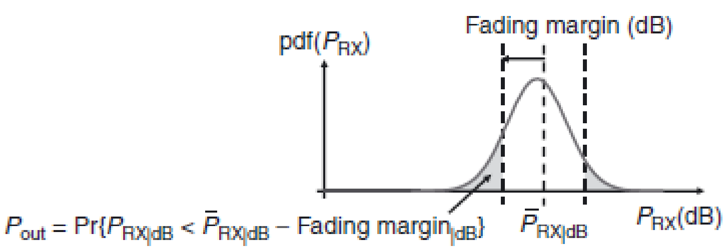
Before showing some examples, the following points should be stressed:

- The attenuation due to propagation effects, between TX and RX. For the purpose of this chapter, we use a simple model, the so-called "breakpoint" model. For distances d

< d break, the received power is proportional to d^{-2}, according to Eq.1. Beyond that point, the power is proportional to d^{-n}, where n typically lies between 3.5 and 4.5. The received power is thus

$$P_{\mathrm{RX}}(d) = P_{\mathrm{RX}}(d_{\mathrm{break}}) \left(\frac{d}{d_{\mathrm{break}}}\right)^{-n} \text{ for } d > d_{\mathrm{break}} \qquad \text{------ (1)}$$

Wireless systems, especially mobile systems, suffer from temporal and spatial variations of the transmission channel (*fading*). In other words, even if the distance is approximately constant, the received power can change significantly with small movements of the TX and/or RX. The power computed from Eq. (1) is only a *mean* value; the ratio of the transmit power to this mean received power is also known as the *path loss* (inverse of the path gain). If the mean received power is used as the basis for the link budget, then the transmission quality will be above the threshold only in approximately 50% of the times and locations. This is completely unacceptable quality of service. Therefore, we have to add a *fading margin*, which makes sure that the minimum received power is exceeded in at least, e.g., 90% of all cases (see Figure). The value of the fading margin depends on the amplitude statistics of the fading



Uplink (MS to BS) and downlink (BS to MS) are reciprocal, in the sense that the voltage and currents at the antenna ports are reciprocal (as long as uplink and downlink use the same carrier frequency). However, the noise figures of BSs and MSs are typically quite different. As MSs have to be produced in quantity, it is desirable to use low-cost components, which typically have higher noise figures. Furthermore, battery lifetime considerations dictate that BSs can emit more power than MSs. Finally, BSs and MSs differ with respect to antenna diversity, how close they are to interferers, etc. Thus, the link budgets of uplinks and downlinks are different.

Small-Scale Multipath Propagation

Multipath in the radio channel creates small-scale fading effects. The three most important effects are:

- Rapid changes in signal strength over a small travel distance or time interval
- Random frequency modulation due to varying Doppler shifts on different multipath signals

• Time dispersion (echoes) caused by multipath propagation delays.

In built-up urban areas, fading occurs because the height of the mobile antennas are well below the height of surrounding structures, so there is no single line-of-sight path to the base station.

Even when a line-of-sight exists, multipath still occurs due to reflections from the ground and surrounding structures. The incoming radio waves arrive from different directions with different propagation delays.

The signal received by the mobile at any point in space may consist of a large number of plane waves having randomly distributed amplitudes, phases, and angles of arrival. These multipath components combine vectorically at the receiver antenna, and can cause the signal received by the mobile to distort or fade. Even when a mobile receiver is stationary, the received signal may fade due to movement of surrounding objects in the radio channel.

If objects in the radio channel are static, and motion is considered to be only due to that of the mobile, then fading is purely a spatial phenomenon.

The spatial variations of the resulting signal are seen as temporal variations by the receiver as it moves through the multipath field.

Due to the constructive and destructive effects of multipath waves summing at various points in space, a receiver moving at high speed can pass through several fades in a small period of time. In a more serious case, a receiver may stop at a particular location at which the received signal is in a deep fade.

Maintaining good communications can then become very difficult, although passing vehicles or people walking in the vicinity of the mobile can often disturb the field pattern, thereby diminishing the likelihood of the received signal remaining in a deep null for a long period of time. Antenna space diversity can prevent deep fading nulls. Due to the relative motion between the mobile and the base station, each multipath wave experiences an apparent shift in frequency. The shift in received signal frequency due to motion is called the Doppler shift, and is directly proportional to the velocity and direction of motion of the mobile with respect to the the direction of arrival of the received multipath wave.

Factors Influencing Small-Scale Fading:

Many physical factors in the radio propagation channel influence small scale fading. These include the following:

Multipath propagation:

The presence of reflecting objects and scatters in the channel creates a constantly changing environment that dissipates the signal energy in amplitude, phase, and time. These effects result in multiple versions of the transmitted signal that arrive at the receiving antenna, displaced with respect to one another in time and spatial orientation. The random phase and amplitudes of the different multipath components cause fluctuations in signal strength, thereby inducing small-scale fading, signal distortion, or both. Multipath propagation often lengthens the time required for the baseband portion of the signal to reach the receiver which can cause signal smearing due to inter-symbol interference.

Speed of the mobile:

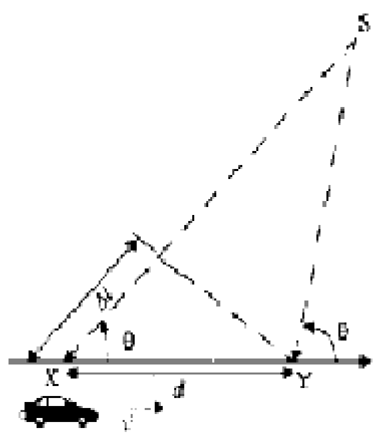
The relative motion between the mobile results in random frequency modulation due to different Doppler shifts on each of the multipath components. Doppler shift will be positive or negative depending on whether the mobile receiver is moving toward or away from the base station.

Speed of surrounding objects:

If objects in the radio channel are in motion, they induce a time varying Doppler shift on multipath components. If the surrounding objects move at a greater rate than the mobile, then this effect dominates the small-scale fading. Otherwise, motion of surrounding objects may be ignored, and only the speed of the mobile need be considered. The transmission bandwidth of the signal — If the transmitted radio signal bandwidth is greater than the "bandwidth" of the multipath channel, the received signal will be distorted, but the received signal strength will not fade much over a local area (i.e., the small-scale signal fading will not be significant). The bandwidth of the channel can be quantified by the coherence bandwidth which is related to the specific multipath structure of the channel. The coherence bandwidth is a measure of the maximum frequency difference for which signals are still strongly correlated in amplitude. If the transmitted signal has a narrow bandwidth as compared to the channel, the amplitude of the signal will change rapidly, but the signal will not be distorted in time. Thus, the statistics of small-scale signal strength and the likelihood of signal smearing appearing over small-scale distances are very much related to the specific amplitudes and delays of the multipath channel, as well as the bandwidth of the transmitted signal.

Doppler Shift :

Consider a mobile moving at a constant velocity v, along a path segment having length d between points X and Y, while it receives signals from a remote source S, as illustrated in Figure 1. The difference in path lengths traveled by the wave from source S to the mobile at points X andY is ΔlΦ= dcosO = vΔtcosθ. where Δt is the time required for the mobile to travel from X to Y, and θ is assumed to be the same at points X and Y since the source is assumed to be very far away. The phase change in the received signal due to the difference in path lengths is therefore



$$\Delta\phi = \frac{2\pi\Delta l}{\lambda} = \frac{2\pi v\Delta t}{\lambda}\cos\theta$$

and hence the apparent change in frequency, or Doppler shift, Doppler shift to the mobile velocity and the spatial angle between the direction of motion of the mobile and the direction of arrival of the wave. Multipath components from a CW signal which arrive from different directions contribute to Doppler spreading of the received signal, thus increasing the signal bandwidth.

Coherence Bandwidth

The delay spread is a natural phenomenon caused by reflected and scattered propagation paths in the radio channel, the coherence bandwidth B_c, is a defined relation derived from the rms delay spread. Coherence bandwidth is a statistical measure of the range of frequencies over which the channel can be considered "flat" (i.e., a channel which passes all spectral components with approximately equal gain and linear phase); In other words, coherence bandwidth is the range of frequencies over which two frequency components have a strong potential for amplitude correlation. Two sinusoids with frequency separation greater than are affected quite differently by the channel. If the coherence bandwidth is defined as the bandwidth over which the frequency correlation function is above 0.9, then the coherence bandwidth is approximately

$B_c = (1 /50\ \sigma_\tau)$ ----------------------- (1)

If the definition is relaxed so that the frequency correlation function is above 0.5. then the coherence bandwidth is approximately

$B_c = (1 /50\sigma_\tau)$-----------------------------(2)

It is important to note that an exact relationship between coherence bandwidth and rms delay spread does not exist, and equations (1) and (2) are "ball park estimates".

Coherence time T_c:

Coherence time T_c, is the time domain dual of Doppler spread and is used to characterize the time varying nature of the frequency dispersiveness of the channel in the time domain. The Doppler spread and coherence time are inversely proportional to one another. That is,

$T_c = (1/ f_m)$ ------------------(3)

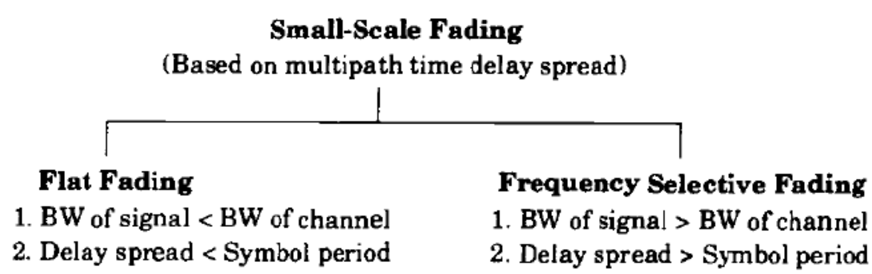
Coherence time is actually a statistical measure of the time duration over which the channel impulse response is essentially invariant, and quantifies the similarity of the channel response at different times. In other words, coherence time is the time duration over which two received signals have a strong potential for amplitude correlation. If the reciprocal bandwidth of the baseband signal is greater than the coherence time of the channel, then the channel will change during the transmission of the baseband message, thus causing distortion at the receiver. If the coherence time is defined as the time over which the time correlation function is above 0.5, then the coherence time is approximately

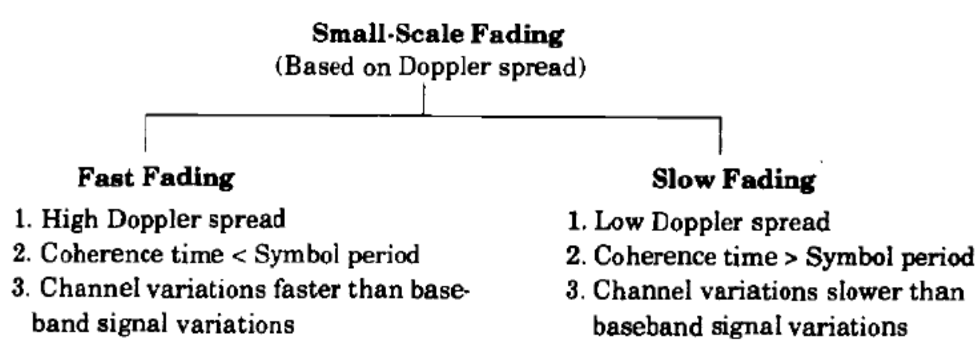
$T_c = (9 / 16\pi f_m)$ ------------------(4)

where f_m is the maximum Doppler shift given by $f_m = v/\lambda$.

<u>Types of Small-Scale Fading:</u>

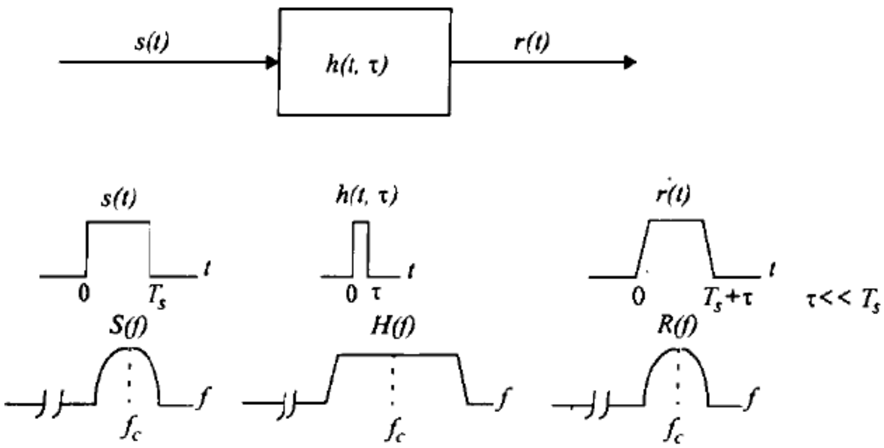
The type of fading experienced by a signal propagating through a mobile radio channel depends on the nature of the transmitted signal with respect to the characteristics of the channel. Depending on the relation between the signal parameters (such as bandwidth, symbol period, etc.) and the channel parameters (such as rms delay spread and Doppler spread), different transmitted signals will undergo different types of fading. The time dispersion and frequency dispersion mechanisms in a mobile radio channel lead to four possible distinct effects, which are manifested depending on the nature of the transmitted signal, the channel, and the velocity. While multipath delay spread leads to time dispersion and frequency selective fading, Doppler spread leads to frequency dispersion and time selective fading. The two propagation mechanisms are independent of one another.





<u>Flat fading:</u>

If the mobile radio channel has a constant gain and linear phase response over a bandwidth which is greater than the bandwidth of the transmitted signal, then the received signal will undergo flat fading. This type of fading is historically the most common type of fading described in the technical literature. In flat fading, the multipath structure of the channel is such that the spectral characteristics of the transmitted signal are preserved at the receiver. However the strength of the received signal changes with time, due to fluctuations in the gain of the channel caused by multipath



Flat fading channel characteristics.

- It can be seen from the above figure that, if the channel gain changes over time, a change of amplitude occurs in the received signal.
- Over time, the received signal r (t) varies in gain, but the spectrum of the transmission is preserved.
- In a flat fading channel, the reciprocal bandwidth of the transmitted signal is much larger than the multipath time delay spread of the channel, $h_b(t, \tau)$ and can be approximated as having no excess delay (i.e., a single delta function with ($\tau = 0$)
- Flat fading channels are also known as amplitude varying channels and are sometimes referred to as narrowband channels, since the bandwidth of the applied signal is narrow as compared to the channel flat fading bandwidth.
- Typical flat fading channels cause deep fades, and thus may require 20 or 30 dB more transmitter power to achieve low bit error rates during times of deep fades as compared to systems operating over non-fading channels.

Frequency Selective Fading

- If the channel possesses a constant-gain and linear phase response over a bandwidth that is smaller than the bandwidth of transmitted signal, then the channel creates frequency selective fading on the received signal.
- Under such conditions the channel impulse response has a multipath delay spread which is greater than the reciprocal bandwidth of the transmitted message waveform. When this occurs, the received signal includes multiple versions of the transmitted waveform which are attenuated (faded) and delayed in time, and hence the received signal is distorted.
- Frequency selective fading is due to tine dispersion of the transmitted symbols within

the channel. Thus the channel induces inter symbol interference (ISI).

- Viewed in the frequency domain, certain frequency components in the received signal spectrum have greater gains than others.
- Frequency selective fading channels are much more difficult to model than flat fading channels since each multipath signal must be modeled and the channel must be considered to be a linear filter.
- It is for this reason that wideband multipath measurements are made, and models are developed from these measurements.
- When analyzing mobile communication systems, statistical impulse response models such as the 2-ray Rayleigh fading model (which considers the impulse response to be made up of two delta functions which independently fade and have sufficient time delay between them to induce frequency selective fading upon the applied signal), or computer generated or measured impulse responses, are generally used for analyzing frequency selective small-scale fading.

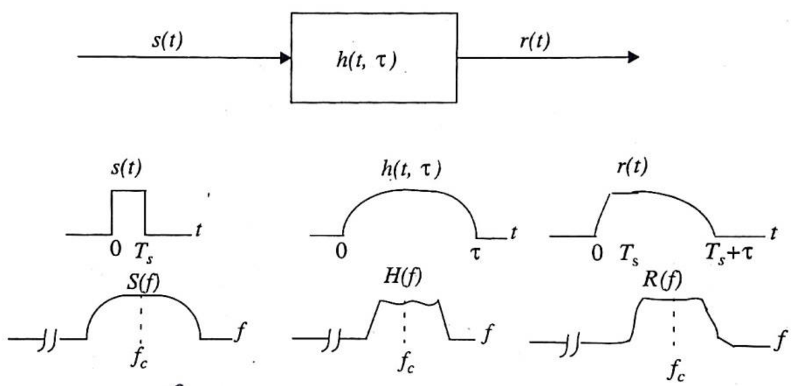


Figure 5.13 Frequency selective fading channel characteristics.

Fading Effects Due to Doppler Spread

Fast Fading

- Depending on how rapidly the transmitted baseband signal changes as compared to the rate of change of the channel, a channel may be classified either as a fast fading or slow fading channel.
- In a fast fading channel, the channel impulse response changes rapidly within the symbol duration. That is, the coherence time of the channel is smaller than the symbol period of the transmitted signal.
- This causes frequency dispersion (also called time selective fading) due to Doppler spreading, which leads to signal distortion. Viewed in the frequency domain, signal distortion due to fast fading increases with increasing Doppler spread relative to the bandwidth of the transmitted signal. Therefore a signal undergoes fast fading if

$T_S > T_C$ & $B_s < B_D$

- It should be noted that when a channel is specified as a fast or slow fading channel, it does not specify whether the channel is flat fading or frequency selective in nature.
- Fast fading only deals with the rate of change of the channel due to motion. In the case of the flat fading channel, we can approximate the impulse response to be simply a delta function (no time delay).
- Hence, a flat fading, fast fading channel is a channel in which the amplitude of the delta function varies faster than the rate of change of the transmitted baseband signal.
- In the case of a frequency selective, fast fading channel, the amplitudes, phases, and time delays of any one of the multipath components vary faster than the rate of change of the transmitted signal. In practice, fast fading only occurs for very low data rates.

<u>Slow Fading</u>

- In a slow fading channel, the channel impulse response changes at a rate much slower than the transmitted baseband signal s(t). In this case, the channel may be assumed to be static over one or several reciprocal bandwidth intervals. In the frequency domain, this implies that the Doppler spread of the channel is much less than the bandwidth of the baseband signals. Therefore, a signal undergoes slow fading if

  $T_S << T_C$ & $B_S >> B_D$

- It should be clear that the velocity of the mobile (or velocity of objects in the channel) and the baseband signaling determines whether a signal undergoes fast fading or slow fading.

<u>Free space propagation model</u>

- The free space propagation model is used to predict received signal strength when the transmitter and the receiver have a clear, unobstructed line-of-sight path between them.
- Satellite communication systems and microwave line-of-sight links typically undergo free space propagation.
- Free space propagation model predicts that received power decays as a function of the T-R separation distance raised to some power.
- The free space power received by a receiver antenna which is separated from a radiating transmitting antenna by a distance d, is given by the Friis free space equation,

$$P_r(d) = \frac{P_t G_t \ G_r \lambda^2}{(4\pi)^2 d^2 L} \qquad ------1$$

- Where P_t is the transmitted power $P_r(d)$ is the received power which is a function of T-R separation , G_t is the Tx antenna gain, G_r is the Rx antenna gain, d is the T-R separation distance in meters, L is the system loss factor which is not related to propagation ($1 \leq L$) and λ is the wavelength in meters.

- A value of L=1 indicates no loss in the system hardware.

The gain of the antenna is related to effective aperture, A_e, by

$$G = \frac{4\pi A_e}{\lambda^2} \quad ------2$$

The effective A_e is related to the physical size of the antenna, and λ is related to the carrier frequency by

$$\lambda = \frac{c}{f_c} = \frac{2\pi c}{\omega_c} \qquad ------3$$

Where f is the frequency in Hertz, ω_c is the carrier frequency in radians per second and c is the speed of light in meters/s.

- The Friis free space equation (1) shows that the received power falls off as the square of the T-R separation distance. This implies that the received power decays with distance at a rate of 20 dB/decade.

Effective Isotropic Radiated Power(EIRP)

The Effective Isotropic Radiated Power(EIRP) is defined as

$$EIRP = P_t G_t \quad ------4$$

EIRP represents the maximum radiated power available from a transmitter in the direction of maximum antenna gain , as compared to an **isotropic radiator**.

An **isotropic radiator** is an ideal antenna which radiates power with unit gain uniformly in all directions, and is often used to reference antenna gains in wireless systems.

Effective Radiated Power(ERP)

ERP represents the maximum radiated power available from a transmitter in the direction of maximum antenna gain , as compared to a **half wave dipole antenna**.

Since a dipole antenna has a gain of 1.64 (2.15 dB above an isotropic antenna), the **ERP** will be 2.15 dB smaller than the EIRP for the same transmission system.

Path loss

The path loss is defined as the difference (in dB) between the effective transmitted power and the received power.

The path loss for the free space model is given by,

$$PL(dB) = 10log\frac{P_t}{P_r} = -10\log\left[\frac{G_t G_r \lambda^2}{(4\pi)^2 d^2}\right] \quad ------5$$

When antenna gains are excluded, the antennas are assumed to have unity gain, and path loss is given by,

$$PL(dB) = 10log\frac{P_t}{P_r} = -10\log\left[\frac{\lambda^2}{(4\pi)^2 d^2}\right] \quad ------6$$

- The Friis space model is a valid predictor for Pr for values of d which are in the far field of the transmitting antenna.
- The far field of Fraunhofer region, of a transmitting antenna is defined as the region beyond the far field distance df, which is related to the largest linear dimension of the TX antenna aperture and the carrier wave length.

The Fraunhofer distance is given by

$d_f = \frac{2D^2}{\lambda}$ $------ 7a$

Where D is the largest physical linear dimension of the antenna. To be in the far field region, df must satisfy

$$d_f \gg D \quad ------ 7b \quad \text{and}$$

$$d_f \gg \lambda \quad ------ 7c$$

The equation (1) does not hold for d = 0. The received power in free space is expressed in terms of reference distance d_0. The reference distance must be chosen such that it lies in the far field region, that is $d_f \leq d_0$, and d_0 is chosen to be smaller than any practical distance used in mobile communication system.

The received power, $P_r(d)$, at any distance $d>d_0$, is related to P_r at d_0. The value $P_r(d_0)$ may be predicted from equation (1) or may be measured.

The received power, $P_r(d)$, at any distance $d>d_0$, is given by,

$$P_r(d) = P_r(d_0)\left(\frac{d_0}{d}\right)^2 \qquad d \geq d_0 \geq d_f \quad ------ 8$$

Because of the large dynamic range of received power levels, often dBm or dBW units are used to express received power levels.

For example, if P_r is in units of dBm, the received power is given by,

$$P_r(d)\ dBm = 10\log\left[\frac{P_r\ (d_0)}{0.001\ W}\right] + 20\log\left(\frac{d_0}{d}\right) \qquad d \geq d_0 \geq d_f ------ 9$$

<u>Received power in terms of electric field E induced at the receiver antenna</u>:

$$P_r(d) = \frac{|E|^2}{120\pi} A_e = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2} \quad ------ 10$$

E= induced electric field at the receiver antenna

A_e = effective aperture of the antenna

<u>Received power in terms of receiver input voltage:</u>

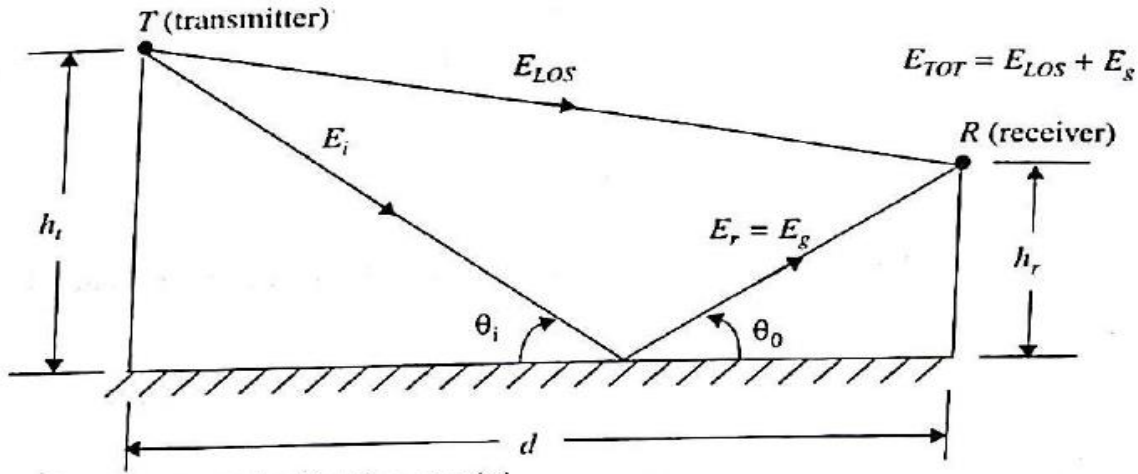
$$P_r(d) = \frac{V^2}{R_{ant}} = \frac{(V_{ant}/2)^2}{R_{ant}} = \frac{V_{ant}^2}{4R_{ant}} \quad ------ 11$$

V= rms voltage at the input of the receiver

R_{ant} = receiver antenna resistance

<u>Ground Reflection (two-Ray) Model</u>

- Two ray ground reflection model shown in figure is a useful propagation model that is based on geometric optics, and considers both the direct path and a ground reflected propagation path between T_X and R_X.
- This model is reasonably accurate for predicting the large scale signal strength over distances of several kilometers.



Two-ray ground reflection model.

In most mobile communication systems, the maximum T-R separation distance is few tens of kilometers and the earth may be assumed to be flat.
The total received field , E_{TOT} is the vector sum of Line-of- sight component E_{LOS} and ground reflected component E_g.
If E_0 is the free space E field (in units of V/m) at a reference distance d_0 from the transmitter, then for $d>d_0$, the free space propagating E-field is given by,

$$E(d,t) = \frac{E_0 d_0}{d} \cos\left(\omega_c\left(t - \frac{d}{c}\right)\right) \qquad (d > d_0) \quad ------1$$

Two propagating waves arrive at the receiver:

(i) direct wave that travels a distance d'

(ii) The reflected wave that travels a distance d''

The E- field due to the line-of-sight component at the receiver can be expressed as,

$$E_{LOS}(d',t) = \frac{E_0 d_0}{d'} \cos\left(\omega_c\left(t - \frac{d'}{c}\right)\right) \qquad ------2$$

The E- field for the ground reflected wave can be expressed as,

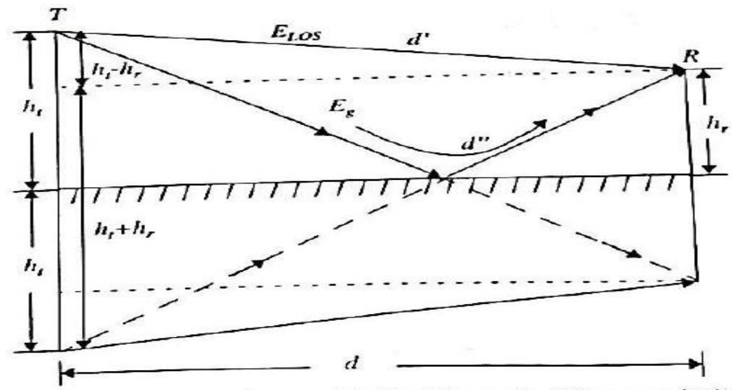
$$E_g(d'',t) = \Gamma\frac{E_0 d_0}{d''} \cos\left(\omega_c\left(t - \frac{d''}{c}\right)\right) \qquad ------3$$

According to laws of reflection in dielectrics,
$\theta_i = \theta_0 \quad ------4$
And $\quad E_g = \Gamma E_i \quad ------5a$
$E_t = (1+\Gamma)E_i \quad ------5b$
Where Γ is the reflection coefficient for ground. For small values of θ_i (i.e., grazing incidence), the reflected wave is equal in magnitude and 180^0 out of phase with the incident wave.
The resultant E- field is the vector sum of E_{LOS} and E_g and the total E-field envelop is given by,

$$|E_{TOT}| = |E_{LOS} + E_g| \quad ------6$$

The electric field $E_{TOT}(d,t)$ can be expressed as the sum of equations (2) and (3)

$$E_{TOT}(d,t) = \frac{E_0 d_0}{d'} \cos\left(\omega_c\left(t - \frac{d'}{c}\right)\right) + (-1)\frac{E_0 d_0}{d''} \cos\left(\omega_c\left(t - \frac{d''}{c}\right)\right) --- 7$$



The method of images is used to find the path difference between the line-of-sight and the ground reflected paths.

From the geometry of figure, the path difference , Δ between the line-of-sight path and ground reflected path can be expressed as,

$$\Delta = d'' - d' = \sqrt{(h_t + h_r)^2 + d^2} - \sqrt{(h_t - h_r)^2 + d^2} \; ------ 8$$

Where, h_t is the height of the transmitter and h_r is the height of the receiver.

When the T-R separation distance d is very large compared to h_t , h_r,

$$d' = (d^2 - 2h_t h_r)^{1/2} = d\left(1 - \frac{2h_t h_r}{d^2}\right)^{1/2} \approx d\left(1 - \frac{2h_t h_r}{2d^2}\right)$$

$$d'' = (d^2 + 2h_t h_r)^{1/2} = d\left(1 + \frac{2h_t h_r}{d^2}\right)^{1/2} \approx d\left(1 + \frac{2h_t h_r}{2d^2}\right)$$

$$\Delta = d'' - d' = d\left(1 + \frac{2h_t h_r}{2d^2} - 1 + \frac{2h_t h_r}{2d^2}\right) = \frac{2h_t h_r}{d} \; ------ 9$$

The phase difference θ_Δ between the two E- field components and the time delay between the arrival of two components is computed as,

$$\theta_\Delta = \frac{2\pi\Delta}{\lambda} = \frac{\Delta\omega_c}{c} \; ------ 10$$

And $\tau_d = \frac{\Delta}{c} = \frac{\theta_\Delta}{2\pi f_c} \; ------ 11$

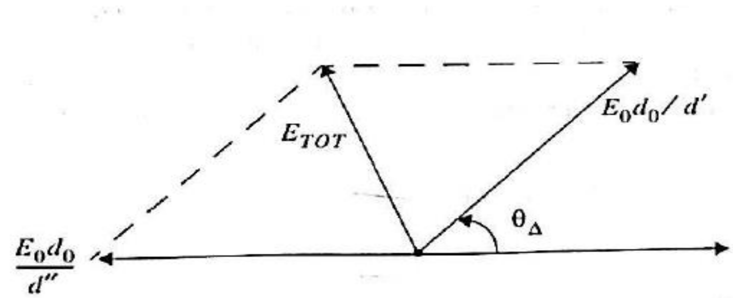
As d becomes large, the difference between d' and d'' becomes very small. And the amplitudes of E_{LOS} and E_g are virtually identical and differ only in phase.

$$\left|\frac{E_0 d_0}{d}\right| \approx \left|\frac{E_0 d_0}{d'}\right| \approx \left|\frac{E_0 d_0}{d''}\right| \; ------ 12$$

If the received field is evaluated at some time $t = \frac{d''}{c}$, equation (7) can be expressed as a phasor sum

$$E_{TOT}\left(d, t = \frac{d''}{c}\right) = \frac{E_0 d_0}{d'} \cos\left(\omega_c\left(\frac{d'' - d'}{c}\right)\right) - \frac{E_0 d_0}{d''} \cos(0^0) \; ------ 13$$

$$E_{TOT}(d,t) = \frac{E_0 d_0}{d'} L\theta_\Delta - \frac{E_0 d_0}{d''} \approx \frac{E_0 d_0}{d}[L\theta_\Delta - 1]$$



Phasor diagram showing the electric field components of the line-of-sight, ground reflected, and total received E-fields

$$|E_{TOT}(d)| = \left|\frac{E_0 d_0}{d}\left[e^{j\theta_\Delta} - 1\right]\right|$$

$$|E_{TOT}(d)| = \left|\frac{E_0 d_0}{d}[cos\theta_\Delta + jsin\theta_\Delta - 1]\right|$$

$$|E_{TOT}(d)| = \sqrt{\left(\frac{E_0 d_0}{d}\right)^2 (cos\theta_\Delta - 1)^2 + \left(\frac{E_0 d_0}{d}\right)^2 sin^2\theta_\Delta} \quad ------14$$

$$|E_{TOT}(d)| = \frac{E_0 d_0}{d}\sqrt{2 - 2cos\theta_\Delta} = \frac{E_0 d_0}{d}\sqrt{2 \times 2sin^2\left(\frac{\theta_\Delta}{2}\right)} = \frac{2E_0 d_0}{d} sin\left(\frac{\theta_\Delta}{2}\right) ----15$$

Equation (15) provides exact received E-field for the two ray ground reflection model.

$$\left(\frac{\theta_\Delta}{2}\right) \text{ is very less (less than } 0.3 \text{ radian)}$$

$$sin\left(\frac{\theta_\Delta}{2}\right) \approx \left(\frac{\theta_\Delta}{2}\right)$$

$$E_{TOT}(d) = \frac{2E_0 d_0}{d} \times \frac{1}{2} \times \frac{2\pi}{\lambda} \times \frac{2h_t h_r}{d} \quad ------16`$$

The received power at a distance d from the transmitter for the two ray ground bounce model can be expressed as,

$$P_r = P_t G_t G_r \frac{{h_t}^2 {h_r}^2}{d^4} \quad ------17$$

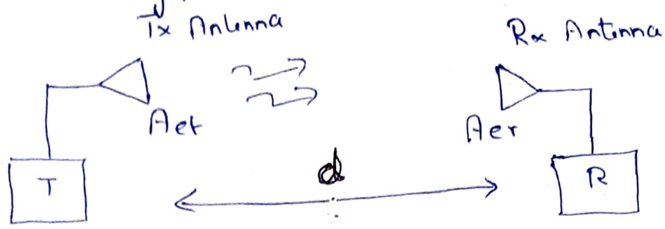
Path loss in dB for the two ray model can be expressed as,

$$PL(dB) = 10 \; log\left[\frac{d^4}{G_t G_r h_t^2 h_r^2}\right]$$

$$PL(dB) = 40 \log d - 10 \; logG_t - 10logG_r - 20logh_t - 20logh_r \quad ------18$$

The radio communication Link - Friis Txn formula

Friis Tx formula - 1946 - Harald T. Friis of Bell Telephone Laboratory.



} Free space propagation model

* Friis Txn formula gives the Power received over a link.
* Assuming lossless, matched antennas
  Let the Txr feed a power P_t to a Tx antenna of effective aperture A_{et}
* At a distance d a receiving antenna of effective aperture A_{er} intercepts some of the power radiated by the transmitting antenna and delivers it to the receiver R.
* Assuming that the transmitting antenna is isotropic, the power per unit area available at the receiving antenna is

$$S_r = P_t / 4\pi d^2 \quad \text{--- (1)}$$

if the antenna has gain G_t, the power per unit area available at the receiving antenna will be increased in proportion as given by,

$$S_r = \frac{P_t G_t}{4\pi d^2} \quad \text{--- (2)}$$

Power collected by the lossless, matched receiving antenna of effective aperture A_{er} is

$$P_r = S_r A_{er} = \frac{P_t G_t A_{er}}{4\pi d^2} \quad \text{--- (3)}$$

The gain of the transmitting antenna can be expressed as

$$G_t = \frac{4\pi A_{et}}{\lambda^2} \quad \text{--- (4)}$$

$$A_{et} = \frac{G_t \lambda^2}{4\pi}$$

$$A_{er} = \frac{G_r \lambda^2}{4\pi}$$

Substituting (4) in (3)

$$P_r = \frac{P_t G_t A_{er}}{4\pi d^2}$$

$$P_r = P_t \cdot \frac{4\pi A_{et}}{\lambda^2} \times \frac{A_{er}}{4\pi d^2}$$

$$\frac{P_r}{P_t} = \frac{A_{er} A_{et}}{d^2 \lambda^2}$$

Friis Txn formula

$$\frac{P_r}{P_t} = \frac{G_t \lambda^2 \, G_r \lambda^2}{(4\pi)^2 d^2 \lambda^2} = \frac{G_t G_r \lambda^2}{(4\pi)^2 d^2}$$

P_r = received Power

P_t = transmitted Power

A_{et} = effective aperture of Tx antenna m^2

A_{er} = effective aperture of Rx antenna m^2

r = distance between antennas m

λ = wavelength, m

<u>example:</u> If a Transmitter produces 50W of power, express the transmit power in units of (a) dBm and (b) dBW. If 50W is applied to a unity gain antenna with a 900 MHz carrier frequency, Find the received power in dBm at a free space distance of 100 m from the antenna. What is P_r (10km). Assume unity gain for the receiver antenna.

(a) $P_t(dBm) = 10\log_{10}\left(\frac{P_t}{1mW}\right) = 10\log_{10}\left(\frac{50}{1\times10^{-3}}\right) = 47.0$ dBm

(b) $P_t(dBW) = 10\log_{10}\left(\frac{P_t}{1W}\right) = 10\log_{10}(50) = 17.0$ dBW

(c) Received power $= \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2} = \frac{50\times1\times1\times(1/3)^2}{(4\pi)^2\times(100)^2}$

$\lambda = \frac{3\times10^8}{900\times10^6} = 1/3$

$= 3.5\times10^{-6}$ W

$P_r(dBm) = 10\log_{10}(3.5\times10^{-3})$

$= -24.5$ dBm

$$P_r(d)\,dBm = 10\log\left[\frac{P_r(d_0)}{0.001W}\right] + 20\log\left(\frac{d_0}{d}\right)$$

$d_0 = 100$ m $\quad d = 10$ km

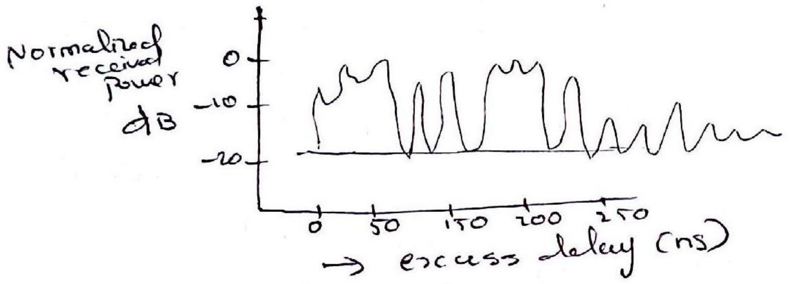
$= 20\log\left(\frac{100}{10000}\right) - 24.5 = -64.5$ dBm

Parameters of Mobile Multipath Channels

Multipath channel parameters are derived from the power delay profile.

Power delay profile — measured
— represented as plots of relative received power as a function of excess delay with respect to a fixed time delay reference.

example



typical (indoor) profile.

Time dispersion Parameters

To compare different multipath channels and to develop some general design guidelines for wireless systems, parameters which grossly quantify the multipath channel are used.

Mean excess delay
rms delay spread
excess delay spread

} parameters that can be determined from a power delay profile.

Time dispersive properties of wide band multipath channels are most commonly quantified by their mean excess dely ($\bar{\tau}$) and rms delay spread (σ_τ)

The mean excess delay is the first moment of the power delay profile

$$\bar{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k a_k^2} = \frac{\sum_k P(\tau_k)\tau_k}{\sum_k P(\tau_k)}$$

rms delay spread is the square root of the second central moment of the power delay profile.

$$\sigma_\tau = \sqrt{\overline{\tau^2} - (\bar{\tau})^2}$$

where,

$$\overline{\tau^2} = \frac{\sum_k a_k^2 \tau_k^2}{\sum_k a_k^2} = \frac{\sum_k P(\tau_k)\tau_k^2}{\sum_k P(\tau_k)}$$

These delays are measured relative to the first detectable signal arriving at the receiver at $\tau_0 = 0$

Typical values of rms delay spread are – order of μs (outdoor mobile channel)

– order of ns (indoor radio channels)

(x) rms delay spread, mean excess delay } defined from a single power delay profile.

__Maximum excess delay__ – time delay during which multipath energy falls to X dB below the maximum.

In other words,

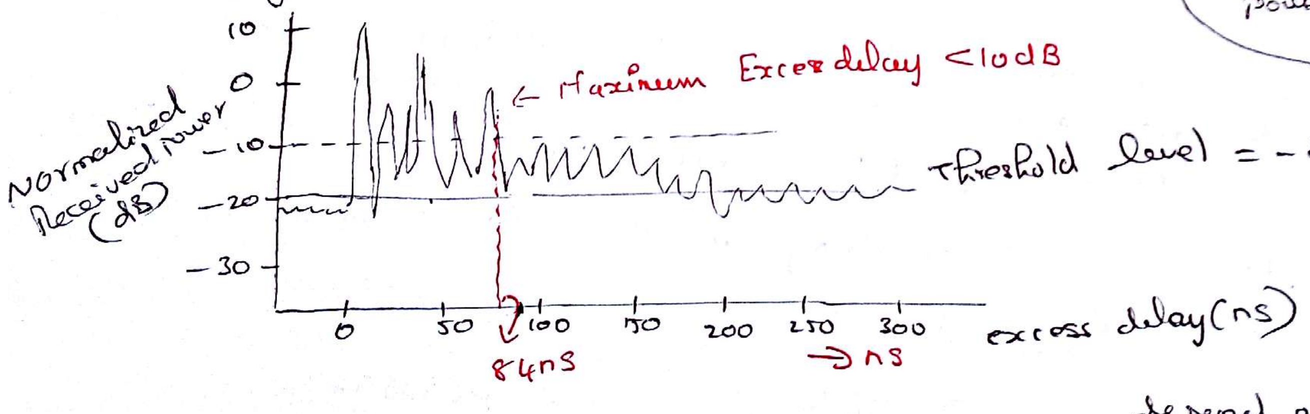
Maximum excess delay is defined as $\tau_x - \tau_0$,

τ_0 is the first arriving signal

τ_x is the maximum delay at which a multipath component is within X dB of the strongest arriving multipath signal.

__Fig__ illustrates __the computation__

typical indoor power delay profile



values for $\bar{\tau}$, $\overline{\tau^2}$ and σ_τ depend on the choice of noise threshold used to process $P(\tau)$

②

example

Compute the rms delay spread for the following power delay profile:

P(τ)

(a) calculate the rms delay spread for the given figure.

(b) If BPSK modulation is used, what is the maximum bit rate that can be sent through the channel without needing an equalizer?

a) $$\bar{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k a_k^2} = \frac{\sum_k P(\tau_k)\tau_k}{\sum_k P(\tau_k)}$$

$$\bar{\tau} = \frac{(1\times 0)\mu s + (1\times 1)\mu s + (1\times 2)\mu s}{1+1+1} = \frac{3\,\mu s}{2+1} = 1.0\,\mu s$$

$$\overline{\tau^2} = \frac{\sum_k P(\tau_k)\tau_k^2}{\sum_k P(\tau_k)} = \frac{1\times 0 + 1\times 1 + 1\times 4}{1+1+1} = \frac{5(\mu s)^2}{3} = 1.67(\mu s)^2$$

$$\sigma_\tau = \sqrt{\overline{\tau^2} - (\bar{\tau})^2} = \sqrt{1.67 - 1} = \sqrt{0.67} = 0.8185\,\mu s$$

Analogous to the delay spread parameters in the time domain, coherence bandwidth is used to characterize the channel in the frequency domain.

* rms delay spread and coherence bandwidth are inversely proportional to one another.

coherence bandwidth — derived from the rms delay spread.

* coherence bandwidth is the ~~rang~~ statistical measure of the range of frequencies over which the channel can be considered "flat"

flat → a channel which passes all spectral components with approximately equal gain and linear phase.

* coherence bandwidth (B_c) is the range of frequencies over which two frequency components have a strong potential for amplitude correlation.

* Two sinusoids with frequency separation greater than B_c are affected quite differently by the channel.

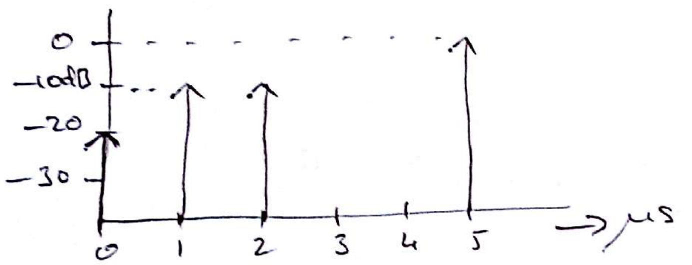
* If the coherence bandwidth is defined as the bandwidth over which the frequency correlation function is above 0.9, then the B_c is approximately

$$B_c \approx \frac{1}{50\,\sigma_\tau}$$

* If the B_c is defined as the bandwidth over which the frequency correlation function is above 0.5, then B_c is approximately,

$$B_c \approx \frac{1}{5\,\sigma_\tau}$$

<u>example</u> calculate the mean excess delay, rms delay spread, and the maximum excess delay (10 dB) for the multipath profile given in the figure below. Estimate the 50% coherence bandwidth of the channel.



$10\log_{10} P =$ dB

~~$10 \log_{10} (P)$~~

$10\log_{10}(P) = -10$

$P = 0.1$

$10\log_{10}(P) = -20$

$P = 0.01$

$10\log_{10}(P) = 0$

$P = 1$

$$\bar{\tau} = \frac{\sum_k P(\tau_k)\tau_k}{\sum_k P(\tau_k)} = \frac{(0\times 0.01) + (1\times 0.1) + (2\times 0.1) + (5\times 1)}{0.01 + 0.1 + 0.1 + 1} = 4.38\,\mu s$$

$$\overline{\tau^2} = \frac{\sum_k P(\tau_k)\tau_k^2}{\sum_k P(\tau_k)} = \frac{(0.01)\times 0^2 + (0.1\times 1^2) + (0.1\times 2^2) + (5^2\times 1)}{0.01 + 0.1 + 0.1 + 1}$$

$$= 21.07\,\mu s^2$$

$$\sigma_\tau = \sqrt{21.07 - (4.38)^2} = 1.37\,\mu s$$

$$B_c \approx \frac{1}{5\sigma_\tau} = \frac{1}{5(1.37\,\mu s)} = 145985$$

$$= 146\ \text{kHz}$$

Doppler spread and coherence Time

* Delay spread and coherence bandwidth are parameters which describe the time dispersive nature of the channel in a local area.

* Doppler spread and coherence time are parameters which describe the time varying nature of the channel in a small scale region.

* B_D – Doppler spread – Measure of the spectral broadening caused by the time rate of change of the mobile radio channel.

→ Defined as the range of frequencies over which the received Doppler spectrum is essentially non-zero.

When a pure sinusoidal frequency f_c is transmitted, the received signal spectrum, called the Doppler spectrum, will have components in the range $f_c - f_d$ to $f_c + f_d$, where f_d is the Doppler shift.

The amount of spectral broadening depends on f_d which is a function of the relative velocity of the mobile, and the angle θ between the direction of motion of the mobile and direction of arrival of the scattered waves.

If the baseband signal bandwidth is much greater than B_D, the effects of Doppler spread are negligible at the receiver. This is a slow fading channel.

Coherence time T_c → time domain dual of Doppler spread.

→ used to characterize the time varying nature of the frequency dispersiveness of the channel in the time domain.

$$T_c \approx \frac{1}{f_m} \quad \text{--- (1)}$$

Doppler spread and coherence time are inversely proportional to one another.

Coherence time is → measure of the time duration over which the channel impulse response is essentially invariant, and quantifies the similarity of the channel responses at different times.

Coherence time is the time duration over which two received signals have a strong potential for amplitude correlation.

If the reciprocal bandwidth of the baseband signal is greater than the coherence time of the channel then the channel will change during the transmission of the baseband message, thus causing distortion at the receiver.

If the coherence time is defined as the time over which the time correlation function is above 0.5, then the coherence time is approximately,

$$T_c \approx \frac{9}{16\pi f_m} \quad \text{— (2)}$$

where f_m is the maximum Doppler shift

$$f_m = v/\lambda$$

A popular rule of thumb for modern digital communication is to define the coherence time as the geometric mean of equations (1) & (2)

$$T_c = \sqrt{\frac{9}{16\pi f_m^2}} = \frac{0.423}{f_m}$$

Types of Small scale fading

Depending on the relation between the signal parameters (bandwidth, symbol period etc.) and the channel parameters (rms delay spread, Doppler spread), different transmitted signals will undergo different types of fading.

The time dispersion and frequency dispersion in a radio channel lead to four possible distinct effects.

mechanisms

flat fading
frequency selective fading } multipath delay spread

fast fading
slow fading } Doppler spread

Fading effects due to Multipath Time Delay spread

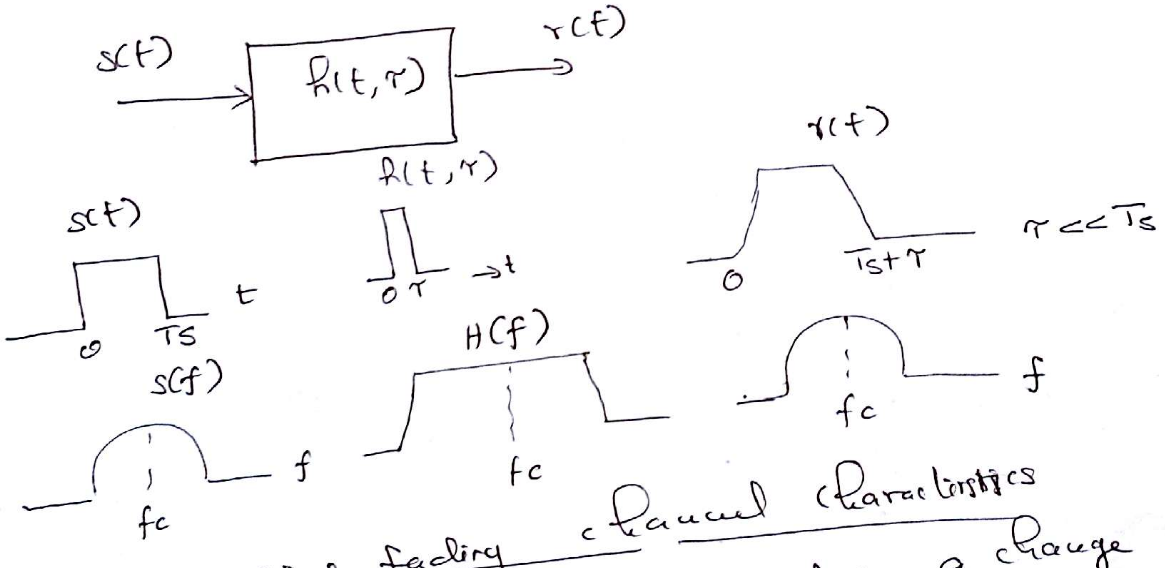
Time dispersion due to multipath causes the transmitted signal to undergo
✓ flat fading (or)
✓ frequency selective fading

flat fading

If the (mobile) radio channel has a constant gain and linear phase response over a bandwidth which is greater than the bandwidth of the transmitted signal then the received signal will undergo flat fading.

In flat fading,

the multipath structure of the channel is such that the spectral characteristics of the transmitted signal are preserved at the receiver. However the strength of the received signal changes with time, due to fluctuations in the gain of the channel (caused by multipath)



flat fading channel characteristics

Note:

* If channel gain changes over time a change of amplitude occurs in the received signal.
* the received signal r(t) varies in gain
* spectrum of the transmitted signal is preserved.

In a flat fading channel, the reciprocal bandwidth of the transmitted signal is much larger than the multipath time delay spread of the channel, and $h_b(t, \tau)$ can be approximated as having no excess delay (i.e., a single delta function with $\tau = 0$)

referred to as
$\left\{ \begin{array}{l} \text{flat fading channel} \\ \text{amplitude varying channels} \\ \text{narrow band channel (}\because\text{ the BW of the applied signal is narrow as compared to the channel flat fading BW)} \end{array} \right\}$

Typical flat-fading channels cause deep fades

$\Rightarrow$ May require 20 or 30 dB more transmitter power to achieve low BER during times of deep fades, as compared to systems operating over non-fading channels.

<u>Summary</u>

A signal undergoes flat fading if

$$B_s \ll B_c$$

&

$$T_s \gg \sigma_\tau$$

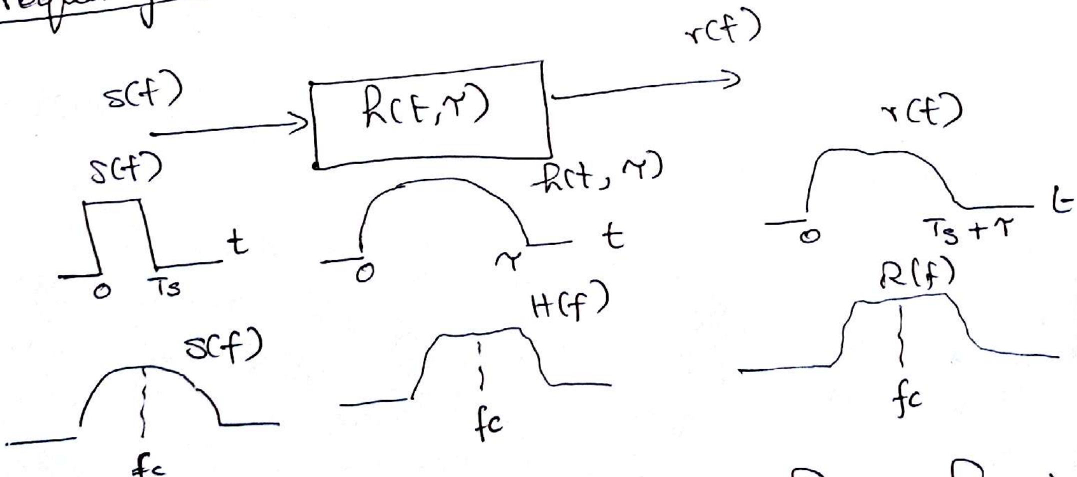
T_s - reciprocal bandwidth

B_s - Bandwidth

σ_τ - rms delay spread

B_c - coherence bandwidth.

<u>Frequency selective fading</u>



frequency selective fading ~~statistics~~ channel characteristics.

If the channel possesses a constant-gain and linear phase response over a bandwidth that is smaller than the bandwidth of transmitted signal, then the channel creates frequency selective fading on the received signal.

Under such condition, the channel impulse response has a multipath delay spread which is greater than the reciprocal bandwidth of the transmitted message waveform.

When this occurs, the received signal includes multiple version of the transmitted waveform which are attenuated and delayed in time, and hence the received signal is distorted.

Frequency selective fading is due to time dispersion of the transmitted symbols within the channel. Thus the channel induces intersymbol interference (ISI).

For frequency selective fading, the spectrum S(f) of the transmitted signal has a bandwidth which is greater than the coherence bandwidth Bc of the channel.

In frequency domain, where the gain is frequency selective, becomes different for different frequency components.

Frequency selective fading is caused by multipath delays which approach or exceed the symbol period of the transmitted symbol.

Frequency selective fading channels are also known as wide band channels, since the bandwidth of the signal S(t) is wider than the bandwidth of the channel impulse response.

Summary: A signal undergoes frequency selective fading if

$B_S > B_C$ &

$T_S < \sigma_\tau$

(x.) Thumb rule,

if $T_s \geq 10\sigma_\tau$ → ~~frequency~~ flat fading

$T_s < 10\sigma_\tau$ → frequency selective

Fading effects due to Doppler spread

Fast fading – In a fast fading channel, the channel impulse response changes rapidly within the symbol duration

i.e, coherence time of the channel is smaller than the symbol period of the transmitted signal

This causes frequency dispersion (also called time selective fading) due to Doppler spreading, which leads to signal distortion.

Viewed in the frequency domain, signal distortion due to fast fading increases with increasing Doppler spread relative to the bandwidth of the transmitted signal.

Therefore, a signal undergoes fast fading if

$$T_s > T_c$$

$$\& \quad B_s < B_D$$

flat fading, fast fading – is a channel in which the amplitude of the delta function varies faster than the rate of change of the transmitted base band signal.

frequency selective, fast fading – the amplitudes, phases, and time delays of any one of the multipath components vary faster than the rate of change of the transmitted signal.

In practice fast fading only occurs for very low data rates.

slow fading

In a slow fading channel, the channel impulse response changes at a rate much slower than the transmitted base band signal.

In the frequency domain, this implies that the Doppler spread of the channel is much less than the bandwidth of the base band signal.

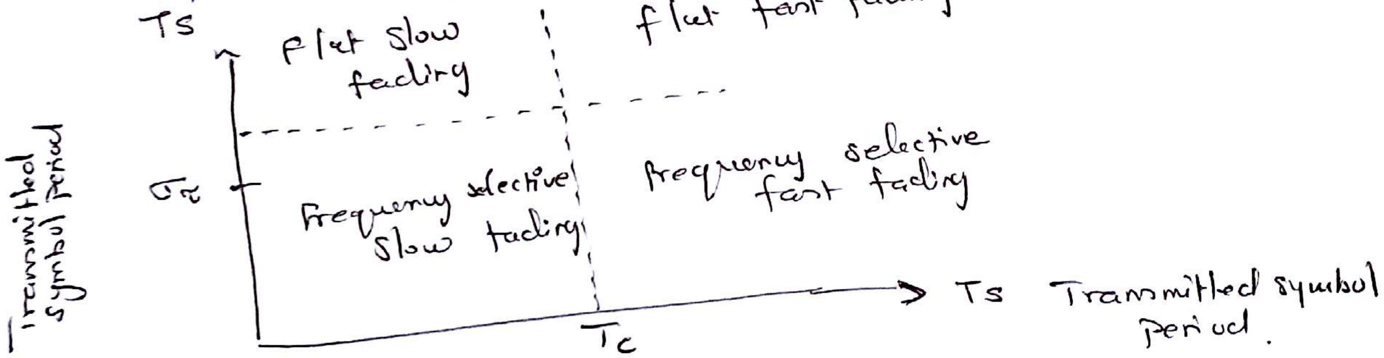
A signal undergoes slow fading if

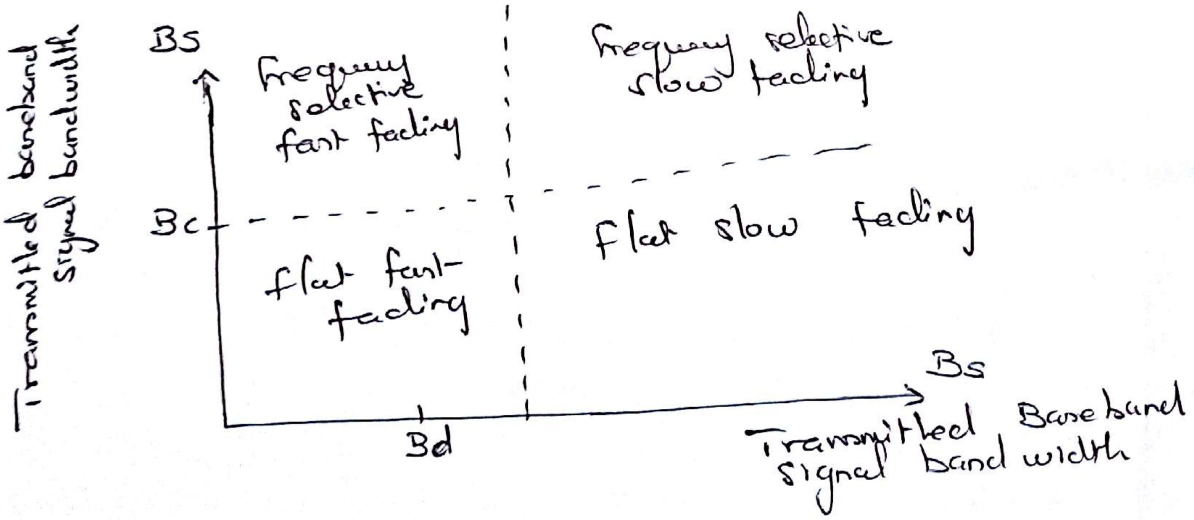
$$T_s << T_c$$

& $$B_s >> B_D$$

Note: velocity of the objects (mobile) in the channel and the base band signalling determines whether a signal undergoes fast fading or slow fading.

Summary (relationship between the various multipath parameters and the type of fading experienced by the signal)





Link Budget for wireless communication system

- When designing a complete, i.e. end to end radio communications system, it is necessary to calculate what is termed the radio link budget.

- The link budget is a summary of the transmitted power along with all the gains and losses in the system and this enables the strength of the received signal to be calculated.

Purpose of link Budget

- Using the knowledge on link budget,
- it is possible to determine whether power and gain levels are
  - sufficient,
  - too high, or
  - too low

and then apply corrective action to ensure the system will operate satisfactorily.

Influence on the cost of the system

- This ensures that once the system is installed and is ready for operation, there will be sufficient signal for it to operate correctly, or whether the signal is too even high and action can be taken to save costs.

- Larger than required antennas, high transmitter power levels and the like can add considerably to the cost, so it is necessary to balance these to minimise the cost of the system while still maintaining performance.

link budget – the basics - calculations

- a radio link budget is a summary of all the gains and losses in a transmission system.
- The radio link budget sums the transmitted power along with the gains and loses to determine the signal strength arriving at the receiver input.
- In essence the link budget will take the form of the equation below:

$$\text{Received power (dBm)} = \text{Transmitted power (dBm)} + \text{Gains (dB)} - \text{Losses (dB)}$$

basic calculation

- The basic calculation to determine the link budget is quite straightforward. It is mainly a matter of accounting for all the different losses and gains between the transmitter and the receiver.

- Once the link budget has been calculated, then it is possible to compare the calculated received level with the parameters for the receiver to discover whether it will be possible to meet the overall system performance requirements of signal to noise ratio, bit error rate, etc.

A typical link budget equation for a wireless communications system

- A typical link budget equation for a wireless communications system may look like the following:

$$P_{RX} = P_{TX} + G_{TX} + G_{RX} - L_{TX} - L_{FS} - L_{P} - L_{RX}$$

- Where:
  - P_{RX} =received power (dBm)
  - P_{TX} =transmitter output power (dBm)
  - G_{TX} = transmitter antenna gain (dBi)
  - G_{RX} = receiver antenna gain (dBi)
  - L_{TX} = transmit feeder and associated losses (feeder, connectors, etc.) (dB)
  - L_{FS} = free space loss or path loss (dB)
  - L_{P} = miscellaneous signal propagation losses (these include fading margin, polarization mismatch, losses associated with medium through which signal is travelling, other losses...) (dB)
  - L_{RX} = receiver feeder and associated losses (feeder, connectors, etc.) (dB)

Equalizers

- Equalization compensates for inter symbol interference (ISI) created by multipath within time dispersive channels.
- If the modulation bandwidth exceeds the coherence bandwidth of the radio channel, ISI occurs and modulation pulses are spread in time.
- An equalizer within a receiver compensates for the average range of expected channel amplitude and delay characteristics. Equalizers must be adaptive since the channel is generally unknown and time varying.

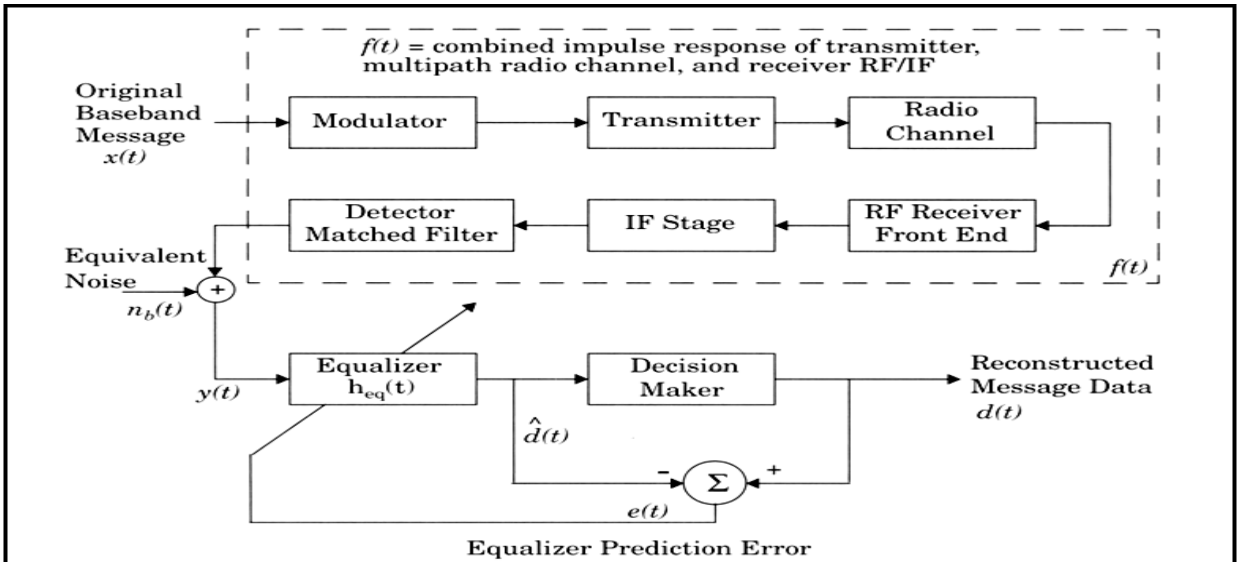


Figure Block diagram of a simplified communications system using an adaptive equalizer at the receiver.

There are two types of equalizers

(i)Linear Equalizer (ii) Non-Linear Equalizer

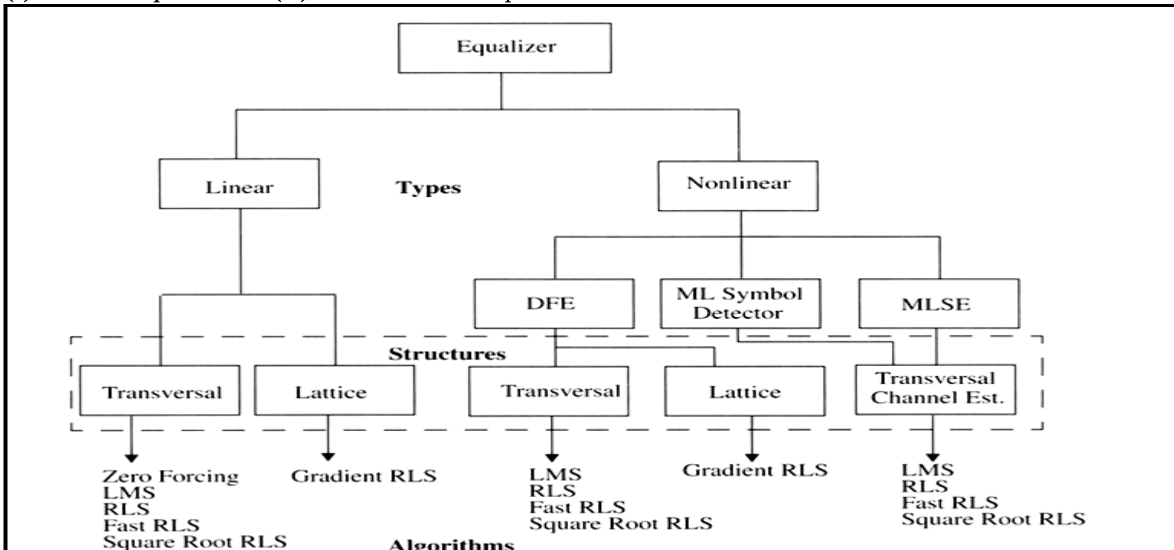


Figure Classification of equalizers.

Linear Equalizers

- A linear equalizer can be implemented as an FIR filter. It is known as transversal filter. This type of equalizer is the simplest form of available equalizer. The adaptive equalizer falls into the category of linear equalizers.
- There are two phases of an adaptive equalizer. The **training phase** and the **testing phase**. The adaptive equalizers consist of an adaptive filter to which the input is given. The adaptive algorithm is fed into the processing block according to the desired output. The weights of the adaptive filter vary so that the error signal generated becomes zero.
- In **the training phase** the equalizer is trained to operate over a known set of inputs. These are the set of sequences that are first sent by the transmitter which is also known to the receiver. So based on these known inputs the error is computed and the weights of the adaptive filter is varied to suite the characteristics of the channel. The channel is estimated during the training phase.
- In **the testing phase** the adaptive filter coefficients are formed and the input is given to the adaptive equalizer to find out the error and the estimated or the detected symbol is given as output from the system.

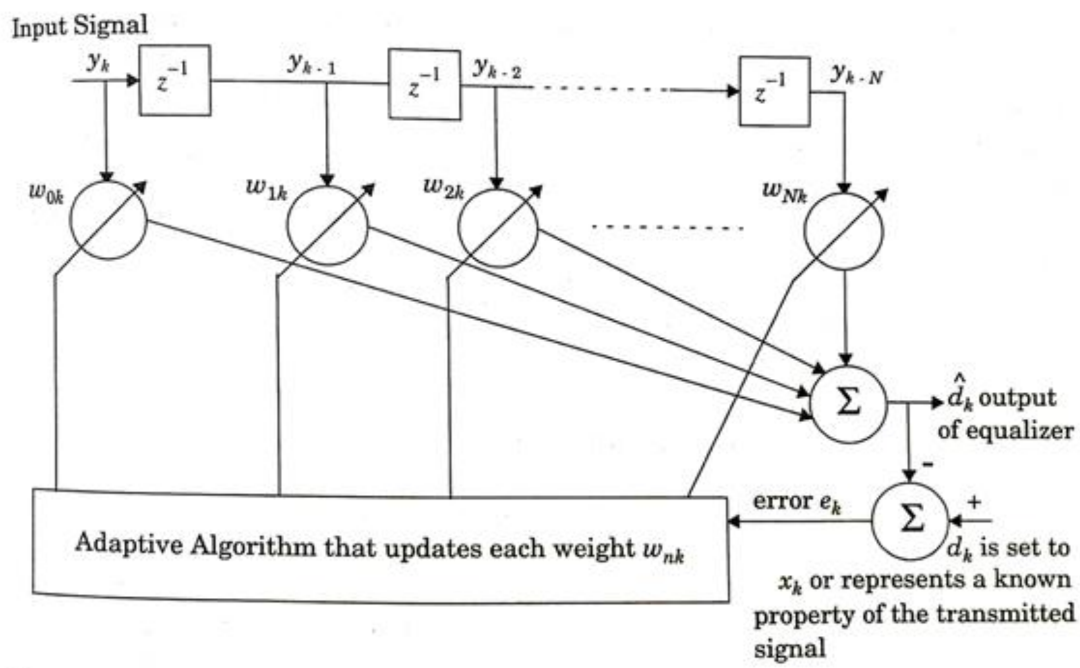


Figure A basic linear equalizer during training.

- Y_k is the input given to the system. W_k is the weights of the adaptive filter. The output of the equaliser and the set of known symbols are subtracted and the error is computed. This error is one of the inputs to the algorithm. The weights are continuously updated by the adaptive algorithm.

New weights = Previous weights + (constant)X(Previous error)X(Current input vector)
Previous error = Previous desired output – previous actual output.

- Different algorithms are run in the processing area according to the requirements. The requirements maybe accuracy in results, time constraint or resources constraint. This is the working of an adaptive equaliser.

Non Linear Equalizers

- Decision Feedback Equalizer
- Maximum likelihood Symbol Detection
- Maximum likelihood Sequence Detection

Linear equalizers do not perform well on channels which have deep spectral nulls in the pass band. This is overcome by non-linear equalizers.

Decision Feedback Equalizer

- The basic idea behind decision feedback equalization is that once an information symbol has been detected and decided upon, ISI that it induces on future symbols can be estimated and subtracted out before detection of subsequent symbols.
- The DFE can be realized in either the direct transversal form or as a lattice filter. It consists of a feed forward filter (FFF) and a feedback filter (FBF).
- The FBF is driven by decisions on the output of the detector, and its coefficients can be adjusted to cancel the ISI on the current symbol from past detected symbols.
- The equalizer has $N_1 + N_2 + 1$ taps in the feed forward filter and N3 taps in the feedback filter, and its output can be expressed as:

$$\widehat{d_k} = \sum_{n=-N_1}^{N_2} c_n^* \; y_{k-n} + \sum_{n=1}^{N_3} F_i \, d_{k-i}$$

where c_n^* and y_n, are tap gains and the inputs, respectively, to the forward filter, F_i are tap gains for the feedback filter, and d_i ($i < k$) is the previous decision made on the detected signal. That is, once the estimate of d_k is obtained using equation, d_k is decided from it. Then, d_k along with previous decisions d_{k-1}, d_{k-2}, are fed back into the equalizer, and estimate of d_{k+1} is obtained using equation.

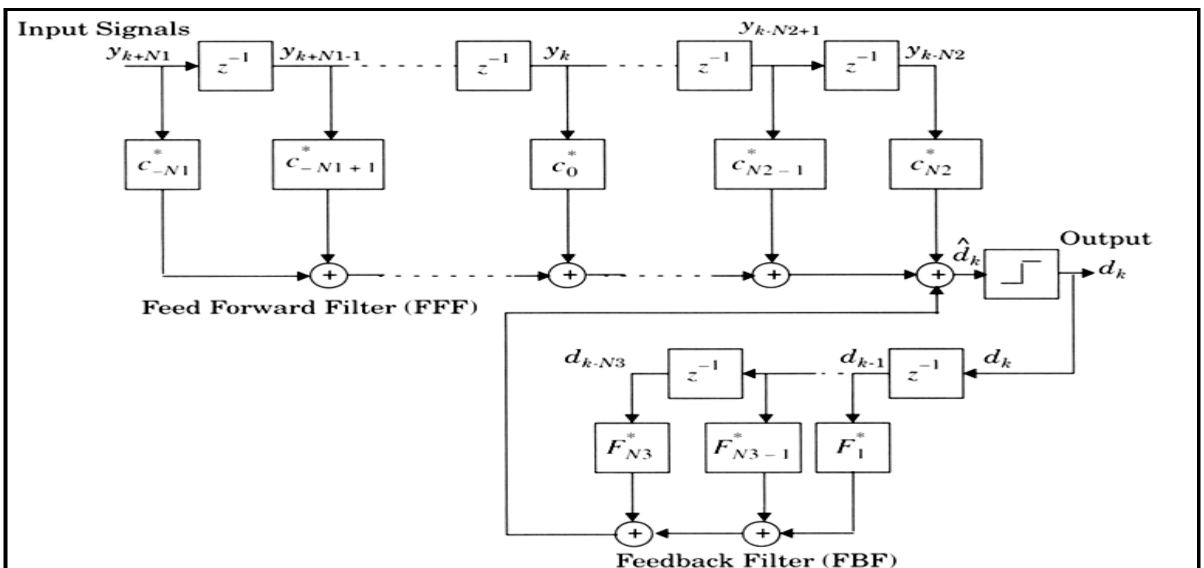


Figure Decision feedback equalizer (DFE).

- The above given is the diagram of the DFE. The DFE contains of two blocks the Feed Forward Filter and the Feedback Filter. The FFF part is same as that of the linear equalizer. But the FBF is the new addition to the DFE. The FBF is driven by the output of the detector and adds accuracy to the output computed.

- The advanced version of the DFE is the predictor DFE. Here the FBF part gets the input from the Output decision device as well as the error signal and thus the FBF acts as noise predictor because it predicts the noise and the residual ISI contained in the signal.

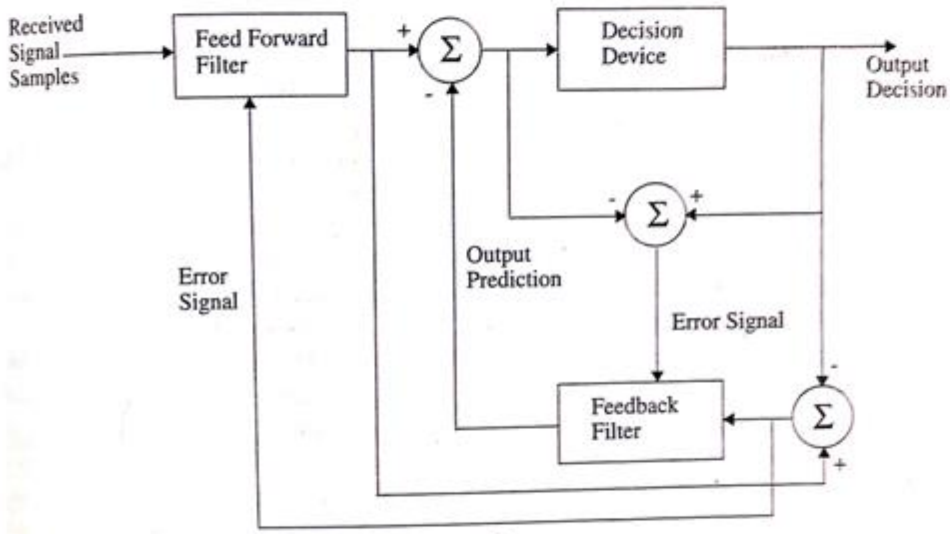


Figure . Predictive decision feedback equalizer.

Maximum Likelihood Sequence Estimation

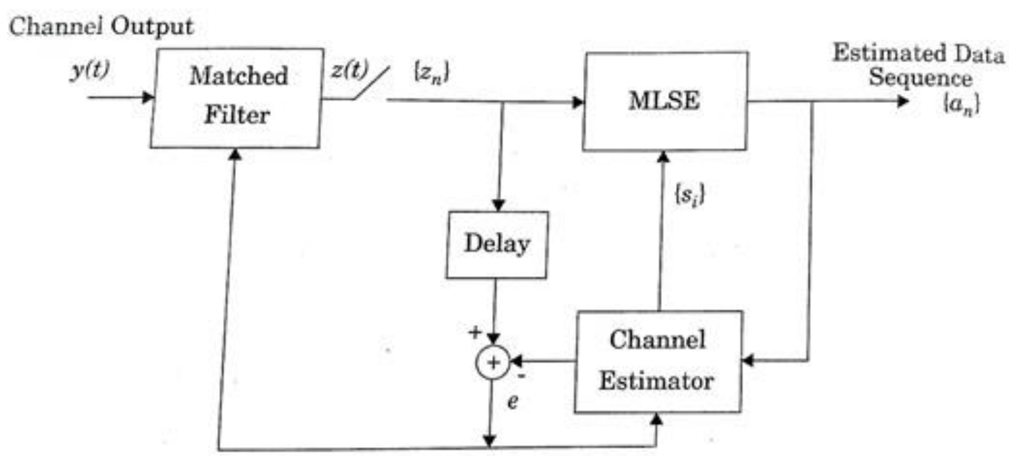


Figure The structure of a maximum likelihood sequence estimator (MLSE) with an adaptive matched filter.

- Using the channel impulse response simulator within the algorithm, the MLSE tests all possible data sequences and chooses the data sequence with the maximum probability as the output.
- The block diagram of the MLSE receiver based on the DFE is shown.
- The MLSE minimizes the probability of a sequence error. The channel characteristics are to be known in order to determine the noise corrupting the signal.
- The matched filter operates on a continuous signal whereas the MLSE on a discrete signal. The conversion of continuous to discrete signal is done by the switch.
- The channel estimator estimates the characteristics of the channel and the error signal is given as output. The delayed version of the input is given to the summer and the output is given as feedback input to the matched filter. This is the working of the MLSE equalizer.

Algorithms for Adaptive Equalization

The performance of an algorithm is determined by various factors such as

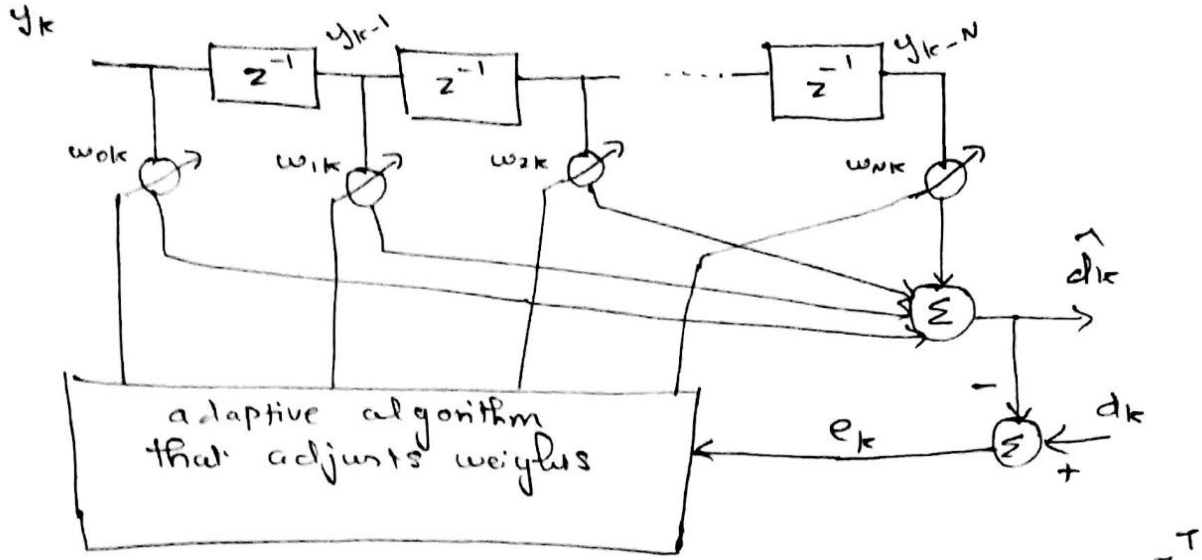
- **Rate of Convergence** – It is defined as the number of iterations required for the algorithm, in response to stationary inputs, to converge closely to the optimum solution
- **Misadjustment** – The level of deviation of the produced result from that of the optimum result.
- **Computational Complexity** – This is the number of operations required for one complete iteration of the algorithm to finish.
- **Numerical properties** – The stability of the algorithm is affected by factors like round-off noise and representation errors in the computer, these are classified as numerical properties.

Different types of algorithms exists for adaptive eualization. The classical algorithms are:

- Zero forcing algorithm
- Least mean square algorithm
- Recursive least square algorithm

equalizer.

Adaptive equalizer : Basic structure is shown in figure.



Input to the equalizer $\mathbf{y}_k = [y_k, y_{k-1}, y_{k-2} \dots y_{k-N}]^T$

$$\hat{d}_k = \sum_{n=0}^{N} w_{nk} y_{k-n}$$

$$\mathbf{w}_k = [w_{0k}, w_{1k}, w_{2k} \dots\dots w_{Nk}]^T$$

$$\hat{d}_k = \mathbf{y}_k^T \mathbf{w}_k = \mathbf{w}_k^T \mathbf{y}_k$$

d_k = desired output = x_k

error signal $e_k = d_k - \hat{d}_k = x_k - \hat{d}_k = x_k - \mathbf{w}_k^T \mathbf{y}_k$ —— (1)

$$|e_k|^2 = x_k^2 + \mathbf{w}_k^T \mathbf{y}_k \mathbf{y}_k^T \mathbf{w}_k - 2 x_k \mathbf{y}_k^T \mathbf{w}_k \quad \text{—— (2)}$$

Mean Square error = MSE $= E[|e_k|^2]$

$$MSE = E[x_k^2] + \mathbf{w}_k^T E[\mathbf{y}_k \mathbf{y}_k^T] \mathbf{w}_k - 2 E[x_k \mathbf{y}_k^T] \mathbf{w}_k$$

$$E[x_k \mathbf{y}_k^T] = E[x_k y_k \;\; x_k y_{k-1} \;\; x_k y_{k-2} \dots x_k y_{k-N}]^T = \mathbf{P} \quad \text{—— (3)}$$

↳ cross correlation vector

$$E[\mathbf{y}_k \mathbf{y}_k^T] = E \begin{bmatrix} y_k^2 & y_k y_{k-1} & y_k y_{k-2} & \dots & y_k y_{k-N} \\ y_{k-1} y_k & y_{k-1}^2 & \dots & \dots & y_{k-1} y_{k-N} \\ \vdots & \vdots & & & \\ & & & \dots & y_{k-N}^2 \\ y_{k-N} y_k & & & & \end{bmatrix} = \mathbf{R}$$

↳ input correlation matrix. —— (4)

If x_k and y_k are stationary, elements of R and P are second order statistics which do not vary with time.

Mean square error $= \xi = E[x_k^2] + W^T R W - 2P^T W$

By minimizing above equation, interms of weight vector W_k, it is possible to adaptively tune the equalizer to provide a flat spectral response (minimal ISI) in the received signal.

To find optimum weight vector $\hat{W}$ for minimum mean square error (MMSE)

$$\nabla = \frac{\partial \xi}{\partial w} = 0 \quad \text{where } \xi = E[x_k^2] + W^T R W - 2P^T W$$

$$\nabla = 2RW - 2P = 0$$

$$RW = P$$

$$\hat{W} = R^{-1} P = \text{optimum weight vector} \quad \text{—— (5)}$$

$$\xi_{min} = E[x_k^2] - P^T \hat{W} = E[x_k^2] = P^T R^{-1} P$$

Adaptation Algorithms

In order to directly solve equation (5) find the optimum equalizer weight, we can

This requires extensive complex operations.

To ease the computational burden, iterative algorithms have been developed.

Important factors that describe the quality of an iterative algorithm are:

Convergence rate: This gives the number of iterations that are required to closely approximate the final result.

It is usually assumed that the channel does not change during the iteration period. If the algorithm converges too slowly, it will never reach a stable state — the channel has changed before the algorithm has converged.

Misadjustment: the size of deviation of the converged state of the iterative algorithm from the exact MSE solution.

Computational complexity - Number of operations required to make one complete iteration of the algorithm.

Stability - Inaccuracies are produced due to round-off of values. These influence the stability of the algorithm.

Least Mean Square Algorithm (Stochastic gradient algorithm)

$\hat{w} = R^{-1} p$ ⟹ optimum weight vector for minimum mean square error.

$\hat{w}_N = R_{NN}^{-1} P_N$

The equalizer weights are updated by the update equations given below:

$$\hat{d}_k(n) = W_N^T(n)\, y_N(n)$$

$$e_k(n) = x_k(n) - \hat{d}_k(n)$$

$$W_N(n+1) = W_N(n) - \alpha\, e_k^*(n)\, y_N(n)$$

where

$n \rightarrow$ sequence of iteration

$N \rightarrow$ number of delay stages in the equalizer

$\alpha \rightarrow$ step size that controls the convergent rate and stability of the algorithm.

To prevent the adaption from being unstable, the value of α is chosen from

$$0 < \alpha < \frac{2}{\sum_{i=1}^{N} \lambda_i}$$

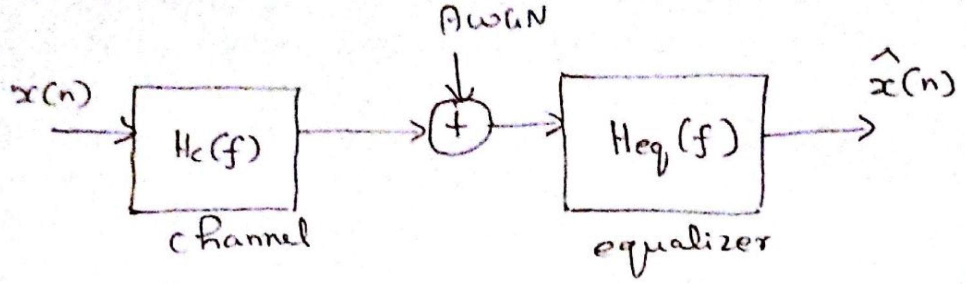
where λ_i is the eigenvalue of the covariance matrix R_{NN}.

since $\sum_{i=1}^{N} \lambda_i = y_N^T(n)\, y_N(n)$, the step size can be controlled by the total input power.

LMS algorithm is the simplest algorithm and requires only $2N+1$ operations per iteration.

Scanned by CamScanner

Zero forcing equalizer



$$H_{eq}(f) = \frac{1}{H_c(f)}$$

$$H_{eq}(f) \cdot H_c(f) = 1$$

If the frequency response for a particular channel is $H_c(f)$, then the zero forcing equalizer is constructed such that,

$$H_{eq}(f) = \frac{1}{H_c(f)}$$

where, $H_{eq}(f)$ is the frequency response of the equalizer.

The combination of channel and equalizer gives a flat frequency response and linear phase.

ie, $H_{eq}(f)\, H_c(f) = 1$

Zero forcing equalizer removes all interference. ZFA works well when channel is noiseless. But when the channel is noisy, it will amplify the noise depending on the channel magnitude.

UNIT II –CELLULAR ARCHITECTURE

Multiple Access Techniques used in Wireless Mobile Communication

Multiple access schemes allow many mobile users to share a finite amount of radio spectrum simultaneously. Sharing of spectrum is done to achieve high capacity. For high quality communications, sharing of the available bandwidth has to be done without severe degradation in the performance of the system.

In wireless communication system, it is often desirable to allow the subscriber to send information to the BS, while receiving information from the BS at the same time.

Major Access Techniques

- **Frequency Division Multiple Access (FDMA)**
- **Time Division Multiple Access (TDMA)**
- **Code Division Multiple Access (CDMA)**

Frequency Division Multiple Access (FDMA):

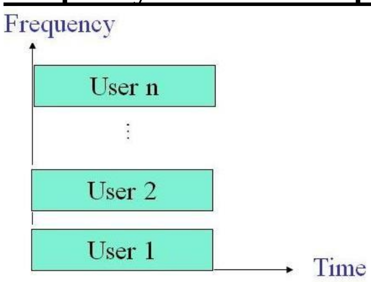


fig. 1. FDMA

FDMA assigns individual channels to individual users**.**

Each user is allocated a unique band or channel.

FDMA/FDD: Users are assigned a channel as a pair of frequencies; one frequency for forward channel and other frequency for reverse channels. Eg.AMPS

FDMA/TDD: 2 simplex time slots on the same frequency; one slot for forward channel and other slot for reverse channel. Eg. DECT

Features of FDMA:

- FDMA channel carries only one phone circuit at a time.
- If FDMA channel is not in use, it is idle and cannot be used by other users- wasted resource.
- After the assignment of channel, BS and mobile transmit simultaneously and continuously.
- Bandwidth of FDMA channels is relatively narrow.
- Symbol time of a narrow band signal is large as compared to the average delay spread (Amount of ISI is low, so no equalization is required).
- Complexity of FDMA mobile system is lower compared to TDMA system.
- Fewer bits are needed for overhead purpose compared to TDMA (synchronization, framing).
- System cost is high compared to TDMA (costly band pass filters).

- FDMA uses duplexers since both transmitter and receiver operate at the same time. This results in increase in the cost of FDMA subscriber units and BS.
- FDMA requires tight RF filtering to minimize adjacent channel interference (ACI).

<u>Number of channels in FDMA system:</u>

$$N = \frac{B_t - B_{guard}}{B_c}$$

B_t = Total spectrum allocated

B_{guard} = spectrum allocated at edge of band

B_c = Channel Bandwidth

Advantages of FDMA:

- ISI is low, so no equalization is required
- Complexity of FDMA mobile system is lower compared to TDMA system.
- Fewer bits are needed for overhead purpose compared to TDMA

Disadvantages in FDMA system:

- ❖ **Power amplifiers** are operated at or near saturation for maximum power efficiency and are **nonlinear. Non linearity's** cause signal spreading, thereby generating **Intermodulation frequencies (IM).** These **IM** frequencies interfere with other channels in FDMA system.
- ❖ **System cost is high.**

<u>Time Division Multiple Access (TDMA):</u>

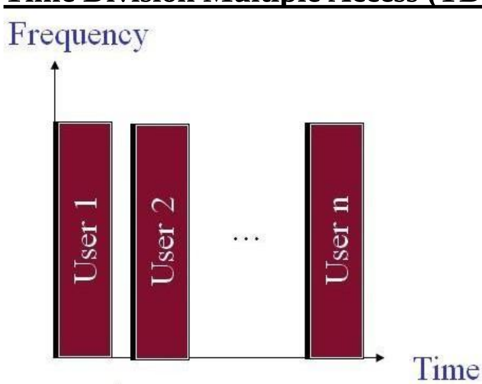


fig.2. TDMA

- ❖ TDMA divides the radio spectrum into time slots and in each slot, only one user is allowed to either transmit or receive.
- ❖ Particular time slot reoccurs every frame.

❖ 'N' time slots comprise a frame.

❖ Transmission for any user is non continous(buffer and burst method)

❖ Each frame is made up of a preamble, tail bits and N time slots.

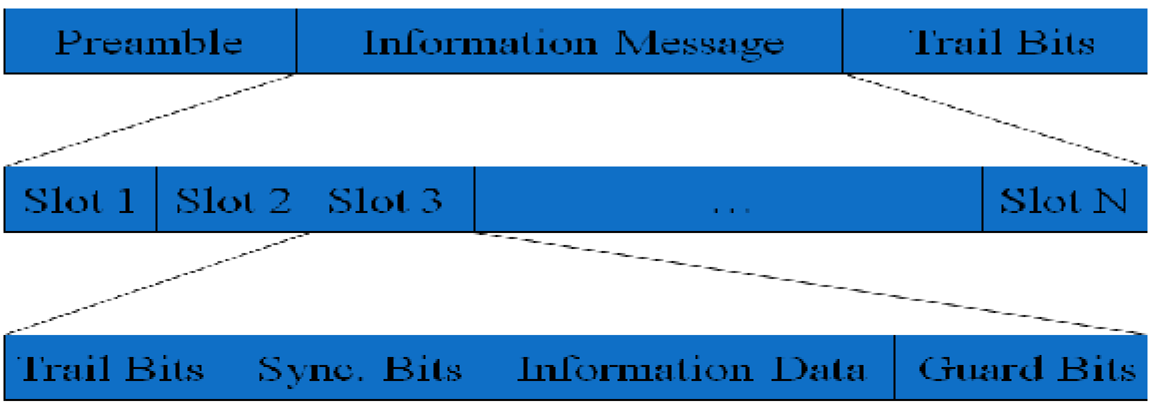
TDMA/TDD:

- Half the time slots used for forward link
- Half the time slots used for reverse link

TDMA/FDD:

- Carrier frequencies are different for forward and reverse link

TDMA frame structure



The frame is cyclically repeated over time
Preamble: contains address and synchronization information (Identification of BS and subscriber)
Guard Time: used to allow synchronization of the receivers between different time slots and frames.

Features of TDMA:

- TDMA shares a single carrier frequency with several users-each user makes use of non-overlapping time slots.
- Data transmission is not continuous. Subscriber transmitter can be turned off when not in use- results in low battery consumption.
- Because of non-continuous transmission, hand off process is much simpler.
- TDMA user's use different time slots for transmission and reception. Hence Duplexers are not required.
- Transmission rates are generally very high compared to FDMA channels. Hence **adaptive Equalization** required.
- Guard time should be minimized to avoid ACI.

- High synchronization overhead required because of burst transmission.

Advantage of TDMA:

It is possible to allocate different number of time slots per frame to different users. Thus, bandwidth can be supplied on demand, by concatenating or reassigning time slots based on priority.

Efficiency of TDMA:

It is a measure of the % of transmitted data that contains information as opposed to the overhead bits

$$\eta_f = \left[1 - \frac{b_{OH}}{b_T}\right] x\, 100\%$$

$$b_{OH} = N_r.b_r + N_t.b_p + N_t.b_g + N_r.b_g$$

N_r = No of reference bursts/frame

b_r = No of overhead bits / reference bursts

N_t = Total no of traffic bursts/frame

b_p = No of overhead bits / preamble in each slot

b_g = No of equivalent bits in each guard time interval

b_T = Total no of bits/frame

Number of channels in TDMA system:

$$N = \frac{m(B_{tot} - B_{guard})}{B_c}$$

B_{tot} = Total spectrum BW

B_{guard} = guard band

B_c = Channel Bandwidth

Spread Spectrum Multiple access (SSMA)

SSMA uses signals that have a transmission bandwidth much greater than the minimum required RF BW.

A pseudo noise (PN) sequence converts NB signal to WB noise like signal before transmission.

It provides immunity to multipath interference.

SSMA is BW efficient in multiple user environments.

SSMA	
Frequency hopped Multiple access	**Direct sequence Multiple access**
Carrier frequencies of individual users are varied in a pseudo random fashion within a WB channel	NB message signal is multiplied by a very large BW signal called spreading signal
Digital data of each user is broken into uniform sized bursts, which are transmitted on different channels within the allocated spectrum band.	Spreading signal is a PN sequence, with a chip rate greater than data rate of message.

If rate of change of carrier frequency is greater than the (Fast frequency hopping system)	Many users of CDMA system share the same frequency. CDMA has a soft capacity limit(no absolute limit on the no of users)
If channel changes at a rate < or equal to symbol rate the (Slow frequency hopping system)	Multipath fading is reduced because signal is spread over a large spectrum. Channel data rates are very high

Disadvantages of CDMA

- **Self-Jamming** arises when spreading sequences of different users are not exactly orthogonal.
- **Near-Far Problem**

Efficient frequency allocation in a cellular Radio system

- As the demand for service increases, the number of base stations may be increased, thereby providing additional radio capacity with no additional increase in radio spectrum.
- For efficient use of the radio spectrum, a frequency reuse scheme which can (i) increase the capacity and (ii) minimize interference is required.
- In order to achieve these objectives we have

  (i) Fixed channel Assignment
  (ii) Demand channel Assignment

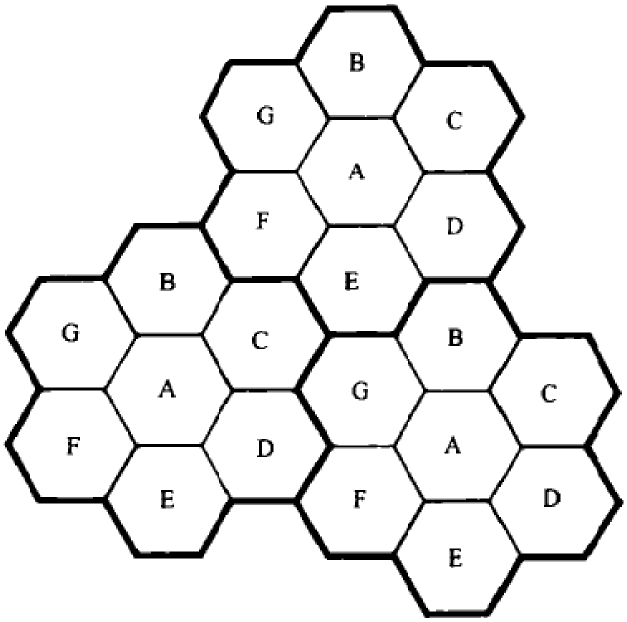
Frequency Reuse

The design process of selecting and allocating channel groups for all cellular BS's within a system is called **Frequency Reuse.**

- A relatively large market area called a coverage zone is divided into smaller section called cells.
- The area occupied by each cell is known as foot print.
- Each cell BS is allocated a group of channel frequencies that are different from those of neighboring cells.
- BS antennas are chosen to achieve a desired coverage pattern within its cell.
- As long as a coverage area is limited to within the cell boundaries, the same group of channel frequencies may be used in different cells, without interfering with each other, provided the 2 cells are at a sufficient distance from one another.

Illustration of Frequency Reuse

Each area is further divided into hexagonal shaped cells, that fit together to form a honey comb pattern.



- Hexagonal shaped cells are chosen because it provides most effective transmission by approximating a circular pattern while eliminating gaps inherently present between adjacent cells.
- When using hexagon to model coverage areas, BS transmitters are shown as either being in the center of the cell (center-excited cell) or on 3 of the 6 cell vertices (edge excited cells)
- Omni directional antennas are used on center-excited cells.
- Sectored antennas are used on corner-excited cells.
- Consider a cellular system which has a total of S duplex channels available. If each cell is allocated a group of 'k' channels(k<S) and if the S channels are divided among N cells into unique and disjoint channel groups , where each channel group have the same number of channels, then the total number of available radio channels is S='k.N
- N cells that collectively use the complete set of available frequencies is called a cluster.
- If a cluster is replicated 'M' times within the system, total number of duplex channels C=M k N=MS
- 'C' is a measure of capacity. Capacity of a cellular system is directly proportional to the no of times a cluster is replicated in a fixed service area.
- 'N' is called cluster size and is equal to 4,7,12.
- If cluster size 'N' is reduced, while cell size is kept constant, more clusters are required to cover a given area and hence 'C' is increased.
- Larger the values of 'N', co-channel cells are much closer leading to greater co-channel interference.

- Frequency reuse factor of a cellular system is 1/N.
- Cells use a hexagonal shape, which provides exactly 6 equidistant neighboring cells.
- Lines joining the centers of any cell with its neighboring cell are separated in multiples of 60º.
- To connect cells without gaps in between, no of cells per cluster can have only values that satisfy the equation $N = i^2 + j^2 + ij$

  N= No of cells/cluster; I &j= non negative integer value

Fixed Channel Assignment

- Each cell is allocated a pre-determined set of voice channels. Any call attempt within the cell can only be served by the unused channels in that cell.
- If all the channels in the cell are occupied, the call is blocked and subscriber does not receive service.
- In borrowing strategy, a cell is allowed to borrow channels from a neighboring cell, if all of its own channels are already occupied.

Dynamic Channel Assignment

- Voice channels are not allocated to different cells permanently.
- Each time a call request is made, the serving BS requests a channel from the MSC. The MSC then allocates a channel to the requesting cell.
- This reduces likelihood of blocking, which increases trunking capacity of the system.

Handoff Scenario at Cell boundary

Handoff: When a mobile moves into a different cell, while conversation is in progress, the MSC automatically transfers the call to a new channel of the new BS. The transfer of a mobile unit from one BS control to another BS control is called **handoff.**

- A handoff consists of 4 stages i) Initiation ii) Resource reservation iii) Execution iv)Completion
- Handoff operation involves
- Identifying new BS
- Requires voice and control signals to be allocated to channels of the new BS.
- Handoff operation must be performed successfully, as infrequently as possible and must be unnoticed to the users.

In order to meet these requirements

- System designers must specify an optimum signal level at which to initiate a handoff.
- A particular signal level is specified as the minimum usable signal for acceptable voice quality at the BS receiver (-90dbm to -100 dbm)

- A slightly stronger signal level than specified above is used as a threshold at which handoff is made.

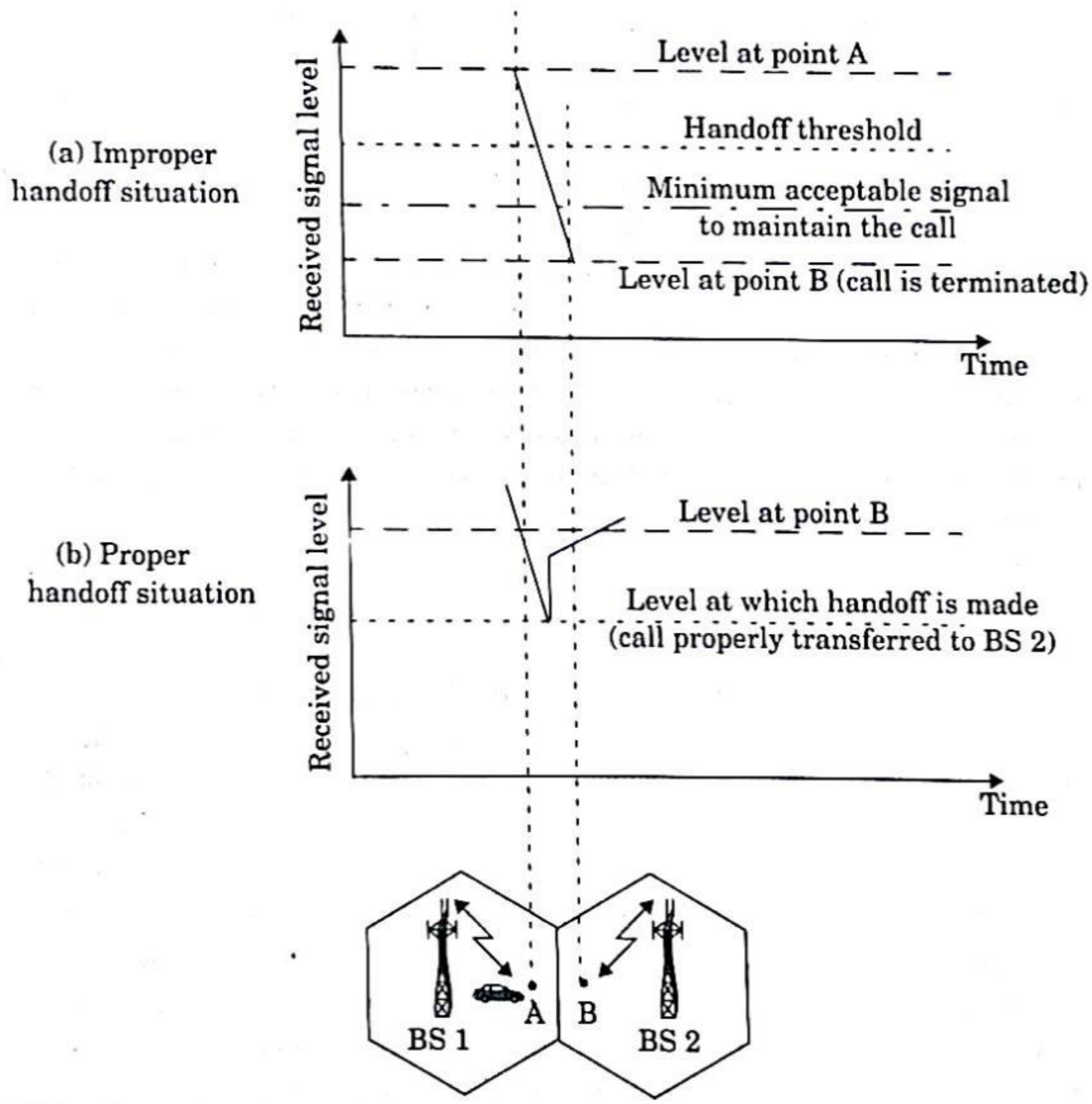


Fig: Illustration of a handoff scenario at cell boundary

- In situation (a), a handoff is not made and the signal drops below the minimum acceptable level. The call gets dropped.

Call can get dropped

- If there is excessive delay by MSC in doing a handoff.
- Excessive delay may occur during high traffic , due to computational loading at MSC.
- If no channels are available on any of the nearby BS's.

When to Handoff

- It is to be ensured that the drop in the measured signal level is not due to momentary fading.

- It is to be ensured that the mobile is actually moving from the serving BS.
- Time required to decide if a handoff is needed, depends on the speed at which vehicle is moving.

Dwell time:
Time over which a call maybe maintained within a cell, without handoff.
Factors that govern Dwell time:

- Propagation
- Interference
- Distance b/w subscriber and BS.

<u>In first generation analog cellular systems</u>

- Signal strength measurements are made by BS and supervised by MSC
- MSC decides if a hand off is necessary or not

<u>In 2G systems</u>
-Mobile assisted Handoff (MAHO)
-Every mobile station measures the received power from the surrounding base stations
-Handoff is initiated when the power received from the BS of a neighboring cell begins to exceed the power from the current BS, by a certain level or for a certain period of time
<u>Intersystem Handoff</u>
During a call, if a mobile moves from one cellular system to a different cellular system controlled by a different MSC, Intersystem Handoff becomes necessary.
<u>Prioritizing Handoffs</u>
Priority to handoffs is done by guard channel concept, whereby a fraction of the total available channels in a cell are reserved exclusively for handoff requests from on going calls which may be handed off into the cell.
<u>Queuing of handoff requests</u>
This is done to decrease the probability of forced termination of a call due to lack of available channels.

<u>Interference</u>

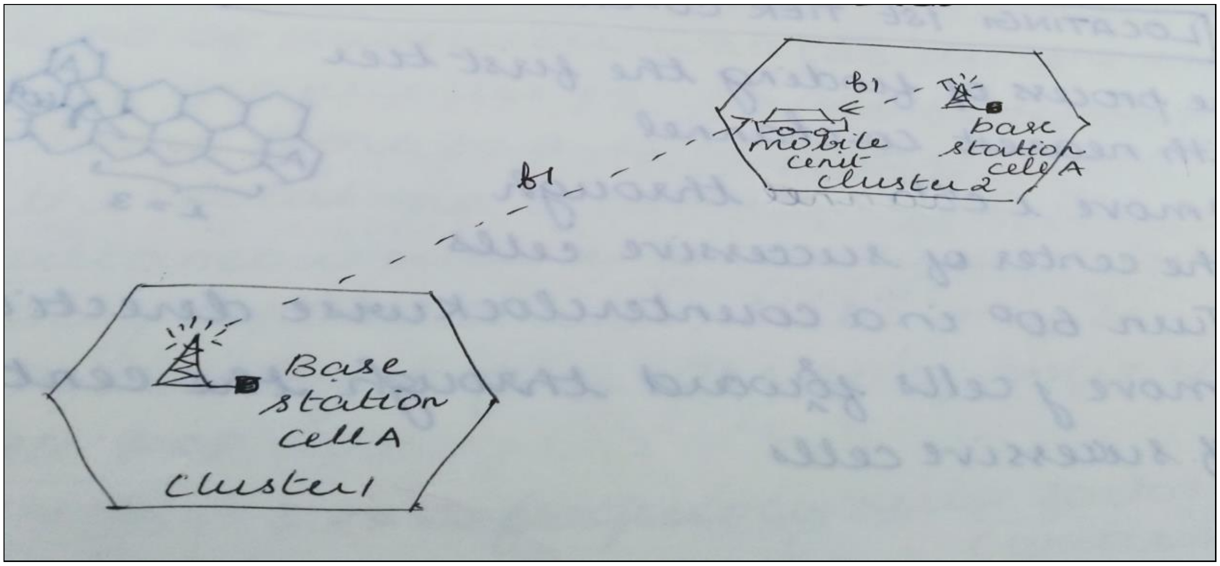
Major limiting factor in the performance of a cellular radio system is interference.
Sources of Interference:

- Another mobile in the same cell
- A call in progress in a neighboring cell
- Other base stations operating in the same frequency band
- Any non-cellular system which inadvertently leaks energy into the cellular band

Two major types of Interference:

1. **Co-Channel Interference**
2. **Adjacent Channel Interference**

<u>Co-Channel Interference</u>: In a given coverage area, there are several cells that use the same set of frequencies. These cells are called co-channel cells and the interference between signals from these cells is Co-Channel Interference.



(i) Base Station (BS) in cell A of cluster 1 is transmitting on frequency ‘f_1’.
(ii) At the same time BS in cell A of cluster 2 is transmitting the same frequency ‘f_1’.
(iii) Mobile in cluster ‘2’ is receiving signals of same frequency from two different BS’s.
(iv) Mobile is under control of BS in cluster ‘2’, but signal from cluster ‘1’ is received at a lower power level as Co-Channel Interference.

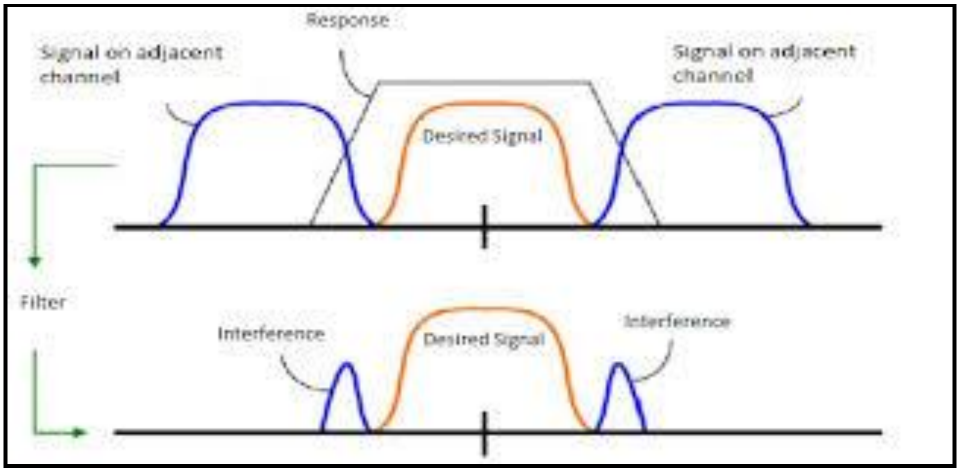
To reduce Co-Channel Interference:
Co-Channel cells must be physically separated by a minimum distance to provide sufficient isolation.

- When the size of each cell is approximately same and the BS’s transmit the same power, the Co-Channel Interference ratio is independent of the transmitted power and becomes a function of Radius of the cell (R) and distance between centers of the nearest Co-Channel cells (D).
- By increasing (D/R) ratio, the spatial separation between Co-Channel cells relative to the coverage distance of a cell is increased. Thus interference is reduced from improved isolation.
- Co-Channel Reuse ratio (Q) = D/R. Q is related to cluster size ‘N’. For hexagonal cell geometry Q = D/R = $\sqrt{3N}$
- Tradeoff Required

(i) Small Q provides larger capacity
(ii) Large Q provides improvement in transmission quality (small level of Co-Channel Interference). Tradeoff must be made between the two.

Adjacent Channel Interference (ACI):

Interference resulting from signals which are adjacent in frequency to the desired signal is called Adjacent Channel Interference.



Adjacent-channel interference (**ACI**) is interference caused by extraneous power from a signal in an adjacent channel. ACI may be caused by inadequate filtering, improper tuning or poor frequency control (in the reference channel, the interfering channel or both).

To reduce Adjacent Channel Interference:

- ❖ By using precise filtering and making careful channel assignments.
- ❖ By maintaining a reasonable frequency separation between channels in a given cell.

Trunking and Grade of Service

- Cellular radio systems rely on trunking to accommodate a large number of users in a limited radio spectrum. The concept of trunking allows a large number of users to share the relatively small number of channels in a cell by providing access to each user, on demand, from a pool of available channels.
- In a trunked radio system, each user is allocated a channel on a per call basis, and upon termination of the call, the previously occupied channel is immediately returned to the pool of available channels.
- Trunking exploits the statistical behavior of users so that a fixed number of channels or circuits may accommodate a large, random user community.
- In a trunked mobile radio system, when a particular user requests service and all of the radio channels are already in use, the user is blocked, or denied access to the system. In some systems, a queue may be used to hold the requesting users until a channel becomes available.

Grade of service (GOS)

The grade of service (GOS) is a measure of the ability of a user to access a trunked system during the busiest hour. The busy hour is based upon customer demand at the busiest hour during a week, month, or year.

It is the wireless designer's job to estimate the maximum required capacity and to allocate the proper number of channels in order to meet the GOS. GOS is typically given as the likelihood that a call is blocked, or the likelihood of a call experiencing a delay greater than a certain queuing time.

Trunking theory – common terms

Set-up Time: The time required to allocate a trunked radio channel to a requesting user.

Blocked Call: Call which cannot be completed at time of request, due to congestion. Also referred to as a lost call.

Holding Time: Average duration of a typical call. Denoted by H (in seconds).

Traffic Intensity: Measure of channel time utilization, which is the average channel occupancy measured in Erlangs. This is a dimensionless quantity and may be used to measure the time utilization of single or multiple channels. Denoted by A.

Load: Traffic intensity across the entire trunked radio system, measured in Erlangs.

Grade of Service (GOS): A measure of congestion which is specified as the probability of a call being blocked (for Erlang B), or the probability of a call being delayed beyond a certain amount of time (for Erlang C).

Request Rate: The average number of call requests per unit time. Denoted by λ seconds^{-1}.

The traffic intensity offered by each user is equal to the call request rate multiplied by the holding time. That is, each user generates a traffic intensity of A_u Erlangs given by

$$A_u = \lambda H$$

where H is the average duration of a call and λ is the average number of call requests per unit time. For a system containing U users and an unspecified number of channels, the total offered traffic intensity A, is given as

$$A = UA_u = U\lambda H$$

In a C channel trunked system, if the traffic is equally distributed among the channels, then the traffic intensity per channel, A_C, is given as,

$$A_C = UA_u / C$$

Types of trunked systems

- Blocked calls cleared
- Blocked Calls Delayed

Blocked calls cleared

The first type(Blocked calls cleared) offers no queuing for call requests. That is, for every user who requests service, it is assumed there is no setup time and the user is given immediate access to a channel if one is available. If no channels are available, the requesting user is blocked without access and is free to try again later. This type of trunking is called **blocked calls cleared** and assumes that calls arrive as determined by' a Poisson distribution.

Furthermore, it is assumed that there are an infinite number of users as well as the following:
(a) there are memoryless arrivals of requests, implying that all users, including blocked users, may request a channel at any time;

(b) the probability of a user occupying a channel is exponentially distributed, so that longer calls are less likely to occur as described by an exponential distribution; and

(c) there are a finite number of channels available in the trunking pool. This is known as an M/M/m/m queue, and leads to the derivation of the Erlang B formula (also known as the blocked calls cleared formula). The Erlang B formula determines the probability that a call is blocked and is a measure of the GOS for a trunked system which provides no queuing for blocked calls.

$$P_r(blocking) = \frac{\frac{A^C}{C!}}{\sum_{k=0}^{C} \frac{A^k}{k!}} = GOS$$

where C is the number of trunked channels offered by a trunked radio system and A is the total offered traffic.

Blocked Calls Delayed

The second kind of trunked system is one in which a queue is provided to hold calls which are blocked. If a channel is not available immediately, the call request may be delayed until a channel becomes available. This type of trunking is called **Blocked Calls Delayed**, and its measure of GOS is defined as the probability that a call is blocked after waiting a specific length of time in the queue. To find the GOS, it is first necessary to find the likelihood that a call is initially denied access to the system. The likelihood of a call not having immediate access to a channel is determined by the Erlang C formula.

$$P_r(delay > 0) = \frac{A^C}{A^C + C!\left(1 - \frac{A}{C}\right)\sum_{k=0}^{C-1} \frac{A^k}{k!}}$$

If no channels are immediately available the call is delayed, and the probability that the delayed call is forced to wait more than t seconds is given by the probability that a call is delayed, multiplied by the conditional probability that the delay is greater than t seconds. The GOS of a trunked system where blocked calls are delayed is given by

$$\begin{aligned} P_r(delay > t) &= P_r(delay > 0) P_r(delay > t \mid delay > 0) \\ &= P_r(delay > 0) \exp(-(C - A)t / H) \end{aligned}$$

The average delay D for all calls in a queued system is given by

$$D = P_r(delay > 0) \frac{H}{(C - A)}$$

Where the average delay for those calls which are queued is given by H/(C-A)

Trunking efficiency is a measure of the number of users which can be offered a particular GOS with a particular configuration of fixed channels. The way in which channels are grouped can substantially alter the number of users handled by a trunked system.

Improving Capacity in Cellular Systems

Techniques such as **cell splitting, sectoring, and coverage zone** approaches are used in practice to expand the capacity of cellular systems.

- Cell splitting allows an orderly growth of the cellular system.
- Sectoring uses directional antennas to further control the interference and frequency reuse

of channels.

- The zone microcell concept distributes the coverage of a cell and extends the cell boundary to hard-to-reach places.

While cell splitting increases the number of base stations in order to increase capacity, sectoring and zone microcells rely on base station antenna placements to improve capacity by reducing co-channel interference.

Cell Splitting

- Cell splitting is the process of subdividing a congested cell into smaller cells, each with its own base station and a corresponding reduction in antenna height and transmitter power.
- Cell splitting increases the capacity of a cellular system since it increases the number of times that channels are reused. By defining new cells which have a smaller radius than the original cells and by installing these smaller cells (called microcells) between the existing cells, capacity increases due to the additional number of channels per unit area.

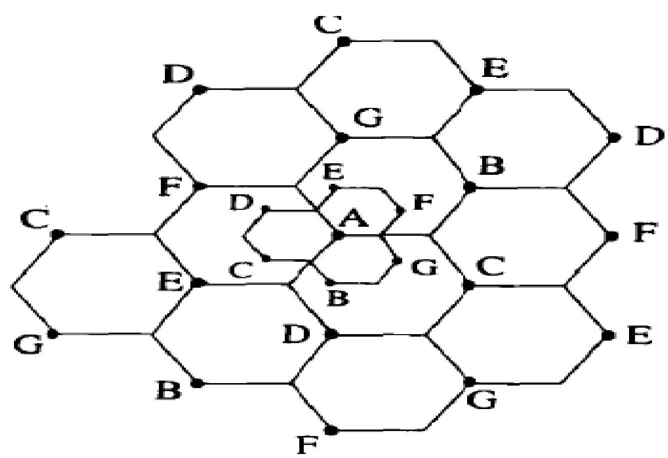


Illustration of cell splitting.

If every cell were reduced in such a way that the radius of every cell was cut in half. In order to cover the entire service area with smaller cells, approximately four times as many cells would be required. This can be easily shown by considering a circle with radius R. The area covered by such a circle is four times as large as the area covered by a circle with radius R/2. The increased number of cells would increase the number of clusters over the coverage region, which in turn would increase the number of channels, and thus capacity, in the coverage area. Cell splitting allows a system to grow by replacing large cells with smaller cells, while not upsetting the channel allocation scheme required to maintain the minimum co-channel reuse ratio Q between co-channel cells.

For the new cells to be smaller in size, the transmit power of these cells must be reduced. The transmit power of the new cells with radius half that of the original cells can be found by examining the received power P_r at the new and old cell boundaries and setting them equal to each other. This is necessary to ensure that the frequency reuse plan for the new microcells behaves exactly as for the original cells.

$$P_r\,(at\,old\,cell\,boundary) \propto P_{t1}\,R^{-n}$$

$$P_r\,(at\ new\,cell\,boundary) \propto P_{t2}\left(\frac{R}{2}\right)^{-n}$$

where P_{t1} and P_{t2} are the transmit powers of the larger and smaller cell base stations, respectively, and n is the path loss exponent. If we take $n = 4$ and set the received powers equal to each other, then

$$P_{t2} = \frac{P_{t1}}{16}$$

In other words, the transmit power must be reduced by 12 dB in order to fill in the original coverage area with microcells, while maintaining the S/I requirement.

In practice, not all cells are split at the same time. It is often difficult for service providers to find real estate that is perfectly situated for cell splitting.Therefore, different cell sizes will exist simultaneously.

In such situations, special care needs to be taken to keep the distance between co-channel cells at the required minimum and handoff issues must be addressed so that high speed and low speed traffic can be simultaneously accommodated.

When there are two cell sizes in the same region we can not simply use the original transmit power for all new cells or the new transmit power for all the original cells.

If the larger transmit power is used for all cells, some channels used by the smaller cells would not be sufficiently separated from co-channel cells. On the other hand, if the smaller transmit power is used for all the cells, there would be parts of the larger cells left unserved. For this reason, channels in the old cell must be broken down into two channel groups, one that corresponds to the smaller cell reuse requirements and the other that corresponds to the larger cell reuse requirements. The larger cell is usually dedicated to high speed traffic so that handoffs occur less frequently.

The two channel group sizes depend on the stage of the splitting process. At the beginning of the cell splitting process there will be fewer channels in the small power groups. However, as demand grows, more channels will be required, and thus the smaller groups will require more channels. This splitting process continues until all the channels in an area are used in the lower power group.

Antenna down tilting, which deliberately focuses radiated energy from the base station towards the ground (rather than towards the horizon), is often used to limit the radio coverage of newly formed microcells.

Sectoring

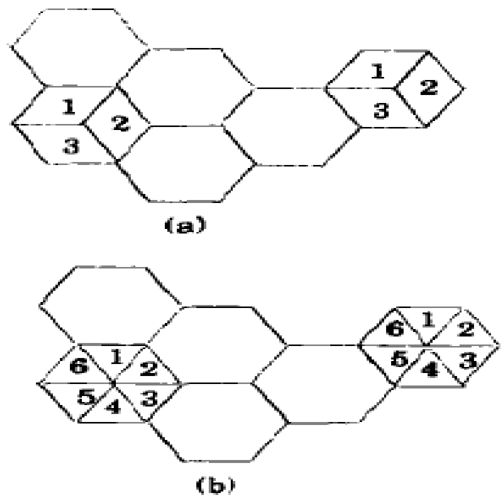
Cell splitting achieves capacity improvement by essentially rescaling the system. By decreasing the cell radius R and keeping the co-channel reuse ratio D/R unchanged, cell splitting increases the number of' channels per unit area.

Another way to increase capacity is to keep the cell radius unchanged and seek methods to decrease the D/R ratio. In this approach, capacity improvement is achieved by reducing the number of cells in a cluster and thus increasing the frequency reuse. However, in order to do this, it is necessary to reduce the relative interference without decreasing the transmit power.

The co-channel interference in a cellular system may be decreased by replacing a single omni-directional antenna at the base station by several directional antennas, each radiating within a specified sector. By using directional antennas, a given cell will receive interference and transmit with only a fraction of the available co-channel cells.

The technique for decreasing co-channel interference and thus increasing system capacity by using directional antennas is called ***sectoring***. The factor by which the co-channel interference is reduced depends on the amount of sectoring used.

A cell is normally partitioned into three 120° sectors or six 60° sectors as shown in Figure (a) and (b).



(a) 120° sectoring.
(b) 60° sectoring.

When sectoring is employed, the channels used in a particular cell are broken down into sectored groups and are used only within a particular sector.

Assuming 7-cell reuse, for the case of 120° sectors, the number of interferers in the first tier is reduced from 6 to 2. This is because only 2 of the 6 co-channel cells receive interference with a particular sectored channel group.

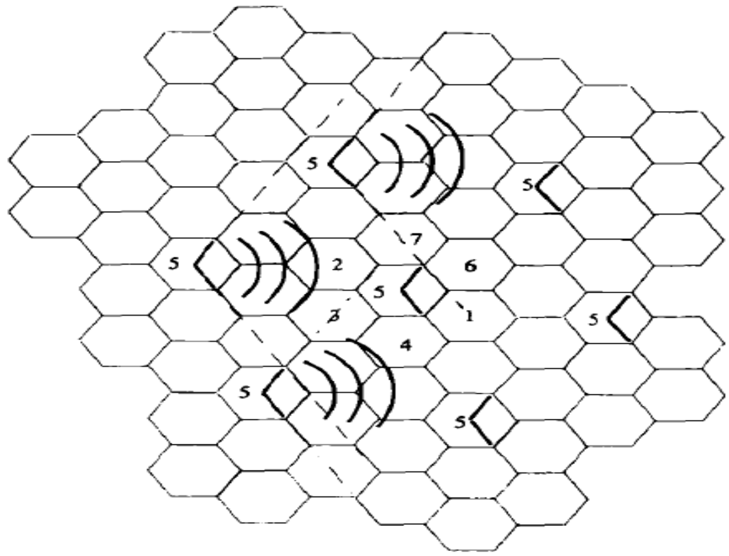


Illustration of how 120° sectoring reduces interference from co-channel cells. Out of the 6 co-channel cells in the first tier, only 2 of them interfere with the center cell. If omni-directional antennas were used at each base station, all 6 co-channel cells would interfere with the center cell.

Referring to Figure above, consider the interference experienced by a mobile located in the right-most sector in the center cell labeled 5. There are 3 co-channel cell sectors labeled "5" to the right of the center cell, and 3 to the left of the center cell. Out of these 6 co-channel cells, only 2 cells have sectors with antenna patterns which radiate into the center cell, and hence a mobile in the center cell will experience interference on the forward link from only these two sectors.

The resulting S/I for this case is found to be 24.2 dB, which is a significant improvement over the omni-directional case where the worst case S/I is 17 dB. In practical systems, further improvement in S/I is achieved by down tilting the sector antennas such that the radiation pattern in the vertical (elevation) plane has a notch at the nearest co- channel cell distance.

The improvement in S/I implies that with 120^o sectoring, the minimum required S/I of 18 dB can be easily achieved with 7-cell reuse, as compared to 12-cell reuse for the worst possible situation in the un sectored case. Thus, sectoring reduces interference, which amounts to an increase in capacity by a factor of 12/7.

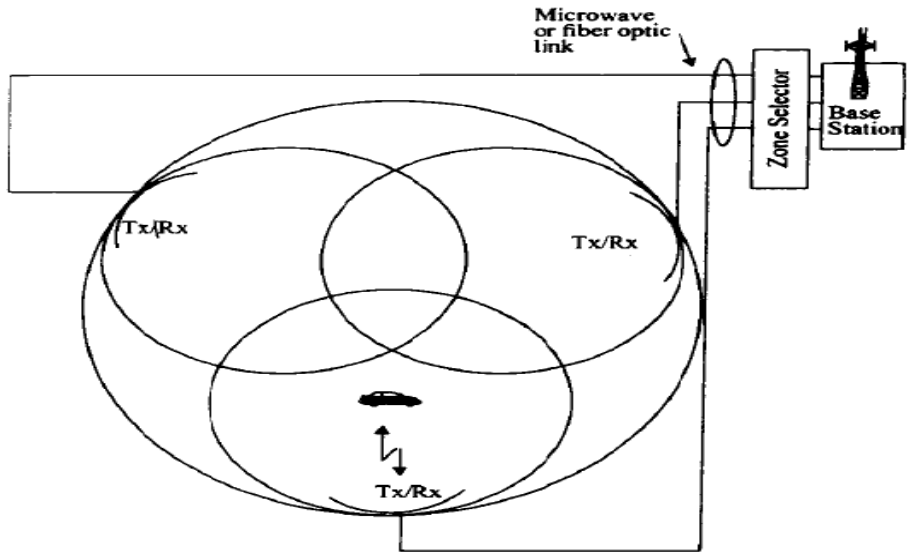
In practice, the reduction in interference offered by sectoring enable planners to reduce the cluster size N, and provides an additional degree of freedom in assigning channels.

<u>Draw Backs</u>

- An increased number of antennas at each base station
- A decrease in trunking efficiency due to channel sectoring at the base station.
- Since sectoring reduces the coverage area of a particular group of channels, the number of handoffs increases.

<u>Microcell Zone Concept</u>

The increased number of handoffs required when sectoring is employed results in an increased load on the switching and control link elements of the mobile system. Microcell zone concept give the solution to this problem.



The microcell concept

In this scheme, each of the three (or possibly more) zone sites (represented as Tx/Rx in Figure) are connected to a single base station and share the same radio equipment. The zones are connected by coaxial cable, fiberoptic cable, or microwave link to the base station. Multiple zones and a single base station make up a cell. As a mobile travels within the cell, it is served by the zone with the strongest signal. This approach is superior to sectoring since antennas are placed at the outer edges of the cell, and any base station channel may be assigned to any zone by the base station.

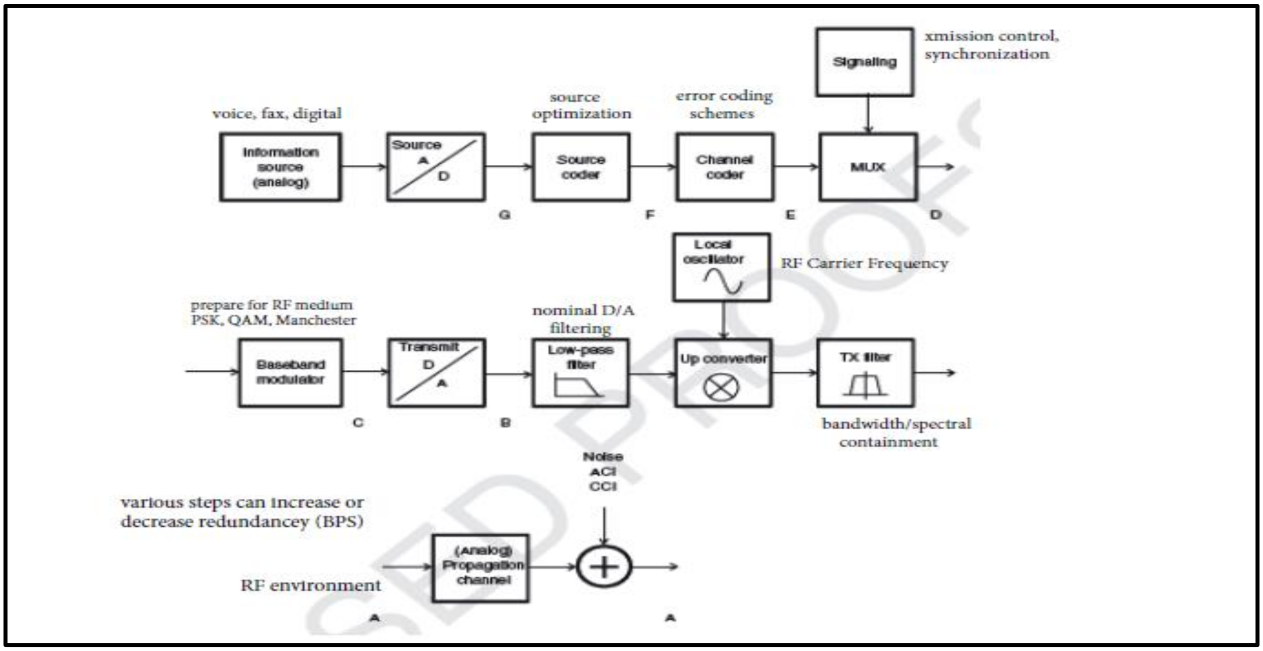
As a mobile travels from one zone to another within the cell, it retains the same channel. Thus, unlike in sectoring, a handoff is not required at the MSC when the mobile travels between zones within the cell. The base station simply switches the channel to a different zone site. In this way, a given channel is active only in the particular zone in which the mobile is traveling, and hence the base station radiation is localized and interference is reduced. The channels are distributed in time and space by all three zones and are also reused in co-channel cells in the normal fashion. This technique is particularly useful along highways or along urban traffic corridors.

The advantage of the zone cell technique is that while the cell maintains a particular coverage radius, the co-channel interference in the cellular system is reduced since a large central base station is replaced by several lower powered transmitters (zone transmitters) on the edges of the cell. Decreased co-channel interference improves the signal quality and also leads to an increase in capacity, without the degradation in trunking efficiency caused by sectoring.

For the same S/I requirement, this system provides a significant increase in capacity over conventional cellular planning.

Structure of a wireless communications link

WIRELESS COMMUNICATION SYSTEM TRANSMITTER

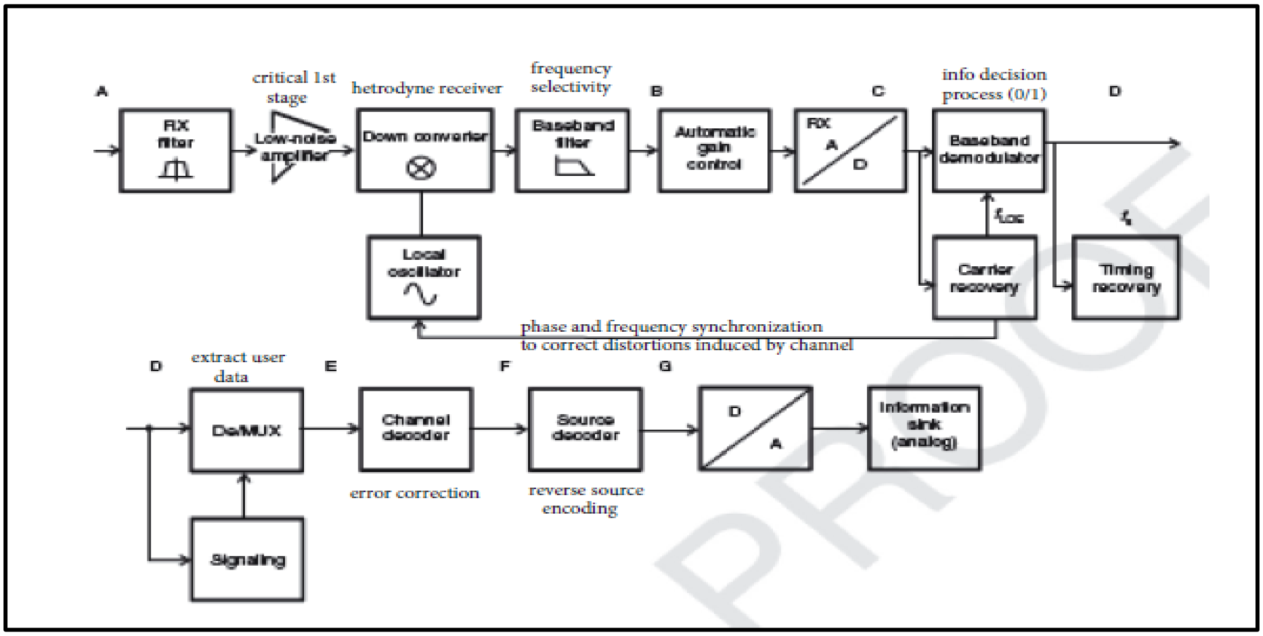


- The ***information source*** provides an analog source signal and feeds it into the *source ADC* (Analog to Digital Converter). This ADC first band limits the signal from the analog information source (if necessary), and then converts the signal into a stream of digital data at a certain sampling rate and resolution (number of bits per sample).
- The ***source coder*** uses a priori information on the properties of the source data in order to reduce redundancy in the source signal. This reduces the amount of source data to be transmitted, and thus the required transmission time and/or bandwidth.
- The ***channel coder*** adds redundancy in order to protect data against transmission errors.
- *Signaling* adds control information for the establishing and ending of connections, for associating information with the correct users, synchronization, etc.
- The ***multiplexer*** combines user data and signaling information. (The multiplexer at a Base Station (BS) combines the data from multiple users for transmission.)
- The *baseband* ***modulator*** assigns the gross data bits to complex transmit symbols in the baseband.
- The *TX* ***Digital to Analog Converter*** **(DAC)** generates a pair of analog, discrete amplitude voltages corresponding to the real and imaginary part of the transmit symbols, respectively.

• The ***analog low-pass filter*** in the TX eliminates the (inevitable) spectral components outside the desired transmission bandwidth.

- The ***TX Local Oscillator*** **(LO)** provides an unmodulated sinusoidal signal, corresponding to one of the admissible center frequencies of the considered system.
- The ***upconverter*** converts the analog, filtered baseband signal to a passband signal by mixing it with the LO signal.
- The ***RF TX filter*** eliminates out-of-band emissions in the RF domain.

WIRELESS COMMUNICATION SYSTEM RECEIVER



- The *(analog)* ***propagation channel*** attenuates the signal, and leads to delay and frequency dispersion.Furthermore, the environment adds noise And co-channel interference.
- The ***RX filter*** performs a rough selection of the received band.

• The ***low-noise amplifier*** amplifies the signal, so that the noise added by later components of the RX chain has less effect on the Signal-to-Noise Ratio (SNR).

- The ***RX LO*** provides sinusoidal signals corresponding to possible signals at the TX LO.

LOs at the TX and the RX produce oscillations with the same frequency and phase.

- The ***RX downconverter*** converts the received signal (in one or several steps) into baseband.
- The ***RX low-pass filter*** provides a selection of desired frequency bands for one specific user. It eliminates adjacent channel interference as well as noise.
- The ***Automatic Gain Control*** **(AGC)** amplifies the signal such that its level is well adjusted to the quantization at the subsequent ADC
- The **RX *ADC*** converts the analog signal into values that are discrete in time and amplitude

- *The* ***Carrier recovery*** determines the frequency and phase of the carrier of the received signal, and uses it to adjust the RX LO.

• The ***baseband demodulator*** obtains *soft-decision* data from digitized baseband data, and hands them over to the decoder.

• ***Symbol-timing recovery*** uses demodulated data to determine an estimate of the duration of symbols, and uses it to fine-tune sampling intervals.

• The ***decoder*** uses soft estimates from the demodulator to find the original (digital) source data.

- ***Carrier recovery*** determines the frequency and phase of the carrier of the received signal, and uses it to adjust the RX LO.

• The ***baseband demodulator*** obtains *soft-decision* data from digitized baseband data, and hands them over to the decoder.

• ***Symbol-timing recovery*** uses demodulated data to determine an estimate of the duration of symbols, and uses it to fine-tune sampling intervals.

• The ***decoder*** uses soft estimates from the demodulator to find the original (digital) source data.

- ***Signaling recovery*** identifies the parts of the data that represent signaling information and controls the subsequent demultiplexer.

• The ***demultiplexer*** separates the user data and signaling information and reverses possible time compression of the TX multiplexer.

• The ***source decoder*** reconstructs the source signal from the rules of source coding.

- If the source data are digital, the output signal is transferred to the data sink. Otherwise, the data are transferred to the DAC, which converts the transmitted information into an analog signal, and hands it over to the information sink.

Digital Modulation—an Overview

- Modern mobile communication systems use **digital modulation** techniques.
- Advancements in very large-scale integration (VLSI) and digital signal processing (DSP) technology have made digital modulation more cost effective than analog transmission systems.
- Digital modulation offers many **advantages** over analog modulation.

Some advantages include

- greater noise immunity and robustness to channel impairments

- easier multiplexing of various forms of information (e.g., voice, data, and video)

- greater security

- Furthermore, digital transmissions accommodate digital error-control codes which detect and/or correct transmission errors

- support complex signal conditioning and pro-cessing techniques such as source coding, encryption, and equalization to improve the performance of the overall communication link.

- New multipurpose programmable digital signal processors have made it possible to implement digital modulators and demodulators completely in software.

- Instead of having a particular modem design permanently frozen as hardware, embedded software implementations now allow alterations and improvements without having to redesign or replace the modem.

- In digital wireless communication systems, the modulating signal (e.g., the message) may be represented as a time sequence of symbols or pulses, where each symbol has **m** finite states.
- Each symbol represents **n** bits of information, where $n = \log_2 m$ bits/symbol.
- Many digital modulation schemes are used in modern wireless communication systems, and many more are sure to be introduced.
- Some of these techniques have subtle differences between one another, and each technique belongs to a family of related modulation methods.
- For example, phase shift keying (PSK) may be either coherently or differentially detected; and may have two, four, eight or more possible levels (e.g., n = 1, 2, 3, or more bits) per symbol, depending on the manner in which information is transmitted within a single symbol.

Factors That Influence the Choice of Digital Modulation

- Several factors influence the choice of a digital modulation scheme.
- A desirable modulation scheme

-provides low bit error rates at low received signal-to noise ratios

- performs well in multipath and fading conditions

- occupies a minimum of bandwidth, and

-is easy and cost-effective to implement.

- Existing modulation schemes do not simultaneously satisfy all of these requirements.
- Some modulation schemes are better in terms of the bit error rate performance, while others are better in terms of bandwidth efficiency.
- Depending on the demands of the particular application, trade-offs are made when selecting a digital modulation.
- The performance of a modulation scheme is often measured in terms of its **power efficiency** and **bandwidth efficiency.**

Power efficiency

Power efficiency describes the ability of a modulation technique to preserve the fidelity of the digital message at low power levels. In a digital communication system, in order to increase noise immunity, it is necessary to increase the signal power. However, the amount by which the signal power should be increased to obtain a certain level of fidelity (i.e., an acceptable bit error probability) depends on the particular type of modulation employed.

The **power efficiency**, η_p(sometimes called energy efficiency) of a digital modulation scheme is a measure of how favourably this tradeoff between fidelity and signal power is made, and is often expressed as the ratio of the signal energy per bit to noise power spectral density (E_b/N_0) required at the receiver input for a certain probability of error (say 10^{-5}).

Bandwidth efficiency

Bandwidth efficiency describes the ability of a modulation scheme to accommodate data within a limited bandwidth. In general, increasing the data rate implies decreasing the pulse width of a digital symbol, which increases the bandwidth of the signal.

Thus, there is an unavoidable relationship between data rate and bandwidth occupancy. However, some modulation schemes perform better than the others in making this tradeoff.

Bandwidth efficiency reflects how efficiently the allocated bandwidth is utilized and is defined as the ratio of the throughput data rate per Hertz in a given bandwidth. If R is the data rate in bits per second, and B is the bandwidth occupied by the modulated RF signal, then bandwidth efficiency is expressed as

$$\eta_B = \frac{R}{B} \quad bps/Hz$$

The **system capacity** of a digital mobile communication system is directly related to the **bandwidth efficiency** of the modulation scheme, since a modulation with a greater value of η_B will transmit more data in a given spectrum allocation.

There is a fundamental upper bound on achievable bandwidth efficiency. Shannon's channel coding theorem states that for an arbitrarily small probability of error, the maximum possible bandwidth efficiency is limited by the noise in the channel, and is given by the channel capacity formula,

$$\eta_{Bmax} = \frac{C}{B} = log_2\left(1 + \frac{S}{N}\right)$$

where C is the channel capacity (in bps), B is the RF bandwidth, and S/N is the signal-to-noise ratio.

- In the design of a digital communication system, very often there is a tradeoff between bandwidth efficiency and power efficiency.

- For example, adding error control coding to a message increases the bandwidth occupancy (and this, in turn, reduces the bandwidth efficiency), but at the same time reduces the required received power for a particular bit error rate, and hence trades bandwidth efficiency for power efficiency.

- On the other hand, higher level modulation schemes (M-ary keying) decrease bandwidth occupancy but increase the required received power, and hence trade power efficiency for bandwidth efficiency.

- While power and bandwidth efficiency considerations are very important, other factors also affect the choice of a digital modulation scheme.

- For example, for all personal communication systems which serve a large user community, the cost and complexity of the subscriber receiver must be minimized, and a modulation which is simple to detect is most attractive.

- The performance of the modulation scheme under various types of channel impairments such as Rayleigh and Rician fading and multipath time dispersion, given a particular demodulator implementation, is another key factor in selecting a modulation.

- In cellular systems where interference is a major issue, the performance of a modulation scheme in an interference environment is extremely important.

- Sensitivity to detection of timing jitter, caused by time-varying channels, is also an important consideration in choosing a particular modulation scheme.

- In general, the modulation, interference, and implementation of the time-varying effects of the channel as well as the performance of the specific demodulator are analyzed as a complete system using simulation to determine relative performance and ultimate selection.

List of common digital modulation techniques

The most common digital modulation techniques are:

- **Phase-shift keying (PSK)**
  - **Binary PSK (BPSK), using M=2 symbols**
  - **Quadrature PSK (QPSK), using M=4 symbols**
  - **8PSK, using M=8 symbols**
  - **16PSK, using M=16 symbols**
  - **Differential PSK (DPSK)**
  - **Differential QPSK (DQPSK)**
  - **Offset QPSK (OQPSK)**
  - **π/4–QPSK**
- **Frequency-shift keying (FSK)**
  - **Audio frequency-shift keying (AFSK)**
  - **Multi-frequency shift keying (M-ary FSK or MFSK)**
  - **Dual-tone multi-frequency (DTMF)**
- **Amplitude-shift keying (ASK)**
- **On-off keying (OOK), the most common ASK form**
  - **M-ary vestigial sideband modulation, for example 8VSB**
- **Quadrature amplitude modulation (QAM), a combination of PSK and ASK**
  - **Polar modulation like QAM a combination of PSK and ASK**
- **Continuous phase modulation (CPM) methods**
  - **Minimum-shift keying (MSK)**
  - **Gaussian minimum-shift keying (GMSK)**
  - **Continuous-phase frequency-shift keying (CPFSK)**
- **Orthogonal frequency-division multiplexing (OFDM) modulation**
  - **Discrete multitone (DMT), including adaptive modulation and bit-loading**
- **Wavelet modulation**
- **Trellis coded modulation (TCM), also known as Trellis modulation**
- **Spread-spectrum techniques**
  - **Direct-sequence spread spectrum (DSSS)**
  - **Chirp spread spectrum (CSS) according to IEEE 802.15.4a CSS uses pseudo-stochastic coding**
  - **Frequency-hopping spread spectrum (FHSS) applies a special scheme for channel release**

UNIT 5 MULTIPLE ANTENNA TECHNIQUES

1.Multiple Input Multiple Output Systems

- MIMO systems are systems with *Multiple Element Antennas* (MEAs) at *both* link ends.
- The MEAs of a MIMO system can be used for four different purposes:

(i) beamforming
(ii) diversity
(iii) interference suppression
(iv) spatial multiplexing (transmission of several data streams in parallel).

- The first three concepts are the same as for smart antennas. Spatial multiplexing, on the other hand, is a new concept, and has thus drawn the greatest attention.
- Spatial multiplexing allows direct improvement of capacity by simultaneous transmission of multiple data streams.

Diversity processing

The transmitter, the receiver or both use multiple antennas to

- Increase the received **signal power**.
- Reduce the amount of **fading**.

Spatial multiplexing

- The transmitter and receiver both use multiple antennas to increase the **data rate**

Beamforming

- Uses multiple antennas at the BS to increase the **coverage** of the cell

Interference suppression

- One of the environmental issues with which communication systems must contend is interference, either unintentional or intentional.
- Because MIMO systems use antenna arrays, localized interference can be mitigated naturally.

MIMO system - System Model

- System model is shown in the block diagram. At the T_X, the data stream enters an encoder, whose outputs are forwarded to N_t transmit antennas.

- From the antennas, the signal is sent through the wireless propagation channel, which is assumed to be quasi-static and frequency-flat. (quasi-static - the coherence time of the channel is so long that "a large number" of bits can be transmitted within this time).

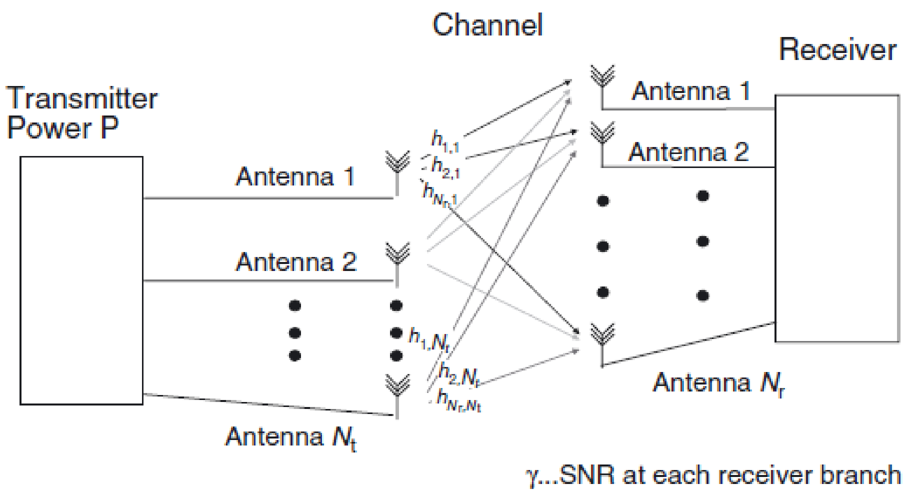


Figure Block diagram of a multiple-input multiple-output system.

- *N*r × *N*t matrix of the channel is **H**

$$H = \begin{pmatrix} h_{11} & h_{12} & \dots & \dots & h_{1Nt} \\ h_{21} & h_{22} & \dots & \dots & h_{2Nt} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{Nr1} & h_{Nr2} & \dots & \dots & h_{NrNt} \end{pmatrix}$$

- h_{ij} is the complex channel gain (transfer function) from the j^{th} transmit antenna to the i^{th} receive antenna.

The received signal vector = **r**

$$\mathbf{r} = \mathbf{Hs} + \mathbf{n} = \mathbf{X} + \mathbf{n}$$

r - contains signal received by Nr antenna elements

s - transmit signal vector

n - noise vector

2. Spatial Multiplexing

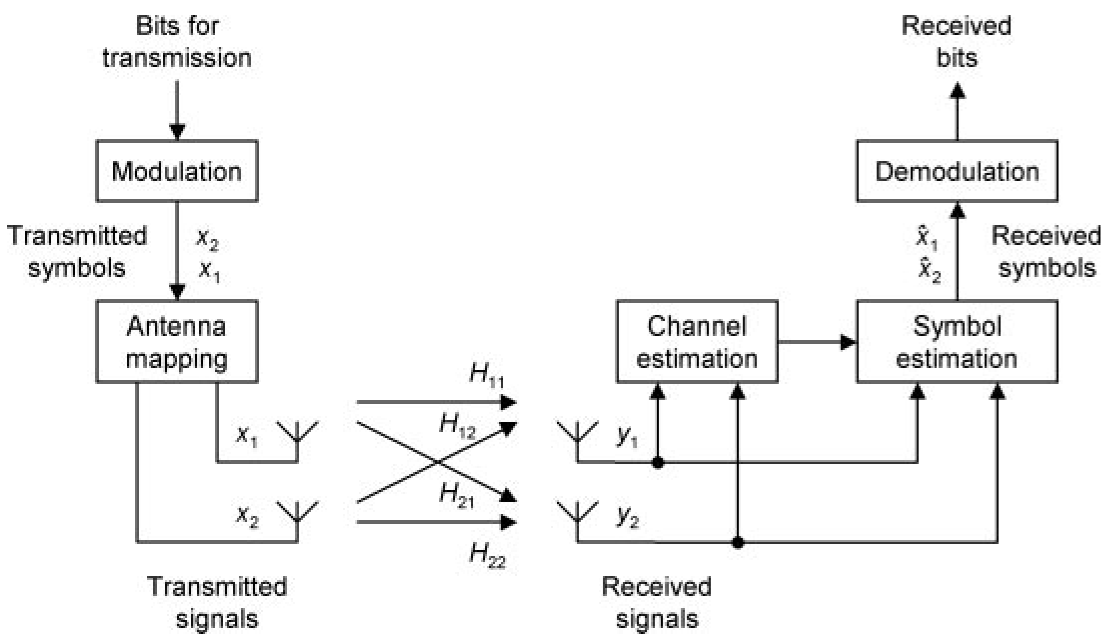
- Spatial multiplexing uses MEAs at the TX for transmission of parallel data streams .
- An original high-rate data stream is multiplexed into several parallel streams, each of which is sent from one transmit antenna element.
- The channel "mixes up" these data streams, so that each of the receive antenna elements sees a combination of them. If the channel is well behaved, the received signals represent *linearly independent* combinations.
- Appropriate signal processing at the RX can separate the data streams.

- A basic condition is that the number of receive antenna elements is at least as large as the number of transmit data streams.
- It is clear that this approach allows the data rate to be drastically increased – namely, by a factor of min (Nt,Nr).

Example:

spatial multiplexing system, in which the transmitter and receiver both have two antennas

- In the transmitter, the antenna mapper takes symbols from the modulator two at a time, and sends one symbol to each antenna.
- The antennas transmit the two symbols simultaneously, so as to double the transmitted data rate.



Basic principle of a 2x2 spatial multiplexing system

- The symbols travel to the receive antennas by way of four separate radio paths, so the received signals can be written as follows:

$$y_1 = H_{11}x_1 + H_{12}\,x_2 + n_1$$
$$y_2 = H_{21}x_1 + H_{22}\,x_2 + n_2$$

- x_1 and x_2 are the signals sent from the two transmit antennas.
  y_1 and y_2 are the signals that arrive at the two receive antennas.
  n_1 and n_2 represent the received noise and interference.
  H_{ij} expresses the way in which the transmitted symbols are attenuated and phase-shifted, as they travel to receive antenna *i* from transmit antenna *j*
- The receiver's first task is to estimate the four channel elements H_{ij}

- The **transmitter** broadcasts reference symbols with one extra feature based on which receiver estimate the channel.
- The receiver estimate the transmitted symbols x_1 and x_2.
- The simplest way is a zero-forcing detector, which operates as follows:
- If we ignore the noise and interference, then the equation is a pair of simultaneous equations for two unknown quantities, x1 and x2 which can be estimated as

$$\hat{x}_1 = \frac{\hat{H}_{22} y_1 - \hat{H}_{12} y_2}{\hat{H}_{11}\hat{H}_{22} - \hat{H}_{21}\hat{H}_{12}}$$
$$\hat{x}_2 = \frac{\hat{H}_{11} y_2 - \hat{H}_{21} y_1}{\hat{H}_{11}\hat{H}_{22} - \hat{H}_{21}\hat{H}_{12}}$$

3. Spatial Diversity

- In spatial diversity, many signal copies are transmitted from different antennas and are received at more than one antenna. This redundancy is provided by employing an array of antennas, with a minimum separation of half wavelength between neighboring antennas.
- Diversity techniques are used to mitigate the effect of **fading** in wireless communication sysyems

Types of spatial diversity:

- **Receiver diversity**
- **Transmitter diversity**

Receiver diversity

- In receive diversity one Tx antenna and many Rx antenna are used.

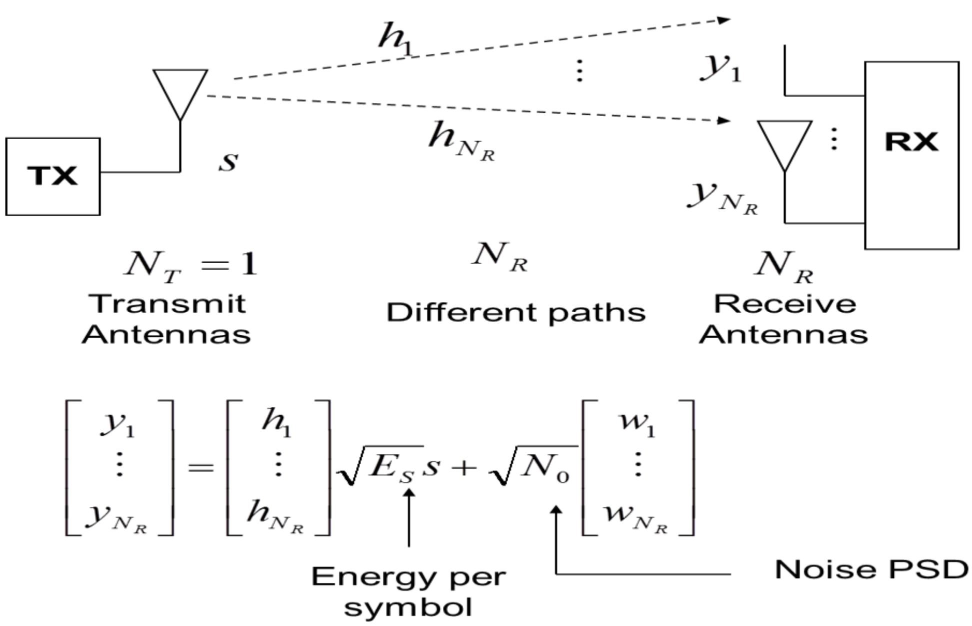


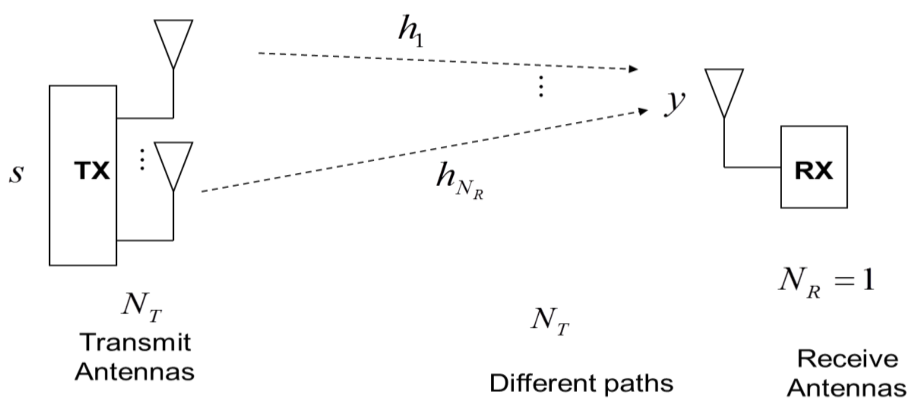
Figure Receiver Diversity

- The energy per bit and the Noise PSD are given.
- Since there is one transmitter and Nr receivers, the signal takes Nr different paths.
- Y is the received component. H is the transfer function of the system. W is the weight factor or the multiplication factor which is multiplied with the signal to get back the original transmitted signal.
- Assuming that the channel transfer function is known at the receiver, the signal is decoded as,

$$y = \sum_{i=1}^{N_R} h_i^* y_i = \underbrace{\boxed{\sqrt{E_S}\sum_{i=1}^{N_R} |h_i|^2 s}}_{\text{signal}} + \boxed{\sqrt{N_0}\sum_{i=1}^{N_R} h_i^* w_i}$$

<u>Transmitter Diversity</u>

- In transmitter diversity one Rx antenna and many Tx antenna are used.
- In transmitter diversity there are multiple transmitters and a single receiver. So the signal takes Nt travel paths.



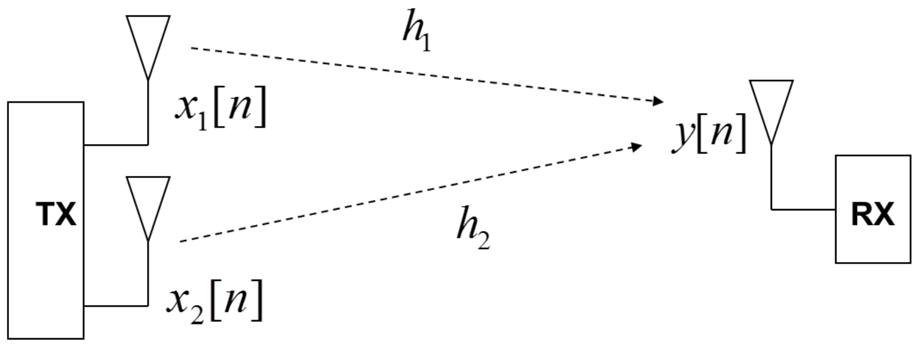
$$y = \left(\sqrt{\frac{E_S}{N_T}}\sum_{i=1}^{N_T} h_i\right)s + \sqrt{N_0}\,w$$

Equivalent to one channel, with no benefit.

Total energy equally distributed on transmit antennas

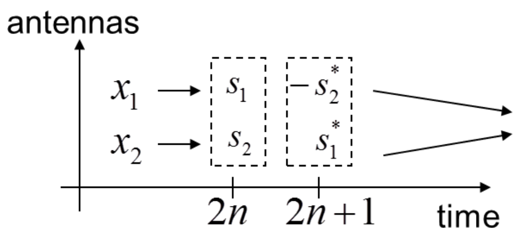
Figure :Transmitter Diversity

Example : Transmitter Diversity with two Tx antennas and one Rx antenna



Given two sequences $s_1[n], s_2[n]$

code them within the two antennas as follows



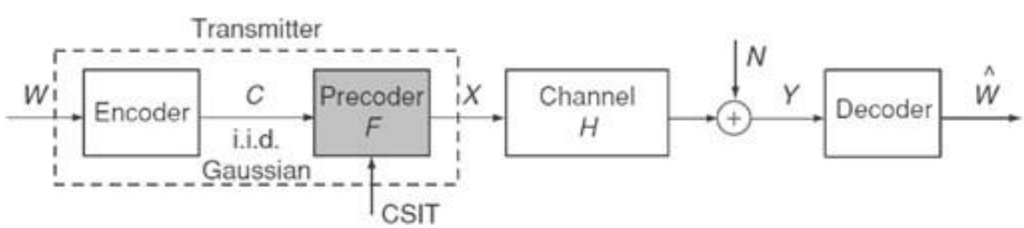
$$y[2n] = \sqrt{\frac{E_S}{2}}\left(h_1 s_1 + h_2 s_2\right) + \sqrt{N_0}\, w_1$$

$$y[2n+1] = \sqrt{\frac{E_S}{2}}\left(-h_1 s_2^* + h_2 s_1^*\right) + \sqrt{N_0}\, w_2$$

4.Precoding

- Precoding is a processing technique that makes use of channel state information of the transmitter (CSIT) before the signal is transmitted.
- Precoding is done inorder to optimize the beams transmitted in intended areas.

Precoding system structure



Precoding for point – to – point MIMO systems

- In point-to-point multiple-input multiple-output (MIMO) systems, a transmitter equipped with multiple antennas communicates with a receiver that has multiple antennas. Most classic precoding results assume narrowband, slowly fading channels, meaning that the channel for a certain period of time can be described by a single channel matrix which does not change faster.

- In practice, such channels can be achieved, for example, through OFDM. The precoding strategy that maximizes the throughput, called channel capacity, depends on the channel state information available in the system.

Statistical channel state information

If the receiver knows the channel matrix and the transmitter has statistical information, eigen beamforming is known to achieve the MIMO channel capacity. In this approach, the transmitter emits multiple streams in eigen directions of the channel covariance matrix.

Full channel state information

- If the channel matrix is completely known, singular value decomposition (SVD) precoding is known to achieve the MIMO channel capacity.
- In this approach, the channel matrix is diagonalized by taking an SVD and removing the two unitary matrices through pre- and post-multiplication at the transmitter and receiver, respectively.
- Then, one data stream per singular value can be transmitted (with appropriate power loading) without creating any interference whatsoever.

5. Beam Forming

- **Beamforming** or **spatial filtering** is a signal processing technique used in wireless communication systems for directional signal transmission or reception.
- This is achieved by combining elements in an antenna array in such a way that signals at particular angles experience constructive interference while others experience destructive interference.
- Beamforming can be used at both the transmitting and receiving ends in order to achieve spatial selectivity. The improvement compared with omni directional reception or transmission is known as the directivity of the array.
- Beamforming techniques can be broadly divided into two categories:

✓ conventional (fixed or switched beam) beam formers
✓ adaptive beam formers or phased array

Conventional beam formers use a fixed set of weightings and time-delays (or phasings) to combine the signals from the antennas in the array, primarily using only information about the location of the antennas in space and the wave directions of interest.

Adaptive beamforming techniques (e.g., MUSIC, SAMV) generally combine this information with properties of the signals actually received by the array, typically to improve rejection of unwanted signals from other directions. This process may be carried out in either the time or the frequency domain.

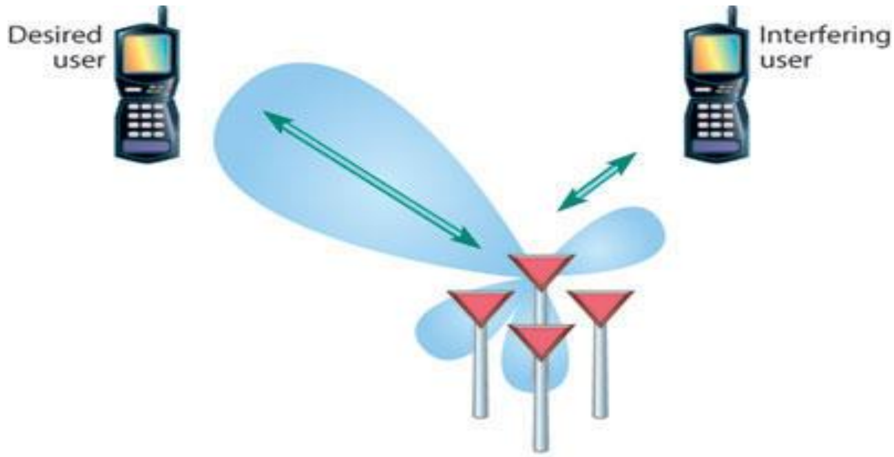


Figure MIMO beam forming

6.Channel state information (CSI)

- In wireless communications , CSI refers to channel properties of a communication link.
- This information describes how a signal propogates from the transmitter to the receiver and represents the combined effect of scattering, fading and power decay with distance.
- The CSI makes it possible to adapt transmissions to current channel conditions , which is critical for achieving reliable communication with high data rates in multi-antenna systems.
- CSI needs to be estimated at the receiver and usually quantized and fed back to the transmitter . Transmitter and Receiver can have different CSI .
- The CSI at the receiver is referred to as CSIR and the CSI at the transmitter is referred to as CSIT.

Two types of CSI

(i) Instantaneous CSI (or short term CSI)

- Instantaneous CSI means current channel condition
- This gives an opportunity to adapt the transmitted signal to the channel impulse response and therby optimize the received signal for spatial multiplexing or to achieve low BER.

(ii) **Statistical CSI (or long term CSI)**

- Statistical characterization of the channel.
- Statistical characterization include , the type of fading distribution, the average channel gain , line of sight component and the spectral correlation .

- Statistical CSI can be used for transmission optimization.
- In fast fading channels where channel conditions vary rapidly , use of statistical CSI is reasonable for transmission adaptation.
- In slow fading channels , instantaneous CSI can be estimated with reasonable accuracy and used for transmission adaptation

Channel state information Estimation

CSI can be estimated using received pilot signals and LMS , MMSE algorithms.

7.Capacity equation for MIMO system in nonfading channels

(Foschini's equation)

Capacity equation for normal single antenna AWGN channel is (as given by Shannon)

$$C_{shannon} = log_2\,(1 + \gamma.\,|H|^2) \qquad -----1$$

$\gamma = SNR\ at\ the\ receiver$

$H = Normalized\ transfer\ function\ from\ the\ Transmitter\ to\ the\ receiver$

For MIMO case, where the channel is represented by matrix H,

$$H = \begin{pmatrix} h_{11} & h_{12} & \ldots. & \ldots. & h_{1Nt} \\ h_{21} & h_{22} & \ldots. & \ldots. & h_{2Nt} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{Nr1} & h_{Nr2} & \ldots. & \ldots. & h_{NrNt} \end{pmatrix}$$

Singular value decomposition of the channel gives,

$$\boldsymbol{H = W\Sigma U^{+}} \qquad \boldsymbol{-----2}$$

Σ – diagonal matrix containing singular values

$\mathbf{W},\boldsymbol{U^{+}}$ - Unitary matrices composed of the left and right singular vectors respectively

The received signal is then,

$$\mathbf{r = Hs\ +\ \ n} \qquad \boldsymbol{-----3}$$

$$\mathbf{r = W\Sigma U^{+}s\ + n} \qquad \boldsymbol{-----4}$$

Then, multiplication of the transmit data vector by matrix **U** and the received signal vector by $\boldsymbol{W^{+}}$ diagonalizes the channel:

$$\boldsymbol{W^{+}}\mathbf{r} = \boldsymbol{W^{+}}\mathbf{W\Sigma U^{+}U}\boldsymbol{\tilde{s}}\ + \boldsymbol{W^{+}}\mathbf{n}$$

$$\boldsymbol{\tilde{r}\ =\ \Sigma\,\tilde{s}\ +\ \tilde{n}} \qquad \boldsymbol{-----5}$$

Since **U** and **W** are unitary matrices, $\boldsymbol{\tilde{n}}$ has the same statistical properties as **n**.

The capacity of the system is same as that of the system $\mathbf{r = Hs\ +\ \ n}$

Computation of the capacity of system represented by equation 5 is given below:

The matrix **Σ** is a diagonal matrix with R_H nonzero entries σ_k, where R_H is the rank of **H** (and thus defined as the number of nonzero singular values), and σ_k is the k^{th} singular value of **H**.

We have therefore R_H *parallel* channels (eigen modes of the channel), and it is clear that the capacity of parallel channels just adds up.

The capacity of channel **H** is thus given by the sum of the capacities of the eigen modes of the channel:

$$C = \sum_{k=1}^{R_H} log_2\,\left[1 + \frac{P_k}{\sigma_n^2}\ \ \sigma_k^2\right]\ -----6$$

$\sigma_n^2\ = noise\ variance$

$P_k = Power\ allocated\ to\ the\ k^{th}\ eigen\ mode$

Equation 6 is equivalent to

$$C =\ log_2\left[det\left(I_{Nr} + \frac{\bar{\gamma}}{N_t}\ \ H\,R_{ss}H^{+}\right)\right]\ -----7$$

$I_{Nr}\ \ is\ N_r \times N_r\ \ identity\ matrix$

$\bar{\gamma}\ \ is\ the\ mean\ SNR\ per\ R_X\ branch$

$R_{ss}\ \ is\ the\ correlation\ matrix\ of\ the\ transmit\ data$

No Channel State Information at the Transmitter and Full CSI at the Receiver

When the R_X knows the channel perfectly, but no CSI is available at the T_X, it is optimum to assign equal transmit power to all T_X antennas, $P_k = P/N_t$, and use uncorrelated data streams. Capacity thus takes on the form:

$$C = \log_2\left[\det\left(I_{Nr} + \frac{\bar{\gamma}}{N_t}\ H\,H^+\right)\right] \quad ------8$$

It is worth noting that, for sufficiently large *N*s(Number of interacting objects), the capacity of a MIMO system increases linearly with $\min(N_t,N_r)$, irrespective of whether the channel is known at the T_X or not.

special cases:

Assume that $N_t = N_r = N$

Case 1

- All transfer functions are identical – i.e., $h_{1,1} = h_{1,2} = \ldots = h_{N,N}$. This case occurs when all antenna elements are spaced very closely together, and all waves are coming from similar directions.
- In such a case, the rank of the channel matrix is unity. Then, capacity is

$$C_{MIMO} = \log_2 (1 + N\bar{\gamma}) \quad ------9$$

- We see that in this case the SNR is increased by a factor of *N* compared with the single antenna case, due to beamforming gain at the R_X. However, this only leads to a logarithmic increase in capacity with the number of antennas.

Case 2

- All transfer functions are different such that the channel matrix is full rank, and has *N* eigenvalues of equal magnitude. This case can occur when the antenna elements are spaced far apart and are arranged in a special way.
- In this case, capacity increases linearly with the number of antenna elements.

$$C_{MIMO} = N \log_2 (1 + \bar{\gamma}) \quad ------10$$

Case 3

- Parallel transmission channels – e.g., parallel cables. In this case, capacity also increases linearly with the number of antenna elements.
- However, the SNR per channel decreases with *N*, so that total capacity is

$$C_{MIMO} = N \log_2 \left(1 + \frac{\bar{\gamma}}{N}\right) \quad ------11$$

Figure shows capacity as a function of *N* for different values of SNR.

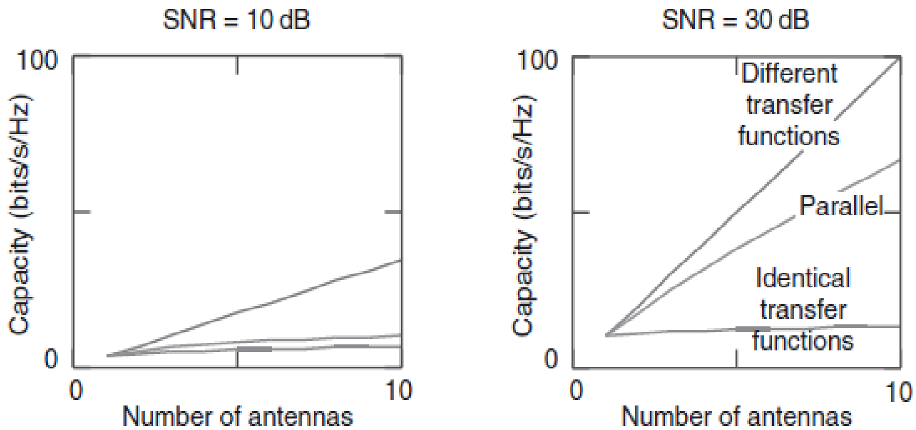


Figure Capacity of multiple-input-multiple-output systems in additive white Gaussian noise channels.

Full Channel State Information at the Transmitter and Full CSI at the Receiver

- Consider the case where both the R_X and T_X know the channel perfectly.
- In such a case, it can be more advantageous to distribute power not uniformly between the different transmit antennas (or eigenmodes) but rather assign it based on the channel state.
- For optimally allocating power to several parallel channels, each of which has a different SNR, **water filling power allocation strategy** is used.

$$C_{waterfill} = \sum_{n=1}^{N} log_2 \left(1 + \frac{P_n \alpha_n^2}{\sigma_n^2}\right)$$

$$\alpha_n = gain(\, inverse\ attenuation) of\ the\ n^{th}\ sub\ channel$$

$$\sigma_n^2 = noise\ variance$$

$$P_n = Power\ allocation\ of\ the\ n^{th}\ sub\ channel$$

8. Capacity in Flat-Fading Channels

- In wireless systems, we have channel fading due to multipath propagation.
- In this case, entries in channel matrix **H** are random variables.
- If the channel is Rayleigh fading, and fading is independent at different antenna elements, the *hij* are iid zero-mean, circularly symmetric complex Gaussian random variables with unit variance – i.e., the real and imaginary part each has variance 1/2.
- The power carried by each *hij* is chi-square-distributed with 2 degrees of freedom. This is the simplest possible channel model.
- Since fading is independent, there is a high probability that the channel matrix is full rank and the eigenvalues are fairly similar to each other; consequently, capacity increases linearly with the number of antenna elements.
- Thus, the existence of heavy multipath, which is usually considered a drawback, becomes a major advantage in MIMO systems.

- Because the entries of the channel matrix are random variables, information-theoretic approach is employed to determine channel capacity.
- Two different definitions of capacity exist for MIMO systems:
  (i) *Ergodic (Shannon) capacity*
  (ii) *Outage capacity*

Ergodic (Shannon) capacity: this is the expected value of the capacity, taken over all realizations of the channel. This quantity assumes an infinitely long code that extends over all the different channel realizations.
• *Outage capacity*: this is the minimum transmission rate that is achieved over a certain fraction of the time – e.g., 90% or 95%. This quantity assume that data are encoded with a near-Shannon-limit achieving code that extends over a period that is much shorter than the channel coherence time.

- Thus, each channel realization can be associated with a (Shannon) capacity value. Capacity thus becomes a random variable (RV) with an associated cumulative distribution function (cdf);

- Figure shows the result for systems at a 21-dB SNR. The (1, 1) curve describes a Single Input Single Output (SISO) system. We find that the median capacity is on the order of 6 bit/s/Hz, but the 5% outage capacity is considerably lower (on the order of 3 bit/s/Hz).
- When using a (1, 8) system – i.e., 1 transmit antenna and 8 receive antennas – the mean capacity does not increase that significantly – from 6 to 10 bit/s/Hz. However, the 5% outage capacity increases significantly from 3 to 9 bit/s/Hz. The reason for this is the much higher resistance to fading that such a diversity system has.
- However, when going to a (8, 8) system – i.e., a system with 8 transmit and 8 receive antennas – both capacities increase dramatically: the mean capacity is on the order of 46 bit/s/Hz, and the 5% outage probability is more than 40 bit/s/Hz.

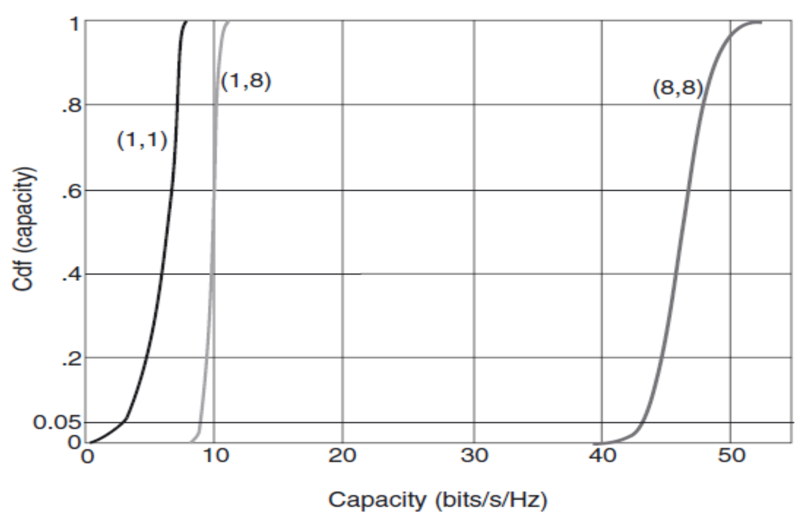


Figure Cumulative distribution function of capacity for 1×1, 1×8, and the 8×8 optimum scheme.

The exact expression for the *ergodic capacity* given as

$$E(C) = \int_0^{\infty} log_2 \left[1 + \frac{\bar{\gamma}}{N_t}\lambda\right] \sum_{k=0}^{m-1} \frac{k!}{(k+n-m)!} [L_k^{n-m}(\lambda)]^2 \ \lambda^{n-m} \exp(-\lambda)\, d\lambda$$

Where,

$$\bar{\gamma} = mean\ SNR\ per\ channel$$
$$N_t = Number\ of\ transmit\ antenna$$
$$N_r = Number\ of\ receive\ antenna$$
$$m = \min(N_t, N_r)$$
$$n = \max(N_t, N_r)$$
$$L_k^{n-m}(\lambda) = Laguerre\ polynomial$$

<u>Perfect Channel State Information at the Transmitter and Receiver</u>

- The capacity gain by waterfilling (compared with equal-power distribution) is rather small when the number of transmit and receive antennas is identical.
- This is especially true in the limit of large SNRs: when there is a lot of water available, the height of "concrete blocks" in the vessel has little influence on the total amount that ends up in the vessels.
- When *N*t is larger than *N*r, the benefits of waterfilling become more pronounced (see Figure).

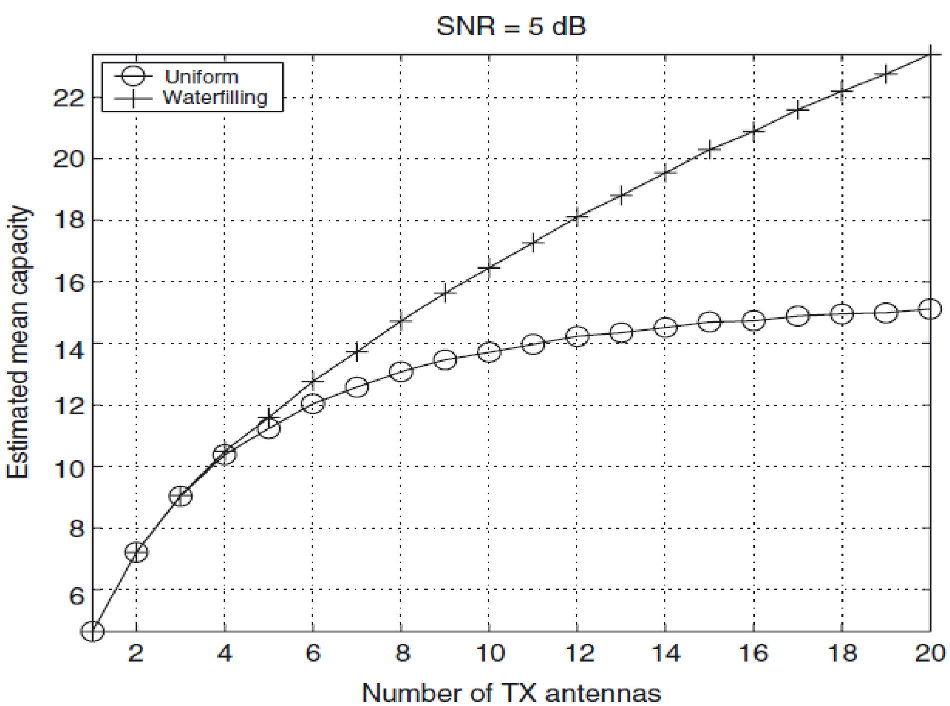


Figure Capacity with and without channel state information at the transmitter with $N_r = 8$ antennas and a signal-to-noise ratio of 5 dB.

- We can interpret this the following way: if the TX has no channel knowledge, then there is little point in having more transmit than receive antennas – the number of data streams is limited by the number of receive antennas.
- Of course, we can transmit the same data stream from multiple transmit antennas, but this does not increase the SNR for that stream at the RX; without channel knowledge at the TX, the streams add up incoherently at the RX.
- On the other hand, if the TX has full channel knowledge, it can perform beamforming, and direct the energy better toward the receive array.
- Thus, increasing the number of TX antennas improves the SNR, and (logarithmically) capacity. Thus, having a larger TX array improves capacity. However, this also increases the demand for channel estimation.

9.Waterfilling Power allocation strategy

- Transmitter has to decide how to select the correct transmission parameter for each channel and it should also decide the amount of power assigned to each channel
- For a given number of parallel sub channels with different attenuations, **Waterfilling Power allocation strategy** gives the distribution of transmission power that maximizes capacity.
- The was given by Shannon in the 1940s, and is known as "waterfilling "

Power allocation *Pn* of the *n*th sub channel is,

$$P_n = max\left(0, \varepsilon - \frac{\sigma_n^2}{|\alpha_n|^2}\right)$$

Where,

$$\alpha_n = gain(\ inverse\ attenuation) of\ the\ n^{th}\ sub\ channel$$

$$\sigma_n^2 = noise\ variance$$

$$\varepsilon =\ threshold\ determined\ by\ the\ constraintof\ the\ total\ transmitted\ power\ P$$

$$P = \sum_{n=1}^{N} P_n$$

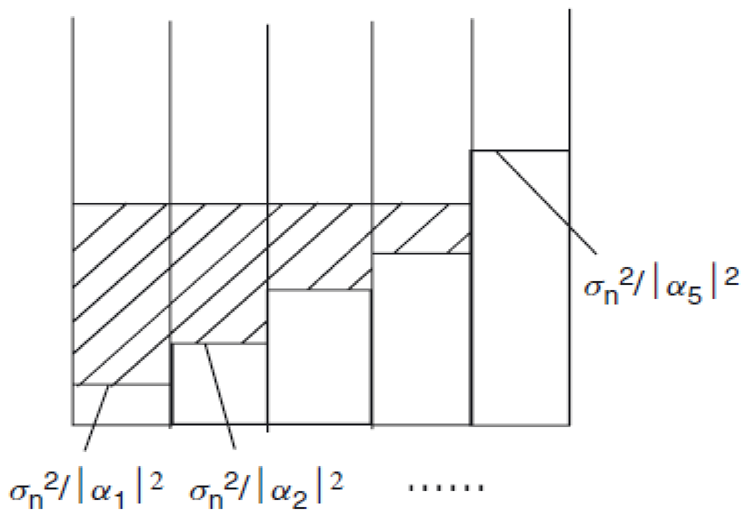


Figure Principle behind waterfilling.

- Waterfilling can be interpreted visually, according to Figure shown.
- Imagine a number of connected vessels. At the bottom of each vessel is a block of concrete with a height that is proportional to the inverse SNR of the sub channel that we are considering.
- Then take water, and pour it into the vessels; the amount of poured water is proportional to the total transmit power that is available.
- Because the vessels are connected, the surface level of the water is guaranteed to be the same in all vessels. The amount of power assigned to each sub channel is then the amount of water in the vessel corresponding to this sub channel.
- Obviously, sub channel 1, which has the highest SNR, has the most water in it. It can also happen that some sub channels that have a poor SNR (like channel 5), do not get any power assigned to them at all (the concrete block of that vessel is sticking out of the water surface).
- Essentially, waterfilling makes sure that energy is not wasted on sub channels that have poor SNR.
- With waterfilling, power is allocated preferably to sub channels that have a good SNR.
- This is optimum from the point of view of theoretical capacity; however, it requires that the transmitter can actually make use of the large capacity on good sub channels.
- In each sub channel (subcarrier), signaling as close to capacity as possible should be performed. This means that the transmitter has to adapt the data rate according to the SNR that is available

Binary Frequency shift Keying

In BFK, frequency of constant amplitude carrier is switched between two values depending on the bit value.

$$S_{FSK}(t) = \sqrt{\frac{2E_b}{T_b}}\, cos[2\pi f_C + 2\pi\Delta f]t \qquad 0 \le t \le T_b \qquad for\ binary\ 1$$

$$S_{FSK}(t) = \sqrt{\frac{2E_b}{T_b}}\, cos[2\pi f_C - 2\pi\Delta f]t \qquad 0 \le t \le T_b \qquad for\ binary\ 0$$

$2\pi\Delta f$ = constant offset from the nominal carrier frequency

Generation of FSK

To generate BFSK signal, switch between two independent oscillators according to whether the data is 0 or 1.

FSK Detection

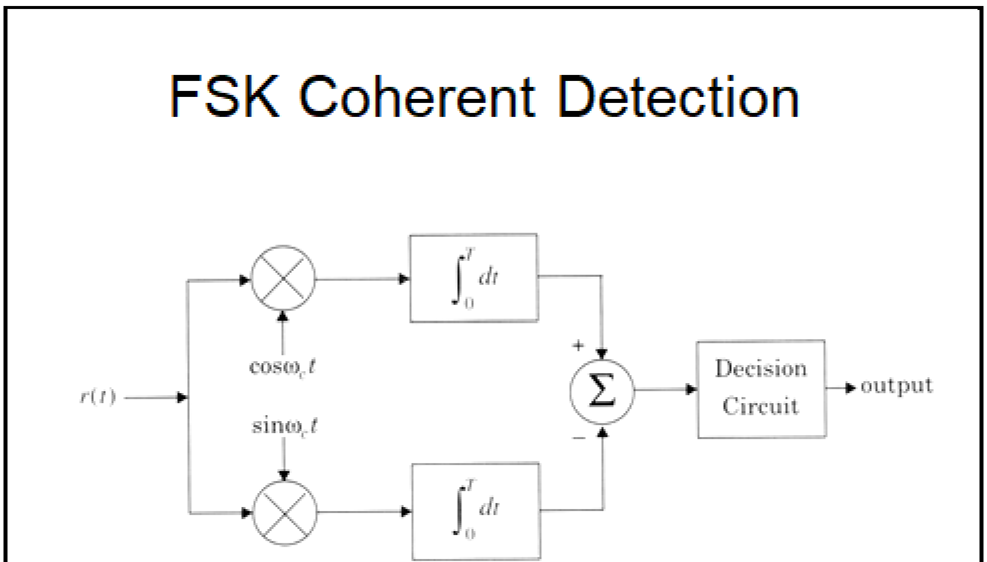


Figure Coherent detection of FSK signals.

- Consists of two correlators which are supplied with locally generated coherent reference signals.
- The difference of the correlator outputs is compared with a threshold comparator.
- If the difference is greater than the threshold, binary 1 is detected, otherwise 0 is detected.

Non coherent detection of BFSK

- Non coherent detection of BFSK is shown in figure below. The two matched filters are band pass filters centered at f_H and f_L.

- Output of envelope detector is sampled at every $t=kT_b$, where k is an integer and their values are compared with threshold to decide 0 or 1.

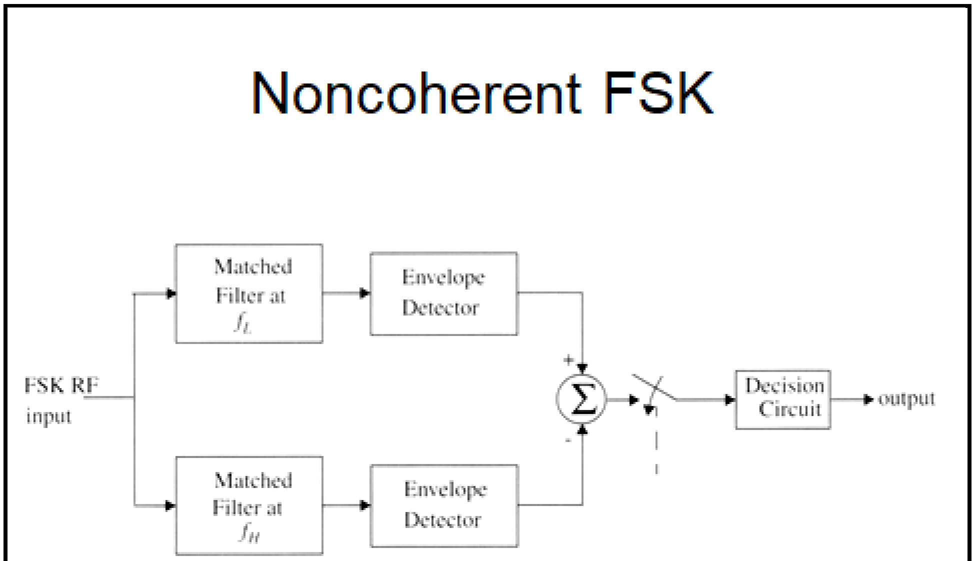


Figure Block diagram of noncoherent FSK receiver.

Minimum Shift Keying

- Minimum shift keying (MSK) is a special type of continuous phase Frequency shift keying (CPFSK) wherein the peak frequency deviation is equal to 1/4 the bit rate.
  - In other words, MSK is continuous phase FSK with a modulation index of 0.5. The modulation index of an FSK signal is similar to the FM modulation index, and is defined as, $modulation\ index = \frac{2\Delta f}{R_b}$ where Δf is the peak RF frequency deviation and R_b is the bit rate.
  - Multiplying a carrier signal with $\cos\left(\frac{\pi t}{2T}\right)$ produces two phase-coherent signals at $f_c + \frac{1}{4T}$ and $f_c - \frac{1}{4T}$. These two FSK signals are separated using two narrow bandpass filters and appropriately combined to form the in-phase and quadrature carrier components x(t) and y(t), respectively. These carriers are multiplied with the odd and even bit streams, $m_I(t)$ and $m_Q(t)$, to produce the MSK modulated signal $S_{MSK}(t)$.

$$S_{MSK}(t) = \sqrt{\frac{2E_b}{T_b}}\ cos\left[2\pi f_c t - m_I(t)m_Q(t)\frac{\pi t}{2T_b} + \emptyset_k\right]$$

    Where $\emptyset_k(t)$= 0 or π depending on whether $m_I(t)$ is 1 or -1

  - A modulation index of 0.5 corresponds to the minimum frequency spacing that allows two FSK signals to be coherently orthogonal, and the name minimum shift keying implies the minimum frequency separation.

MSK modulation

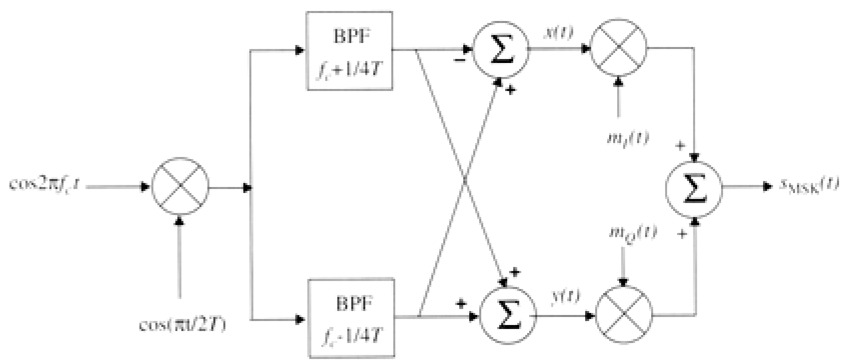


Figure Block diagram of an MSK transmitter. Note that $m_I(t)$ and $m_Q(t)$ are offset by T_b.

Minimum Shift Keying spectra

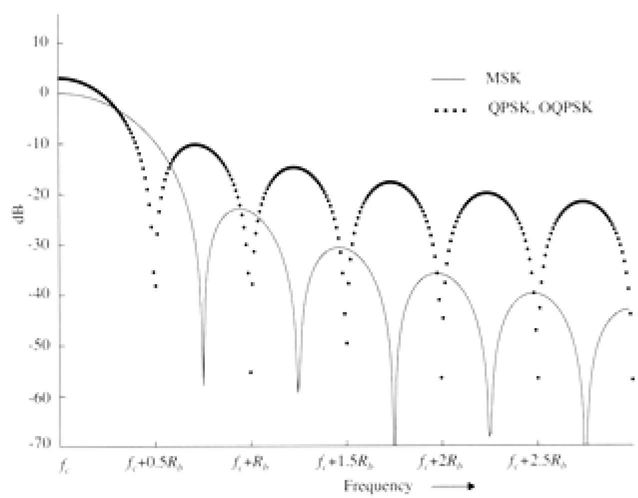


Figure Power spectral density of MSK signals as compared to QPSK and OQPSK signals.

MSK Receiver

- The received signal $S_{MSK}(t)$ (in the absence of noise and interference) is multiplied by the respective in-phase and quadrature carriers x(t) and y(t).
- The output of the multipliers are integrated over two bit periods and dumped to a decision circuit at the end of each two bit periods.
- Based on the level of the signal at the output of the integrator, the threshold detector decides whether the signal is 0 or 1. The output data streams correspond to $rn_I(t)$ and $m_Q(t)$, which are offset combined to obtain the demodulated signal.

MSK reception

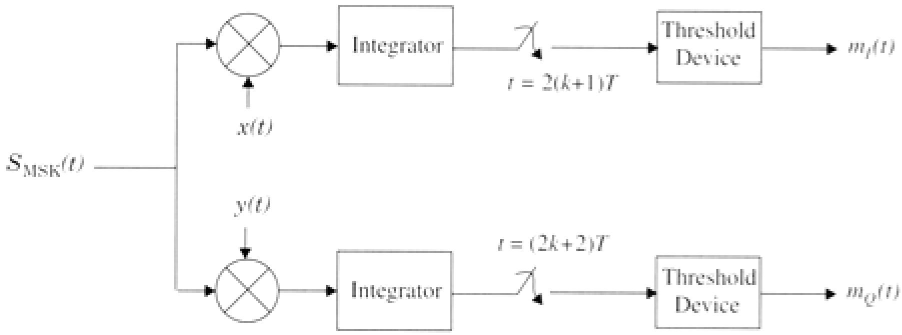


Figure Block diagram of an MSK receiver.

GMSK (Gaussian Minimum Shift Keying)

- The spectral efficiency of MSK is further enhanced by filtering the baseband signal of square pulses with a **Gaussian filter**. In GMSK the side lobe levels of the spectrum are further reduced.
- Gaussian pulse shaping smooths the phase trajectory of the MSK signal and hence stabilizes the instantaneous frequency variation over time. This has the effect of considerably reducing the side lobe levels in the transmitted spectrum.

Simple GMSK generation

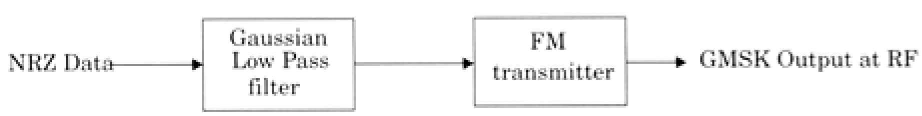


Figure Block diagram of a GMSK transmitter using direct FM generation.

GMSK spectral shaping

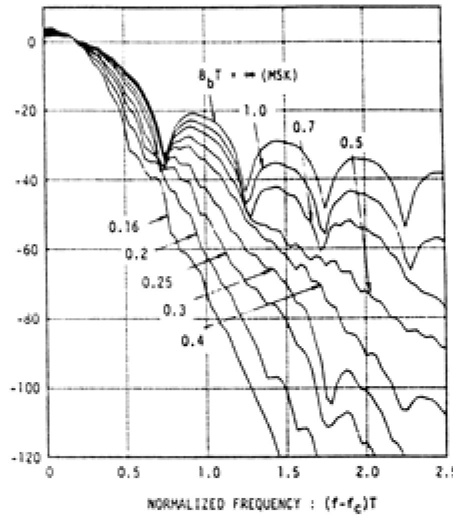


Figure Power spectral density of a GMSK signal [from [Mur81] © IEEE].

<u>GMSK Detection</u>

GMSK signals can be detected using orthogonal coherent detectors as shown below:

GMSK Demodulator

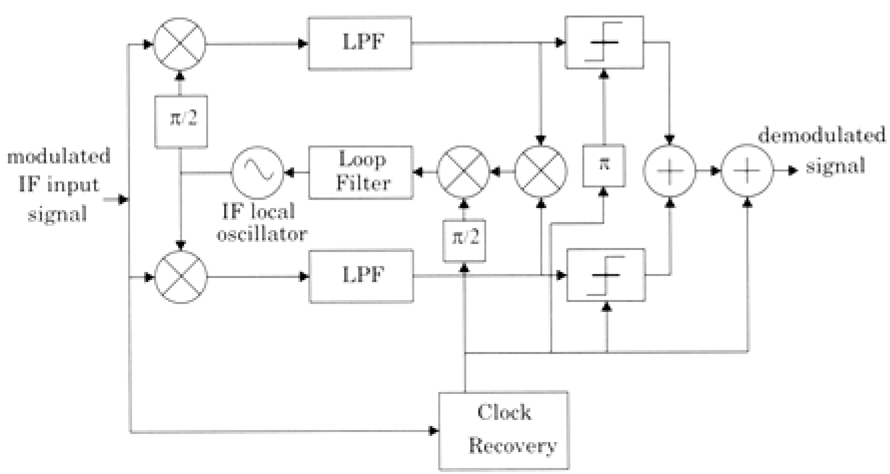


Figure Block diagram of a GMSK receiver.

DPSK (Differential Phase Shift Keying)

- Differential PSK is a non-coherent form of phase shift keying which avoids the need for a coherent reference signal at the receiver.
- Non-coherent receivers are easy and cheap to build, and hence are widely used in wireless communications.
- In DPSK systems, the input binary sequence is first differentially encoded and then modulated using a BPSK modulator.

The differentially encoded sequence { d_k is generated from the input binary sequence I } by complementing the modulo-2 sum of m_k and d_k - The effect is to leave the symbol d_k unchanged from the previous symbol if the incoming binary symbol m_k is 1, and to toggle d_k if m_k is 0.

Differential PSK encoding

Table Illustration of the Differential Encoding Process

$\{m_k\}$		1	0	0	1	0	1	1	0
$\{d_{k-1}\}$		1	1	0	1	1	0	0	0
$\{d_k\}$	1	1	0	1	1	0	0	0	1

DPSK modulation

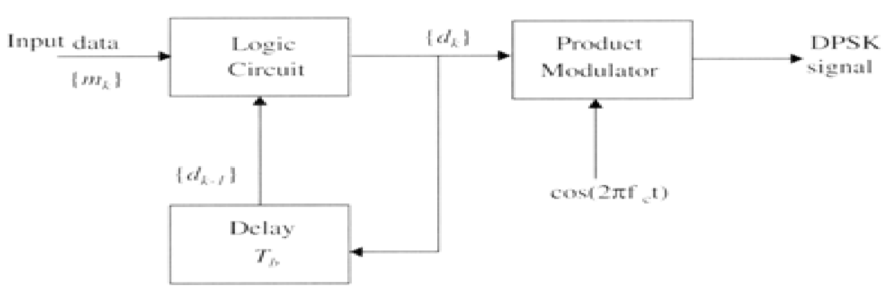


Figure Block diagram of a DPSK transmitter.

DPSK receiver

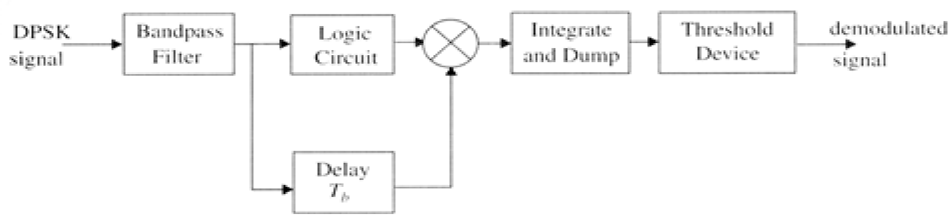


Figure Block diagram of DPSK receiver.

Quadrature phase shift keying (QPSK)

Quadrature phase shift keying (QPSK) has twice the bandwidth efficiency of BPSK, since 2 bits are transmitted in a single modulation symbol. The phase of the carrier takes on 1 of 4 equally spaced values, such as 0, π/2, π, and 3π/2, where each value of phase corresponds to a unique pair of message bits. The QPSK signal for this set of symbol states may be defined as

$$S_{QPSK(t)} = \sqrt{\frac{2E_S}{T_S}} cos\left[2\pi f_c t + (i-1)\frac{\pi}{2}\right] \qquad 0 \leq t \leq T_S$$

Where i=1,2,3,4

T_S = symbol duration $=2T_b$

T_b= Bit duration

Compared to BPSK, QPSK provides twice the spectral efficiency with exactly the same energy efficiency.

QPSK constellation diagrams

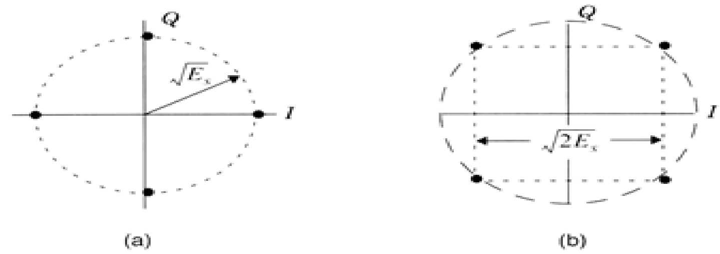


Figure (a) QPSK constellation where the carrier phases are 0, $\pi/2$, π, $3\pi/2$; (b) QPSK constellation where the carrier phases are $\pi/4$, $3\pi/4$, $5\pi/4$, $7\pi/4$.

<u>QPSK Transmitter</u>

QPSK modulation

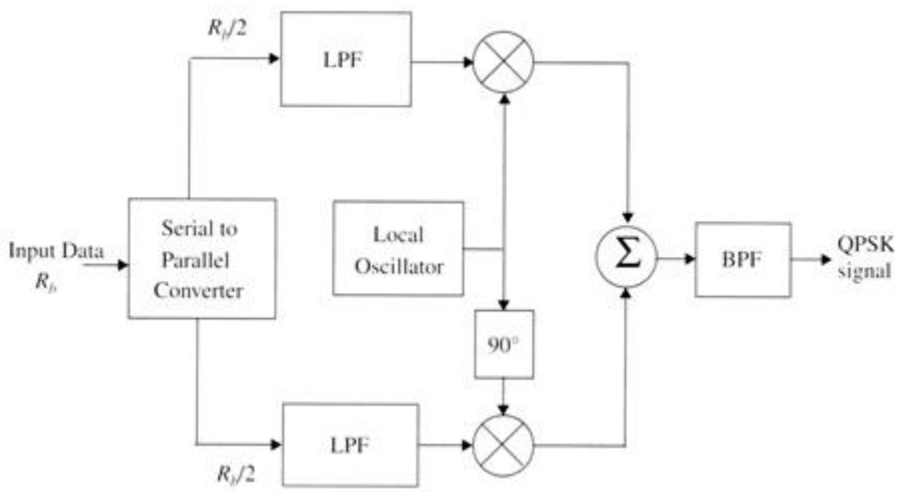


Figure Block diagram of a QPSK transmitter.

- The unipolar binary message stream has bit rate R_b and is first converted into a bipolar non-return-to-zero (NRZ) sequence using a unipolar to bipolar converter.
- The bit stream m(t) is then split into two bit streams m_I (t) and m_Q (t) (in-phase and quadrature streams), each having a bit rate of $R_b/2$. The bit stream $m_I(t)$ is called the "even" stream and m_Q (t) is called the "odd" stream.
- The two binary sequences are separately modulated by two carriers which are in quadrature.
- The two modulated signals, each of which can be considered to be a BPSK signal, are summed to produce a QPSK signal.

- The filter at the output of the modulator confines the power spectrum of the QPSK signal within the allocated band. This prevents spill-over of signal energy into adjacent channels and also removes out-of-band spurious signals generated during the modulation process.
- In most implementations, pulse shaping is done at baseband to provide proper RF filtering at the transmitter output.

QPSK Receiver

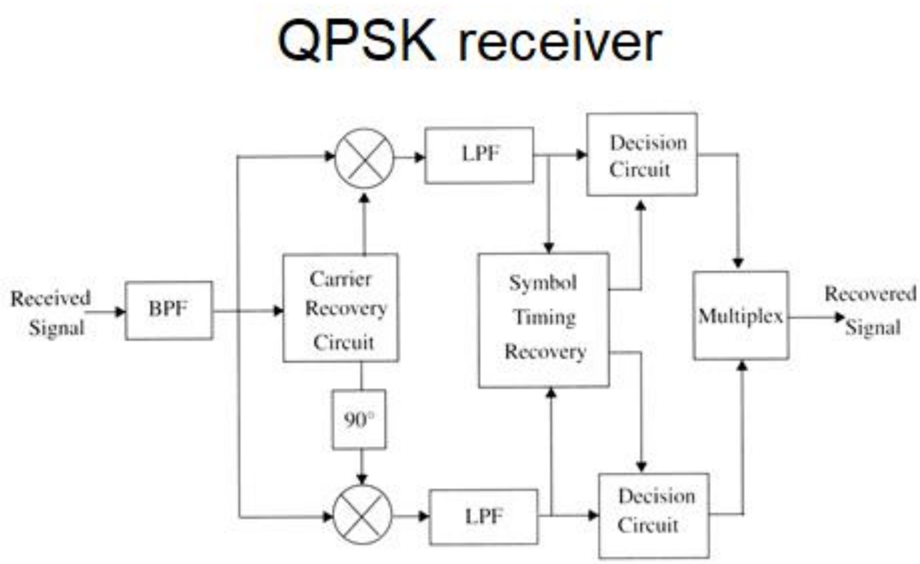


Figure Block diagram of a QPSK receiver.

- The frontend band pass filter removes the out-of-band noise and adjacent channel interference.
- The filtered output is split into two parts, and each part is coherently demodulated using the in-phase and quadrature carriers.
- The coherent carriers used for demodulation are recovered from the received signal using carrier recovery circuits of the type described in Figure.
- The outputs of the demodulators are passed through decision circuits which generate the in-phase and quadrature binary streams. The two components are then multiplexed to reproduce the original binary sequence.

OQPSK (Offset Quadrature Phase Shift Keying)

• In QPSK phase shift of π radians will cause the amplitude to fluctuations. This will lead to generation of side lobes and spectral widening.
• To reduce this 180^{o} phase shift we use O-QPSK. In this, maximum phase shift of the transmitted signal at any given time is limited to ±90°
• This is achieved by delaying one channel by T_b sec.
• In QPSK the even and odd bit streams occur at the same time instants.

• But in OQPSK signaling, the even and odd bit streams, $m_I(t)$ and $m_Q(t)$ are offset in their relative alignment by one bit period (half-symbol period).
• Because of this at any given time only one of the two bit streams can change values.
• So the maximum phase shift of the transmitted signal at any given time is limited to ±90°

Offset QPSK waveforms

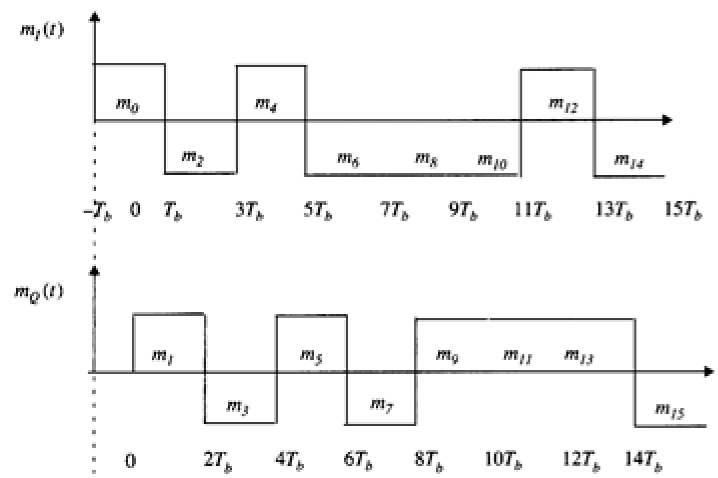


Figure _ The time offset waveforms that are applied to the in-phase and quadrature arms of an OQPSK modulator. Notice that a half-symbol offset is used.

- Since 180° phase transitions have been eliminated, band limiting of (i.e., pulse shaping) OQPSK signals does not cause the signal envelope to go to zero.
- Obviously, there will be some amount of ISI caused by the band limiting process, especially at the 90 phase transition points. But the envelope variations are considerably less, and hence hard limiting or nonlinear amplification of OQPSK signals does not regenerate the high frequency side lobes as much as in QPSK.
- Thus, spectral occupancy is significantly reduced, while permitting more efficient RF amplification. The spectrum of an OQPSK signal is identical to that of a QPSK signal, hence both signals occupy the same bandwidth.

π/4 DQPSK

• In π/4 DQPSK the maximum phase change is limited to ±135°.
• So it has less amplitude fluctuations than QPSK.
• It can be detected non coherently .
• The phase shift between successive symbols is an integer multiple of $\pi/4$ radians.
• There exists two sets of signal constellations (0,90,180,270) & (45,135,225,315).

- In wireless channels in the presence of multipath spread and fading, π/4 QPSK performs better than OQPSK. When π/4 QPSK signals are differentially encoded, the scheme is referred to as π/4D QPSK.

Pi/4 QPSK signaling

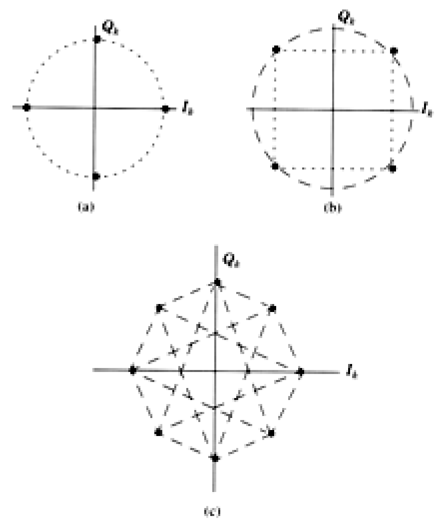


Figure Constellation diagram of a π/4 QPSK signal: (a) possible states for θ_k when $\theta_{k-1} = n\pi/4$; (b) possible states when $\theta_{k-1} = n\pi/2$; (c) all possible states.

Pi/4 QPSK phase shifts

Table Carrier Phase Shifts Corresponding to Various Input Bit Pairs

Information bits m_{Ik}, m_{Qk}	Phase shift ϕ_k
1 1	$\pi/4$
0 1	$3\pi/4$
0 0	$-3\pi/4$
1 0	$-\pi/4$

Pi/4 QPSK transmitter

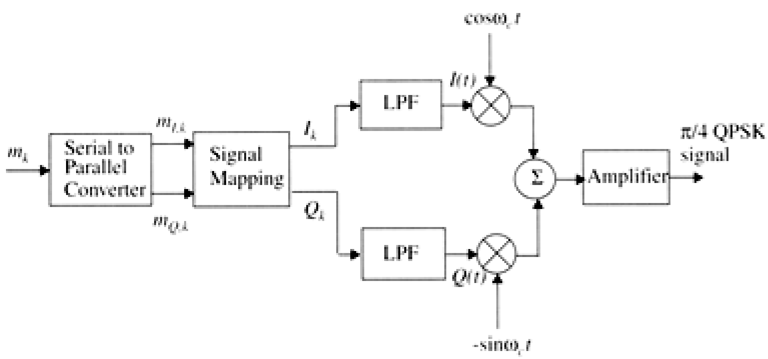


Figure Generic π/4 QPSK transmitter.

Differential detection of pi/4 QPSK

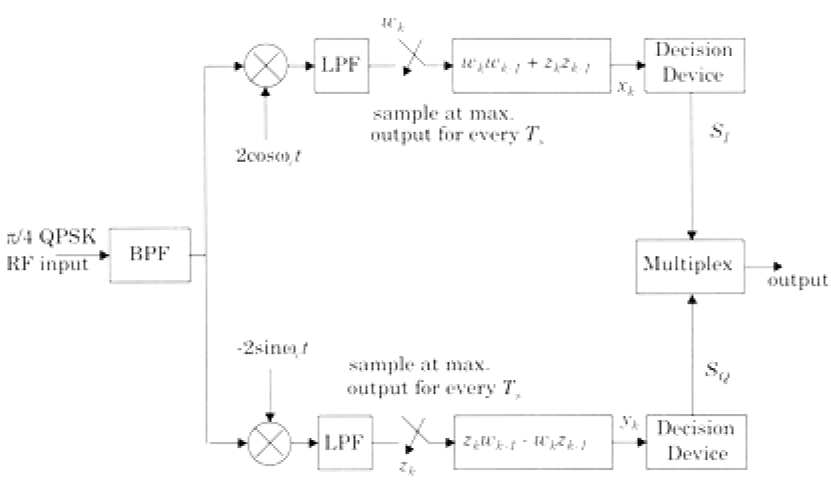


Figure Block diagram of a baseband differential detector

IF Differential Detection

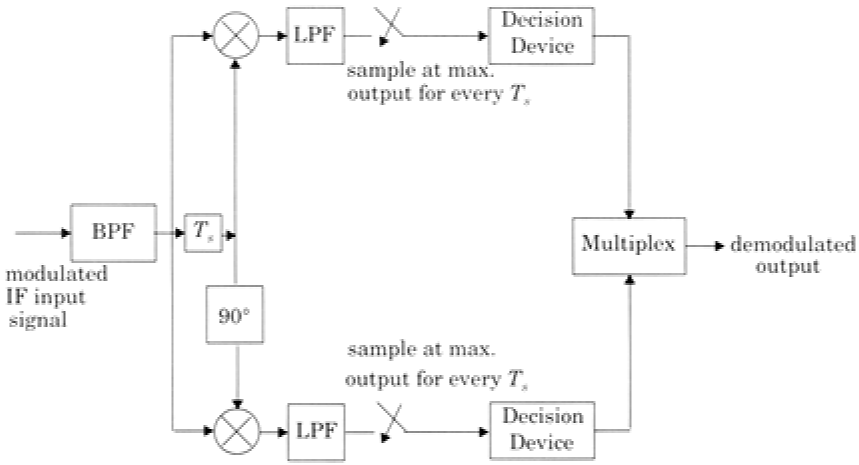


Figure Block diagram of an IF differential detector for π/4 QPSK.

FM Discriminator detector

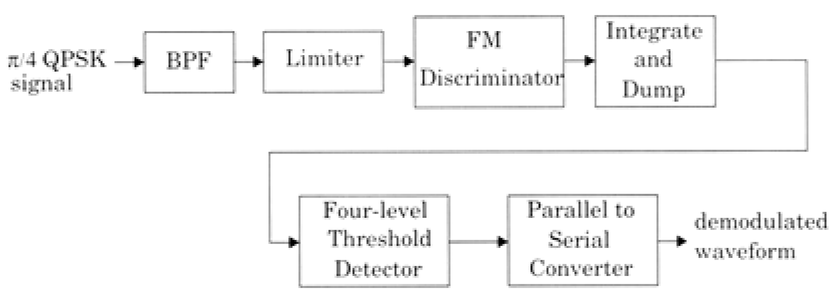


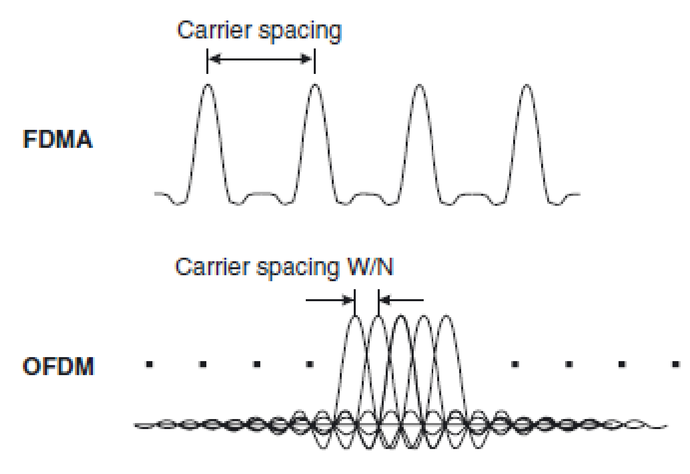
Figure FM discriminator detector for π/4 DQPSK demodulation.

Orthogonal Frequency Division Multiplexing (OFDM)

- **Basis for 4G (LTE, WiMAX)**
- **WLAN standards such as 802.11/a/g/n are based on OFDM**
- **Key broad band wireless technology which support data rates in excess of 100Mbps.**

Orthogonal Frequency Division Multiplexing

- (OFDM) is a modulation scheme that is especially suited for high-data-rate transmission in delay-dispersive environments.
- It converts a high-rate data stream into a number of low-rate streams that are transmitted over parallel, narrowband channels that can be easily equalized.
- OFDM splits a high-rate data stream into N parallel streams, which are then transmitted by modulating N distinct carriers (henceforth called *subcarriers* or *tones*).
- In order for the receiver to be able to separate signals carried by different subcarriers, they have to be orthogonal. Conventional Frequency Division Multiple Access (FDMA), can achieve this by having large (frequency) spacing between carriers. This, however, wastes precious spectrum.
- A much narrower spacing of subcarriers can be achieved. Specifically, let subcarriers be at the frequencies $f_n = nW/N$, where n is an integer, and W the total available bandwidth; in the most simple case, $W = N/T_s$.
- Due to the rectangular shape of pulses in the time domain, the spectrum of each modulated carrier has a $\sin(x)/x$ shape. The spectra of different modulated carriers overlap, but each carrier is in the spectral nulls of all other carriers. Therefore, as long as the receiver does the appropriate demodulation, the data streams of any two subcarriers will not interfere.

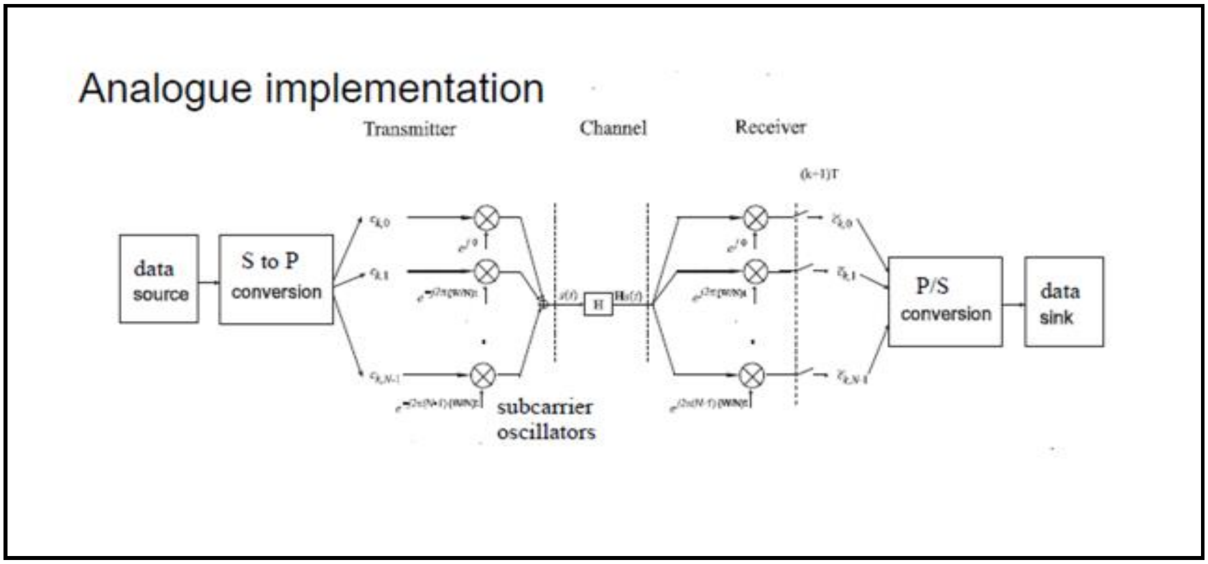


Principle behind orthogonal frequency division multiplexing: N carriers within a bandwidth of W.

Implementation of OFDM Transceivers

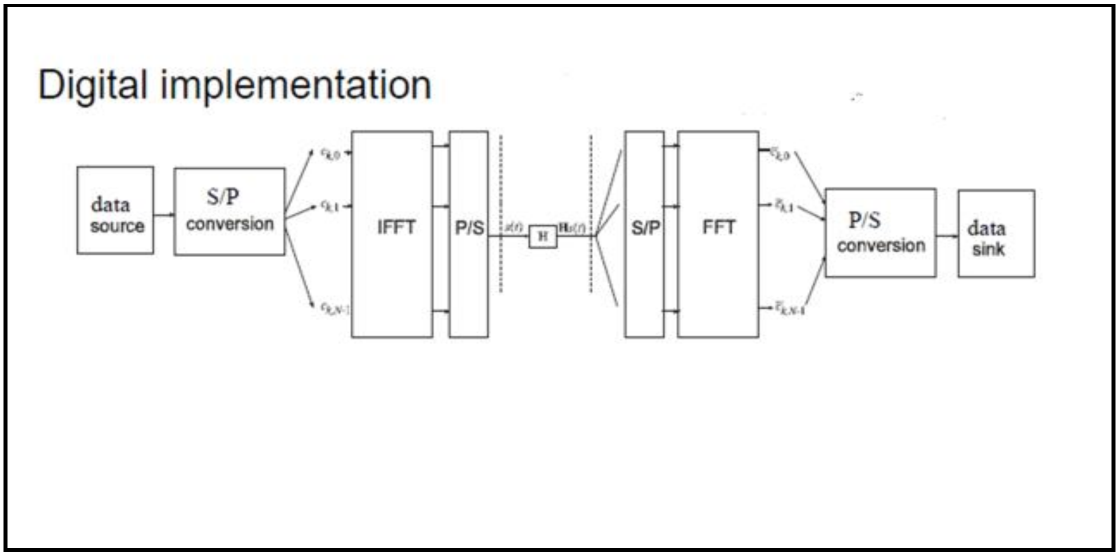
Analog implementation

- Original data stream is first split into *N* parallel data streams, each of which has a lower data rate. We furthermore have a number of local oscillators (LOs) available, each of which oscillates at a frequency $f_n = nW/N$, where $n = 0, 1, \ldots, N-1$.
- Each of the parallel data streams then modulates one of the carriers. This picture allows an easy understanding of the principle, but is not suitable for actual implementation, since the hardware required is too high (multiple local oscillators).



Digital Implementation

An alternative implementation is *digital* . It first divides the transmit data into blocks of *N* symbols. Each block of data is subjected to an *Inverse Fast Fourier Transformation* (IFFT), and then transmitted. This approach is much easier to implement with integrated circuits.



Frequency-Selective Channels

- Intuitively, we would anticipate that delay dispersion will have only a small impact on the performance of OFDM
- we convert the system into a parallel system of narrowband channels, so that the symbol duration on each carrier is made much larger than the delay spread.
- Delay dispersion can lead to appreciable errors. Delay dispersion also leads to a loss of orthogonality between the subcarriers, and thus to *Inter Carrier Interference* (ICI). Both these negative effects can be eliminated by a special type of guard interval, called the *cyclic prefix (CP)*.

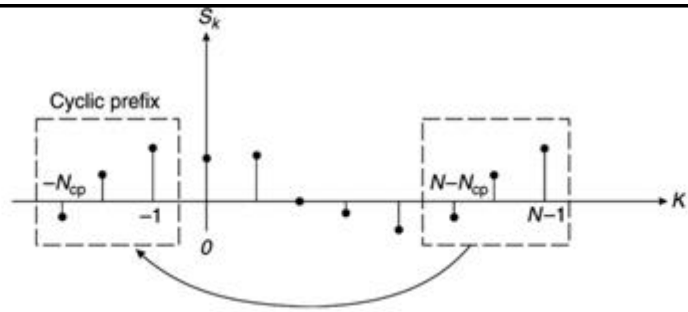


Figure Principle of the cyclic prefix. $N_{cp} = NT_{cp}/(N/W)$ is the number of samples in the cyclic prefix.

- The block diagram of an OFDM system, including the cyclic prefix, is given in Figure.
- The original data stream is S/P converted. Each block of *N* data symbols is subjected to an IFFT, and then the last NT_{cp}/T_S samples are prepended.
- The resulting signal is modulated onto a (single) carrier and transmitted over a channel, which distorts the signal and adds noise.
- At the receiver, the signal is partitioned into blocks. For each block, the cyclic prefix is stripped off, and the remainder is subjected to an FFT.
- The resulting samples (which can be interpreted as the samples in the frequency domain) are "equalized" by means of one-tap equalization – i.e., division by the complex channel attenuation – on each carrier.

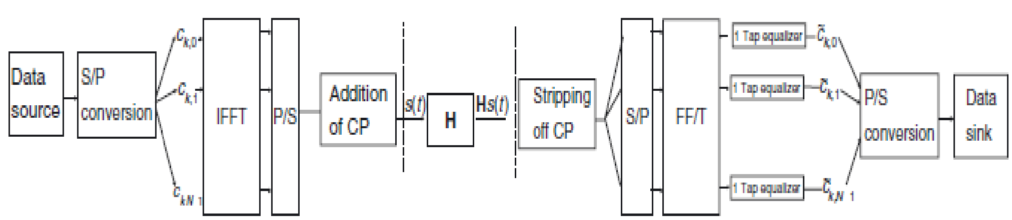


Figure 19.4 Structure of an orthogonal-frequency-division-multiplexing transmission chain with cyclic prefix and one-tap equalization.

Peak-to-Average Power Ratio

- One of the major problems of OFDM is that the peak amplitude of the emitted signal can be considerably higher than the average amplitude.
- This *Peak-to-Average Ratio* (PAR) issue originates from the fact that an OFDM signal is the superposition of N sinusoidal signals on different subcarriers.
- On average the emitted power is linearly proportional to N. However, sometimes, the signals on the subcarriers add up constructively, so that the *amplitude* of the signal is proportional to N.

Three main methods to deal with the Peak-to-Average Power Ratio (PAPR):

1. Put a power amplifier into the transmitter that can amplify linearly up to the possible *peak* value of the transmit signal.

   This is usually not practical, as it requires expensive and power-consuming class-A amplifiers. The larger the number of subcarriers N, the more difficult this solution becomes.

2. Use a nonlinear amplifier, and accept the fact that amplifier characteristics will lead to distortions in the output signal.

- Those nonlinear distortions destroy orthogonality between subcarriers, and also lead to increased out-of-band emissions (*spectral regrowth* – similar to third-order intermodulation products – such that the power emitted outside the nominal band is increased).
- The first effect increases the BER of the desired signal
- The second effect causes interference to other users and thus decreases the cellular capacity of an OFDM system

3. Use PAR reduction techniques.

Peak-to-Average power Ratio Reduction Techniques

1. *Coding for PAR reduction*
2. *Phase adjustments*
3. *Correction by multiplicative function*
4. *Correction by additive function*

1.Coding for PAR reduction

- Under normal circumstances, each OFDM symbol can represent one of 2^N codewords (assuming BPSK modulation).
- Now, of these codewords only a subset of size 2^K is acceptable in the sense that its PAR is lower than a given threshold.
- Both the transmitter and the receiver know the mapping between a bit combination of length K, and the codeword of length N that is chosen to represent it, and which has an admissible PAR.
- The transmission scheme is thus the following: (i) parse the incoming bitstream into blocks of length K; (ii) select the associated codeword of length N; (iii) transmit this codeword via the OFDM modulator.
- The coding scheme can guarantee a certain value for the PAPR.

2.Phase adjustments

- This scheme first defines an ensemble of phase adjustment vectors ,that are known to both the transmitter and receiver.
- The transmitter then multiplies the OFDM symbol to be transmitted by each of these phase vectors .
- Then select which gives the lowest PAPR.
- The receiver can then undo phase adjustment and demodulate the OFDM symbol.

3.Correction by multiplicative function

- Another approach is to multiply the OFDM signal by a time-dependent function whenever the peak value is very high.
- If the signal attains a level $sk > A_0$, it is multiplied by a factor A_0/sk.

4.Correction by additive function

- We can choose an additive, instead of a multiplicative, correction function.
- The correction function should be smooth enough not to introduce significant out-of-band interference.
- Furthermore, the correction function acts as additional pseudo noise, and thus increases the BER of the system.

Comparison of different PAPR reduction methods

- When comparing the different approaches to PAPR reduction, we find that there is no single "best" technique.
- The coding method can guarantee a maximum PAPR value, but requires considerable overhead, and thus reduced throughput.
- The phase adjustment method has a smaller overhead (depending on the number of phase adjustment vectors), but cannot give a guaranteed performance.
- Neither of these two methods leads to an increase in either ICI or out-of-band emissions.
- The correction by multiplicative functions can guarantee performance – up to a point .Also, it can lead to considerable ICI, while out-of-band emissions are fairly well controlled.

UNIT IV-MULTIPATH MITIGATION TECHNIQUES

- Multipath propagation give rise to ISI, Fading and error in the received signal.
- Techniques employed to mitigate the effect of multipath propagation:
  - (i) Diversity Technique – to mitigate fading
  - (ii) Equalization – to mitigate ISI
  - (iii) Coding – to mitigate error

PRINCIPLE OF DIVERSITY

• **Diversity** is a method for improving the reliability of a message signal by using two or more communication channels with different characteristics.
• It is based on the fact that individual channels experience different levels of fading and interference.
• The main concept of diversity is that if one radio path undergoes a deep fade, another independent path may have a strong signal.
• *Diversity* is usually implemented by using two or more receiving antennas.
• Multiple versions of the same signal may be received and combined in the receiver
• Diversity plays an important role in combating fading, co-channel interference and avoiding error bursts.

TYPES OF DIVERSITY

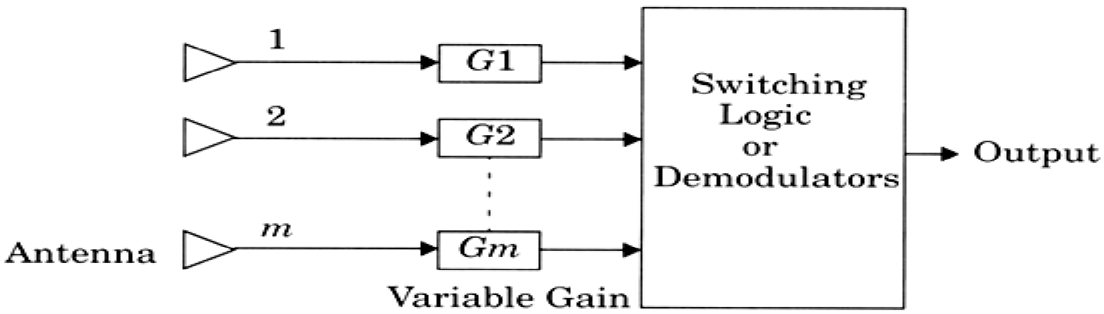
- ❖ **Micro Diversity**
- ❖ **Macro Diversity**

Micro Diversity

Methods that can be used to combat small-scale fading, are called "microdiversity." The five most common methods are as follows:
1. *Spatial diversity*: several antenna elements separated in space.
2. *Temporal diversity*: transmission of the transmit signal at different times.
3. *Frequency diversity*: transmission of the signal on different frequencies.
4. *Angular diversity*: multiple antennas (with or without spatial separation) with different antenna patterns.
5. *Polarization diversity*: multiple antennas with different polarizations (e.g., vertical and horizontal).

Spatial Diversity

- The transmit signal is received at several antenna elements, and the signals from these antennas are then further processed.
- But, irrespective of the processing method, performance is influenced by correlation of the signals between the antenna elements.



Generalized block diagram for space diversity.

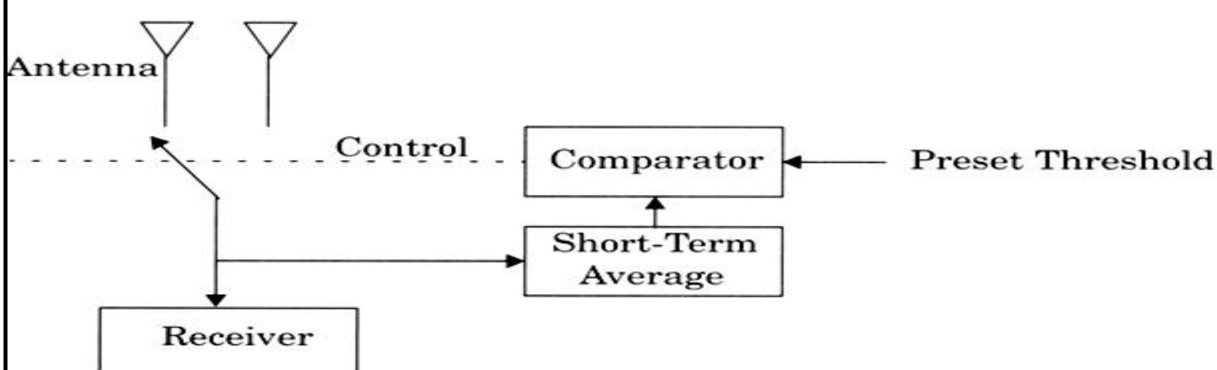
- Space diversity reception can be classified in to four categories:
  (i) Selection diversity
  (ii) Feed back diversity (or) Scanning diversity
  (iii) Maximal Ratio Combining
  (iv) Equal gain Combining

<u>(i) Selection diversity</u>

- m diversity branches are employed
- Gains are adjusted to provide the same average SNR for each branch
- The receiver branch having the highest instantaneous SNR is connected to the demodulator for further processing

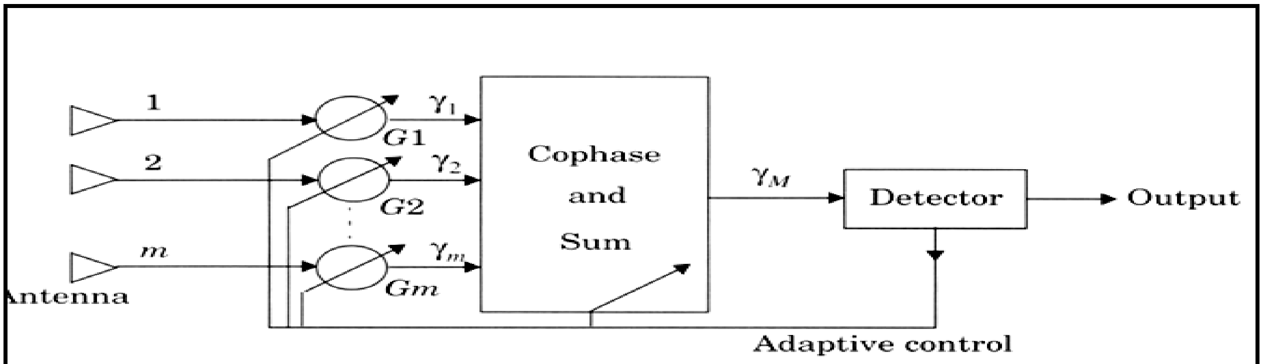
<u>(ii) Feed back diversity (or) Scanning diversity</u>

- Instead of always using the best of **m** signals, the **m** signals are scanned in a fixed sequence until one is found to be above a predetermined threshold.
- This signal is then received until it falls below threshold and the scanning process is again initiated.
- **Advantage:**This technique is very simple to implement
- **Disadvantage**:Performance is inferior to other methods



Basic form of scanning diversity.

(iii) Maximal Ratio Combining



Maximal ratio combiner.

- Individual signals are cophased before being summed. This requires individual receiver and phasing circuit for each antenna.
- MRC produces an output SNR equal to the sum of the individual SNRs.
- This method has a advantage of producing an output with an acceptable SNR even when none of the individual signals are acceptable.

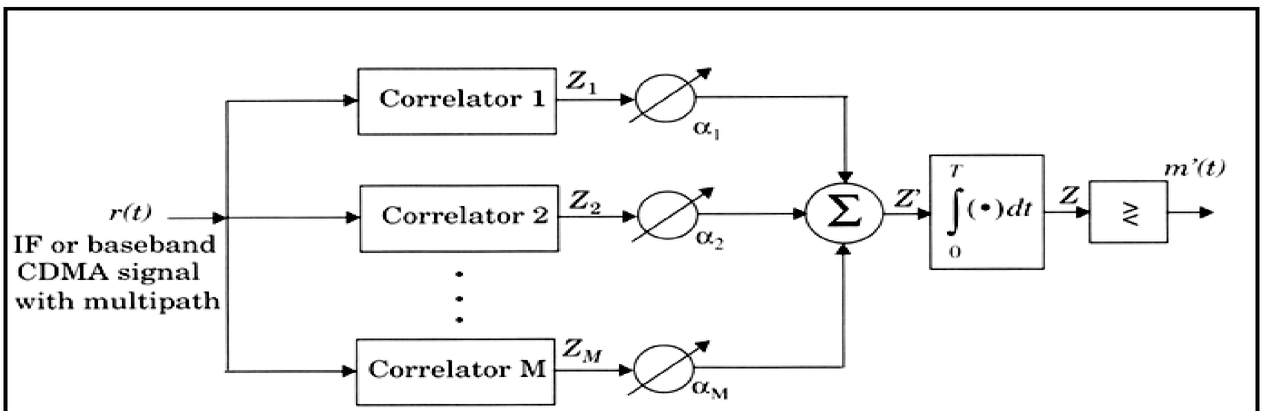
(iv) Equal gain Combining

In equal gain combining, branch weights are all set to unity. The signals from each branch are co-phased to provide equal gain combining diversity(No weighting, but just a phase correction).

Temporal Diversity

- As the wireless propagation channel is time variant, signals that are received at different times are uncorrelated.
- For "sufficient" decorrelation, the temporal distance must be at least $1/(2\nu_{max})$, where ν_{max} is the maximum Doppler frequency.
- In a static channel, where neither transmitter (TX), RX, nor the IOs are moving, the channel state is the same at all times. Such a situation can occur, e.g., for WLANs. In such a case, the correlation coefficient is $\rho = 1$ for all time intervals, and temporal diversity is useless.

Rake Receiver



· Rake receiver is designed to counter the effects of multipath fading.

· If multipath components are delayed in time by more than one chip duration, they appear like uncorrelated noise

· It is mainly used in reception of CDMA signals where conventional equalization won't work.

· Multipath results in multiple versions of the transmitted signal at the receiver.

· Each component has some information in it.

· The RAKE receiver uses a **multipath time diversity principle**.

· It uses several "sub-receivers" called fingers, that is, several correlators each assigned to a different multipath component.

· Each multipath component is extracted by using a single correlator. In all we use several correlators which independently decodes a single multipath component.

· The outputs of each correlator are weighted to provide better estimate of the transmitted signal than is provided by a single component.

· The Integrator is used to provide the average for a specific time period.

· The decision maker is used to regenerate digital signals from the incoming weak signals.

· Outputs of the M correlators are denoted as Z_1, Z_2,…, and $Z_{M.}$

· The weighting coefficients are based on the power or the SNR from each correlator output.

· If the power or SNR is small out of a particular correlator, it will be assigned a small weighting factor, the total output Z^1 is given by.

$$Z^1 = \sum_{m=1}^{M} \propto_m Z_m$$

$$\propto_m = \frac{{Z_m}^2}{\sum_{m=1}^{M} {Z_m}^2}$$

·

- RAKE receiver has to know

- Multipath delays -> time delay synchronization
- Phases of the multipath components -> carrier phase synchronization
- Amplitudes of the multipath components -> amplitude tracking
- Number of multipath components

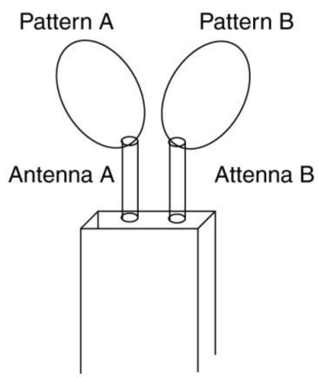
- The main challenge is receiver synchronization.

Frequency Diversity

- In frequency diversity, the same signal is transmitted at two (or more) different frequencies.
- If these frequencies are spaced apart by more than the coherence bandwidth of the channel, then their fading is approximately independent, and the probability is low that the signal is in a deep fade at both frequencies simultaneously.

Angle Diversity(*pattern diversity*).

- A fading dip is created when MPC(Multi Path Component)s, which usually come from different directions, interfere destructively. If some of these waves are attenuated or eliminated, then the location of fading dips changes.
- In other words, two collocated antennas with different patterns "see" differently weighted MPCs, so that the MPCs interfere differently for the two antennas. This is the principle of *angle diversity* (also known as *pattern diversity*).



Polarization Diversity

- Horizontally and vertically polarized MPCs propagate differently in a wireless channel, as the reflection and diffraction processes depend on polarization.
- Even if the transmit antenna only sends signals with a single polarization, the propagation effects in the channel lead to depolarization so that both polarizations arrive at the RX.
- The fading of signals with different polarizations is statistically independent. Thus, receiving both polarizations using a dual-polarized antenna, and processing the signals separately, offers diversity. This diversity can be obtained without any requirement for a minimum distance between antenna elements.

Macrodiversity

- Methods that can be used for combating large-scale fading, which is created by shadowing effects.
- Shadowing is almost independent of transmit frequency and polarization, so that frequency diversity or polarization diversity are not effective.
- Spatial diversity (or equivalently, temporal diversity with moving TX/RX) can be used, but we have to keep in mind that the correlation distances for large-scale fading are on the order of tens or hundreds of meters.
- In other words, if there is a hill between the TX and RX, adding antennas on either the BS or the MS does not help to eliminate the shadowing caused by this hill. Rather, we should use a separate base station (BS2) that is placed in such a way that the hill is not in the connection line between the MS and BS2. This in turn implies a large distance between BS1 and BS2, which gives rise to the word *macro diversity*.

Diversity Technique

Selection diversity

Consider M independent diversity branch. Assume that each branch has the same average SNR

$$SNR = \Gamma = \frac{E_b}{N_0}\alpha^2 \qquad \text{—— (1)}$$

Note: For Rayleigh fading channels, the fading amplitude α has a Rayleigh distribution. The fading power α^2 and consequently Γ have a chi-square distribution, with two degrees of freedom.

If each branch has an instantaneous SNR = γ_i, the probability density function of γ_i is

$$p(\gamma_i) = \frac{1}{\Gamma} e^{-\frac{\gamma_i}{\Gamma}} \qquad \gamma_i \geq 0 \qquad \text{—— (2)}$$

The probability that a single branch has an instantaneous SNR less than some threshold γ is

$$P_r[\gamma_i \leq \gamma] = \int_0^{\gamma} p(\gamma_i)\, d\gamma_i = \int_0^{\gamma} \frac{1}{\Gamma} e^{-\frac{\gamma_i}{\Gamma}}\, d\gamma_i = 1 - e^{-\frac{\gamma}{\Gamma}} \qquad \text{—— (3)}$$

The probability that all M independent diversity branches receive signals which are simultaneously less than some specific SNR threshold γ is

$$P_r[\gamma_1, \gamma_2, \ldots \gamma_M \leq \gamma] = \left(1 - e^{-\frac{\gamma}{\Gamma}}\right)^M = P_M(\gamma) \qquad \text{—— (4)}$$

$P_M(\gamma)$ = probability of all branches failing to achieve an instantaneous SNR $= \gamma$

Probability that SNR $> \gamma$ for one or more branches is given by,

$$P_r(\gamma_i > \gamma) = 1 - P_M(\gamma) = 1 - (1 - e^{-\frac{\gamma}{\Gamma}})^M \quad \text{—— (5)}$$

Equation (5) is the expression for the probability of exceeding a threshold when selection diversity is used.

SNR improvement offered by selection diversity:

$P_M(\gamma) \longrightarrow$ CDF

$p_M(\gamma) \longrightarrow$ pdf of γ, the instantaneous SNR when M branches are used.

$$p_M(\gamma) = \frac{d}{d\gamma} P_M(\gamma) = \frac{d}{d\gamma}\left[1 - e^{-\frac{\gamma}{\Gamma}}\right]^M$$

$$= \frac{M}{\Gamma}(1 - e^{-\frac{\gamma}{\Gamma}})^{(M-1)} e^{-\frac{\gamma}{\Gamma}}$$

$$\bar{\gamma} = \text{mean SNR} = \int_0^\infty \gamma \, p_M(\gamma)\, d\gamma$$

let $x = \frac{\gamma}{\Gamma}$

$\gamma = x\Gamma$

$d\gamma = dx\,\Gamma$

$$\bar{\gamma} = \frac{M}{\Gamma}\int_0^\infty (1 - e^{-x})^{M-1} e^{-x}\, dx\, \Gamma\, \gamma$$

$$\bar{\gamma} = \Gamma \int_0^\infty M x (1 - e^{-x})^{M-1} e^{-x}\, dx$$ $\Rightarrow$ this yields the result

$$\frac{\bar{\gamma}}{\Gamma} = \sum_{k=1}^{M}\left(\frac{1}{k}\right) \quad \text{—— (6)}$$

Γ is the average SNR for a single branch (when no diversity is used).

equation (6) shows the SNR improvement offered by selection diversity.

Example

Assume 5 branch diversity is used, where each branch receives an independent Rayleigh fading signal. If the average SNR is 20dB, determine the probability that the SNR will drop below 10 dB. Compute the mean SNR. Compare this with the case of a single receiver without diversity.

Solution

Five branch diversity

$\gamma = 10 dB$ $\qquad$ $\Gamma = 20 dB$

$\frac{\gamma}{\Gamma} = 0.1$ $\qquad$ $P_M(\gamma) = (1 - e^{-\frac{\gamma}{\Gamma}})^M$

$P_1(10dB) = (1 - e^{-0.1})^1 = 0.095$ when $M=1$

$P_5(10dB) = (1 - e^{-0.1})^5 = 0.0000078$ when $M=5$

Mean SNR $= \bar{\gamma} = \Gamma(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}) = 20 \times 2.28$

$= 45.6$ dB

$45.6 dB \gg 20 dB$

i.e., due to selection diversity average SNR improves multifold.

MRC improvement

In MRC, the voltage signals r_i from each of the M diversity branches are co-phased to provide coherent voltage addition and are independently weighted to provide optimal SNR.

If each branch has gain G_i, then the resulting signal envelope applied to the detector is,

$$r_{MT} = \sum_{i=1}^{M} G_i r_i \quad \text{--- (1)}$$

Assuming that each branch has the same average noise power N, the total noise power N_T applied to the detector is simply the weighted sum of the noise in each branch.

$$N_T = N \sum_{i=1}^{M} G_i^2 \quad \text{--- (2)}$$

SNR applied to the detector γ_M is given by

$$\gamma_M = \frac{r_{MT}^2}{2 N_T} \quad \text{--- (3)}$$

using Chebychev inequality, γ_M is maximized when $G_i = \frac{r_i}{N}$ which lead to

$$\gamma_M = \frac{1}{2} \frac{\sum \left(\frac{r_i^2}{N}\right)^2}{N \sum \left(\frac{r_i^2}{N^2}\right)} = \frac{1}{2} \sum_{i=1}^{M} \frac{r_i^2}{N}$$

$$\gamma_M = \sum_{i=1}^{M} \gamma_i \quad \text{--- (4)}$$

From equation (4) it follows that, the average SNR, γ_M is simply the sum of the individual SNR γ_i from each branch.