EC8652 WIRELESS COMMUNICATION L T P C 3 0 0 3

OBJECTIVES:

The student should be made to:

- Know the characteristic of wireless channel
- Learn the various cellular architectures
- Understand the concepts behind various digital signaling schemes for fading channels
- Be familiar the various multipath mitigation techniques
- Understand the various multiple antenna systems

UNIT I WIRELESS CHANNELS 9

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.

UNIT II CELLULAR ARCHITECTURE 9

Multiple Access techniques - FDMA, TDMA, CDMA – Capacity calculations–Cellular concept- Frequency reuse - channel assignment- hand off- interference & system capacity- trunking & grade of service – Coverage and capacity improvement.

UNIT III DIGITAL SIGNALING FOR FADING CHANNELS 9

Structure of a wireless communication link, Principles of Offset-QPSK, p/4-DQPSK, Minimum Shift Keying, Gaussian Minimum Shift Keying, Error performance in fading channels, OFDM principle – Cyclic prefix, Windowing, PAPR.

UNIT IV MULTIPATH MITIGATION TECHNIQUES 9

Equalisation – Adaptive equalization, Linear and Non-Linear equalization, Zero forcing and LMS Algorithms. Diversity – Micro and Macrodiversity, Diversity combining techniques, Error probability in fading channels with diversity reception, Rake receiver,

UNIT V MULTIPLE ANTENNA TECHNIQUES 9

MIMO systems – spatial multiplexing -System model -Pre-coding - Beam forming - transmitter diversity, receiver diversity- Channel state information-capacity in fading and non-fading channels.

TOTAL: 45 PERIODS

UNIT I

WIRELESS CHANNELS

LARGE SCALE PATH LOSS

FREE SPACE PROPOGATION MODEL:

The free space propagation model is used to predict received signal strength when the transmitter and receiver have a clear, unobstructed line of the sight path between them. Satellite communication systems and microwave line – of- sight radio links typically undergo free space propagation.

In large – scale radio wave propagation models, the free space model predicts that received power decays as a function of T-R separation distance raised to some power. The free space power received by receiver antenna which is separated from a radiating transmitter antenna by a distance d, is given by 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, Pt – transmitted power

Pr (d) – received power which is a function of T–R separation

Gt – Transmitted antenna gain

d – T-R separation distance in meters

L – System loss factor not related to propagation (L≥1)

λ- Wavelength in meters

The gain of an antenna is related to its effective aperture, Ae, by

$$G = \frac{4\pi A_e}{\lambda^2} \qquad (2)$$

The effective aperture Ae is related to the physical size of the antenna, λ is related to the carrier frequency b

$$\lambda = \frac{c}{f} = \frac{2\pi c}{\omega_c} \longrightarrow (3)$$

Where, f – carrier frequency in Hertz

Wc – carrier frequency in radians / second

C- speed of light in meter / second

The values for Pt and Pr must be expressed in the same units, and Gt and Gr is dimensionless quantities. The various losses L(L≥1) are usually due to transmission line attenuation system. A value of L=1 indicates no loss in the system hardware.

The 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.

The isotropic radiation 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.

$$EIRP = P_t G_t \longrightarrow (4)$$

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

The path loss, represents the signal attenuation as a positive quantity measured in dB is clear as the difference (in dB) between the effective transmitted power and the received power, may or may not be included the effect of the antenna gains.

The path loss for the free space model when antenna gains are included is given by,

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

\

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

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

The Friis free space model is only a valid prediction for Pr for values of which are in the far–field of the transmitting antenna. The far – field of (or) fraunhofer region, of a transmitting antenna is defined as the region beyond the largest linear dimensions of transmitter antenna aperture and the carrier wave length.

The fraunhofer distance is given by

$$d_f = \frac{2D^2}{\lambda} \quad \longrightarrow 7$$

Where,
D – Largest physical linear dimension of the antenna.
Equation 1 does not hold for d=0. So, large scale propagation models use a close – in distance do as a known received power reference point.

The received power in free space at a distance greater than do is given by

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

In mobile radio systems, it is unusual to find that Pr may change by many orders of magnitude over a typical coverage area of several square kilometers. Because of the large dynamic range of received power levels, often dBm or dBw units are used to express received power levels

Equation 8 can be expressed in dBm or dBw by simply taking logarithm of both sides and multiplying by 10. For E.g., If Pr is in units of dBm, the received power is given by,

$$\Delta = d'' - d' \approx \frac{2h_t h_r}{d}$$

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

where , Pr(d0) is in wat

Two-Ray Ground Reflection Model :

In a mobile radio channel, a single direct path between the base station and a mobile is rarely occurs for propagation and hence the free space propagation model is in most cases inaccurate when used alone.

The two-ray ground reflection model is a useful propagation model that is based on geometric options, and considers both the direct path and a ground reflected propagation path between transmitter and receiver.

The total received E-field; ETOT is then a result of the direct line of sight component, ELOS and the ground reflected component Eg.

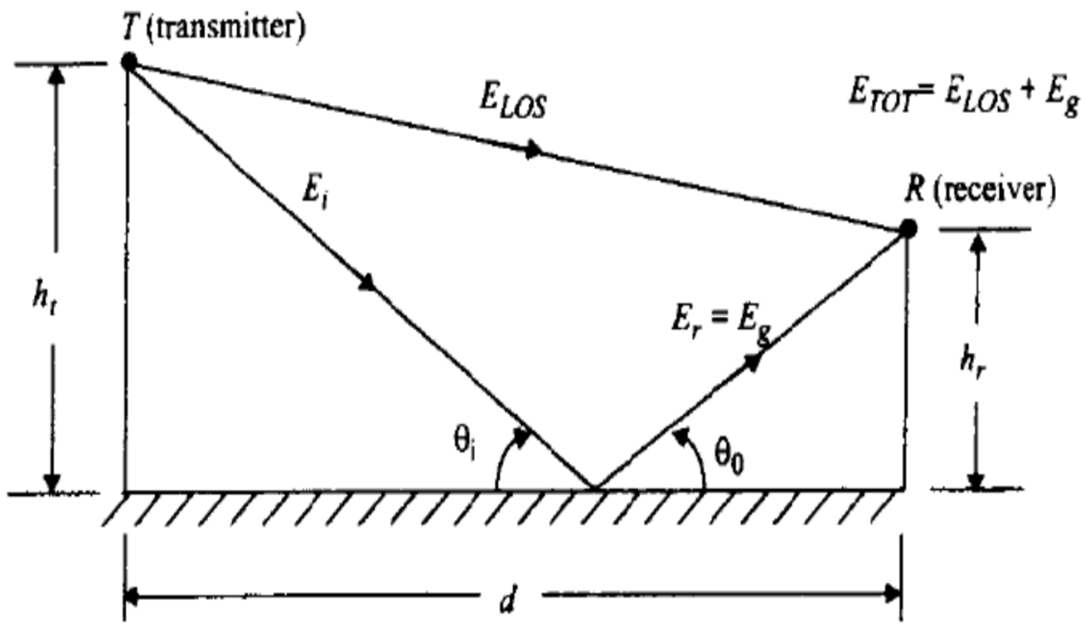


Fig1. Two-ray ground reflection model

In this above figure.1, ht is the heights of transmitter and hr is the height of the receiver. If Eo is free space E-field (units V/m)at a reference distance do from the transmitter, then for d>do, the free space propagating E-field is given by,

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

Where E(d,t) = Eodo/d represents the envelope the E-field at d meters from the transmitter.

Two propagating waves arrive at the receiver. The direct wave that travels a distance d′ and the reflected wave that travels distance d″ . The E field due to the Los component at the receiver can be expressed as,

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

The E-field for the ground reflected wave, which has a propagation distance of d″ can be expressed as

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

According to laws of reflection in dielectrics is expressed as,

$$\theta_i = \theta_0 \longrightarrow [4]$$

$$E_g = \Gamma E_i$$

$$E_t = (1+\Gamma)E_i \longrightarrow [4a] \longrightarrow [4b]$$

Where Γ́ is the reflection co-efficient for ground. For small values of θi , the reflected wave is equal in magnitude and 180° out of phase with the incident wave.

The resultant E-field, assuming perfect horizontal E-field polarization and ground reflection is the vector sum of ELOS and Eg, and the resultant total E-field envelope is given

$$|E_{TOT}| = |E_{LOS} + E_g| \longrightarrow [5]$$

The electric field ETOT (d,t) can be expressed as by using the equ 2&3,

$$E_{TOT}(d,t) = \frac{E_0 d_0}{d'}\cos\left(\omega_c\left(t-\frac{d'}{c}\right)\right) + (-1)\frac{E_0 d_0}{d''}\cos\left(\omega_c\left(t-\frac{d''}{c}\right)\right) \longrightarrow (6)$$

By using the method of images and using the geometric figure 2, the path difference Δ,between the LOS and ground reflected paths can be given as

$$\Delta = d'' - d' = \sqrt{(h_t+h_r)^2+d^2} - \sqrt{(h_t-h_r)^2+d^2} \longrightarrow (7)$$

When the T-R separation distance d is very large compared to ht+hr, the above equation 7 can be simplified by using taylor's series approximation,

$$\Delta = d'' - d' \approx \frac{2h_t h_r}{d} \longrightarrow (8)$$

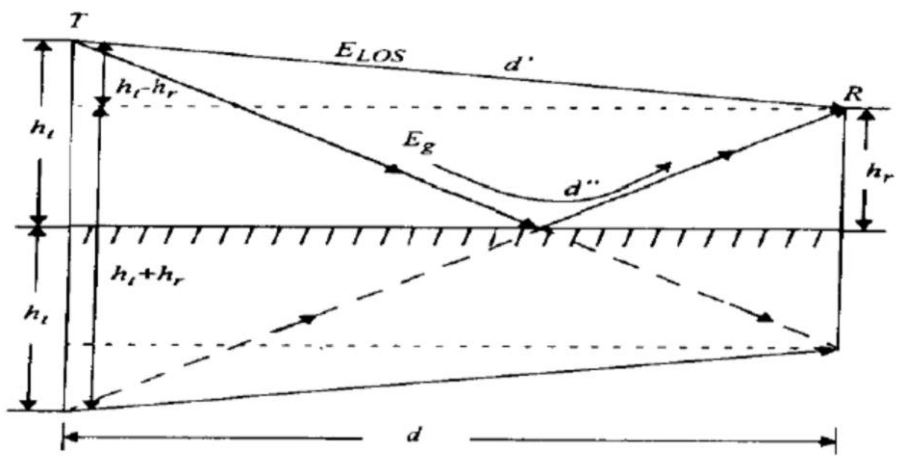


Fig2. Geometry for find the path difference between LOS and ground reflected paths.

Once the path difference is known, the phase difference $\theta\Delta$ between the two E-field components and time delay ζd between the arrivals of the two components can be easily computed using the following relations

$$\theta_\Delta = \frac{2\pi\Delta}{\lambda} = \frac{\Delta\omega_c}{c} \longrightarrow (9)$$

$$\tau_d = \frac{\Delta}{c} = \frac{\theta_\Delta}{2\pi f_c} \longrightarrow (10)$$

If d becomes large, the difference between the distance d’ and d’’ becomes very small and the amplitudes of ELOS and Eg are virtually identical and differ only in phase i.e.

$$\left|\frac{E_0 d_0}{d}\right| \approx \left|\frac{E_0 d_0}{d'}\right| \approx \left|\frac{E_0 d_0}{d''}\right| \qquad \rightarrow (11)$$

If received E-field is evaluated at some time say at t=d”/c, then equ.6 becomes,

$$\begin{aligned} E_{TOT}\left(d, t = \frac{d''}{c}\right) &= \frac{E_0 d_0}{d'}\cos\left(\omega_c\left(\frac{d''-d'}{c}\right)\right) - \frac{E_0 d_0}{d''}\cos 0^o \\ &= \frac{E_0 d_0}{d'}\cos\theta_\Delta - \frac{E_0 d_0}{d''} \qquad \rightarrow (12) \\ &\approx \frac{E_o d_0}{d}\left[\cos\theta_\Delta - 1\right] \end{aligned}$$

Where d-distance over flat earth between the bases of transmitter and receiver antennas. The below phasor diagram 3 represents how the direct and ground reflected rays combine the electric field at a distance d from the transmitter can be written

$$|E_{TOT}(d)| = \sqrt{\left(\frac{E_0 d_0}{d}\right)^2 (\cos\theta_\Delta - 1)^2 + \left(\frac{E_0 d_0}{d}\right)^2 \sin^2\theta_\Delta} \qquad \rightarrow (13)$$

as,

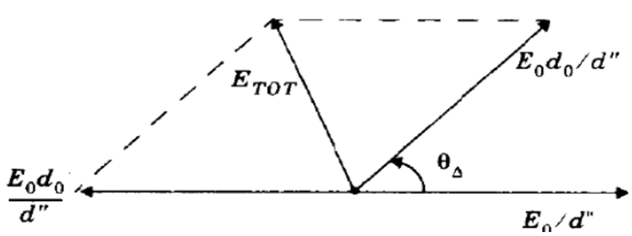


Fig3.Phasor diagram of LOS, ground reflected and total received E-field

$$|E_{TOT}(d)| = \frac{E_0 d_0}{d}\sqrt{2 - 2\cos\theta_\Delta} \qquad \rightarrow (14)$$

By using trigonometric identities, above equation 14 can be expressed as

$$|E_{TOT}(d)| = 2\frac{E_0 d_0}{d}\sin\left(\frac{\theta_\Delta}{2}\right) \qquad \rightarrow (15)$$

The above equ 15 provides the exact received E-field for the two ray ground reflection model.For increasing distance form transmitter ETOT decays in an oscillatory fashion, with local maxima being 6dB greater than free space value and local maxima plummeting to -∞ dB.The equ.15 may be simplified whenever $\sin(\theta\Delta /2) \approx \theta\Delta /2$. It occurs when $\theta\Delta /2$ is < 0.3. Using equ. 8&9

$$\frac{\theta_\Delta}{2} \approx \frac{2\pi h_t h_r}{\lambda d} < 0.3 \text{ rad} \qquad \rightarrow (16)$$

and equation 15 simplified

$$d > \frac{20\pi h_t h_r}{3\lambda} \approx \frac{20 h_t h_r}{\lambda} \qquad \rightarrow (17)$$

Thus as long as d satisfies the above equ,17 then received E-field is given by,

$$E_{TOT}(d) \approx \frac{2E_0 d_0}{d}\frac{2\pi h_t h_r}{\lambda d} \approx \frac{k}{d^2} \text{ V/m} \qquad \rightarrow (18)$$

K-constant related to Eo, antenna heights and wavelength λ. This asymptotical behavior is identical for both E-field in the plane of incidence or normal to plane of incidence.

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

$$P_r = P_t G_t G_r \frac{h_t^2 h_r^2}{d^4} \qquad \rightarrow (19)$$

At a large distance, $(d \gg \sqrt{h_t h_r})$, the received power falls off with distance raised to fourth power or at rate of 40 dB/decade. This is a much more rapid path loss than in free space.At large values of d, the received power and path loss become independent of frequency.

Thus the path loss of two-ray model can be expressed in dB as,

$$PL(dB) = 40\log d - (10\log G_t + 10\log G_r + 20\log h_t + 20\log h_r) \rightarrow (20)$$

At small T-R separation distance, equ 6 is used to calculate total E-field. When $\theta_\Delta = \Pi$, then $d=(4h_t h_r)/\lambda$ is the ground appears in the first Fresnel zone between the transmitter and receiver. The first Fresnel zone distance is a useful parameter in microcell path loss models.

Advantages:

The two-ray ground reflection model is a useful propagation model that is based on geometric options, and considers both the direct path and a ground reflected propagation path between transmitter and receiver.

Disadvantages:

If the phase of the two paths are identical, then they add together and no fading occurs. When the phase of the two paths differ by 180 degrees, then they cancel each other and fading occurs.

Parameters of Mobile Multipath Channels

Many multipath channel parameters are derived from the power delay profile which is measured using the technique that are generally represented as plots of relative received power as a function of excess delay with respect to a fixed time delay reference.

Power delay profiles are found by averaging instantaneous power delay profile measurements over a local area in order to determine an average small-scale power delay profile. Figure 4.9 shows typical power delay profile plots from outdoor and indoor channels, determined from a large number of closely sampled instantaneous profiles.

Some important parameters of Mobile Multipath Channels are:

1. Time Dispersion Parameters
2. Coherence Bandwidth
3. Doppler Spread and Coherence Time

Time Dispersion Parameters

Time Dispersion Parameters are used o to compare different multipath channels and to develop some general design guidelines for wireless systems.

The following parameters have to be measured accurately:

- Mean Excess Delay
- Rms Delay Spread
- Excess Delay Spread

The mean excess delay is the first moment of the power delay profile and is defined to be,

$$\bar{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k a_k^2} = \frac{\sum_k P(\tau_k)\tau_k}{\sum_k P(\tau_k)} \tag{1}$$

The rms delay spread is the square root of the second central moment of the power delay profile and is defined to be,

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

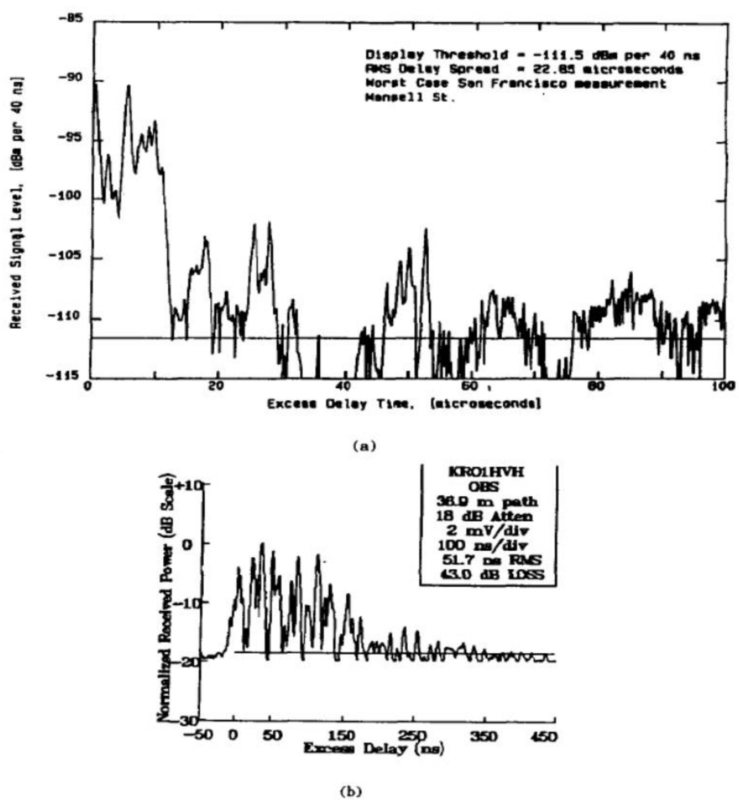


Figure 4.9

Measured multipath power delay profiles

a) From a 900 MHz cellular system in San Francisco [From [Rap90] © IEEE].

b) Inside a grocerv store at 4 GHz [From [Haw91] © IEEE].

$$\overline{\tau^2} = \frac{\sum_k a_k^2 \tau_k^2}{\sum_k a_k^2} = \frac{\sum_k P(\tau_k)\, \tau_k^2}{\sum_k P(\tau_k)} \longrightarrow (3)$$

These delays are measured relative to the first detectable signal arriving at the receiver at $\tau_0 = 0$. Equations 1 - 3 do not rely on the absolute power level of $P(\tau)$, but only the relative amplitudes of the multipath components within $P(\tau)$.

It is important to note that the rms delay spread and mean excess delay are defined from a single power delay profile which is the temporal or spatial average of consecutive impulse response measurements collected and averaged over a local area. The maximum excess delay (X dh) of the power delay profile is defined to be the time delay during which multipath energy falls to X dB below the maximum and it is given by,

$$\tau_X - \tau_0 \longrightarrow (4)$$

where τ_0 is the first arriving signal and τ_X is the maximum delay at which a multipath component is within X dB of the strongest arriving multipath signal.

Figure 4.10 illustrates the computation of the maximum excess delay for multipath components within 10 dB of the maximum. The maximum excess delay (X dB) defines the temporal extent of the multipath that is above a particular threshold. The value of is sometimes called the excess delay spread of a power delay profile, but in all cases must be specified with a threshold that relates the multipath noise floor to the maximum received multi-path component.

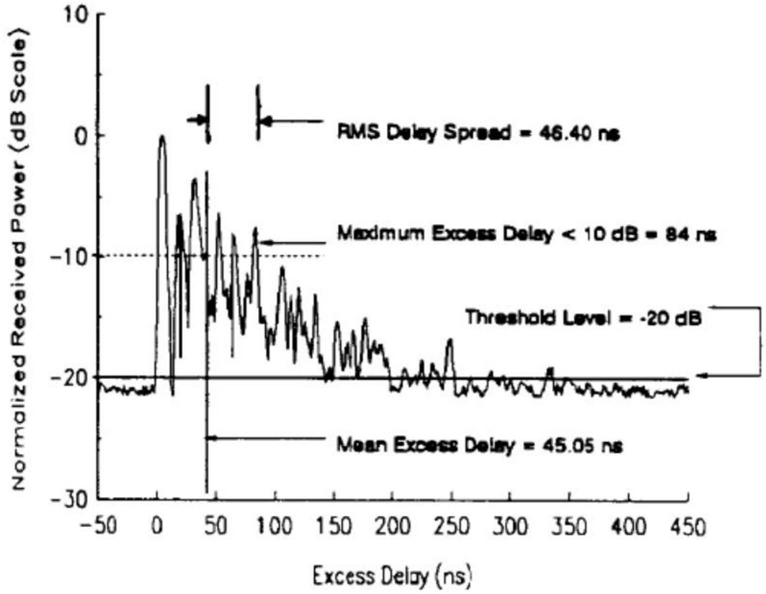


Figure 4.10
Example of an indoor power delay profile; rms delay spread, mean excess delay, maximum excess delay (10 dB), and threshold level are shown.

Coherence Bandwidth

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).It is a defined relation derived from the rms delay spread. 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.

$$B_c \approx \frac{1}{50\sigma_\tau} \longrightarrow 5$$

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

$$B_c \approx \frac{1}{5\sigma_\tau} \longrightarrow 6$$

where, σ_τ is Average r.m.s delay spread

It is important to note that an exact relationship between coherence band-width and rms delay spread does not exist, and equations 5 and 6 are "ball park estimates". Thus coherence bandwidth depends on frequency correlation function.

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. But the time varying nature of the channel in a small-scale region is described by the parameter Doppler spread and coherence time.

Doppler spread B_D is defined as the set of frequencies over which the Doppler spectrum at the receiver end is essentially non-zero. When a pure sinusoidal tone of 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 . If the baseband signal bandwidth is much greater than BD, the effects of Doppler spread are negligible at the receiver. This is a slow fading channel.

The Doppler spread and coherence time are inversely proportional to one another. That is,

$$T_C \approx \frac{1}{f_m}$$

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

$$T_C \approx \frac{9}{16\pi f_m}$$

Where, f_m is the maximum Doppler shift given by $f_m = v/\lambda$. The term v refers to velocity and λ refers to wavelength .The quality of received signal thus depends on coherence time factor.

Coherence Bandwidth

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).It is a defined relation derived from the rms delay spread. 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.

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

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

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

where, σ_τ is Average r.m.s delay spread.

It is important to note that an exact relationship between coherence band-width and rms delay spread does not exist, and the above equations are "ball park estimates". Thus coherence bandwidth depends on frequency correlation function.

Types of small scale fading:

Fading effects due to Multipath Time delay spread

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

Small-Scale Fading

(Based on multipath time delay spread)

Flat Fading

1. BW of signal < BW of channel

2.Delay spread < Symbol period

Frequency Selective Fading

1.BW of signal > BW of channel

2.Delay spread>Symbol period

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. The characteristics of a flat fading channel are illustrated in Figure 4.12

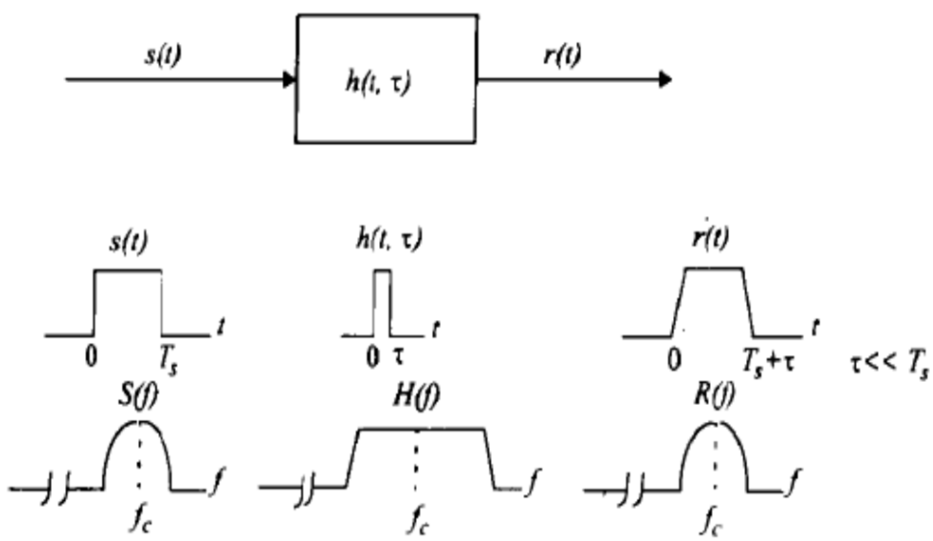


Figure 4.12
Flat fading channel characteristics.

It can be seen from Figure 4.12 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.

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.

The distribution of the instantaneous gain of flat fading channels is important for designing radio links, and the most common amplitude distribution is the Rayleigh distribution.

To summarize, a signal undergoes flat fading if, $B_S \ll B_C$ and

$$T_S \gg \sigma_\tau$$

where, Ts is the reciprocal bandwidth

Bs is the bandwidth

σ_τ and Bc are the rms delay spread and coherence bandwidth respectively.

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 time dispersion of the transmitted symbols within the channel. Thus the channel induces intersymbol interference (ISI).Frequency selective fading channels are much more difficult to model than flat fading channels due to the following reasons,

- Each Multipath signal must be modeled and
- The channel must be considered to be a linear filter

The statistical impulse response models such as the 2-ray Rayleigh fading model or computer generated or measured impulse responses, are generally used for analyzing frequency selective small-scale fading.

Figure 4.13 illustrates the characteristics of a frequency selective fading channel. In frequency selective fading gain is different for different frequency components but the channel becomes frequency selective when Viewed in the frequency domain

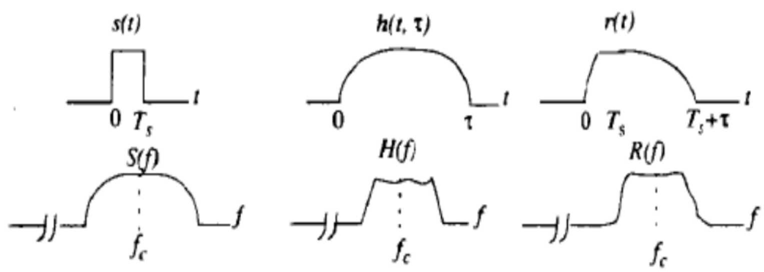


Figure 4.13
Frequency selective fading channel characteristics.

.

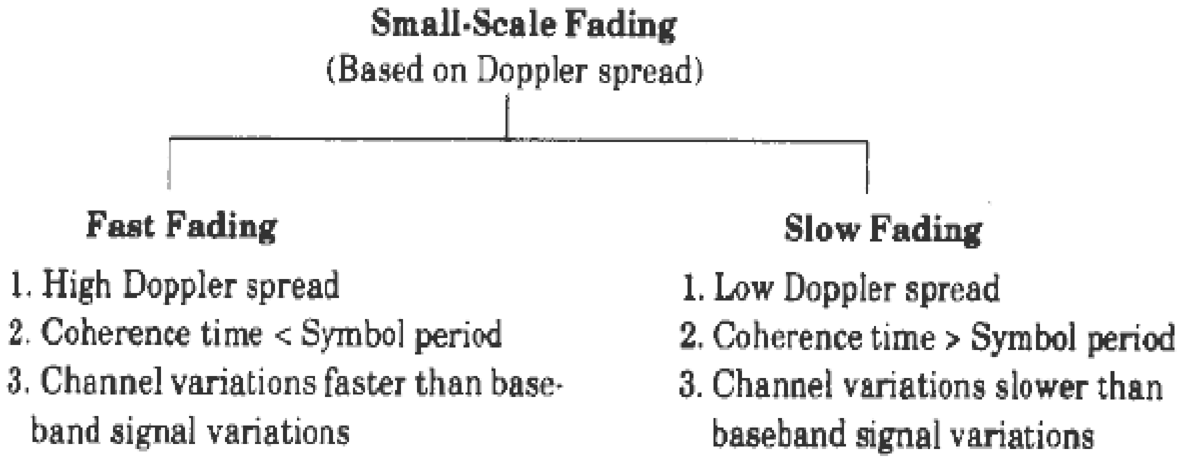
Frequency selective fading is caused by multipath delays' which approach or exceed the symbol period of the transmitted symbol. Frequency selective fading channels are also known as wideband channels since the bandwidth of the signal s{t) is wider than the bandwidth of the channel impulse response.

To summarize, a signal undergoes frequency selective fading if

$B_S > B_C$ and

$T_S < \sigma_\tau$

Fading effects due to Dopper Spread



Fast Fading

When comparing the rate of change of the radio channel the baseband signal transmitted will change .Depending on how fast it takes place 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. Therefore, a signal undergoes fast fading if

$$T_S > T_C \text{ and}$$

$$B_S < B_D$$

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 fiat 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.

Slow Fading

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. Therefore, a signal undergoes slow fading if

$$T_S \ll T_C \text{ and}$$

$$B_S \gg B_D$$

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.The relation between the various multipath parameters and the type of fading experienced by the signal are summarized in Figure 4.14.

UNIT II

CELLULAR ARCHITECTURE

Frequency Division Multiple Access (FDMA)

FDMA describes scheme to sub divide the bandwidth into several non-over lapping frequecny bands.FDMA assigns individual channls to individual users each user allocates a unique frequency band or channelIn FDMA the users are assigned a channles as a pair of frequecies. One is forward channel and other for reverse channel.

FDMA is distinct from frequency division duplexing (FDD). While FDMA allows multiple users simultaneous access to a transmission system, FDD refers to how the radio channel is shared between the uplink and downlink (for instance, the traffic going back and forth between a mobile-phone and a mobile phone base station).

Frequency- division multiplexing (FDM) is also distinct from FDMA. FDM is a physical layer technique that combines and transmits low-bandwidth channels through a high-bandwidth channel. FDMA, on the other hand, is an access method in the data link layer.

Features of FDMA are as follows

(i) Channel only one phone circuit at a time.

(ii) If the FDMA channel is not in use, it can’t be used by other user to increase the capacity.

(iii) Bandwidth of FDMA channel is very narrow about 30khz as each channel supports one circuit per carrier.

(iv) After the assignment of a voice channel base station and mobile station transmit simultanously

(v) FDMA completly eliminates co channel interferecne

(vi) The complexcity of FDMA mobile system is lower than TDMA system

(vii) FDMA requires exact RF filtering to minimize the adjcent channel interference

(viii) In FDMA all users share the satellite transponder or frequency channel simultaneously but each user transmits at single frequency.

(ix) FDMA can be used with both analog and digital signa

(x) FDMA requires high-performing filters in the radio hardware, in contrast to TDMA and CDMA.

(xi) FDMA is not vulnerable to the timing problems that TDMA has. Since a predetermined frequency band is available for the entire period of communication, stream data (a continuous flow of data that may not be packetized) can easily be used with FDMA.

(xii) Due to the frequency filtering, FDMA is not sensitive to near-far problem which is pronounced for CDMA.

(xiii) Each user transmits and receives at different frequencies as each user gets a unique frequency slots.

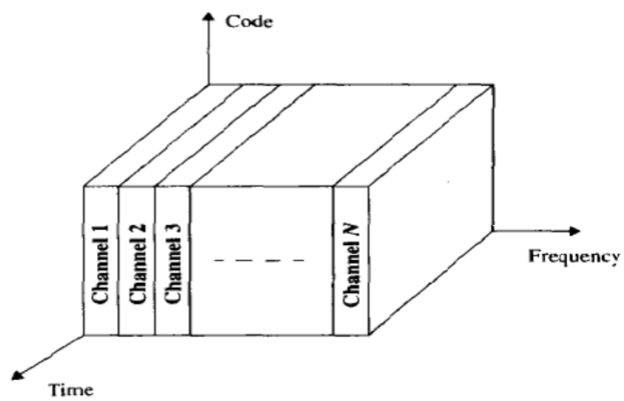


Figure 8.2
FDMA where different channels are assigned different frequency bands.

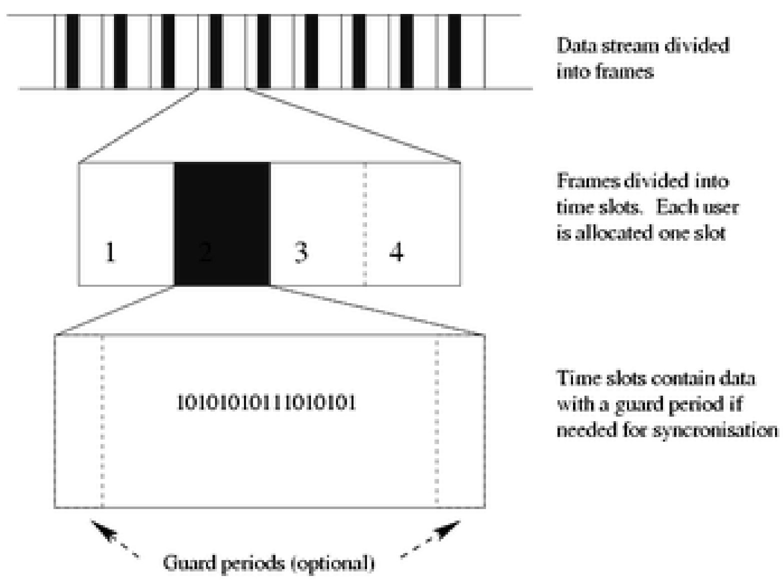
<u>Time Division Multiple Access</u>(TDMA)

TDMA system divides the radio spectrum into time slot and in each slot only one user is allowed to either transmit or to receive. TDMA system transmit data in buffer and burst method thus the transmission of any user is not continous.

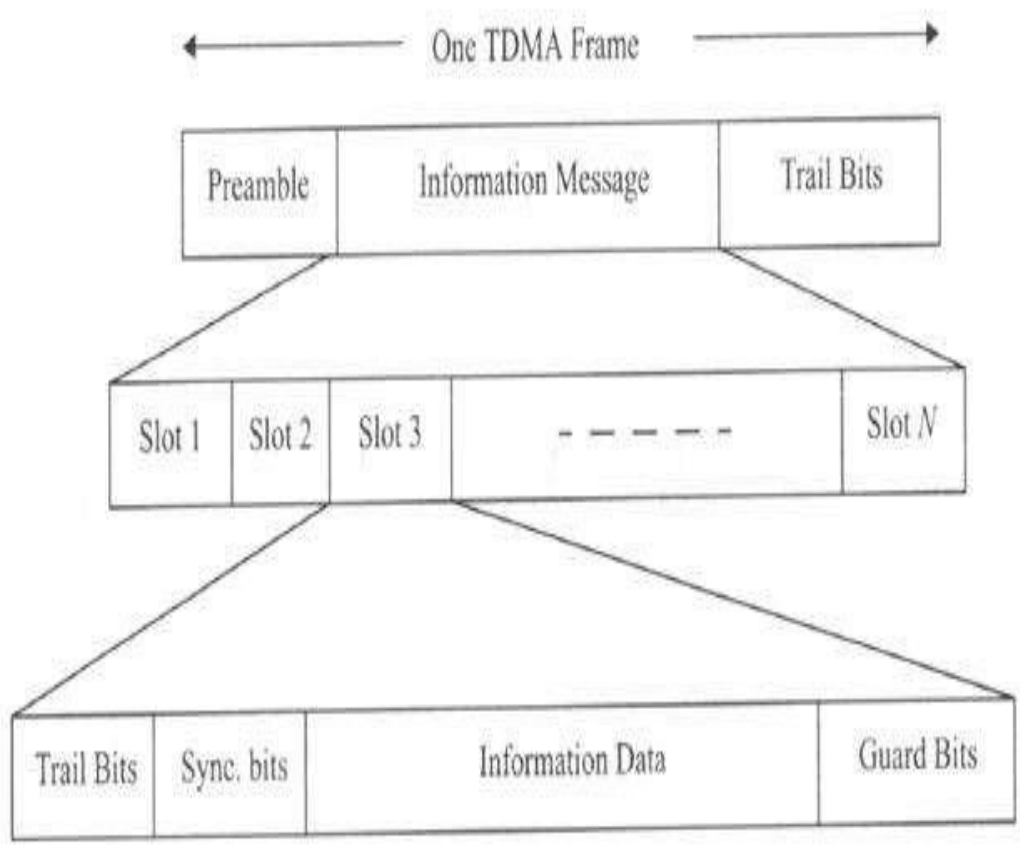
This type of transmission unlike in FDMA systemEach frame is made up of preamble, information message and trails bits.In TDMA/TDD half of the time slot in the frame information message used for forward link and another half is for reverse link but the carrier frequency is different for forward link and reverse link.

Time division multiple access (TDMA) is a channel access method for shared medium networks. It allows several users to share the same frequency channel by dividing the signal into different time slots. The users transmit in rapid succession, one after the other, each using its own time slot. This allows multiple stations to share the same transmission medium (e.g. radio frequency channel) while using only a part of its channel capacity.

TDMA is used in the digital 2Gcellular systems such as Global System for Mobile Communications (GSM), IS-136, Personal Digital Cellular (PDC) and in the Digital Enhanced Cordless Telecommunications (DECT) standard for portable phones. It is also used extensively in satellite systems, combat-net radio systems, and PON networks for upstream traffic from premises to the operator. For usage of Dynamic TDMA packet mode communication, given below in fig.



TDMA frame structure showing a data stream divided into frames and those frames divided into time slots.TDMA is a type of Time-division multiplexing, with the special point that instead of having one transmitter connected to one receiver, there are multiple transmitters. In the case of the uplink from a mobile phone to a base station this becomes particularly difficult because the mobile phone can move around and vary the timing advance required to make its transmission match the gap in transmission from its peers.



Features of TDMA are as follows:

(i) TDMA stores a single carrier frequency with several users without overlapping of time slots

(ii) Data transmission for users of a TDMA system is not continous but occurs in buffer and brust . this results in low battery consumption.

(iii) TDMA uses different time slot for transmission and reception thus duplexers are not required

(iv) High synchronization overhead is required in TDMA systems in brust transmission

(v) TDMA has an advantage in that it is possible to allocate different no of time slots per frame to different users. Thus bandwidth can be supplied on demand to different users

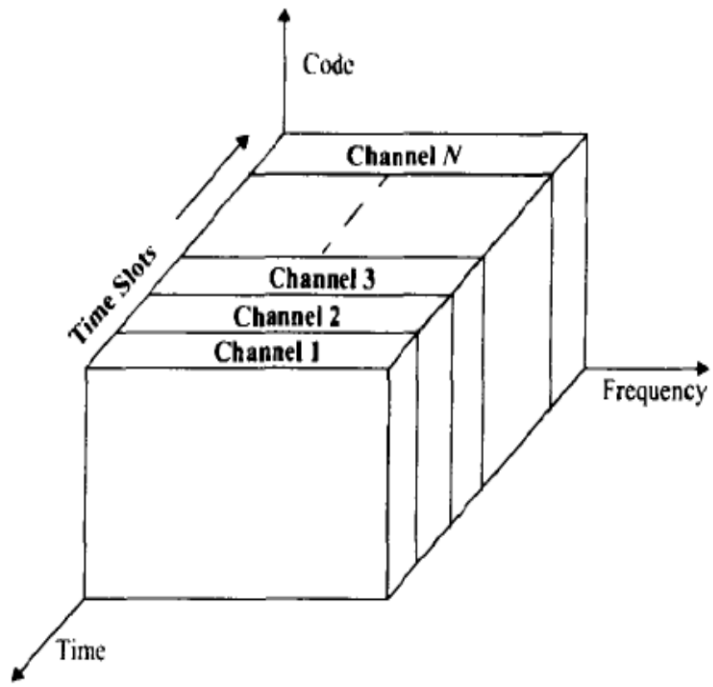


Figure 8.3
TDMA scheme where each channel occupies a cyclically repeating time slot.

<u>Code division multiple access</u>(CDMA):

In CDMA system, a narrow band message signal in multipiled by a large bandwidth signal called spreading signal .The spreading signal is PSEUDO RANDOM sequence that has chip rate is much higher than the data rate of the message signal. All the users in CDMA use the same channel frequency and may transmit simultanously.

Each and every users has its own PSEUDO RANDOM code which is orthogonal to all other user's code. The near far problem occurs when many mobile users share the same channel stronger received signal acts as noise for the weaker signal. There by decreasing the probability that weaker signal will be received to combat near far problem, power control is used in CDMA system.

Power control is implemented at the base station by rapidly sampling the radio signal strengh indicator (RSSI)level of each mobile and then sending a power change command over forward link Features of CDMA are as follows:Many users of CDMA system share the same frequency Unlike in TDMA and FDMA, CDMA has soft capacity limit there is no absoulte limit on no of users in CDMA.RAKE RECEIVER can be used to improve reception by collecting time delayed versions of the desired signal. The near far problem occurs at a CDMA receiver if an undesired user has a high detected power as compared to desired user.

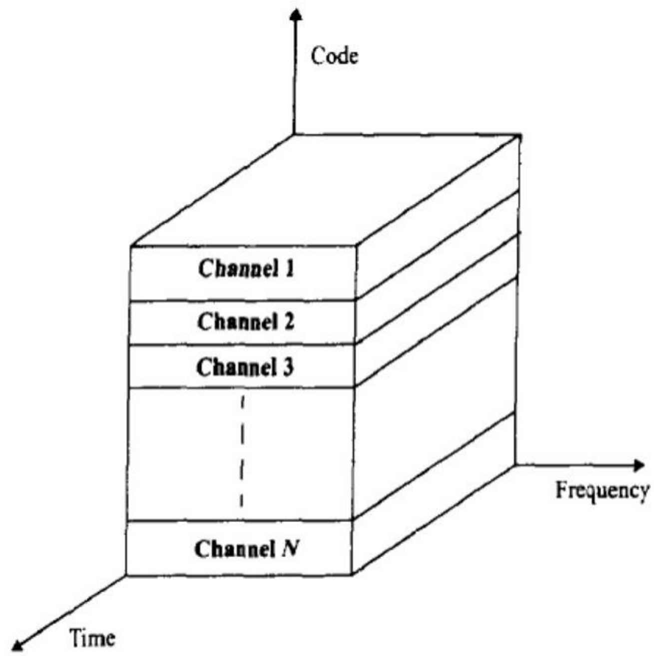


Figure 8.5
CDMA in which each channel is assigned a unique PN code which is orthogonal to PN codes used by other users.

Frequency reuse and Channel Assignment Strategies

Frequency reuse:

Cellular radio systems rely on an intelligent allocation and reuse of channels throughout a coverage region. Each cellular base station is allocated a group of radio channels to be used within a small geographic area called a cell.

The base station antennas are designed to achieve the desired coverage within the particular cell. By limiting the coverage area to within the boundaries of a cell, the same group of channels may be used to cover different cells that are separated from one another by distances. The design process of selecting and allocating channel groups for all of the cellular base stations within a system is called frequency reuse or frequency planning.

Figure 2.1 illustrates the concept of cellular frequency reuse, where cells labeled with the same letter use the same group of channels.

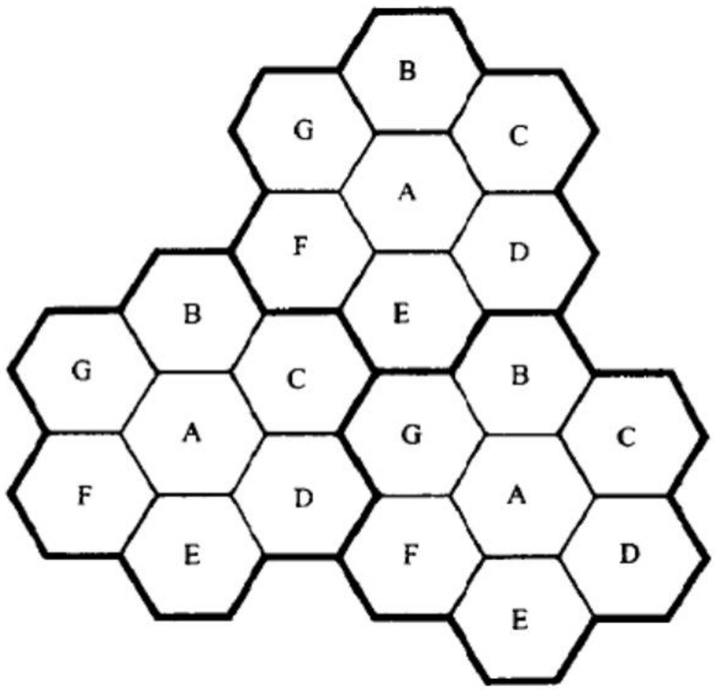


Fig 2.1: Frequency reuse(N=7)

Figure 2.1 is conceptual and is a simplistic model of the radio coverage for each base station, but it has been universally adopted since the hexagon permits easy and manageable analysis of a cellular system. The actual radio coverage of a cell is known as the footprint and is determined from field measurements or propagation prediction models.

To understand the frequency reuse concept, consider a cellular system which has a total of S duplex channels available for use. 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 which each have the same number of channels, the total number of available radio channels can be expressed as

$$S = kN$$

The N cells which collectively use the complete set of available frequencies is called a cluster. If a cluster is replicated M times within the system, the total number of duplex channels, C, can be used as a measure of capacity and is given

$$C = MkN = MS$$

As seen from equation 2, the capacity of a cellular system is directly proportional to the number of times a cluster is replicated in a fixed service area. The factor N is called the duster size and is typically equal to 4, 7, or 12.

The frequency reuse factor of a cellular system is given by l/N, since each cell within a cluster is only assigned I /N of the total available channels in the system. To connect without gaps between adjacent cells the geometry of hexagons is such that the number of cells per cluster, N can only have values which satisfy equation .

$$N = i^2 + ij + j^2$$

where i and j are non-negative integers. To find the nearest co-channel neighbors of a particular cell, one must do the following (1) move i cells along any chain of hexagons and then

(2) turn 60 degrees counter-clockwise and move / cells.

Channel Assignment Strategies:

A variety of channel assignment strategies have been developed to achieve efficient utilization of available radio spectrum. Channel assignment strategies can be classified as

- Fixed channel assignment strategy

- Dynamic channel assignment strategy.

The choice of channel assignment strategy impacts the performance of the system, particularly as to how calls are managed when a mobile user is handed off from one cell to another.

Fixed channel assignment strategy:

In a fixed channel assignment strategy, each cell is allocated a predetermined set of voice channels. Any call attempt within the cell can only be served by the unused channels in that particular cell. If all the channels in that cell are occupied, the call is blocked and the subscriber does not receive service.

Several variations of the fixed assignment strategy exist. In one approach, called the borrowing strategy, a cell is allowed to borrow channels from a neighboring cell if all of its own channels are already occupied. The mobile switching center (MSC) supervises such borrowing procedures and ensures that the borrowing of a channel does not disrupt or interfere with any of the calls in progress in the donor cell.

Dynamic channel assignment strategy:

In a dynamic channel assignment strategy, voice channels are not allocated to different cells permanently. Instead, each time a call request is made, the serving base station requests a channel from the MSC. The switch then allocates a channel to the requested cell following an algorithm that takes into account the likelihood of future blocking within the cell, the frequency of use of the candidate channel, the reuse distance of the channel, and other cost functions.

Dynamic channel assignment reduce the likelihood of blocking, which increases the trunk-ing capacity of the system, since all the available channels in a market are acces-sible to all of the cells. Dynamic channel assignment strategies require the MSCDynamic channel assignment reduce the likelihood of blocking, which increases the trunk-ing capacity of the system, since all the available channels in a market are acces-sible to all of the cells. Dynamic channel assignment strategies require the MSC.

Handoff strategies

When a mobile station moves from one cell site to another cell site while crossing boundary frequency switch over from one base station to another base station without affecting the call which is in progress

This hand off not only involves identifying the new base station but also requires that the voice and control signals be allocated to channel associated with in the new base stationOnce a particular signal level is specified as the minimum usable signal for acceptable voice quality at the base station receiver {-90dBm or -100dBm}.The threshold level is given by,

$$\Delta = P_{r\ handoff} - P_{r\ minimum\ usable}$$

The Δ cannot be too large or too small. If Δ is too large, unnecessary handoffs which burden the MSC may occur, and if Δ is too small, there may be insufficient time to complete a handoff before a call is lost due to weak signal conditions. Figure 2.2 illustrates a handoff situation. Figure
2.2(a) demonstrates the case where a handoff is not made and the signal drops below the minimum acceptable level to keep the channel active.

This dropped call event can happen when there is an excessive delay by the MSC(Mobile switching Center) in assigning a handoff, or when the threshold A is set too small for the handoff time in the system. Excessive delays may occur during high traffic conditions due to computational loading at the MSC or due to the fact that no channels are available on any of the nearby base stations (thus forcing the MSC to wait until a channel in a nearby cell becomes free).

The base station monitors the signal level for a certain period of time before a hand-off is initiated. This running average measurement of signal strength should be optimized so that unnecessary handoffs are avoided the length of time needed to decide if a handoff is necessary depends on the speed at which the vehicle is moving.

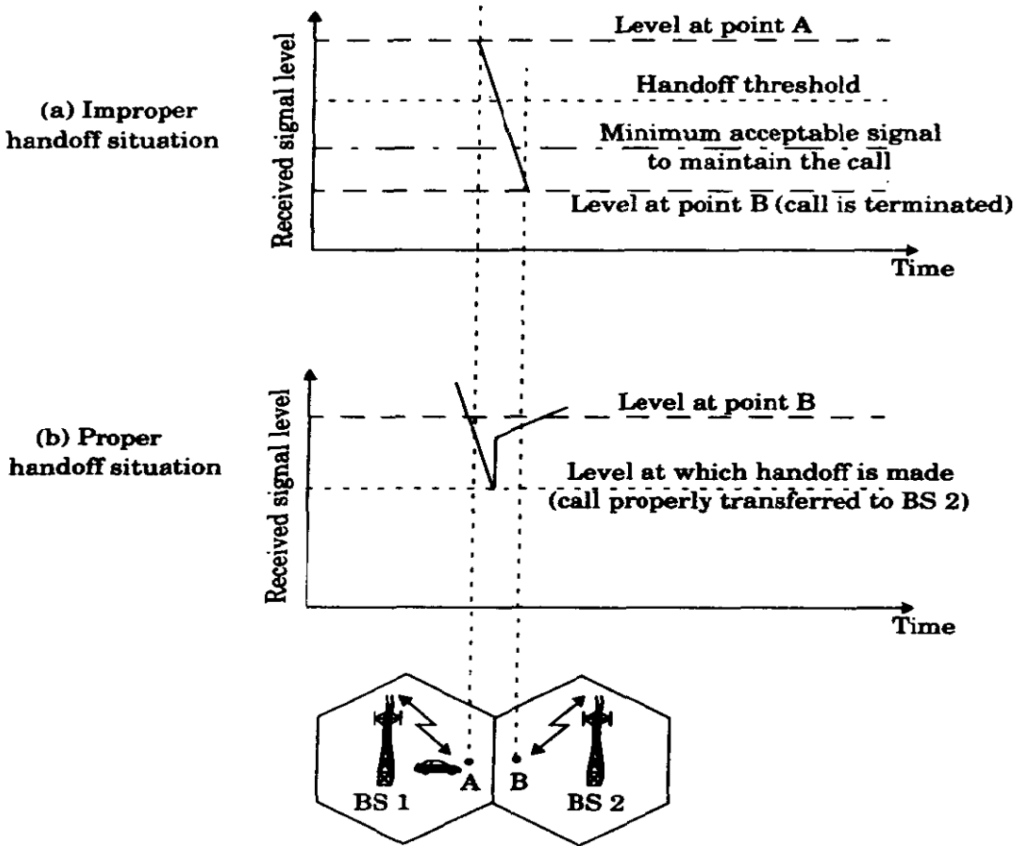


Fig 2.2 Illustration of handoff strategies

The time over which a call may be maintained within a cell, without hand-off, is called the dwell time. The dwell time of a particular user is governed by a number of factors

- Propagation,
- Interference,
- Distance between the subscriber and the base station, and other time varying effects.

In first generation analog cellular systems, signal strength measurements are made by the base stations and supervised by the MSC. Each base station constantly monitors the signal strengths of all of its reverse voice channels to determine the relative location of each mobile user with respect to the base station tower In second generation systems that use digital TDMA technology, handoff decisions are mobile assisted. In mobile assisted handoff(MAHO), every mobile station measures the received power from surrounding base stations and continually reports the results of these measurements to the serving base station.

The MAHO method enables the call to be handed over between base stations at a much faster rate than in first generation analog systems since the handoff measurements are made by each mobile, and the MSC no longer constantly monitors signal strengths. During the course of a call, if a mobile moves from one cellular system to a different cellular system controlled by a different MSC, an intersystem handoff becomes necessary.

An MSC engages in an intersystem handoff when a mobile signal becomes weak in a given cell and the MSC cannot find another cell within its system to which it can transfer the call in progress.

Prioritizing Handoffs

One method for giving priority to handoffs is called the guard channel concept. Guard channels, however, offer efficient spectrum utilization when dynamic channel assignment strategies, which minimize the number of required guard channels by efficient demand-based allocation, are used.

Queuing of handoff requests is another method to decrease the probability of forced termination of a call due to lack of available channels. There is a tradeoff between the decrease in probability of forced termination and total carried traffic.

Practical Handoff Considerations In practical cellular systems, several problems arise when attempting to design for a wide range of mobile velocities. Using different antenna heights (often on the same building or tower) and different power levels, it is possible to provide "large" and "small" cells which are co-located at a single location. This technique is called the umbrella cell approach and is used to provide large area coverage to high speed users while providing small area coverage to users traveling at low speeds.

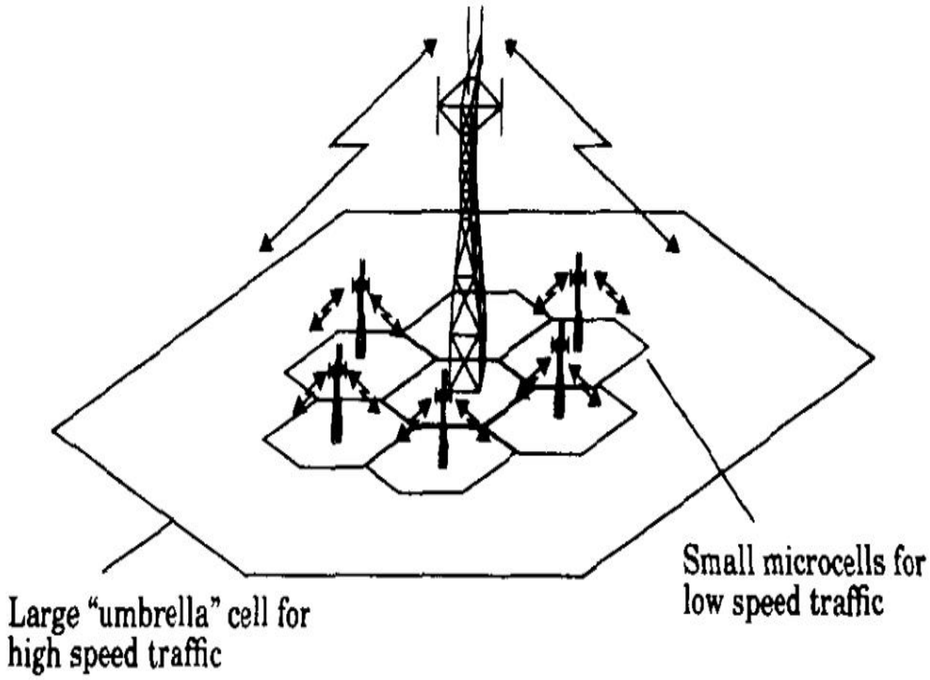


Fig 2.3 Umbrella Cell Pattern

Figure 2.3 illustrates an umbrella cell which is co-located with some smaller microcells. The umbrella cell approach ensures that the number of handoffs is minimized for high speed users and provides additional microcell channels for pedestrian users.

Another practical handoff problem in microcell systems is known as cell dragging. Cell dragging results from pedestrian users that provide a very strong signal to the base station. Such a situation occurs in an urban environment when there is a line-of-sight (LOS) radio path between the subscriber and the base station.

As the user travels away from the base station at a very slow speed, the average signal strength does not decay rapidly.

Trunking and Grade of Service

In cellular mobile communication the two important aspects that has to be considered with more care are,

1) Trunking

2) Grade of Service (GoS)

These aspects have to be well planned so that it will lead to a better system performance.

Trunking

The 'trunking' deals with accommodation of larger number of mobile users in minimum radio spectrum. By using this trunking concept it is possible to allow many users to share smaller number of mobile channels in a cell. It is done by assigning channels on demand basis and allocating a channel from a pool of channels available. That is if an user want to access a channel for establishing a cell then from the pool of channels the required channel will be assigned to the user.

Once the call progress is terminated at the end of the call then the channel used so far will return to the pool and will be ready for any next new access to come. The concept finds application in telephone circuitry, mobile radio communication in a larger way.

For designing trunked radio system that is capable of handling a particulai! capacity at a 'grade of service'. Some fundamental points regarding trunking theory is required. It was developed by a Danish mathematician Erlang. In the 19^{th} century the concept of accommodating large users with limited servers was dealt. The amount of traffic intensity that is carried by one channel is expressed as "one Erlang".

For example, if a mobile radio channel was occupied for 15 minutes of an hourthen it is said as,

$$\text{Traffic intensity} = \frac{1}{4} = 0.25 \text{ Erlangs}$$

Thus trunking is a main concept used to improve system efficiency. Also the term grade of service is closely linked with representation of grade of service.

Grade of Service (GoS)

The Grade of Service (GoS) is another important measure that express the ability of mobile user to access a trunked system mainly in the busiest hour of the day. The busiest hour of day is statistically studied and considered as 4 pm to 6 pm on the Thursday and Friday of a week. The busiest hour is also decided with respect to customer demand in particular hours.

The linked system's performance is defined by its grade of service. Only if the trunked system permits its user to access if even in the busiest hour then such a trunked system is said to be an ideal system.

To meet out an appropriate GoS, the estimation of maximum capacity required for allocating enough number of radio channels in the design is a must. The GoS is also a measure of congestion that is specified as the probability of delaying a call beyond a time limit. The call request rate multiplied with holding time has to be equal to the traffic intensity which is offered by an user. If the traffic intensity in Erlangs generated by user is T_u then,

$$T_u \quad \lambda H$$

Where, $\lambda \rightarrow$ Average number of the call requests/unit time for an user.

$H \rightarrow$ Average call duration.

In case if the traffic is distributed equally among all the channels then the traffic intensity for a single channel. A_c in the trunked system 'C' will be,

$$A_c = NA_u/C$$

Where, $N \rightarrow$ Number of users in a trunked system with an unspecified number of channels in it.

Then there are two category of trunking system is available namely,

1) Trunking system with no queuing for the call requests.

Trunking system with queue provision for holding calls that are blocked. This is given in the tree classification as shown in following Fig

The GoS parameter will be different for the various cellular systems. For example, the AMPS cellular system is mainly designed for GoS of 2 % blocking status, which implies that 2 calls out of 100 calls allotted will be blocked.

The two broad classes of the trunked radio system is also called as Lost Call Cleared (LCC) and Lost Call Delayed (LCD) systems.

In Lost Call Cleared (LCC) system queuing is allowed for the call requests made. If a user wants service it will be served in case of availability of channel within a minimum time. Otherwise the call will be blocked if there is no channel available to assign.

The grade of service is described by Erlang B formulae expressing the probability of an user experiencing blocked call status in the lost call cleared system.

In the second lost call delayed system the queues are made use to hold the callrequests which were initially blocked. In this type if a user makes a call request it will served if channel is available or if the channel is not available at that moment the call will be delayed till the channel becomes available. It is determined by Erlang C formula.

The grade of service in this system will express the probability of till delaying time for a call. There is large number of users available in the second type.

Thus in trunked system there are lost call cleared (no queues used) and lost c delayed (queues used) and the second type of trunking system is widely used.

Summary of some important terms related to trunking theory are as follows :

1.	GoS	The grade of service is a measure of congestion which is a probability of a call which is being blocked.
2.	Request rate	It represents the average number of the call requests made in a unit time and it is denoted by λ/sec.
3.	Load	The available traffic density across the complete trunked radio system.
4.	Setup time	It represents the time needed to allocate a trunked channel for an user who have made request.
5.	Holding time	It represents the average time duration taken for a call and it is denoted by 'H'.

<u>Various cellular techniques used in improving coverage capacity .</u>

To increase the coverage area in a cellular system it is very important to assign more number of radio channels to a cell so as to meet the mobile traffic.

More number of channels, higher will be the coverage range (distance) in the cell thus leading to a higher coverage capacity.

For enhancing the cellular coverage capacity there are many techniques available and some important cellular techniques are discussed below in detail.

i) Cell splitting
ii) Cell sectoring
iii) Micro zone method
iv) Repeaters for extending range.

Cell Splitting

Cell splitting is a technique of subdividing the congested (high traffic) cell into smaller sized cells. The parent cell which was originally congested is called as "macro cells" and the smaller cells are called as micro cells". The main objective of "cell splitting process" is to increases the cellular capacity of the system where frequency reuse technique can be efficiently implemented.

For example, a congested cell is subdivided into smaller cells shown in Fig. 1.31. Each smaller (micro) cell has a base station antenna exclusively and the micro cell radius will be half the radius value of the macro cell

The transmit power (P_t) will be less for the micro cells. Assuming P_{r_0} as the received power at old cell boundary and P_{r_N} as the received power at new cell boundary.

In cell splitting process generally the larger are dedicated to high speed traffic.

The reason for this is the number of 'hand offs' will be less in larger cell and call progress will be smoothly continued in larger cells.

Also the channels in old cells have to be broken into two groups due to following points. i)

If larger transmit power is used for all the available cells then some of the channels used by smaller cells may not be completely separated by co-channel cells. This may lead to interference.

ii) In case if smaller transmit power is used for the available cells then there is chance of 'unserved' problem. That is some parts of the larger cells would be left out as 'unserved'. This is also not acceptable.

Hence the channels of the macro/larger cell has to be divided into two groups. The larger cells for high speed traffic and micro/smaller cells for low speed traffic regions. Antenna down tilting :

The process of an antenna down tilting is done mainly to focus the energy radiated from the Base Station (BS) towards the ground and not towards the horizon so that radio coverage of new micro cells will be properly limited.

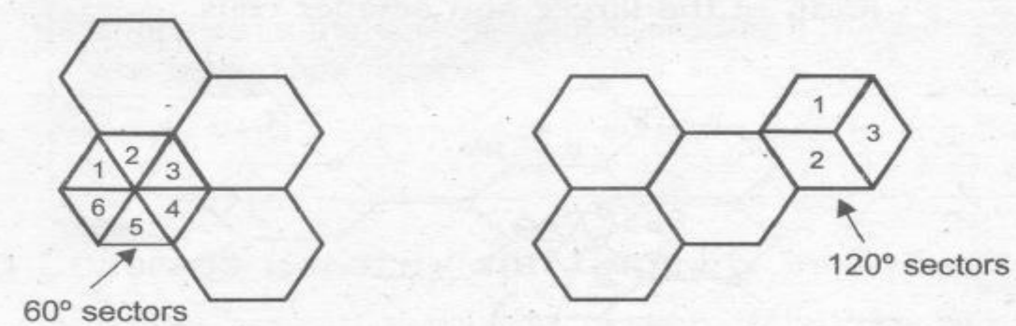
Cell Sectoring

The cell sectoring is again a technique to increase the capacity but it keeps the radius of cell as constant. The size of clusters in cellular region may be reduced because the cell sectoring increases the Signal to Interference Ratio (SIR) value.

The method of decreasing the co-channel interference value and enhancing thel system performance by using the directional antennas is known as "cell sectoring".

Sectors :

A cell in the cellular region is generally divided into 120° sectors or 60° sectors. If the sectoring is 120° a cell of hexagon type consists of three (3) sectors and if the sectoring is 60° sectoring the hexagonal type cell consists of six (6) sectors as shown in Fig. 1.32.



(a) 60° sectoring (b) 120° sectoring

Fig. 1.32 Cell sectoring

If cell sectoring is employed then the channel used in a cell will be divided into groups i.e. called as sectored groups and they are used only within a sector.

For example, in a seven cell reuse pattern with 120° sectoring, the possible number of interferers in the first tier will be only two. It means that only two cells of the six co-channel cells get interfered and it is better to apply cell sectoring in the design aspects of mobile communication to increase cellular capacity with less interference.

Micro Zone Method

Micro cell zone principle :

One of the problem associated with cell sectoring is that "requirement of more number of handoffs" thereby increased load status at the switching and control link components of the cellular mobile system.

There is another new concept known as 'micro zone concept' to minimize this problem.

In Fig. 1.33 there are three zones 1 to 3 shown with T/R set up. But they are connected to a single Base Station (BS) so that they are sharing the same equipment. For establishing connection between these three zones with the common base station microwave link, coaxial cable or fiber optic cable are used. Such an arrangement of several cellular zones with one base station constitutes a cell.

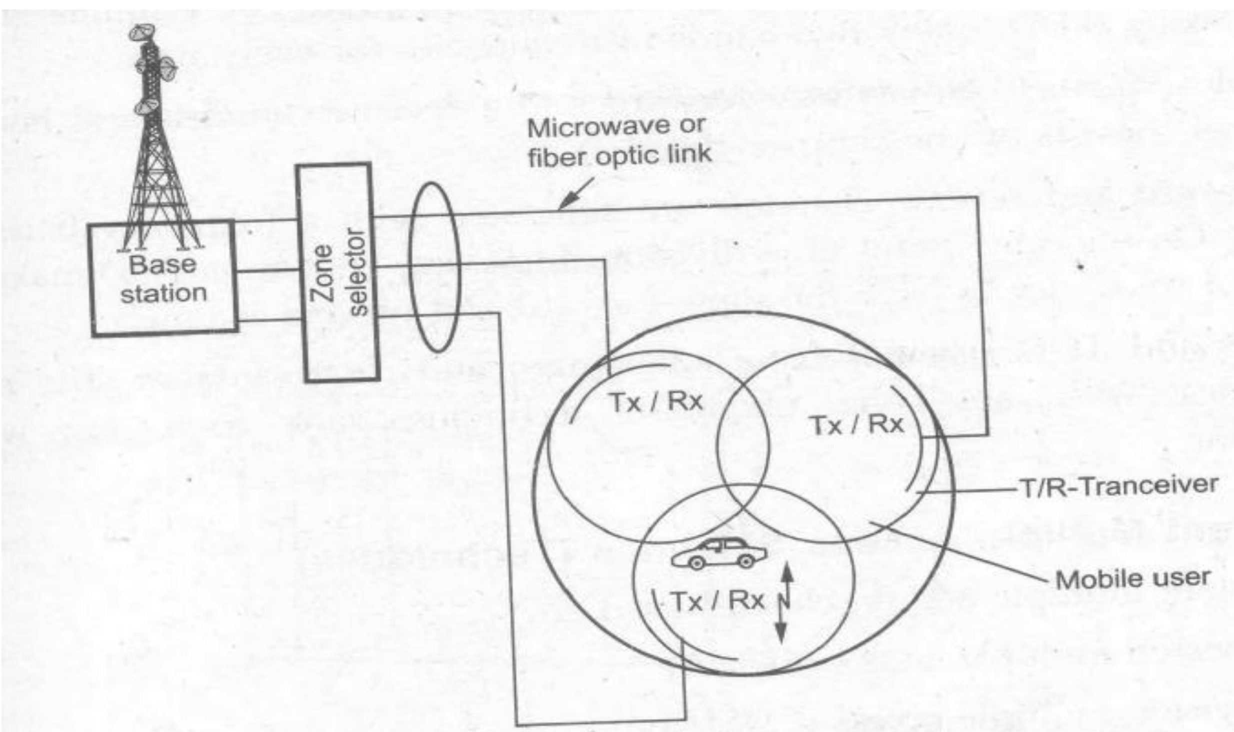


Fig. 1.33 Micro zone concept

Within the cell when a mobile user roams from one zone say 1 to another zone 2 then zone 1 will have strongest signal with respect to the base station.0 In this micro cell zone concept the antennas are placed at the edges of each zone such that when the user moves from one zone to another zone the signal strength does not reduce like other methods. The number of hand offs are less when compared to cell sectoring method when a cell is in progress.The merits of this technique are listed below. Advantages of micro zone cell concept :

i) When the mobile user travels from one zone to another zone within the same, cell the same channel is still maintained for the call progress.

ii) Since low power transmitters are used in each zone apart from control base station the effect of interference is highly reduced.

iii) Improved signal quality is possible.

iv) Reduced number of hand offs when a call is in progress

Unit-III
Digital Signaling For Fading Channels

Wireless communication link with the neat block diagram for transmitter and receiver

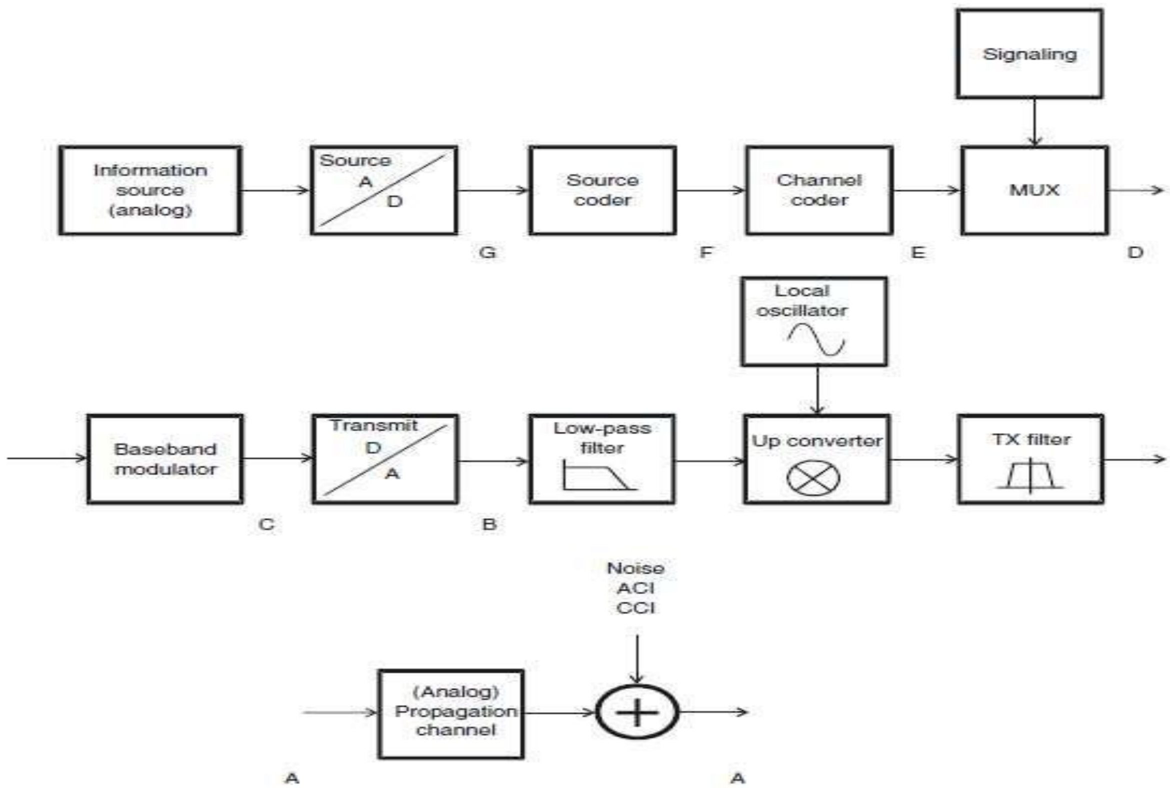


Fig:Block diagram of a digital transmitter chain for mobile communications

- The information source provides an analog source signal and feeds it into the source ADC and then converts the signal into a stream of digital data
- The source coder uses a priori information on the properties of the source data in order to reduce redundancy in the source signal.Eg:(GSM) speech coder reduces the source data rate from 64 kbit/s mentionedabove to 13 kbit/s.
- The channel coder adds redundancy in order to protect data against transmission errors. This increases the data rate that has to be transmitted at interface E – e.g., GSM channel coding increases the data rate from 13 to 22.8 kbit/s.
- Signalingadds 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, and combines the data from multiple users.
- In GSM, multiaccess multiplexing increases the data rate from 22.8 to 182.4 kbit/s
- The baseband modulator assigns the gross data bits Spectral properties, intersymbol interference, peak to-average ratio, and other properties of the transmit signal are determined by this step.
- 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.
- 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.
- The upconverterconverts 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.
- The (analog) propagation channel attenuates the signal, and leads to delay and frequency dispersion. The environment adds noise (Additive White Gaussian Noise – AWGN) and co-channel interference

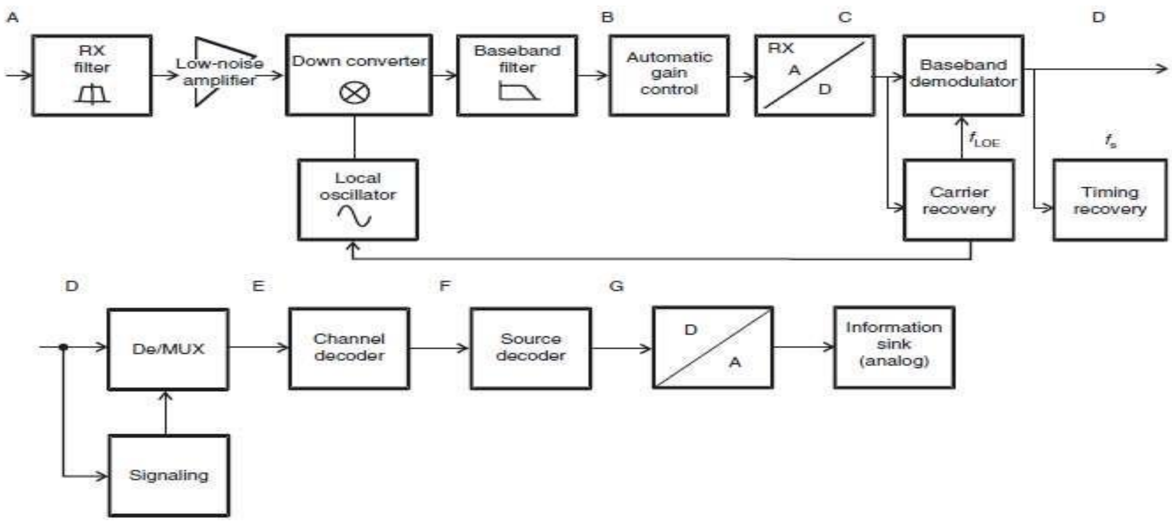


Fig:Block diagram of a digital receiver chain for mobile communications

- 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. The frequency of the LO can be fine-tuned by a carrier recovery algorithm.
- The RX down converter 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.
- 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.
- multiple antennas, then the RX either selects the signal from one of them for further processing or the signals from all of the antennas have to be processed
- 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.

Offset Quadrature Phase Shift Keying signal and its advantages

The amplitude of QPSK signal is ideally constant. When QPSK signals are pulse shaped, they lose constant envelope property. QPSK signals the phase shift of $\prod$ radians can cause the signal envelope to pass through zero.

To hardlimiting or non linear amplification of the Zero-crossings brings back to the filtered sidelobes since the fidelity of the signal at small voltage levels is lost in transmission.To prevent the regeneration sidelobes and spectral widening, QPSK signals be amplified only using linear amplifiers which is less efficient.

A modified form of QPSK called offset QPSK (OQPSK) or staggered QPSK is less susceptible damaging effects and supports more efficient amplification.

OQPSK signaling is similar to QPSK signaling is except for the time alignment of the even and odd bit streams. In QPSK signaling, the bit transitions of 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.

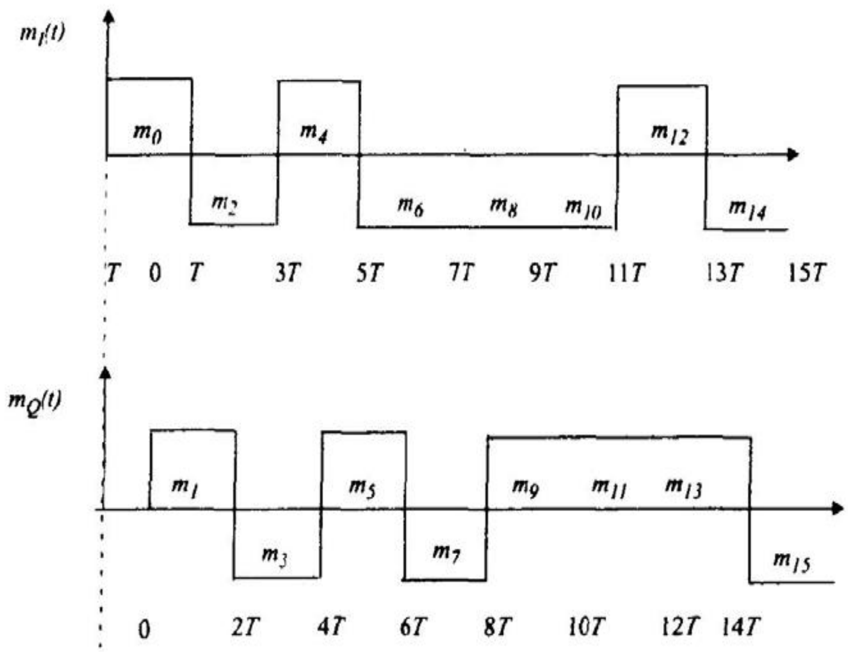


Fig.3.1 Offset waveforms

Due to the time alignment of $m_I(t)$ and $m_q(t)$ in standard QPSK, phase transitions occur only once every $T_s = 2T_b$ s, and will be a maximum of 180° if there is a change in the value of both $m_I(t)$ and $m_q(t)$. However, in OQPSK signaling, bit transitions phase transitions) occur every T_b s. Since the transitions instants are offset $m_I(t)$ and $m_q(t)$, at any given time only one of the two bit streams can change values. This implies that the maximum phase shift of the transmitted signal at any given time is limited to ±90°. Hence, by switching phases more frequently (i.e., every T_b s instead of $2T_b$ s) OQPSK signaling eliminates 180° phase transitions.Since 180° Phase transitions have been eliminated, bandlimiting of QPSK signal does not cause the signal envelope to go to zero. Ultimately some amount of ISI caused by the bandlimitng process expecially at the 90° phase transition points.

But the envelope variations are considerably less, and hence hardlimiting or non linear amplification of OQPSK signals does not generate the high frequency sidelobes as much as in QPSK. The spectral efficiency is significantly reduced, while permitting over efficient RF bandwidth.

The spectrum of an OQPSK signal is identical to that of a QPSK signal, hence both signals occupy the same bandwidth. The staggered alignment of the even and odd bit streams does not change the nature of the spectrum. OQPSK retains its bandlimited nature even after nonlinear amplification, and therefore is very attractive for mobile communication systems where bandwidth efficiency and efficient nonlinear amplifiers are critical for low power drain. Further, OQPSK signals is to perform better than QPSK in the presence of phase jitter due to noisy reference signals at the receiver.

Minimum shift keying (MSK) Derive an expression for MSK and its power spectrum.

Minimum shift keying:

MSK is a special type of continuous phase frequency shift keying (CPFSK). Where the peak frequency deviation is equal to ¼ the bit rate. In other words, MSK is a continuous phase FSK with a modulation index of 0.5 reference signal.

The modulation index of FSK signal is similar to FM modulation index, and is defined as $K_{FSK}=(2\Delta F)/R_b$..

Where ΔF-> Peak RF frequency deviation R_b-> Bit rate

The modulation index of 0.5 corresponds to minimum frequency spacing, that allows two FSK signals to be coherently orthogonal and mainly MSK is minimum frequency separation (B/W) allows orthogonal detection.

Two signals $v_H(t)$ and $v_L(t)$ are said to be orthogonal if,

$$\int_0^T v_H(t) v_L(t) dt = 0 \quad \text{------>1}$$

MSK is sometimes referred as fast FSK, as the frequency spacing used is only half as much as that used in conventional non-coherent FSK.

MSK is spectrally efficient modulation and its particularly attractive for using mobile radio communication. It possesses properties such as constant envelope, spectral efficiency, good BER performance, and self-synchronizing capability.

MSK signal is special form of OQPSK where the baseband rectangular pulses are replaced with half-sinusoidal pulses. If half-sinusoidal pulses are used instead of rectangular pulses, the modified signal can be defined as MSK and for an N-bit stream is given by

$$S_{MSK}(t) = \sum_{i=0}^{N-1} m_I(t)p(t-2iT_b)\cos 2\pi f_c t + \sum_{i=0}^{N-1} m_Q(t)p(t-2iT_b-T_b)\sin 2\pi f_c t$$

$$\text{where } p(t) = \begin{cases} \sin\left(\frac{\pi t}{2T_b}\right) & 0 \le t \le 2T_b \\ 0 & elsewhere \end{cases} \quad \text{----->2}$$

where $m_I(t)$ and $m_Q(t)$ are the "odd" and "even" bits of the bipolar data stream are values of ±1 and which feed the in-phase and quadrature arms of the modulator at a rate of $R_b/2$. There are a number of variations of MSK. For example, while one version of MSK uses only positive half-sinusoids as the basic pulse shape, another version uses alternating positive and negative half-sinusoids as the basic pulse shape.

However, all variations of MSK are continuous phase FSK employing different techniques to achieve spectral efficiency.

The MSK waveform is a special type of a continuous phase FSK if equation 2 is rewritten by using trigonometric identities as

$$S_{MSK}(t) = \sqrt{\frac{2E_b}{T_b}}\cos\left[2\pi f_c t - m_I(t)m_Q(t)\frac{\pi t}{2T_b} + \phi_k\right] \quad \text{----->3}$$

Where Φ_k is 0 or Π depending on whether $M_I(t)$ is 1 or -1.

The Phase continuity at the bit transition periods is ensured by choosing the carrier frequency to be an integral multiple of one fourth the bit rate, 1/4T

MSK Power Spectrum

For MSK the baseband pulse shaping function is given by,

$$p(t) = \begin{cases} \cos\left(\dfrac{\pi t}{2T}\right) & |t| < T \\ 0 & \text{elsewhere} \end{cases} \quad \text{----->4}$$

The normalized power spectral density for MSK is given by

$$P_{\text{MSK}} = \frac{16}{\pi^2}\left(\frac{\cos 2\pi(f+f_c)T}{1.16f^2T^2}\right)^2 + \frac{16}{\pi^2}\left(\frac{\cos 2\pi(f-f_c)T}{1.16f^2T^2}\right)^2 \quad \text{----->5}$$

Figure 3.2 shows the power spectral density of an MSK signal. The spectral density of both QPSK and OQPSK are also drawn for comparison in Figure 3.2, the MSK spectrum has lower side lobes than QPSK and OQPSK. Ninety-nine percent of the MSK power is contained within a bandwidth B = 1.2 / T , while for QPSK and OQPSK, the 99 percent bandwidth B is equal to 8/T. Figure 3.2 also shows that the main lobe of MSK is wider than QPSK and OQPSK, and hence when compared in terms of first null bandwidth, MSK is less spectrally efficient than the phase-shift keying techniques.

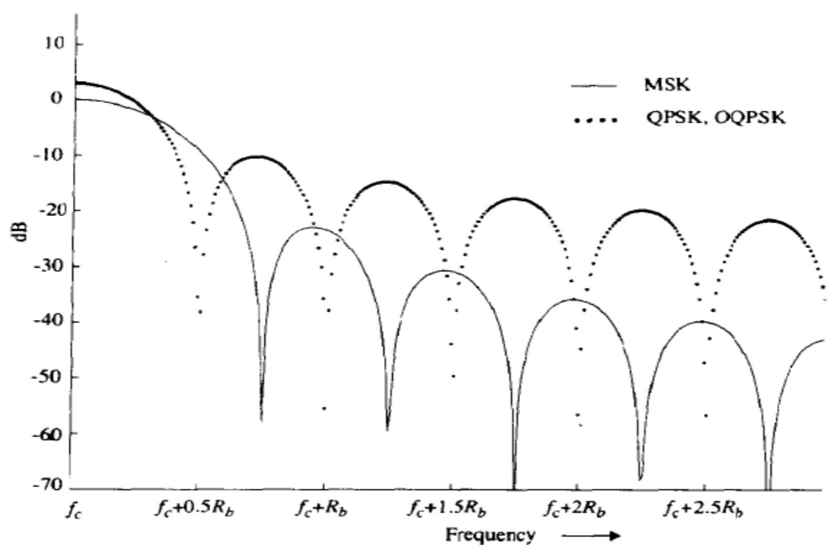


Fig.3.2 power spectrum of MSK

MSK (Minimum shift keying) modulation and demodulation technique.

MSK Transmitter:

Figure 3.3 shows a typical MSK modulator. To Multiplying a

carrier signal with

- 1/ 4T . These two FSK

signals are separated using two narrow band pass 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).

MSK receiver:

The block diagram of an MSK receiver is shown in Figure 3.4. 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 discarded 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 a 0 or a 1. The output data streams correspond to m $_I$ (t) and m $_Q$ (t) .

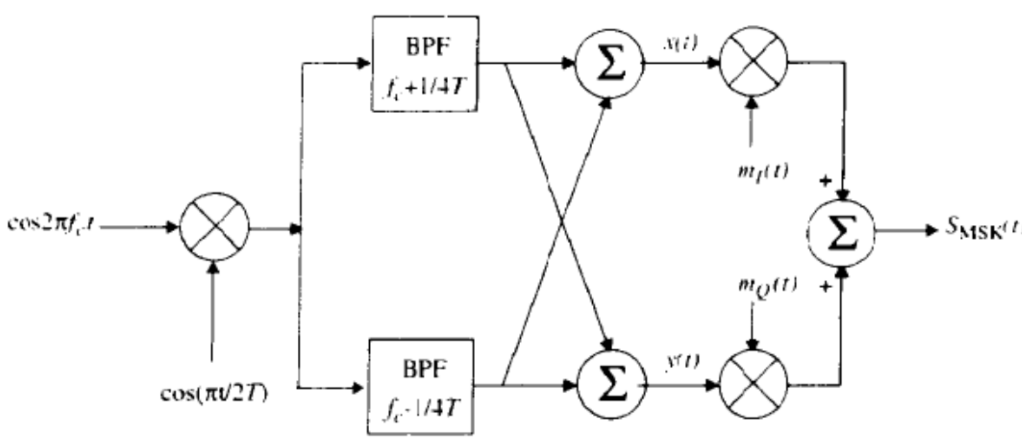


Fig.3.3 Block diagram of MSK Transmitter

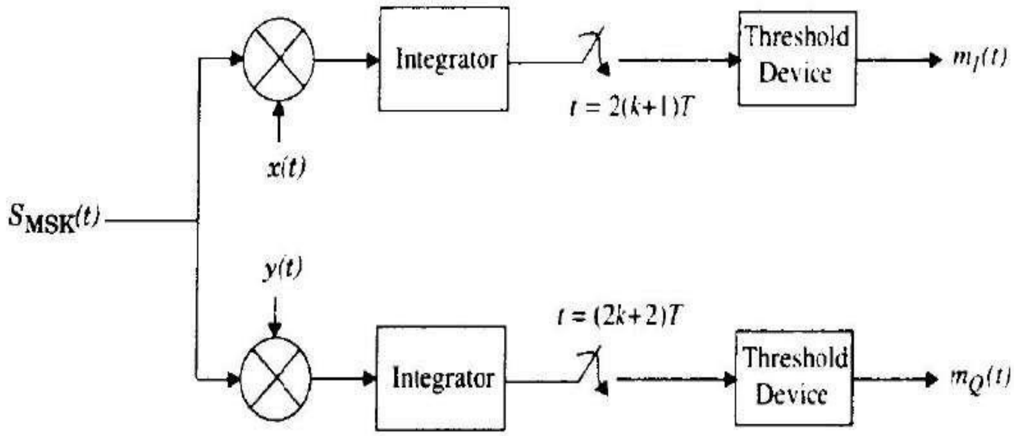


Fig.3.4 Block diagram of MSK Receiver.

<u>GMSK (Gaussian Minimum Shift Keying) and mention its advantages.</u>

GMSK is a simple binary modulation scheme may be derivative of MSK. In GMSK the side lobe levels of the spectrum are further reduced by passing the modulating NRZ data waveform through a pre-modulation Gaussian pulse shaping filter. Baseband Gaussian pulse shape smooth the phase route of MSK signal and tends stabilizes the instantaneous frequency variations over time.

Premodulation Gaussian filtering converts the full response message signal into a partial response scheme where each transmitted symbol spans several bit periods. GMSK can be coherently detected just as an MSK signal, or non-coherently detected as simple FSK. GMSK is most attractive for its excellent power efficiency (due to the constant envelope) and its excellent spectral efficiency.

The pre modulation Gaussian filtering introduces ISI (Inter Symbol Interference) in the transmitted signal, but it can be shown that the degradation is not severe if the 3dB bandwidth bit duration (BT) product of the filter is greater than 0.5.

In GMSK pre modulation filter bas an impulse response given by

$$h_G(t) = \frac{\sqrt{\pi}}{\alpha}\exp\left(-\frac{\pi^2}{\alpha^2}t^2\right) \text{ -----}>1$$

Transfer function is given by,

$$H_G(f) = \exp(-\alpha^2 f^2) \text{ ------}>2$$

The parameter a is related to B, the 3 dB baseband bandwidth of H $_G$ { f) , by

$$\alpha = \frac{\sqrt{ln 2}}{\sqrt{2}B} = \frac{0.5887}{B} \text{ ------}>3$$

GMSK filter may be completely defined from B and the baseband symbol duration T . Therefore GMSK its B T product

Figure 3.5 shows the simulated RF power spectrum of the GMSK signal for various values of B T . The power spectrum of MSK, which is equivalent to GMSK with a B T product of infinity, is also shown for comparison purposes. In this graph that as the B T product decreases, the side lobe levels fall off very rapidly.

For example, for a B T =0.5, the peak of the second lobe is more than 30dB below the main lobe, whereas for simple MSK, the second lobe is only 20 dB below main lobe. However, reducing B T increases the irreducible error rate produced by the low pass filter due to ISI. GMSK Bit Error Rate

The bit error rate for GMSK was determined for AWGN channels, and performance within 1 dB of optimum MSK when BT=0.25. The bit error probability is a function of BT, and the probability for GMSK is given by

$$P_e = Q\left\{\sqrt{\frac{2\gamma E_b}{N_0}}\right\} \quad ---->4$$

where γ is a constant related to *BT* by

$$\gamma \cong \begin{cases} 0.68 & \text{for GMSK with } BT = 0.25 \\ 0.85 & \text{for simple MSK } (BT = \infty) \end{cases} \quad ------>5$$

GMSK TRANSMITTER:

The simplest way to generate a GMSK signal is to pass a NRZ message bit stream through a Gaussian baseband filter having an impulse response given by equ 1 followed by an FM modulator. This modulation technique is shown in Figure 3.6 and is currently used in a variety of analog and digital implementations for the U.S. Cellular Digital Packet Data (CDPD) system as well as for the Global System for Mobile (GSM) system.

Figure 3.6 may also be implemented digitally using a standard I/Q modulator.

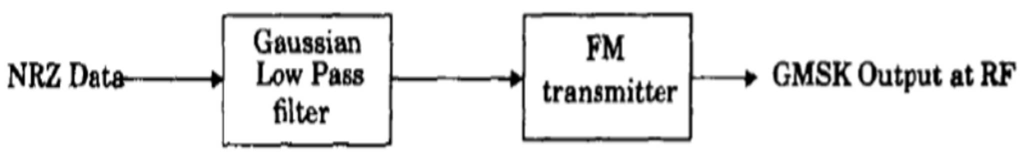


Fig.3.6 Block diagram of GMSK transmitter using direct FM generation

GMSK RECEIVER

GMSK signals can be detected using orthogonal coherent detectors as shown in Figure 3.7, or with simple non coherent detectors such as standard FM discriminators. The Carrier recovery is performed by using the sum of two discrete frequency components contained at the output of a frequency doubler is divided by four.

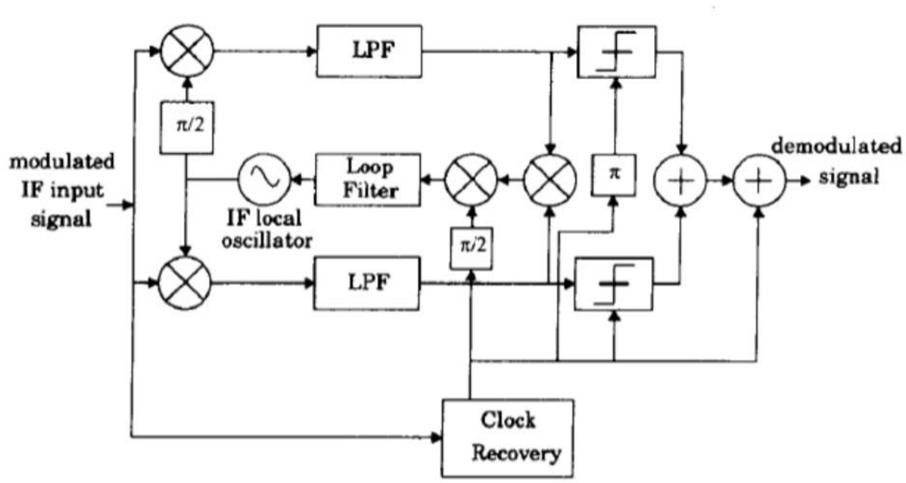


Fig.3.7 Block diagram of GMSK receiver

This type of receiver can be easily implemented using digital logic as shown in Figure 3.8. The two D flip-flops act as a quadrature product demodulator and the XOR gates act as baseband multipliers. The mutually orthogonal reference carriers are generated using two D flip-flops, and the VCO center frequency is set equal to four times the carrier center frequency.

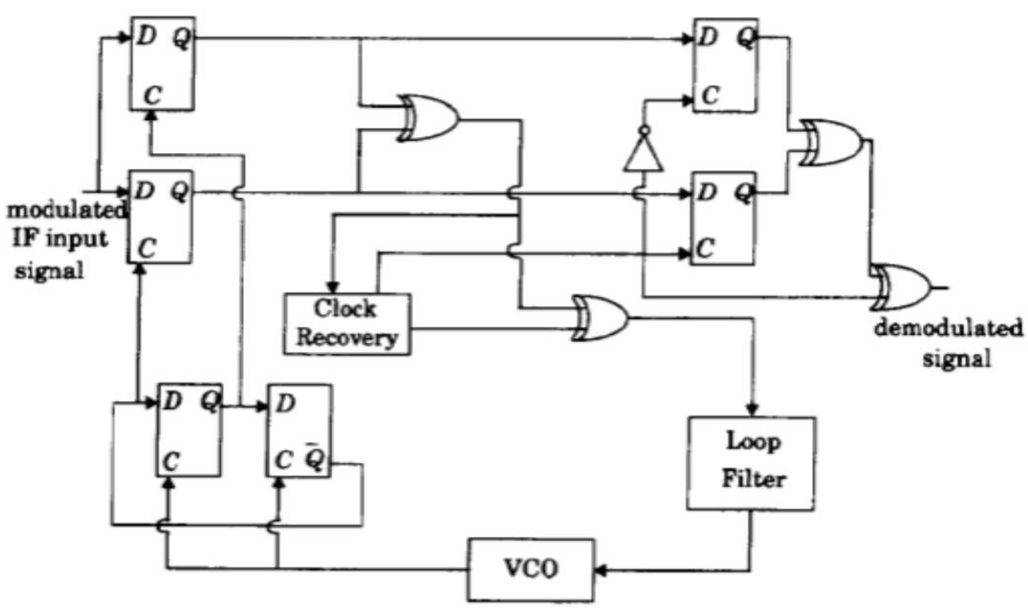
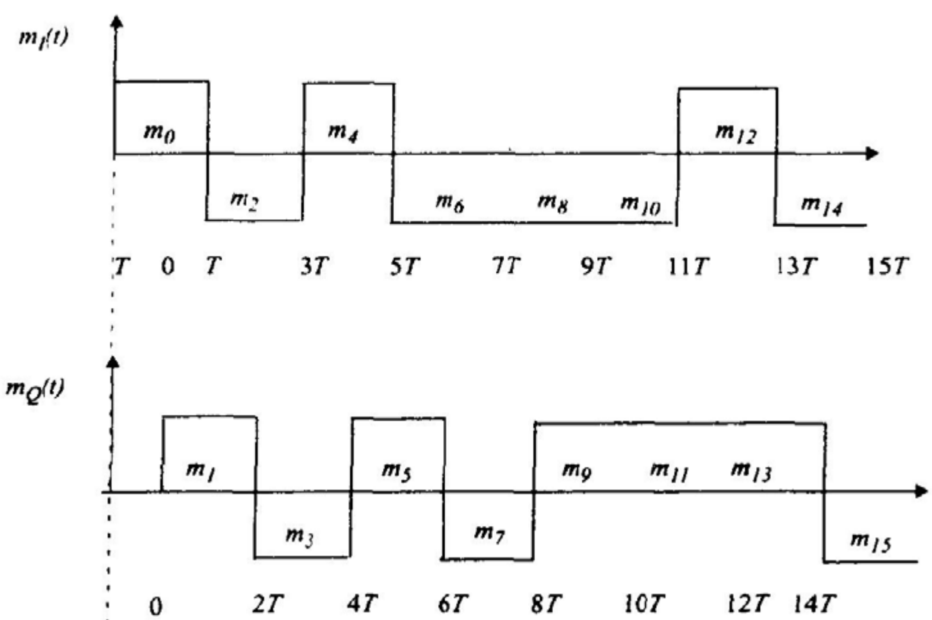


Fig.3.8 Digital logic circuit for GMSK demodulation

Offset Quadrature Phase Shift Keying signal and its advantages.

- Offset QPSK is a modified form of QPSK where the bit waveforms on the I and Q channels are offset or shifted in phase from each other by one-half of a bit time
- The advantage of OQPSK is the limited phase shift that must be imparted during modulation and it supports more efficient amplification.
- OQPSK signaling eliminates the 180 degree phase transitions. With 90 degree phase transitions it has some ISI.
- The spectrum of the OQPSK is identical to QPSK and both signals occupies same bandwidth.
- The disadvantages is that changes in the output phase occur at twice the data rate in either the I and Q channels.
- OQPSK signaling is similar to QPSK signaling is except for the time alignment of the even and odd bit streams.
- To improving the peak-to-average ratio in QPSK.



- Thus the transmitted pulse are

$$\left.\begin{aligned} p1_D(t) &= \sum_{i=-\infty}^{\infty} b1_i g(t - iT_S) = b1_i * g(t) \\ p2_D(t) &= \sum_{i=-\infty}^{\infty} b2_i g\left(t - \left(i+\tfrac{1}{2}\right)T_S\right) = b2_i * g\left(t - \frac{T_S}{2}\right) \end{aligned}\right\}$$

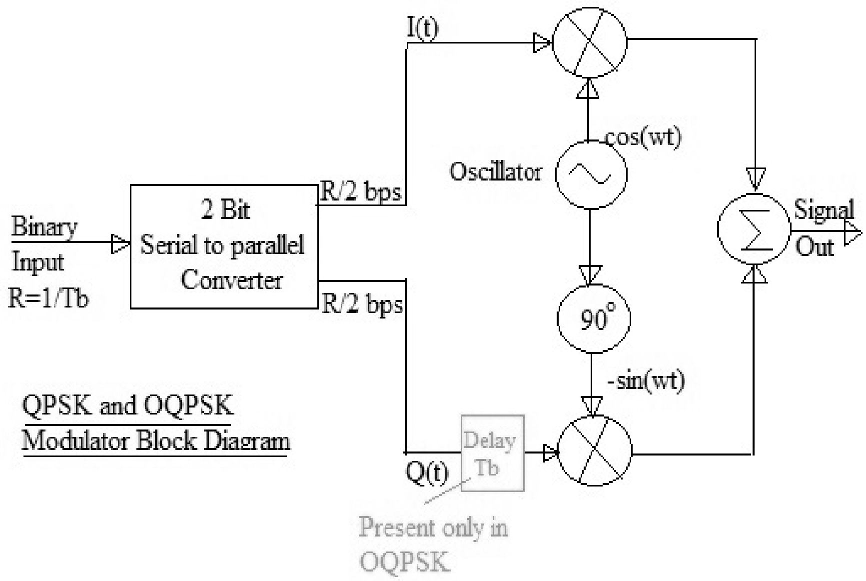


Fig: Difference b/w QPSK and OQPSK

pi/4 Differential Quadrature Phase Shift Keying.

- The principle of π/4-DQPSK can be understood from the signal space diagram of DQPSK .
- There exist two sets of signal constellations: (0, 90, 180, 270◦) and (45, 135, 225, 315◦). All symbols with an even temporal index i are chosen from the first set, while all symbols with odd index are chosen from the second set.
- The duration of the dips is longer when non-rectangular basis pulses are used. Such variations of the signal envelope are undesirable. One possibility for reducing these problems lies in the use of π/4-DQPSK

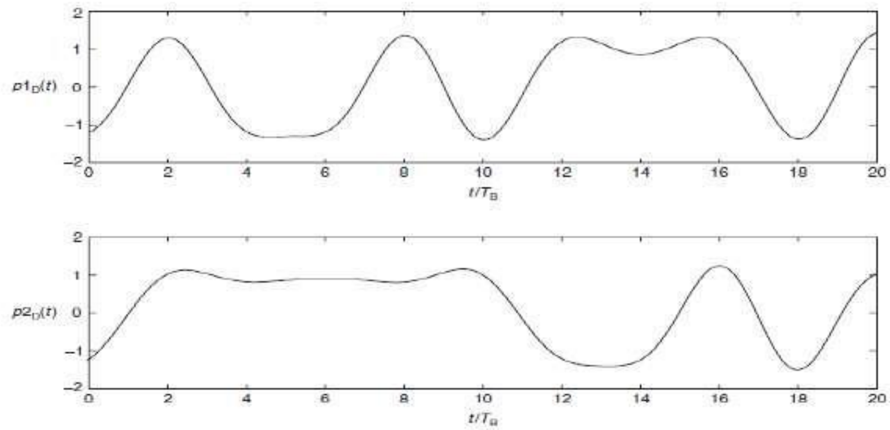
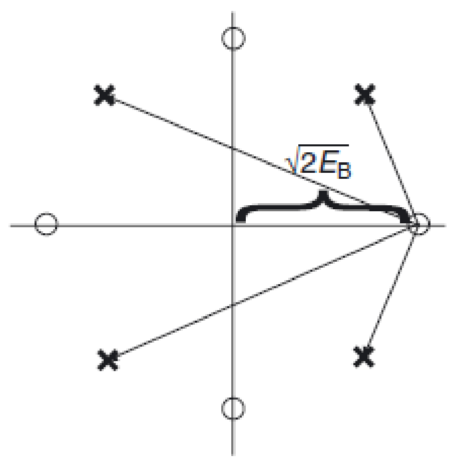


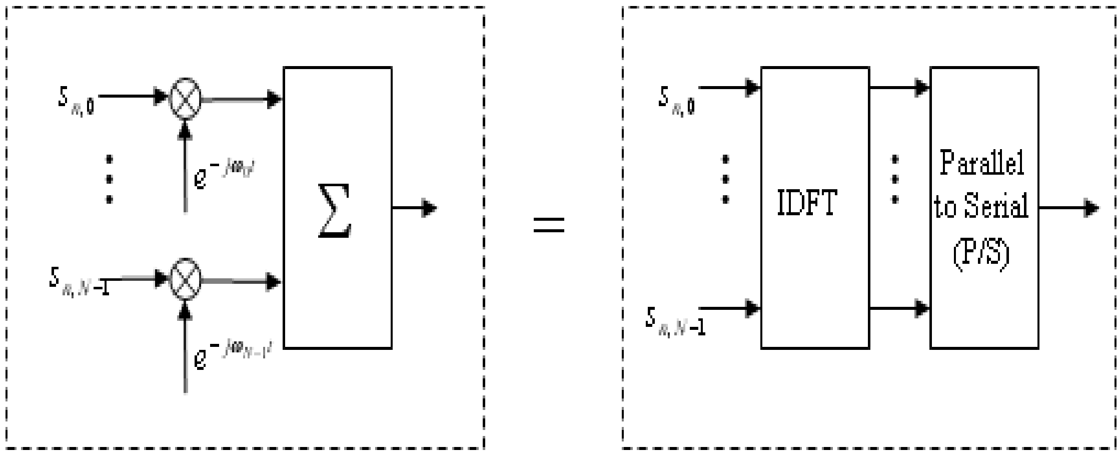
Fig: QAM Pulse Sequence



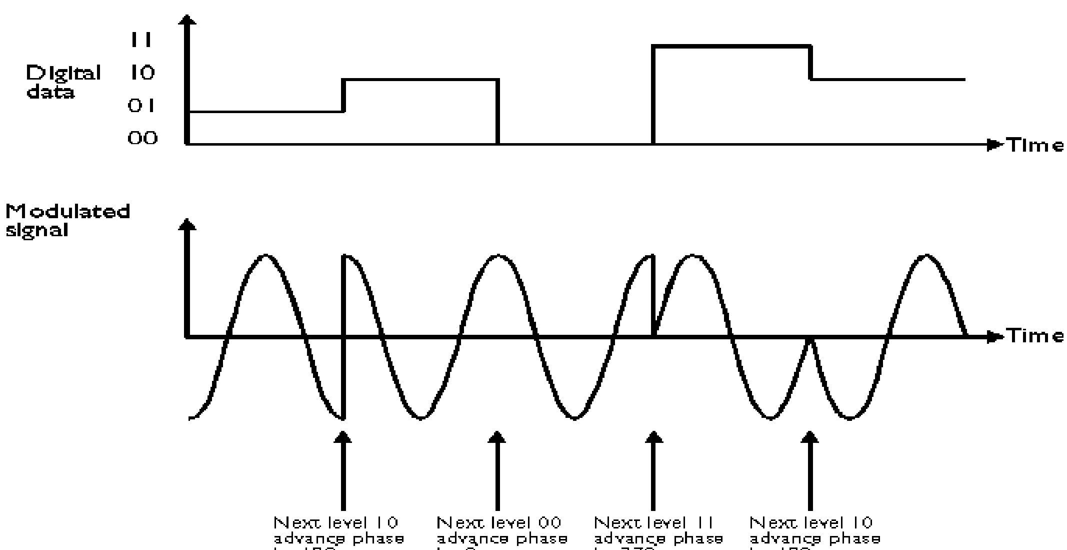
Allowed transitions in the signal space diagram of $\pi/4$ differential quadrature-phase shift keying.

Principles of OFDM and note on cyclic prefix.

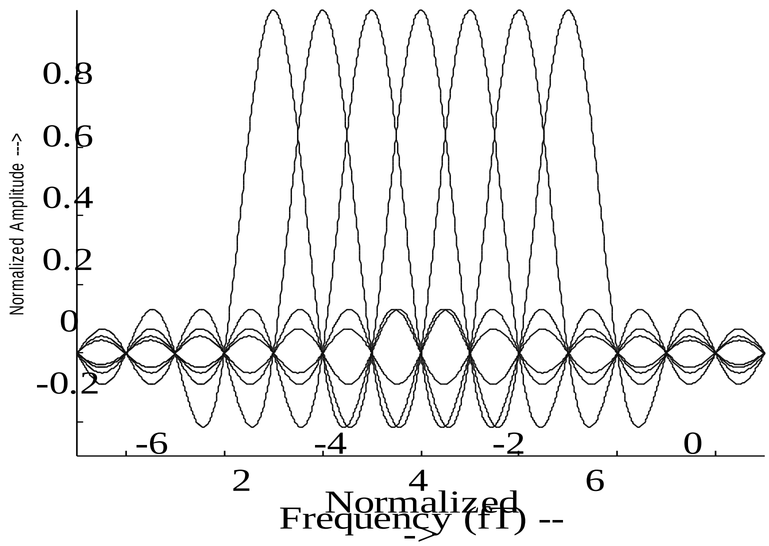
Main idea: split data stream into N parallel streams of reduced data rate and transmit each on a separate subcarrier.OFDM modulation is equivalent to the IDFT



- OFDM Mechanism:
- Parallel Data Streams
- The available frequency spectrum is divided into several sub-channels
- low-rate bit stream is transmitted over one sub-channel by modulating a sub- carrier using a standard modulation scheme, for example: 4-QAM



OFDM Power Spectrum:



OFDM Transmitter :

An OFDM carrier signal is the sum of a number of orthogonal sub-carriers, with base band data on each sub-carrier being independently modulated commonly using some type of quadrature amplitude modulation (QAM) or phase-shift keying (PSK).

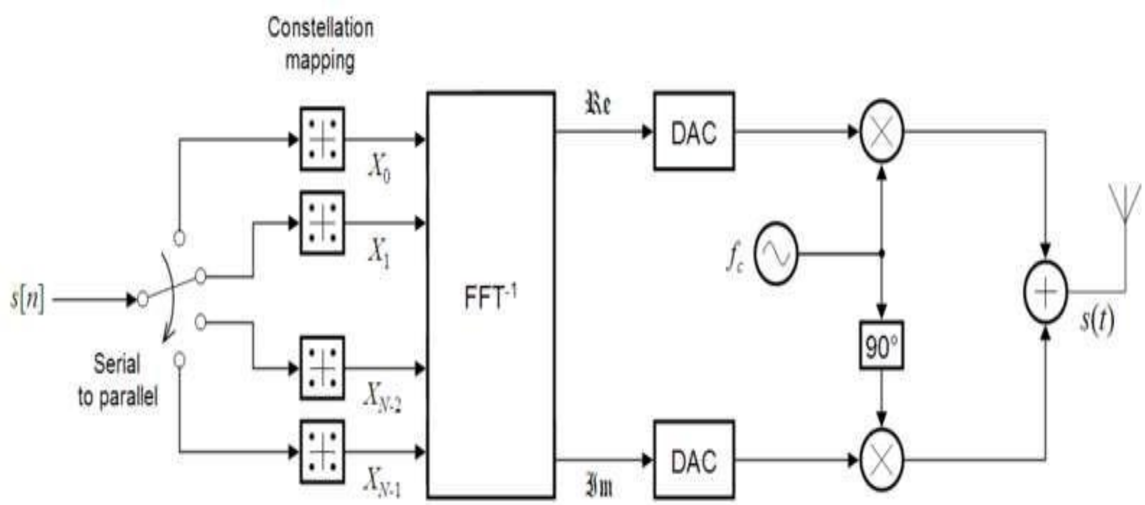


Fig: OFDM Transmitter block diagram

OFDM Receiver:

The receiver picks up the signal r(t), which is then quadrature-mixed down to baseband using cosine and sine waves at the carrier frequency. This returns N parallel streams, each of which is converted to a binary stream using an appropriate symbol detector. These streams are then re-combined into a serial stream, which is an estimate of the original binary stream at the transmitter.

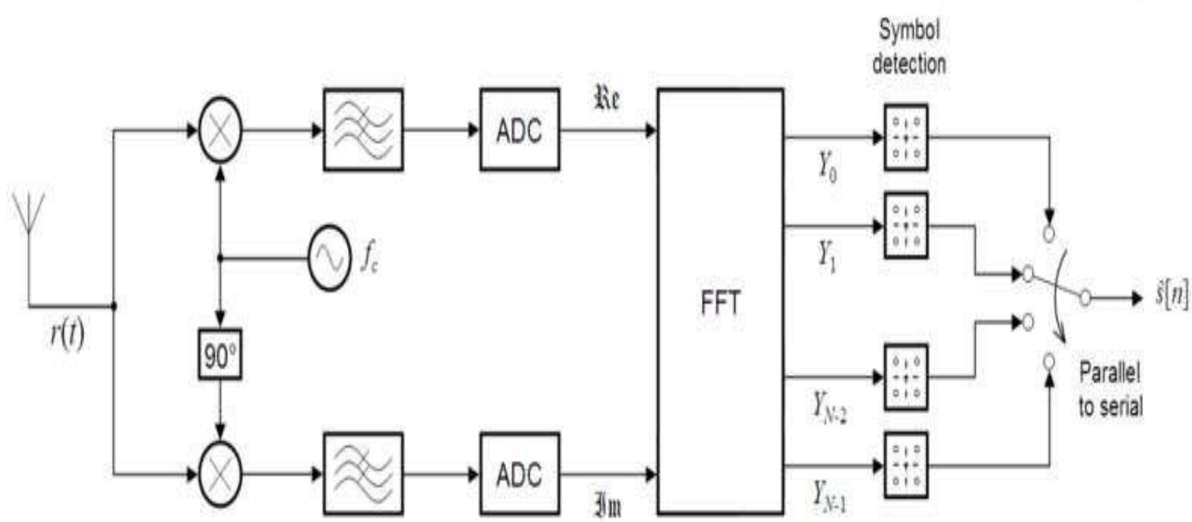


Fig: OFDM Receiver block diagram

CYCLIC PREFIX:

- Convert a linear convolution channel into a circular convolution channel.
- Energy is wasted in the cyclic prefix samples.
- Zeros used in the guard time can alleviate interference (ISI) between OFDM symbols
- Orthogonality of carriers is lost when multipath channels are involved.

OFDM APPLICATIONS:

- Digital Video Broadcasting
- Digital Audio Broadcasting

- ADSL
- Wireless LANs
- OFDMA -Multiple Access.

OFDM ADVANTAGES

- IFFT/FFT operation ensures that sub-carriers do not interfere with each other.
- Information from the affected subchannels can be erased and recovered by the forward error correction (FEC) codes.
- Equalization is very simple compared to Single-Carrier systems
- Cyclic prefix allows the receiver to capture multi- path energy more efficiently.

OFDM DRAWBACKS:

- High sensitivity inter-channel interference, ICI
- OFDM is sensitive to frequency, clock and phase offset
- The OFDM time-domain signal has a relatively large peak-to-average ratio
- Reduce the power efficiency of the RF amplifier

Peak to Average Power Ratio (PAPR) and the methods used in PAPR reduction.

- OFDM has been chosen for high data rate communications and has been widely deployed in many wireless communication standards such as Digital Video Broadcasting (DVB) and based mobile worldwide.
- One of the major problems of OFDM is that the peak amplitude of the emitted signal can be considerably higher than the average amplitude.
- 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, and the power thus goes with N^2

- Due to the large number of sub carriers, OFDM systems have a large dynamic signal range with a very high PAPR.
- As a result, the OFDM signal will be clipped when passed through a non linear power amplifier at the transmitter end.
- Clipping degrades the bit-error-rate (BER) performance and causes spectral spreading. One way to solve this problem is to force the amplifier to work in its linear region.

There are 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.
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.
- The first effect increases the BER of the desired signal

3. Use PAR reduction techniques.

Peak-to-Average Ratio Reduction Techniques:

1. Coding for PAR reduction:

- 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
- The transmission thus follows

  (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.

2.Phase adjustments:

This scheme first defines an ensemble of phase adjustment vectors φl, l =1, . . . , L, that are known to both the transmitter and receiver.

$$\{\hat{c}_n\}_l = c_n \exp[j(\phi_n)_l]$$

$$\hat{l} = \arg\min_l (PAR(\{\hat{c}_n\}_l))$$

3. Correction by multiplicative function:

Multiply the OFDM signal by a time-dependent function whenever the peak value is very high.

$$\hat{s}(t) = s(t)\left[1 - \sum_k \max\left(0, \frac{|s_k| - A_0}{|s_k|}\right)\right]$$

$$\hat{s}(t) = s(t)\left[1 - \sum_n \max\left(0, \frac{|s_k| - A_0}{|s_k|}\right) \exp\left(-\frac{t^2}{2\sigma_t^2}\right)\right]$$

4.Correction by additive function:

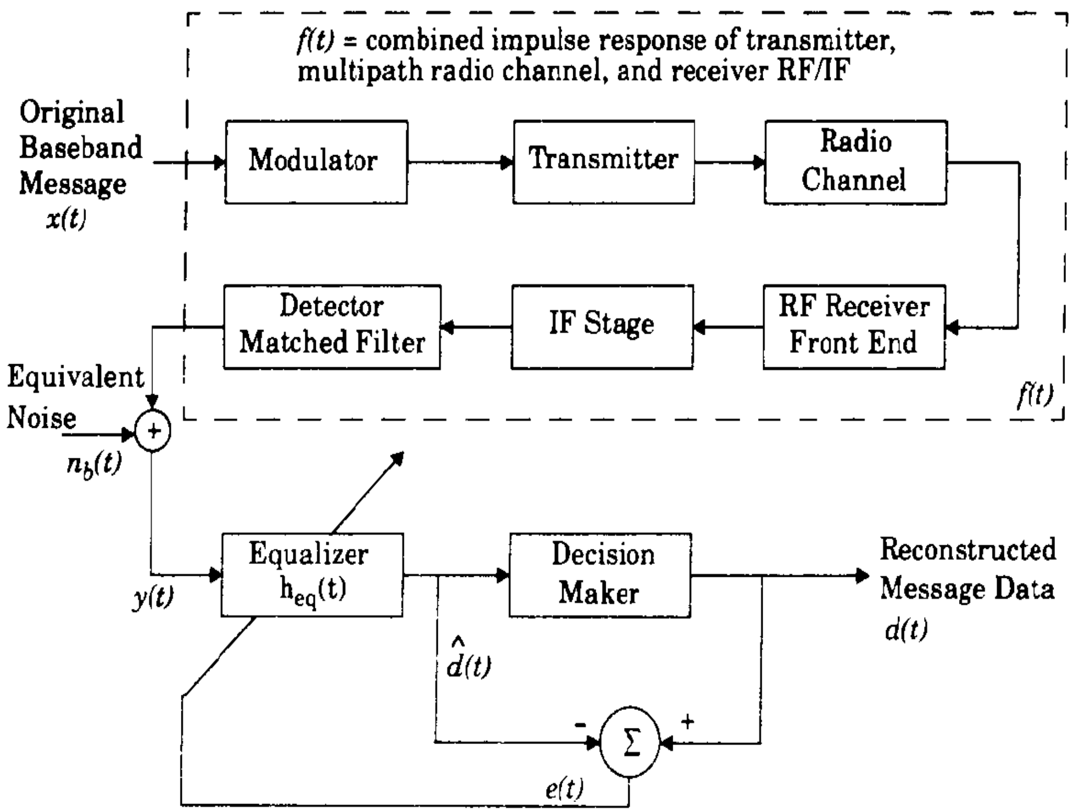
- We can choose an additive, instead of a multiplicative, correction function.
- The correction function should be smooth enough not to introduce significant out-of-band interference.
- Furthermore, the correction function acts as additional pseudo noise, and thus increases the BER of the syste

Unit-IV

Multipath Mitigation Techniques

Equalization:
Fundamentals of Equalisation

Since the mobile fading channels are random and time varying the equalizer must be adaptive in nature.General operating modes of adaptive equalizer is training and tracking.A known fixed length training sequence is sent by the transmitter so that the receiver equalizer adapt to proper BER detection.Immediately following training sequence the message data is sent.At the receiver tracking(recursive algorithm) is used to evaluate the channel and filter coefficients



Equalizer Prediction Error

Reconstructed message data dt that you saw last time is an estimate and should match x(t)f we use dt and use it in a feedback loop so adjust the weights of the equalizer we get into the domain of non-linear equalizer.Feedback is not there we just have a feed forward loop then it is a linear equalizer.So we realize that equalizer is in fact an inverse filter of the channel but in the absence of noise.

Types of equalization

Linear equalizers
Non-linear equalizers

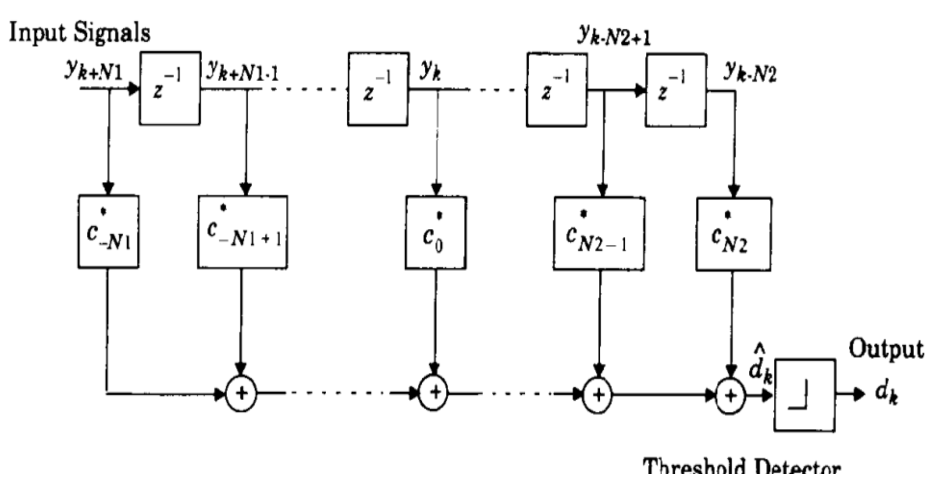
Equalization compensates for Intersymbol interference (ISI) created by multipath within time dispersive channels. If the modulation bandwidth exceeds the coherence bandwidth of the radio channel, ISI occurs and modulation pulses are spread into adjacent symbols. equalizers must be adaptive since the channel is generally unknown and time varying.

A linear equalizers can be implemented as FIR filters, otherwise known as the transversal filter.

This type of equalizer is the simplest type available. In such an equalizer, the current and past values of the received signal are linearly weighted by the filter co-efficient and summed to produce the output as shown in fig1.

If the delays and the tap gains are analog the continuous output of the equalizer is sampled at the symbol rate and the samples are applied to the decision device.

This implementation is usually carried out in the digital domain where the samples of the received signals are stored in a shift register.The output of this transversal filter before decision is made



Linear equalizers:

A linear equalizers can be implemented as FIR filters, otherwise known as the transversal filter.

This type of equalizer is the simplest type available. In such an equalizer, the current and past values of the received signal are linearly weighted by the filter co-efficient and summed to produce the output as shown in fig1.

If the delays and the tap gains are analog the continuous output of the equalizer is sampled at the symbol rate and the samples are applied to the decision device.

This implementation is usually carried out in the digital domain where the samples of the received signals are stored in a shift register

.The output of this transversal filter before decision is made is

Where ,

Cn* - represents the complex filter co-efficient or tap weights d^k

– output at time index k

yi – input received signal at time to+it

to – equalizer starting time

$N = N_1 + N_2 + 1$ is the number of taps .

The values N1 and N2 denote the number of taps are used in the forward and reverse portions of the equalizer.

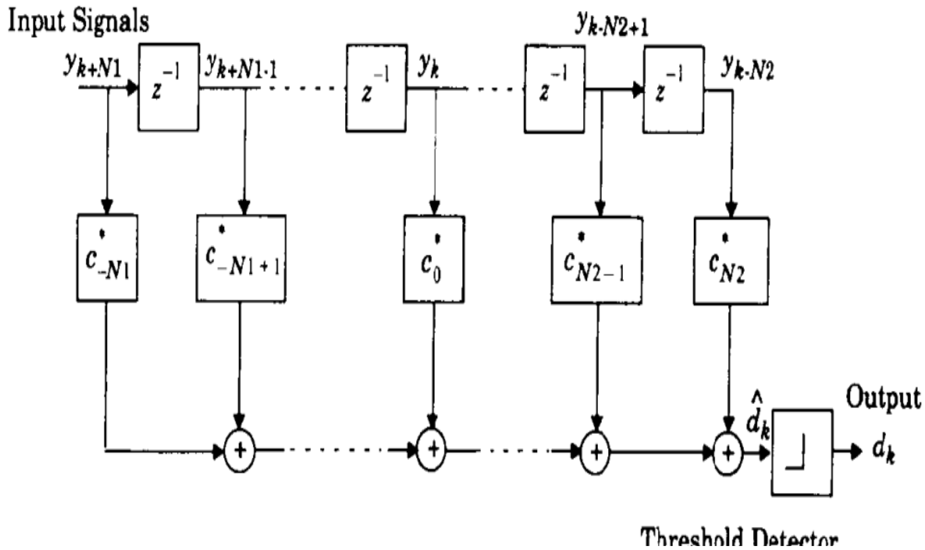


Fig.1 Structure of linear transversal equalizer

The minimum mean squared error $E[|e(n)|^2]$ that a linear transversal equalizer can achieve

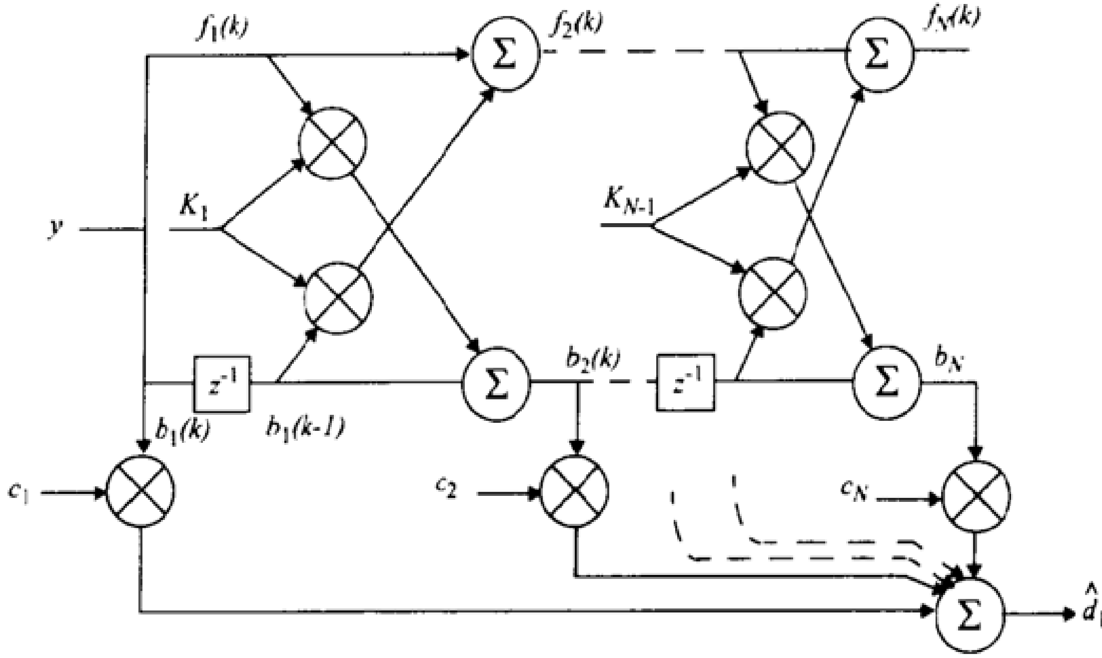
$$E[|e(n)|^2] = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \frac{N_0}{|F(e^{jwT})|^2 + N_0} d\omega \qquad \text{------>2}$$

Where, $F(e^{jwt})$ is frequency response of channel

N_0 - noise power spectral density.

The linear equalizer can also be implemented as a lattice filter, whose structure is shown in fig 2. In lattice filter, the input signal y_k is transformed into a set of N intermediate forward and backward error signals, $f_n(k)$ and $b_n(k)$ respectively, which are used as inputs to the tap multipliers and are used to calculate the updated coefficients. Each stage of the lattice is determined by following recursive equation.

The backward error signals b_n are used as inputs to the tap weights and the output of the equalizer is given by



The structure of a lattice equalizer however is more complicated than a linear transversal equalizer

Advantages of the lattice equalizer its

Numerical stability

faster convergence.

The unique structure of the lattice filter allows the dynamic assignment of the most effective length of the lattice equalizer. Hence, if the channel is not very time dispersive, only a fraction of the stages are used. When the channel becomes more time dispersive, the length of the equalizer can be increased by the algorithm without stopping the operation of the equalizer.

(ii) Non-linear equalizers:

Nonlinear equalizers are used in applications where the channel distortion is too severe for a linear equalizer to handle. So, Linear equalizers do not perform well on channels which have deep spectral nulls in the passband. In an attempt to compensate for the distortion, the linear equalizer places too much gain in the vicinity of the spectral null, thereby enhancing the noise present in those frequencies.Three very effective nonlinear methods have been developed which offer improvements over linear equalization techniques and used mostly in 2G and 3G wireless applications.

1. Decision Feedback Equalization (DFE)
2. Maximum Likelihood Symbol Detection
3. Maximum Likelihood Sequence Estimation (MLSE)

1. Decision Feedback Equalization (DFE)

The basic idea behind decision feedback equalization is that once an information symbol has been detected and decided upon, the ISI includes on future symbols can be estimated and subtracted out before detection of subsequent symbols. The DFE can be realized in either the direct transversal from or as a lattice filter. The direct form is shown in fig3. It consists of a feed forwarded filter (FFF) and a Feedback filter (FBF). The FBF is driven by decision 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 equalizers has N1+ N2+1 taps in the feed forward filter and N3 taps in the feedback
filter and its output can be expressed as,

$$\hat{d}_k = \sum_{n=-N_1}^{N_2} c_n^* y_{k-n} + \sum_{i=1}^{N_3} F_i d_{k-i} \quad \text{------>1}$$

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.
The minimum mean squared error a DFE can achieve is

$$E[|e(n)|^2]_{min} = \exp\left\{\frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} \ln\left[\frac{N_0}{|F(e^{j\omega T})|^2 + N_0}\right]d\omega\right\} \quad \text{----->}$$

The minimum MSE for a DFE in above equation is always smaller than that of an LTE in equation 2 of linear equalizer unless F(ejωT) is constant. If there are nulls in F(ejωT) a DFE has significantly smaller minimum MSE than an LTE.

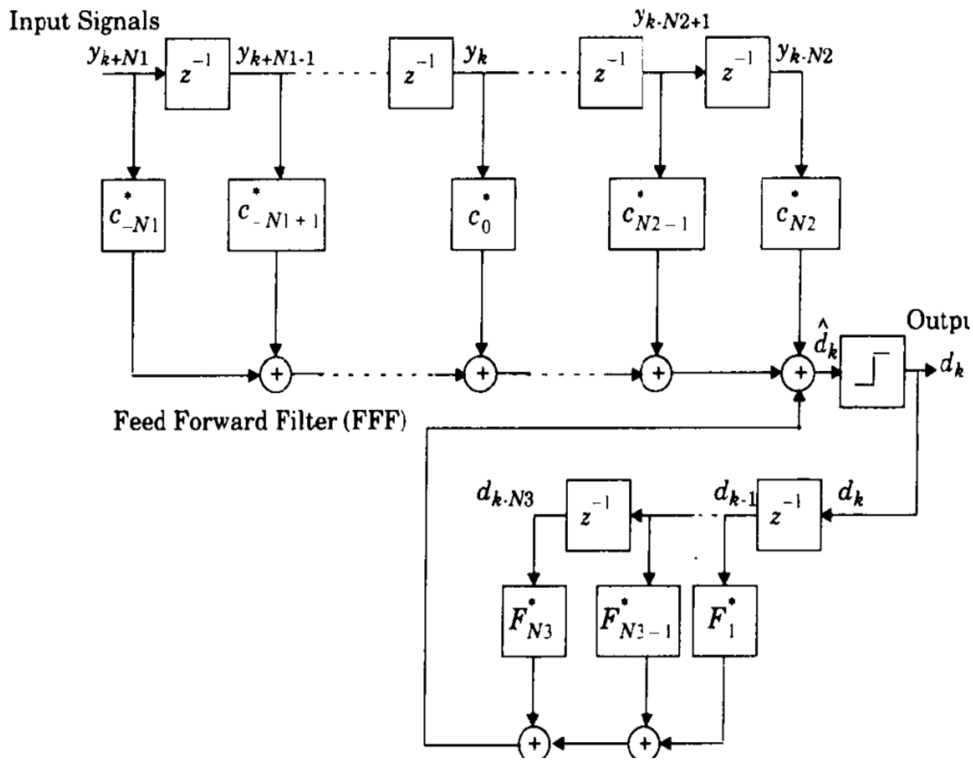


Fig.3 Direct form of Decision feedback equalizer

Therefore, an LTE is well behaved when the channel spectrum is comparatively flat, but if the channel is severely distorted or exhibits nulls in the spectrum, the performance of an LTE deteriorates and the mean squared error of a DFE is much better than a LTE. Also, an LTE has difficulty to equalizing a non minimum phase channel, where the strongest energy arrives after the first arriving signal component. Thus, a DFE is more appropriate for severely distorted wireless channels

The lattice implementation of the DFE is equivalent to a transversal DFE having a feed forward filter of length N1 and a feedback filter of length N2, where N1 >N2.The other name is shown in Figure 4. It also consists of a feed forward filter of (FFF) as in the conventional DFE. However, the feedback filter DFE is called a predictive DFE. (FBF) is driven by an input sequence formed by the difference of the output of the detector and the output of the feed forward filter. Hence, the FBF is called a noise predictor because it predicts the noise and the residual ISI contained in the signal at the FFF output and subtracts from it the detector output after some feedback delay.

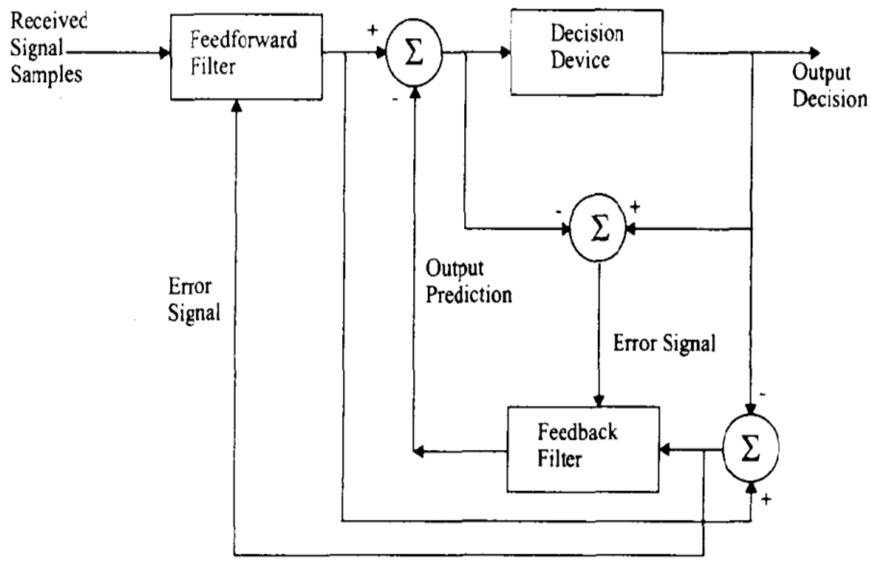


Fig.4 Predictive decision feedback equalizer

The predictive DFE performs as well as the conventional DFE as the limit in the number of taps in the FFF and the FBF approach infinity.

(ii) Maximum Likelihood Sequence Estimation (MLSE) Equalizer

The MSE-based linear equalizers are optimum with respect to the measure of minimum probability of symbol error when the channel does not introduce any amplitude distortion. However this is precisely the condition in which an equalizer is needed for a mobile communications link. These equalizers use various forms of the classical maximum likelihood receiver structure.

Using a channel impulse response simulator within the algorithm, the MLSE tests all possible data sequences (rather than decoding each received symbol by itself), and chooses the data sequence with the maximum probability as the output. An MLSE usually has a large computational requirement, especially when the delay spread of the channel is large. Using the MLSE as an equalizer to predict the MLSE estimator by using the Viterbi algorithm. It has recently been implemented successfully for equalizers in mobile radio channels.

The MLSE can be viewed as a problem to estimating the state of a discrete-time finite state machine, which in this case happens to be the radio channel with coefficients fk, and with a channel state which at any instant of time is estimated by the receiver based on the L most recent input samples. Thus the channel has ML states, where M is the size of the symbol alphabet of the modulation. That is, an ML trellis is used by the receiver to model the channel over time. The Viterbi algorithm tracks the state of the channel by the paths through the trellis and gives at stage k a rank ordering of the most probable sequences terminating in the most recent L symbols.

The block diagram of a MLSE receiver based on the DFE is shown in Figure 5. The MLSE is optimal to minimizes the probability of a sequence error.

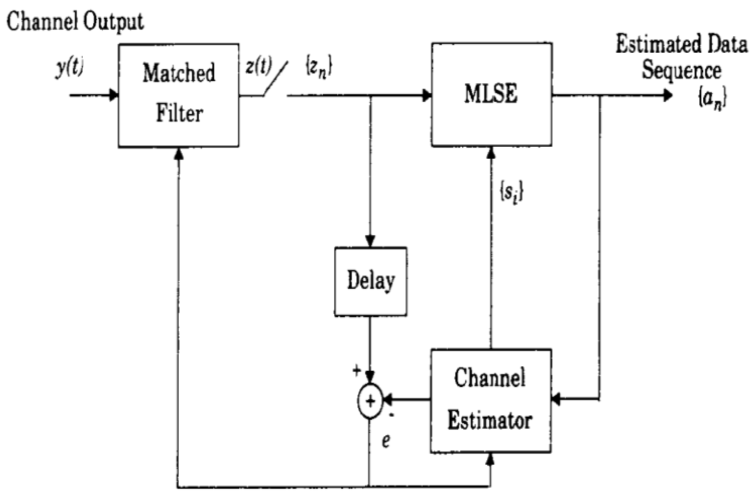


Fig.5 Structure of MLSE with adaptive matched filter

The MLSE requires knowledge of the channel characteristics in order to compute the metrics for making decisions. The MLSE also requires knowledge of the statistical distribution of the noise corrupting the signal.

Thus, the probability distribution of the noise determines the form of the metric for optimum demodulation of the received signal.

Zero Forcing Algorithm

In a zero forcing equalizer, the equalizer coefficientsC_n are chosen to force the samples of the combined channel and equalizer impulse response to zero at all but one of the NT spaced sample points in the tapped delay line filter. By the number of coefficients increase without bound, an infinite length equalizer with zero ISI at the output can be obtained.

When each of thedelay elements provide a time delay equal to the symbol duration T, the frequency response Heq (f) of the equalizer is periodic with a period equal to the symbol rate 1/T. The combined response of the channel with the equalizer must satisfy Nyquist's first criterion.

$$H_{ch}(f)H_{eq}(f) = 1, |f| < 1/2T \text{------->1}$$

where $H_{ch}(f)$ is the folded frequency response of the channel. Thus, an infinite length, zero, ISI equalizer is simply an inverse filter which inverts the folded frequency response of the channel. This infinite length equalizer is usually implemented by a truncated length version.The zero forcing algorithm was developed for wire line communication.

The zero forcing equalizer has the disadvantage that the inverse filter may excessively amplify noise at frequencies where the folded channel spectrum has high attenuation. The ZF equalizer neglects the effect of noise altogether, and is not often used for wireless links.

Least Mean Square Algorithm

A more robust equalizer is the LMS equalizer where the criterion used is the minimization of the mean square error (MSE) between the desired equalizer output and the actual equalizer output.

The prediction error is given by

$$e_k = d_k - \hat{d}_k = x_k - \hat{d}_k \qquad \text{---->1.}$$

And above equ becomes using vector notation

To compute the mean square error $\backslash ek\backslash^2$at time instant k , equation becomes

For a specific channel condition, the prediction error e_k is dependent on the tap gain vector w_N, so the MSE of an equalizer is a function of w_N. Let the function J (wN) denote he mean squared error as a function of tap gain vector w_N

.

$$\hat{d}_k(n) = \boldsymbol{w}_N^T(n)\boldsymbol{y}_N(n) \quad \text{------>6a}$$

$$e_k(n) = x_k(n) - \hat{d}_k(n) \quad \text{------>6b}$$

$$\boldsymbol{w}_N(n+1) = \boldsymbol{w}_N(n) - \alpha e_k^*(n)\boldsymbol{y}_N(n) \quad \text{------>6c}$$

The above equ. is called the normal equation, since the error is minimized and is made orthogonal. When above equation is satisfied, the MMSE of the equalizer is

$$J_{opt} = J(\hat{\boldsymbol{w}}_N) = \boldsymbol{E}[x_k x_k^*] - \boldsymbol{p}_N^T \hat{\boldsymbol{w}}_N \quad \text{------>5 a}$$

To obtain the optimal tap gain vector , equation in (5a) must be solved iteratively as the equalizer converges to an acceptably small value of J_{opt}. The minimization of the MSE is carried out recursively by use of the stochastic gradient algorithm. This is more commonly called the Least Mean Square (LMS) algorithm. LMS is computed iteratively,

$$\hat{d}_k(n) = \boldsymbol{w}_N^T(n)\boldsymbol{y}_N(n) \quad \text{------>6a}$$

$$e_k(n) = x_k(n) - \hat{d}_k(n) \quad \text{------>6b}$$

$$\boldsymbol{w}_N(n+1) = \boldsymbol{w}_N(n) - \alpha e_k^*(n)\boldsymbol{y}_N(n) \quad \text{------>6c}$$

whereN denotes the number of delay stages in the equalizer, and α is the step size which controls the convergence rate and stability of the algorithm.

The LMS equalizer maximizes the signal to distortion ratio at its output within the constraints of the equalizer filter length. If an input signal has a time dispersion characteristic that is greater than the propagation delay through the equalizer, then the equalizer will be unable to reduce distortion.

<u>Principle of diversity</u>

Diversity is a powerful communication receiver technique that provides wireless link improvement at relatively low cost. Unlike equalization, diversity requires no training overhead since a training sequence is not required by the transmitter.

Diversity exploits the random nature of radio propagation by finding independent (or at least highly uncorrected) signal paths for communication. In virtually all applications, diversity decisions are made by the receiver, and are unknown to the transmitter.

The diversity concept can be explained simply, if one radio path undergoes a deep fade, another independent path may have a strong signal. By having more than one path to select from, both the instantaneous and average SNRs at the receiver may be improved, often by as much as 20 dB to 30 dB.

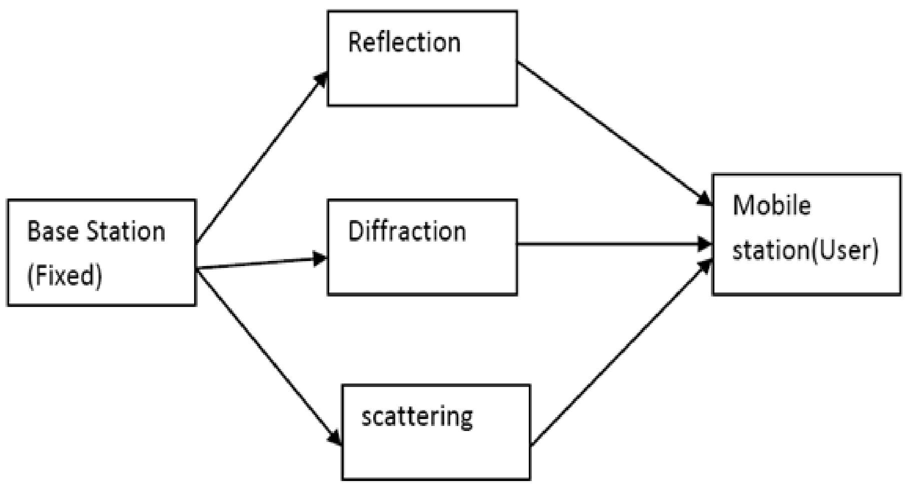


Fig 6• Basic Block diagram of Diversity Receiver

There are two types of fading

Small-scale fading

Large-scale fading

Small-scale fades are characterized by deep fade and rapid amplitude fluctuations which occurs as mobile moves over distances of just few wavelengths. These fades are caused by multiple reflections from the surroundings in the vicinity of the mobile. For narrow band signals small scale fading results in Rayleigh fading distribution of signal strength over small distances. In order to prevent deep fades from occurring, microscopic diversity techniques can exploit the rapid changing of the signal. By selecting the best signal at all times, a receiver can mitigate small-scale fading effects (antenna diversity or space diversity).

Micro diversity

Methods that can overcome small scale fading are therefore called as “micro diversity”. The five most common methods are

(i) Spatial diversity – several antenna elements are separated in space.

(ii) Temporal diversity – representation of transmit signal at different times. (iii) Frequency diversity – Transmission of signal on different frequency.

(iv) Angular diversity – Multiple antenna with different antenna patterns are used. (v) Polarization diversity – Multiple antenna receiving different polarization.

(i) Spatial diversity:

It is the simplest form of diversity.

- Transmit signal is received at different antenna elements.
- Large correlations of signals between the antennas are undesirable.
- The first important step in designing diversity antenna is to establish the relationship

between antenna spacing and correlation coefficient.

Ex: mobile station in cellular and cordless systems, it is standard assumption that waves are incident from all direction at mobile station.Thus points of the –ve interference of mpc are spaced approximately $\lambda/4$ apart.This is the distance required for decorrelation. Minimum distance for antenna elements in 900MHz GSM is 8cm and 1800MHz cordless is 4cm.

(ii)Temporal diversity:

These are 3 types.

(a) Repetition coding.

(b) ARQ (Automatic Repeat Request)

(c) combination of interleaving & coding.

Repetition coding:

The signal is repeated several times, where the repetition intervals are long enough to achieve declaration.

ARQ:

This method receives sense the message, the transmitted to indicate whether it receive the data sufficient quality. If not repeat request send by the receiver. ARQ is better than repetition coding because of higher spectral efficiency.

Combination of interleaving & coding: This is the more advanced reason repetition coding. The different symbols of a code word are transmitted at different times which increase the probability that at least some of them arrived with a good SNR.

(iii) Frequency diversity:

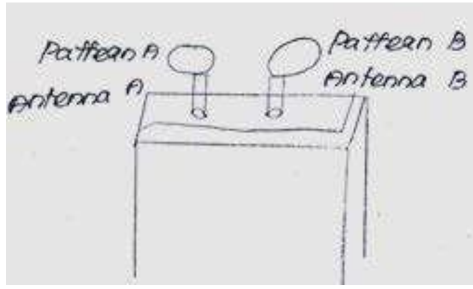
This method some signal is transmitted at different frequency. If these frequencies are apart by more than coherent bandwidth then their fading is approximately independent.

Coherent bandwidth: It is the statistical measurement of a range of frequencies area which the channel can be considered flat or in other words the approximate maximum bandwidth area which two frequency of signal lightly to experience same fading.

It is not common to actually repeat the same information at two different frequencies rather information spread area large bandwidth.

Spreading can be done by CDMA, OFDM, multi carrier CDMA etc., (iv)

Angular diversity:



It is used in conjuction with spatial diversity. This method enhances decorrelation. Different antenna patterns can be achieved ready easily. This diagram two antenna are used these two antennas are identical but can have a different radiation pattern when mounted closed to each other. This effect is due to mutual coupling. Antenna B acts as a reflector for antenna A where pattern is therefore skewed to the left. Similarly, pattern of antenna – B is skewed to the right due to the reflections form the antenna – A. Thus two patterns are different, this will lead to decorrelation.

(v) Polarization diversity:

Horizontally and vertically polarized multipath component (MPcs) propagates differently in a wireless channel. As the reflection and diffraction mainly depends on polarization.

Thus receiving both polarizations using a dual polarized antenna and processing the signal separately offers diversity. This diversity can be obtained without any requirement for a minimum distance between antenna elements.

Macro diversity

The correlation distance is for large scales fading are on the order of tens or hundreds of meters. So the spatial or temporal diversity cannot be used. For example, there is a hill between the transmitter and receiver, adding antenna on either the BS or MS doesn't help to eliminate the shadowing caused by this hill.

The simplest method for macro diversity is the use of "on frequency repeaters", that receivers the signal and retransmit an amplified position of it.

Simulcast:

Simulcast is very similar to ON frequency repeaters in which the same signal is transmitter simultaneously from different BS. The cellular application to base station should be synchronized. It is also widely used for broadcast application especially digital TV. The disadvantages of Simulcast is the large amount of signaling information that has to be carried on landlines or synchronization information as well as transmit data have to be transported on microwave hills of landlines to the BS.

Diversity combining Techniques

There are two ways of exploiting signals from the multiple diversity branches.

i. Selection diversity where the best signal copy is selected and processed (demodulated & decoded).

While all other copies are discarded. There are different criteria for what contributes the best signal.

ii. Combining diversity, where all copies of the signal are combined (before or after the demodulate), and the combined signal is decoded. Again there are different algorithms for combination of the signals. Combining diversity leads to better performance, as all available information is exploited, on the downside, it requires a more complex receiver than selection diversity. In most receivers all processing is done in the baseband.

Thus an receiver with combining diversity needs to down convert all available signals, and combine then approximately in the baseband. Thus it requires Nr antenna element as well as Nr complete radio frequency (RF) chains. An receiver with selection diversity requires only one RF chain, as it process only a single received signal at a time.

Types of combining techniques and space diversity methods:

- Selection diversity
- Feedback diversity
- Maximal ratio combining
- Equal gain diversity

1. Selection Diversity

Selection diversity is the simplest diversity technique. A block diagram of this method is shown in Figure 7, where m demodulators are used to provide m diversity branches whose 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.

The antenna signals themselves could be sampled and the best one sent to a single demodulator. In practice, the branch with the largest (S + N)/N is used, since it is difficult to measure SNR.

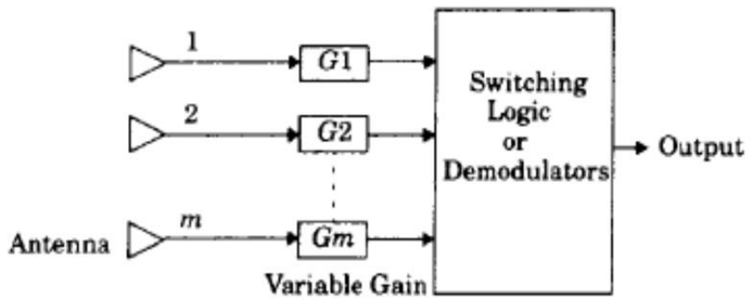


Fig 7.Block diagram of selection diversity

2. Feedback or Scanning Diversity

Scanning diversity is very similar to selection diversity except that 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. The resulting fading statistics are somewhat inferior to those obtained by the other methods but the advantage of this method is very simple to implement — only one receiver is required. A block diagram of this method is shown in Figure 8.

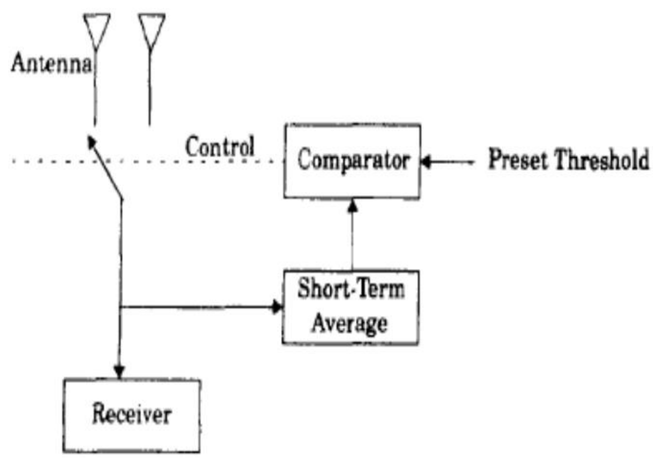


Fig.8 lock diagram of scanning diversity

3. Maximal Ratio Combining

In this method, the signals from all of the M branches are weighted according to their individual signal voltage to noise power ratios and then summed. Figure 9 shows a block diagram of the technique. Here, the individual signals must be co-phased before being summed(unlike selection diversity) which generally requires an individual receiver and phasing circuit for each antenna element.

Maximal ratio combining produces an output SNR equal to the sum of the individual SNRs. Thus, it has the advantage of producing an output with an acceptable SNR even when none of the individual signals are themselves acceptable. This technique gives the best statistical reduction of fading of any known linear diversity combiner.

Modern DSP techniques and digital receivers are now using this optimal form of diversity practical applications.

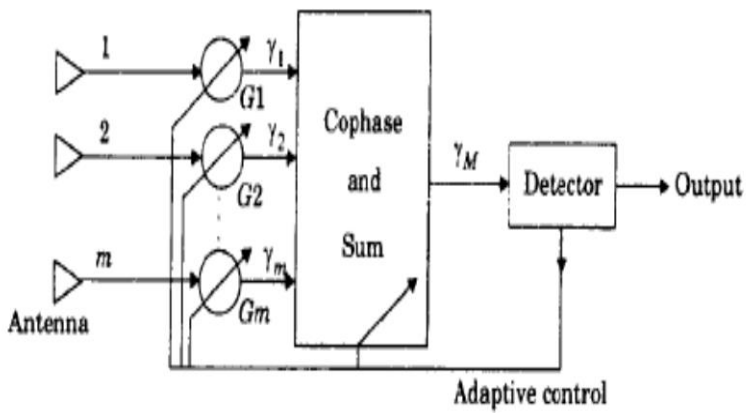


Fig.9 Maximal ratio combiner

RAKE Receiver.

Propagation delay spread in the radio channel merely provides multiple versions of the transmitted signal at the receiver. If these multipath components are delayed in time by more than a chip duration, they appear like uncorrected noise at a CDMA receiver, and equalization is not required.

However, since there is useful information in the multipath components, CDMA receivers may combine the time delayed versions of the original signal transmission in order to improve the signal to noise ratio at the receiver.

A RAKE receiver does just this it attempts to collect the time-shifted versions of the original signal by providing a separate correlation receiver for each of the multipath signals.

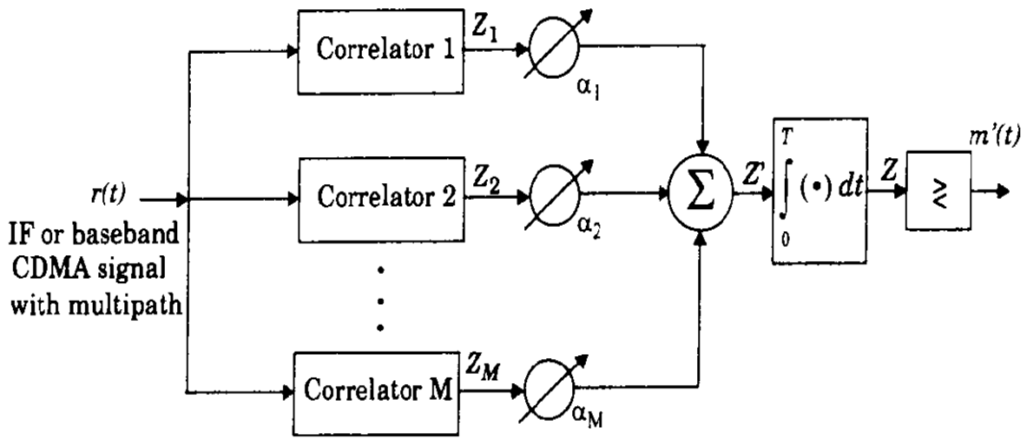


Fig:Rake Receiver

The RAKE receiver, shown in Figure , is essentially a diversity receiver designed specifically for CDMA, where the diversity is provided by the fact that the multipath components are practically uncorrelated from one another when their relative propagation delays exceed a chip period.

A RAKE receiver utilizes multiple correlators to separately detect the M strongest multipath components. The outputs of each correlator are weighted to provide a better estimate of the transmitted signal than is provided by a single component. Demodulation and bit decisions are then based on the weighted outputs of the Mcorrelators.

The basic idea of a RAKE receiver was first proposed by Price and Green [Pri58]. In outdoor environments, the delay between multipath components is usually large and, if the chip rate is properly selected, the low autocorrelation properties of a CDMA spreading sequence can assure that multipath components will appear nearly uncorrelated with each other.

Assume M correlators are used in a CDMA receiver to capture the M strongest multipath components. A weighting network is used to provide a linear combination of the correlator output for bit detection. Correlator 1 is synchronized to the strongest multipath m,. Multipath component m2 arrives t, later than component m].

The second correlator is synchronized to m2 . It correlates strongly with m2 but has low correlation with m, Note that if only a single correlator is used in the receiver ,once the output of the single correlator is corrupted by fading, the receiver cannot correct the value. Bit decisions based on only a single correlation may produce a large bit error rate. In a RAKE receiver, if the output from one correlator is corrupted by fading, the others may not be, and the corrupted signal may be discounted through the weighting process.

Decisions based on the combination of the M separate decision statistics offered by the RAKE provide a form of diversity which can overcome fading and thereby improve CDMA reception.

The M decision statistics are weighted to form an overall decision statistic as shown in Figure The outputs of the M correlators are denoted as Z_{19} Z_{Z1}... and Z_M. They are weighted by a,, a_2,... and a_M, respectively.

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.

Just as in the case of a maximal ratio combining diversity scheme, the overall signal Z' is given by

The weighting coefficients, a_m ,are normalized to the output signal power of the correlator in such a way that the coefficients sum to unity, as shown in following equation

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

Choosing Weighting Coefficients based on the actual outputs of the correlators yield better RAKE performance.

Unit-V

Multiple Antenna Technique

Multiple input multiple output (MIMO) system and spatial multiplexing

Wireless communication using multiple-input multiple-output (MIMO) systems enables increased spectral efficiency for a given total transmit power. Increased capacity is achieved by introducing additional spatial channels that are exploited by using space-time coding. MIMO systems are systems with Multiple Element Antennas (MEAs) at both link ends.

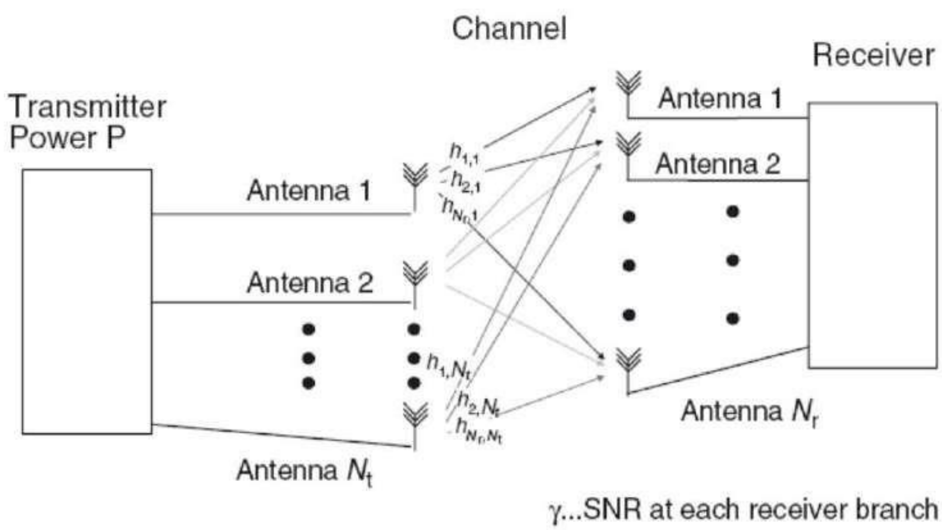


Figure 20.10 Block diagram of a multiple-input multiple-output system.

The MEAs of a MIMO system can be used for four different purposes: (i)

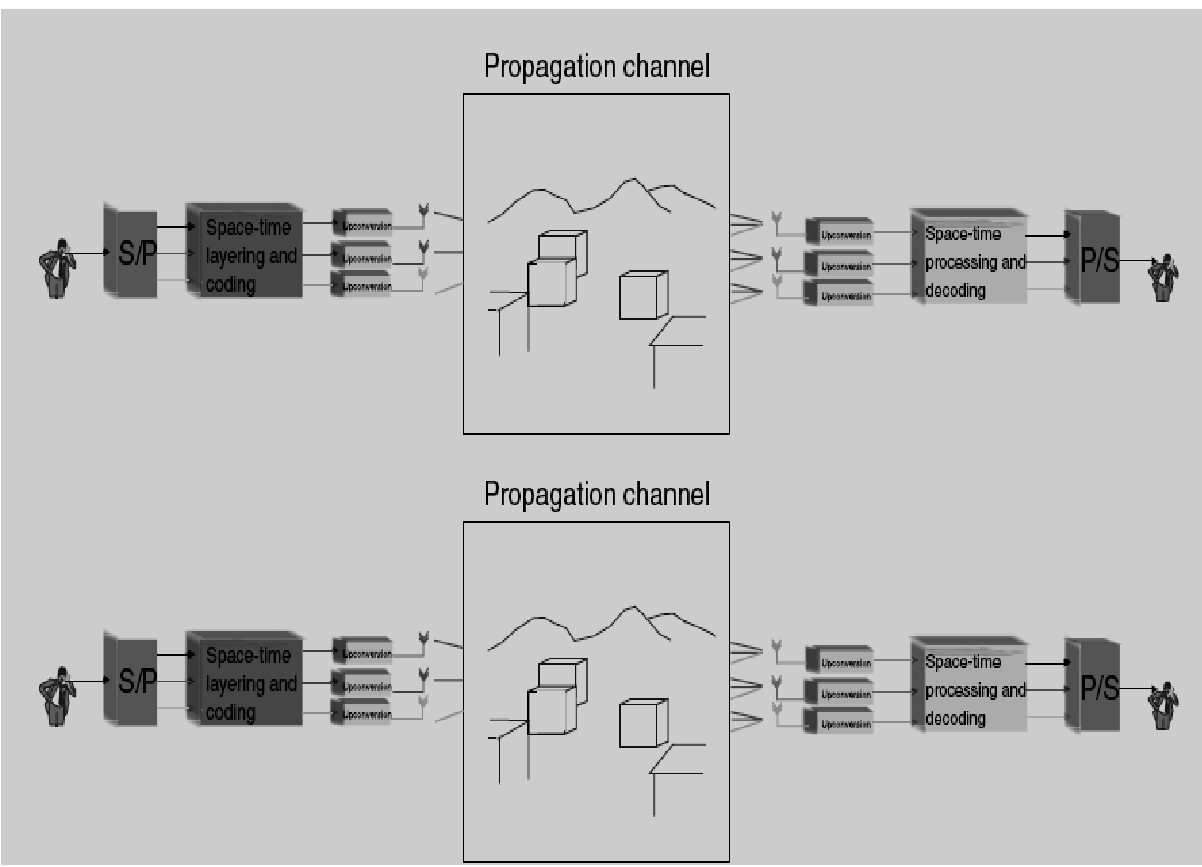
beamforming,

(ii) diversity,

(iii) interference suppression, and

(iv) spatial multiplexing (transmission of several data streams in parallel). The

first three concepts are the same as for smart antennas. Having multiple antennas at both link ends leads to some interesting new technical possibilities, but does not change the fundamental effects of this approach. Spatial multiplexing, on the other hand, is a new concept, and has thus drawn the greatest attention. It allows direct improvement of capacity by simultaneous transmission of multiple data streams



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. In this case, 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).

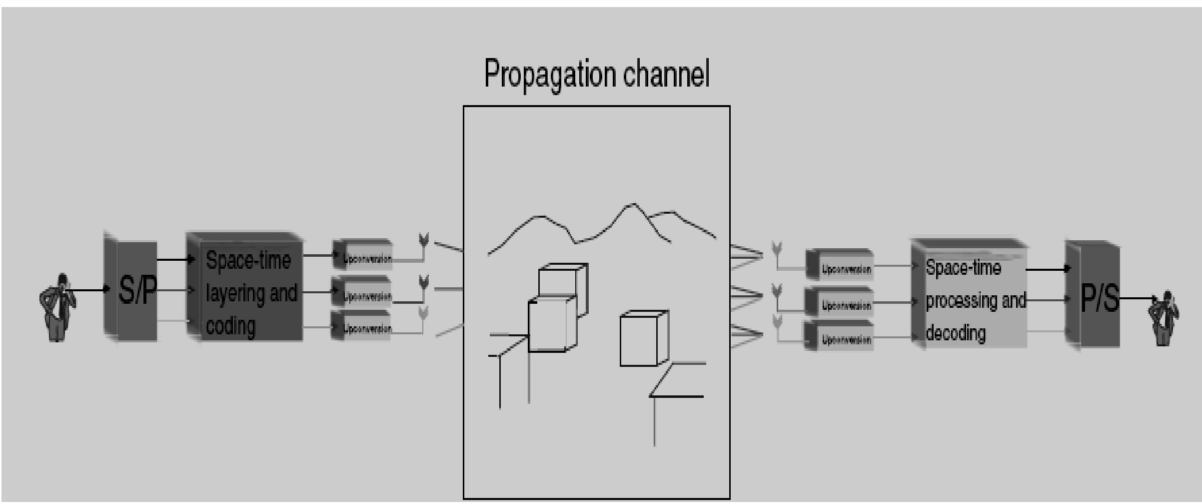


Fig:Principle behind spatial multiplexing

With Nt transmit antennas, we can form Nt different beams. We point all these beams at different Interacting Objects (IOs), and transmit different data streams over them.

At the RX, we can use Nr antenna elements to form Nr beams, and also point them at different IOs. If all the beams can be kept orthogonal to each other, there is no interference between the data streams;in other words, we have established parallel channels.

The IOs (in combination with the beam spointing in their direction) play the same role as wires in the transmission of multiple data streamson multiple wires.

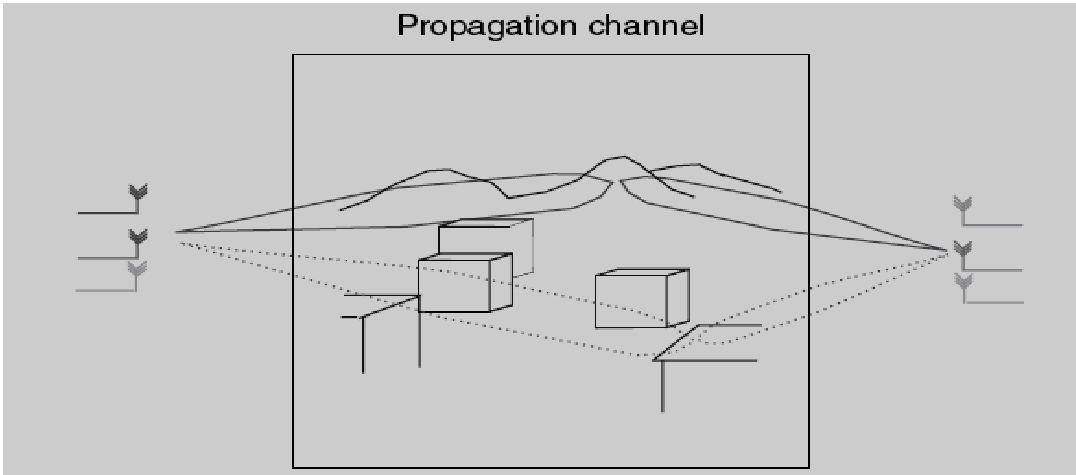


Fig:Transmission of different data streams via different interacting objects.

From this description, we can also immediately derive some important principles: the number of possible data streams is limited by min(Nt,Nr,Ns), where Ns is the number of (significant) IOs.

We have already seen above that the number of data streams cannot be larger than the number oftransmit antenna elements, and that we need a sufficient number of receive antenna elements (atleast as many as data streams) to form the receive beams and, thus, be able to separate the data streams. But it is also very important to notice that the number of IOs poses an upper limit: if two data streams are transmitted to the same IO, then the RX has no possibility of sorting them out by forming different beams. Operation

✓ Spatial multiplexing techniques simultaneously transmit independent information sequences, often called layers, over multiple antennas.

✓ High-rate signal is split into multiple lower-rate streams and each stream is transmitted from a different transmit antenna in the same frequency channel.

Received signal vector r = HS + n = x + n is received by Nr antenna elements , where S is the transmit signal vector and n is the noise vector.

Layered Space-Time Receiver Structures

The layered space time schemes, known as the Bell Laboratories layered space time (BLAST) schemes, were developed to achieve transmission rates above one symbol per channel.

Horizontal BLAST (H-BLAST)

✓ The data stream is first converted into N, parallel streams, each being encoded separately and then submitted to a diffferent transmit antenna. Nt =Nl.

Number of transmit antennas = Number of layers

$$X = \begin{bmatrix} x_1(1) & \dots & x_1(N_x) \\ x_2(1) & \dots & x_2(N_x) \\ x N_t(1) & \dots & x N_t(N_x) \end{bmatrix}$$

Joint ML decoding can be applied on each column. This leads to complexity, this complexity can be further decreased by using the DFE decoder. is DFE — Decision Feedback Equalizer.

✓ At the receiver side, these data streams are separated based on optimum combining.

Fig. 5.5. Assignment of bit streams to different antennas

Diagonal BLAST

✓ D-BLAST is a better solution than H-BLAST

✓ The generated N, substreams are subjected to a stream interleaver.

✓ Each sub stream is subdivided into many sub-blocks and these sub-blocks are transmitted by different antennas according to a round robin schedule.

Space-time code matrix of D-BLAST is given by

$$X = \begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_{N_l}(1) & & & \\ & x_1(2) & x_2(2) & \cdots & N_{N_l}(2) & & \\ & & \ddots & & & & \\ & & & x_1(N_t) & x_2(N_t) & \cdots & x_{N_l}(N_t) \end{bmatrix}$$

where

$x_l(n) \rightarrow n^{th}$ symbol of the l^{th} layer

$N_l \rightarrow$ Number of layers

The code rate of uncoded D-BLAST is given by

$$\eta_{D\text{-}BLAST} = \frac{N_l N_t}{N_l + N_t - 1} \text{ symbols per channel use.}$$

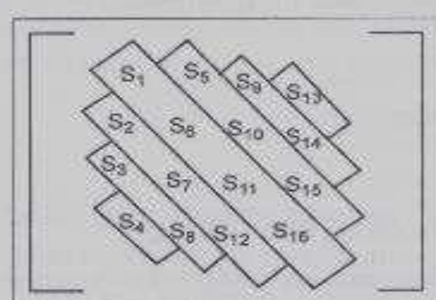


Fig. 5.6. Assignment of bit stream

At receiver side,

ML decoding + linear ZF or MMSE equalization can be used. ZF -

-zero forcing

MMSE —> Minimum Mean Square Error ML Maximum likelihood.

Vertical BLAST

✓ In V-BLAST the bit stream is first temporally encoded, interleaved and symbol mapped.

✓ The resulting Ns symbols are then demultiplexed into N, substreams, and transmitted over the antennas.

Advantages

1. Detection of V-BLAST is simpler than diagonal BLAST

2. An array gain of Nr can be achieved.

Drawback

1. Complex decoding scheme
2. 2. Not practical in cellular environments

Precoding and MIMO-Beam forming in MIMO Architecture

Precoding:

Spatial multiplexing with BLAST structure is used as a capacity-achieving technique. It is not practical in cellular environments. It depends on high SNR, so error propogation will occur. So, to overcome this drawbacks, precoding technique is used.

Precoding is motivated by the concept known as writing on dirty paper. The principle of dirty paper coding (DPC) states that the effect of the interference can be cancelled by proper coding.

Received signal is given by y=s+i+n

where s-- transmitted signal

i --> interference

n --AWGN —Additive White Gaussiain Noise

If there is no interference, capcity of the system is same. If there is interference, in MIMO channel, precoding is used to reduce it.

SVD Precoding:

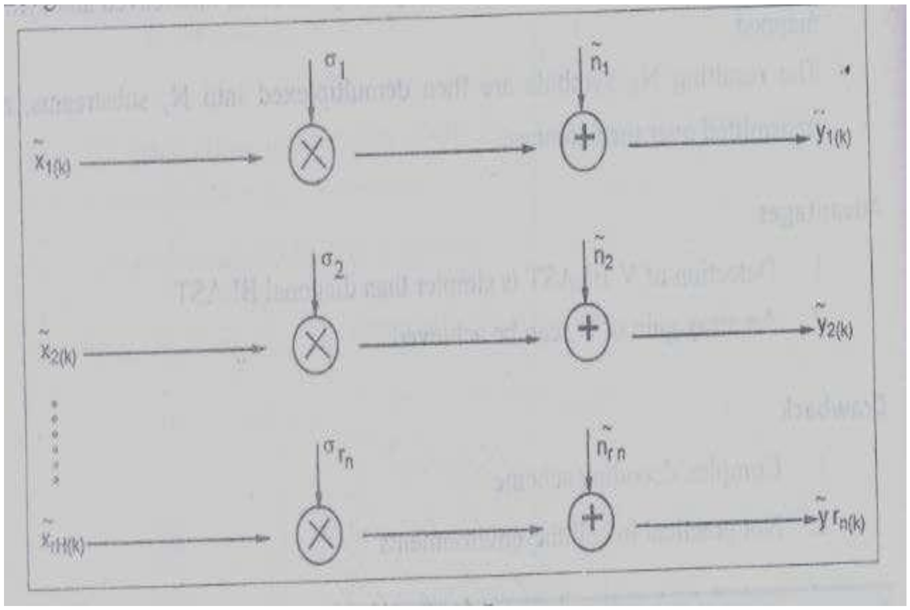


Fig:SVD Precoding x—channel input y—channel output

Shaping channel output

Linear transformation on the channel input x and output y is known as transmit precoding. Transmit precoding and receiver shaping transform the MIMO channel into r_H parallel scalar channels with input

Features

1. SVD precoding does not introduce noise enhancement.

2. If N r>= Nt, the complexity of the channel H is at O(Nr Nt^2,).

3. Closed loop spatial multiplexing can achieve high performance.

Linear Precoding

✓ Linear precoding is based on ZF beamforming

✓ It achieves high capacity

✓ As the number of user goes to infinity ZF beamforming capacity = dirty paper coding capacity .

The linear precoder decouples the input signal into orthogonal signal modes in the form of eigen-beams.

(i) In the case of perfect CSI [channel state information] the precoded orthogonal spatial modes match the channel eigen-directions. There is no interference between these signal streams.
(ii) With spatial CSI, precoder design must reduce the interference among signals
(iii) For perfect CSI at the transmitter, a diversity gain can also be delivered

<u>MIMO-Beam forming</u>

Multiple antennas are used to obtain beamforming or diversity gain. Same information is transmitted by more than one antenna, in the receiver side, symbol is weighted by a complex scale factor.

By MRC-maximal ratio combining technique, the largest eigenvalue of the channel is maximized

✓ Full diversity gain can be achieved by transmit beamforming and receive combining. The beamforming startegy corresponds to transmitter precoding and receiver shaping with Implementation of transmit beamforming requires CSI at the transmitter.

One solution is quantized beamforming. The receiver quantizes the beamforming vector using a fixed codebook. Codebook is availble at both the transmitter and the receiver

System Model and channel state information

Let us first establish the generic system that will be considered for capacity computations. Figure exhibits a block diagram. At the TX, the data stream enters an encoder, whose outputs are forwarded to Nt transmit antennas. From the antennas, the signal is sent through the wireless propagation channel, which is assumed to be quasi-static and frequency-flat if not stated otherwise.

By quasi-static we mean that the coherence time of the channel is so longthat “a large number” of bits can be transmitted within this time.We denote the Nr × Nt matrix of the channel as

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1N_t} \\ h_{21} & h_{22} & \cdots & h_{2N_t} \\ \vdots & \vdots & \cdots & \vdots \\ h_{N_r1} & h_{N_r2} & \cdots & h_{N_rN_t} \end{pmatrix}$$

whose entries hij are complex channel gains (transfer functions) from the j th transmit to the ithreceive antenna. The received signal vector

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{n} = \mathbf{x} + \mathbf{n}$$

contains the signals received by Nr antenna elements, where **s** is the transmit signal vector and **n**is the noise vector. Channel State information Algorithms for MIMO transmission can be categorized by the amount of CSI that they require. We distinguish the following cases:

1. Full CSI at the TX (CSIT) and full CSI at the RX (CSIR): in this ideal case, both the TX and the RX have full and perfect knowledge of the channel. This case obviously results in the highest possible capacity.

2. Average CSIT and full CSIR: in this case, the RX has full information of the instantaneouschannel state, but the TX knows only the average CSI – e.g., the correlation matrix of H or theangular power spectrum. To achieve anddoes not require reciprocity or fast feedback; however, it does require calibration (to eliminatethe nonreciprocity of transmit and receive chains) or slow feedback.

3. No CSIT and full CSIR: this is the case that can be achieved most easily, without any feedback or calibration. The TX simply does not use any CSI, while the RX learns the instantaneouschannel state from a training sequence or using blind estimation.

4. Noisy CSI : when we assume "full CSI" at the RX, this implies that the RX has learned thechannel state perfectly. However, any received training sequence will be affected by additivenoise as well as quantization noise. It is thus more realistic to assume a "mismatched RX,"where the RX processes the signal based on the observed channel Hobs, while in reality the signals pass through channel Htrue

5. No CSIT and no CSIR: it is remarkable that channel capacity is also high when neither the TX nor the RX have CSI. We can, e.g., use a generalization of differential modulation.

Capacity in non fading channels

The first key step in understanding MIMO systems is the derivation of the capacity equation for MIMO systems in nonfading channels, often known as "Foschini's equation" [Foschini and Gan

1998]. Let us start with the capacity equation for “normal” (single-antenna) AdditiveWhite GaussianNoise (AWGN) channels.

$$C_{\text{shannon}} = \log_2\left(1 + \gamma \cdot |H|^2\right)$$

Where γ is the SNR at the RX, and H is the normalized transfer function from the TX to the RX. The key statement of this equation is that capacity increases only logarithmically with the SNR, so that boosting the transmit power is a highly ineffective way of increasing capacity. Consider a singular value decomposition11 of the channel:

$$\mathbf{H} = \mathbf{W}\boldsymbol{\Sigma}\mathbf{U}^{\dagger}$$

where∑is a diagonal matrix containing singular values, and **W** and **U**† are unitary matrices composed of the left and right singular vectors, respectively. The received signal is then

$$\begin{aligned}\mathbf{r} &= \mathbf{H}\mathbf{s} + \mathbf{n} \\ &= \mathbf{W}\boldsymbol{\Sigma}\mathbf{U}^{\dagger}\mathbf{s} + \mathbf{n}\end{aligned}$$

Then, multiplication of the transmit data vector by matrix **U** and the received signal vector by **W**†diagonalizes the channel:

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

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

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

$$C = \sum_{k=1}^{R_{\mathrm{H}}} \log_2\left[1 + \frac{P_k}{\sigma_{\mathrm{n}}^2}\sigma_k^2\right]$$

Where σ2 is noise variance, and Pkis the power allocated to the eigenmode; we assume that

Pk= P is independent of the number of antennas. This capacity expression can be shown to be equivalent to

Where $\mathbf{I}_{Nr}$ is the $Nr \times Nr$ identity matrix, γ is the mean SNR per RX branch, and $\mathbf{R}_{ss}$is the

correlation matrix of the transmit data (if data at the different antenna elements are uncorrelated, it is a diagonal matrix with entries that describe the power distribution among antennas).

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

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

$$C = \log_2 \left[\det \left(\mathbf{I}_{N_r} + \frac{\overline{\gamma}}{N_t} \mathbf{H}\mathbf{H}^{\dagger} \right) \right]$$

It is worth noting that (for sufficiently large Ns) the capacity of a MIMO system increases linearly with min(Nt,Nr), irrespective of whether the channel is known at the TX or not.

Let us now look at a few special cases. To make the discussion easier, we assume that
$N_r = N$:

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

$$C_{\mathrm{MIMO}} = \log_2(1 + N\overline{\gamma})$$

We see that in this case the SNR is increased by a factor of N compared with the single antenna case, due to beam forming gain at the RX. However, this only leads to a logarithmic increase incapacity with the number of antennas.

2. All transfer functions are different such that the channel matrix is full rank, and has N eigen values 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 is

$$C_{\mathrm{MIMO}} = \log_2(1 + N\overline{\gamma})$$

and, thus, increases linearly with the number of antenna elements.

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

$$C_{MIMO} = N \log_2\left(1 + \frac{\overline{\gamma}}{N}\right)$$

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

Let us next consider the case where both the RX and TX know the channel perfectly. In such a case, it can be more advantageous to distribute power not uniformly between the different transmit antennas (or eigen modes) but rather assign it based on the channel state. In other words, we are faced with the problem of optimally allocating power to several parallel channels, each of which has a different SNR, and therefore the answer is the same: water filling.

Transmitter diversity

Transmit Diversity

Multiple antennas can be installed at just one link end (usually the BS). For the uplink transmission from the MS to BS, multiple antennas can act as receive diversity branches.

For the downlink, any possible diversity originates at the transmitter. we will thus discuss ways of transmitting signals from several TX antennas and achieve a diversity effect with it.

Time diversity and frequency diversity inherently involve the TX, and thus need not be discussed again here.

Transmitter Diversity with Channel State Information

The first situation we analyze is the case where the TX knows the channel perfectly. the optimum transmission scheme linearly weights signals transmitted from different antenna elements with the complex conjugates of the channel transfer functions from the transmit antenna elements to the single receive antenna. This approach is known as maximum ratio transmission.

Transmitter Diversity Without Channel State Information

In many cases, Channel State Information (CSI) is not available at the TX. We then cannot simply transmit weighted copies of the same signal from different transmit antennas, because we cannot know how they would add up at the RX. It is equally likely for the addition of different components to be constructive or destructive; in other words, we would just be adding up MPCs with random phases, which results in Rayleigh fading.

We thus cannot gain any diversity (or beam forming).In order to give benefits, transmission of the signals from different antenna elements has to be done is such a way that it allows the RX to distinguish different transmitted signal components. One way is delay diversity. In this scheme, signals transmitted from different antenna elements are delayed copies of the same signal. This makes sure that the effective impulse response is delay dispersive, even if the channel itself is flat fading.

So, in a flat-fading channel, we transmit data streams with a delay of 1 symbol duration (relative to preceding antennas) from each of the transmit antennas. The effective impulse response of the channel then becomes

$$h(\tau) = \frac{1}{\sqrt{N_t}} \sum_{n=1}^{N_t} h_n \delta(\tau - nT_s)$$

where the hn are gains from the nth transmit antenna to the receive antenna, and the impulse response has been normalized so that total transmit power is independent of the number of antenna elements. The signals from different transmit antennas to the RX act effectively as delayed MPCs. If antenna elements are spaced sufficiently far apart, these coefficients fade independently. If the channel from a single transmit antenna to the RX is already delay dispersive, then the scheme still works, but care has to be taken in the choice of delays for different antenna element