Chapter 1. Overview of Wireless Communications

1.1 Introduction

Wireless personal communications:
- any person (properly equipped)
- anywhere
- anytime
- in any format chosen

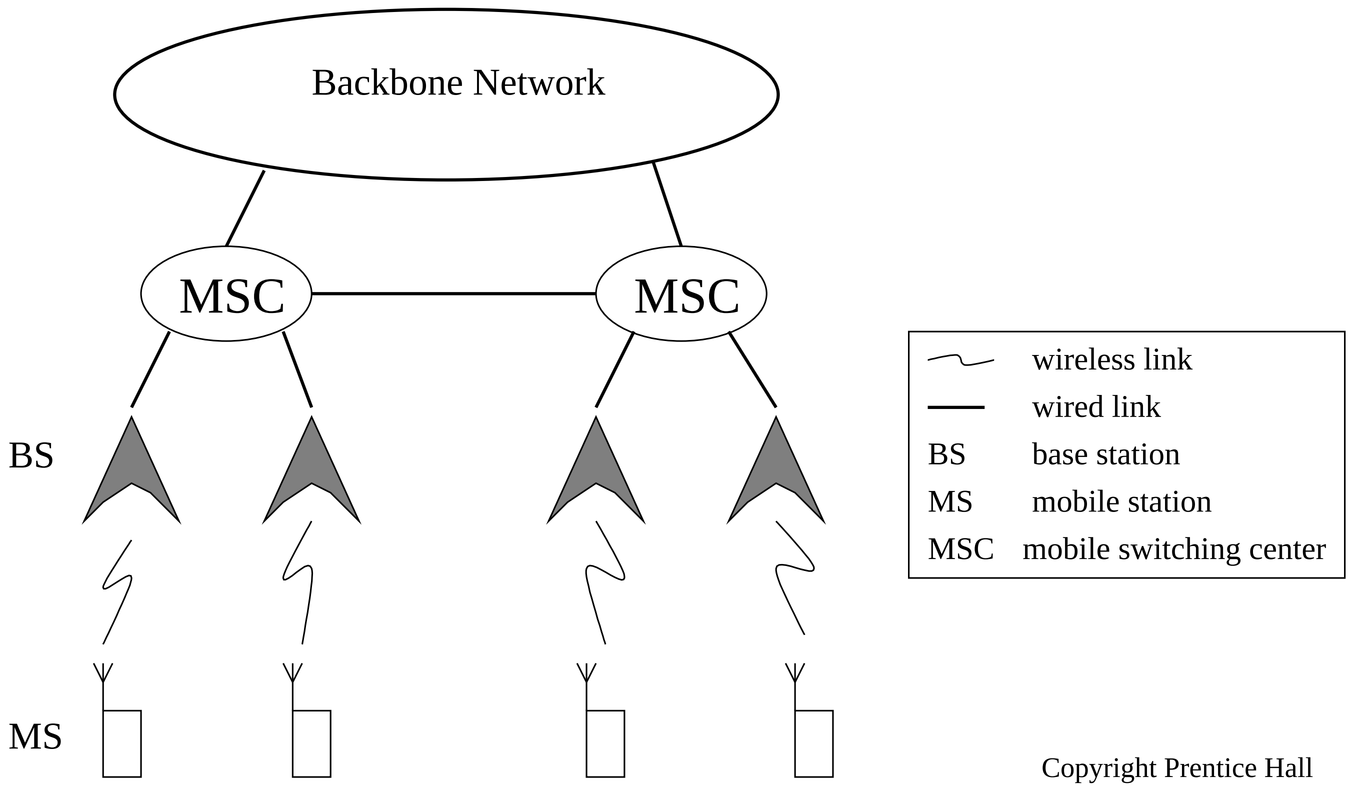


Figure 1: An illustration of wireless communications network

Downlink (forward link): from BS to MS

Uplink (reverse link): from MS to BS

Definition of radio cell

A radio cell is a geographical area served by a single base station supporting the services of many mobile stations.

Depending on the size of the area covered, radio cells are categorized into picocells, microcells, and macrocells.

The main problems in wireless communications come from
- the hostile wireless propagation medium
- user mobility

The wireless channel:

- Multiplicative interference and distortion are normally signal-dependent, including fading, intersymbol interference, etc. They cause attenuation, mutilation, etc., of the transmitted signal. The net result is a reduction in usable frequency spectrum. This form of disturbance cannot be suppressed using filtering.

- Additive noise is not as severe as multiplicative noise, but still reduces signal detectability. Out-of-band noise can be suppressed by filtering, but in-band noise will still penetrate through the filter.

1.2 Wireless Communications Standards

First Generation Analog Cellular Systems:

Region	America	Europe	Japan
Parameter	AMPS	ETACS	NTT
Multiple Access	FDMA	FDMA	FDMA
Duplexing	FDD	FDD	FDD
Forward Channel	869-894 MHz	935-960 MHz	870-885 MHz
Reverse Channel	824-849 MHz	890-915 MHz	925-940 MHz
Channel Spacing	30 kHz	25 kHz	25 kHz
Data Rate	10 kbps	8 kbps	0.3 kbps
Spectral Efficiency	0.33 bps/Hz	0.33 bps/Hz	0.012 bps/Hz
Capacity	832 channels	1000 channels	600 channels

AMPS: Advanced Mobile Phone System (America and Australia)

ETACS: European Total Access Communications System

NTT: Nippon Telephone and Telegraph System

Second Generation Digital Cellular Systems:

Region	U.S.	Europe	Japan	U.S.
Parameter	IS-54	GSM	PDC	IS-95
Multiple Access	TDMA/FDD	TDMA/FDD	TDMA/FDD	CDMA/FDD
Modulation	$\pi/4$ DQPSK	GMSK	$\pi/4$ DQPSK	QPSK/OQPSK
Forward Channel	869-894 MHz	935-960 MHz	810-826 MHz	869-894 MHz
Reverse Channel	824-849 MHz	890-915 MHz	940-956 MHz	824-849 MHz
Channel Spacing	30 kHz	200 kHz	25 kHz	1,250 kHz
Data/Chip Rate	48.6 kbps	270.833 kbps	42 kbps	1.2288 Mcps
Speech Codec Rate	7.95 kbps	13.4 kbps	6.7 kbps	1.2/2.4/ 4.8/9.6 kbps

IS (Interim Standard)-54: North-America TDMA Digital Cellular System

GSM: Groupe Special Mobile $\longrightarrow$ Global System for Mobile

PDC: Personal Digital Cellular

IS-95: CDMA Cellular System

Third Generation Wireless Communications Networks:

○ support high rate and variable rate transmission

○ provide voice, data, and low-rate video multimedia services

○ use convolutional coding for voice and Turbo coding for data

○ use wideband CDMA (WCDMA) or multi-carrier CDMA (MC-CDMA)

○ operate in TDD or FDD

○ fast power control and reverse link pilot signals

○ synchronous or asynchronous transmission

○ will be equipped with the infrastructure to support Personal Communications Systems (PCS)

○ The network infrastructure support will likely include

- public land mobile networks (PLMNs),
- Mobile Internet Protocol (Mobile IP),
- wireless asynchronous transfer mode (WATM) networks, and
- low earth orbit (LEO) satellite networks.

3G Air Interface Options:

Air interface	IMT class	Technology	Standard	Name
CDMA	IMT-DS	direct spread	UTRA-FDD	WCDMA
CDMA	IMT-MC	multicarrier	IS-95C/cdma2000	Multicarrier
CDMA	IMT-TC	time code	UTRA-TDD	TD-SCDMA
TDMA	IMT-SC	single carrier	UWC-136	n.a.

Partnership Projects:

- Third-Gneration Partnership Project (3GPP)
  - comprised primarily of those bodies that were responsible for GSM standards (mainly Europe and Asia)
  - concerned with UTRA-FDD and UTRA-TDD
  - support GSM-MAP approach for network signaling
  - support WCDMA air interface
- Third-Generation Partnership Project 2 (3GPP2)
  - comprised primilarily of those bodies that were responsible for the IS-136 and IS-95 standards (mainly US)
  - concerned with cdma2000 and UWC-136
  - support ANSI-41 approach for network signaling
  - support multicarrier CDMA

High-speed wireless data technologies as of early 2004:

Technologies	Gener-ation	Multiple Access	Carrier Bandwidth	Peak Data Rate	Modulation	Data/Voice Support
GPRS	2.5G	TDMA	0.2MHz	115kbps	GMSK	data
	Designed as overlay on GSM / TDMA Networks					
EDGE	2.5G	TDMA	2.4Mbps	384kbps	8PSK/GMSK	data
	Designed as overlay on GSM / TDMA Networks					
cdma2000 1x Rel. 0	3G	CDMA	1.25Mbps	628kbps	BPSK/QPSK	data+voice
IS-856 (1xEV-DO, HDR)	3G+	CDMA/ TDMA	1.25Mbps	2.4576Mbps	QPSK/8PSK/ 16QAM	data
	Designed as high data rate extension to cdma2000					
cdma2000 1x Rel. D (1xEV-DV)	3G+	CDMA	1.25Mbps	3.0912Mbps	QPSK/8PSK 16QAM	data+voice
	Designed as high data rate extension to cdma2000 1x with voice support					
UMTS Rel. 5 (HSDPA)	3G	CDMA	3.84MHz	1.96Mbps	QPSK/16QAM	data+voice
	Designed as high throughput, high peak data rate extension to UMTS Release 99					

1.3 What we will study in this course?

Focus: principles of wireless communications

- Characterization of the propagation channel (Chapter 2)
- Bandpass signaling techniques (Chapter 3)
- Diversity (Chapter 4)
- Fundamentals of cellular communications (Chapter 5)
- Multiple access techniques (Chapter 6)
- Mobility management (Chapter 7)
- Wireless/wireline interworking (Chapter 8)

Ch.2. Characterization of the Wireless Channel

2.1 Multipath Propagation

Multipath propagation: A radio signal arrives at a receiver via two or more of many possible routes with the result that the arriving signals. although having a common time origin at the transmitter, arrive "out of step".

Channel characterization:

• **Time dispersion**: different propagation paths have different propagation delays $\Longrightarrow$ received signal pulse is wider than the transmitted pulse, which introduces inter-symbol interference (ISI) to the received signal (mainly to wideband signals)

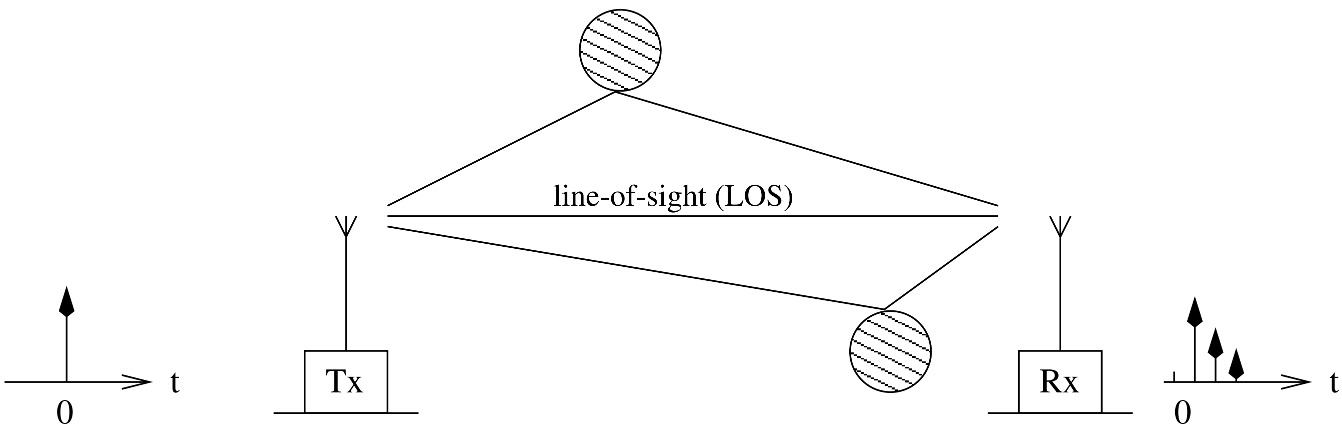


Figure 1: An illustration of multipath propagation

Many different propagation paths can have a common propagation delay as long as the paths have the same length.

All the scatterers located on an ellipse with the transmitter and receiver as the foci corresponds to a *distinct* propagation delay, and is therefore called a *distinct scatterer*.

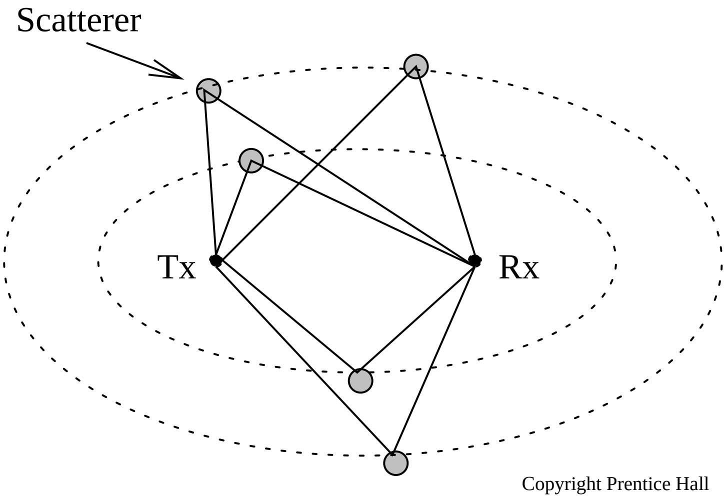


Figure 2: Ellipsoidal portrayal of scatterer location

The receiver cannot distinguish the individual received signal component coming from path 1 from the component coming from path 2, etc.

$\Longrightarrow$ The received signal component with propagation delay $\tau_1 = d/c$ (where c is the velocity of light) is a result of the multiple propagation paths.

• **Fading**: Consider transmitting a single-tone sinusoid. For simplicity, consider a two-path channel, where the delay of the line-of-sight (LOS) or direct path is assumed to be zero, and the delay of the non-line-of-sight (NLOS) or reflected path is τ.

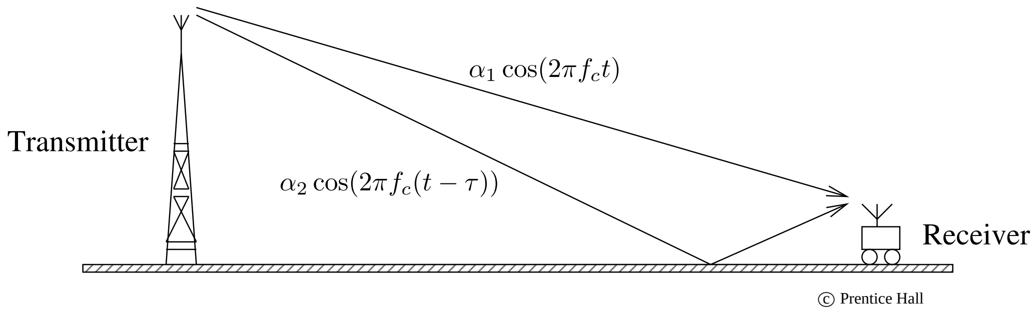


Figure 3: A channel with two propagation paths

The received signal, in the absence of noise, can be represented as

$$r(t) = \alpha_1 \cos(2\pi f_c t) + \alpha_2 \cos(2\pi f_c(t-\tau)), \tag{1}$$

where α_1 and α_2 are the amplitudes of the signal components from the two paths respectively. The received signal can also be represented as

$$r(t) = \alpha \cos(2\pi f_c t + \phi), \tag{2}$$

where

$$\alpha = \sqrt{\alpha_1^2 + \alpha_2^2 + 2\alpha_1\alpha_2 \cos(2\pi f_c \tau)}$$

and

$$\phi = -\tan^{-1}[\frac{\alpha_2 \sin(2\pi f_c \tau)}{\alpha_1 + \alpha_2 \cos(2\pi f_c \tau)}]$$

are the amplitude and phase of the received signal. Both α and ϕ are functions of α_1, α_2, and τ.

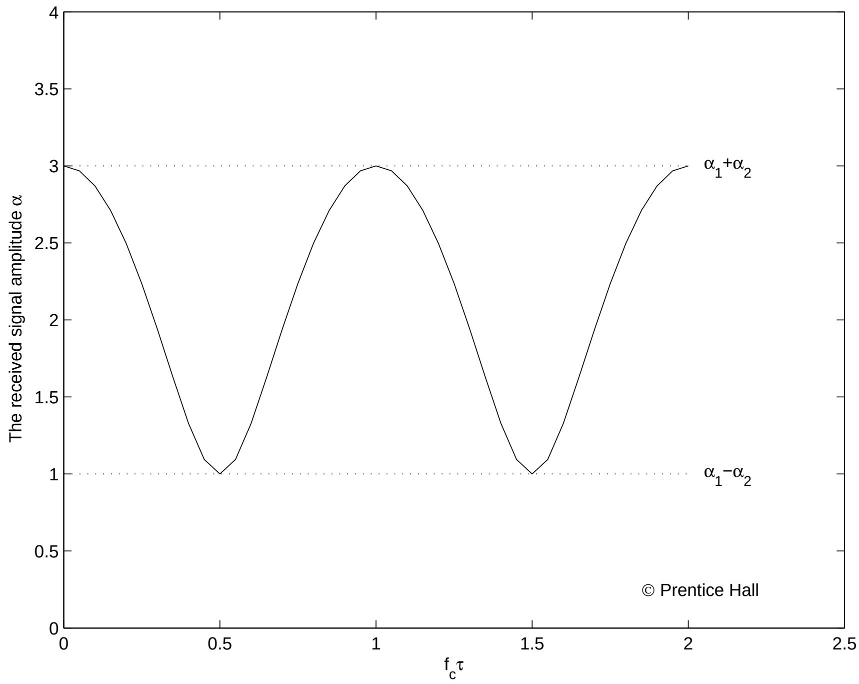


Figure 4: The amplitude fluctuation of the two-path channel with $\alpha_1 = 2$ and $\alpha_2 = 1$

As the mobile user and/or scatterer moves, $f_c\tau$ changes $\rightarrow$ the two received components add constructively (when $f_c\tau = 0, 1, 2, \ldots$) or destructively (when $f_c\tau = 0.5, 1.5, 2.5, \ldots$).

$\Longrightarrow |r(t)|$ is a function of t, referred to as amplitude fading.

At each deep fade, instantaneous SNR $\downarrow \longrightarrow$ BER $\uparrow$.

In summary, multipath propagation in the wireless mobile environment results in a fading dispersive channel.

2.2 The Linear Time-Variant Channel Model

Consider a multipath propagation environment with N distinct scatterers, each characterized by amplitude fluctuation $\alpha_n(t)$ and propagation delay $\tau_n(t)$, $n = 1, 2, \ldots, N$.

Consider a narrowband signal $\tilde{x}(t)$ transmitted over the wireless channel at a carrier frequency f_c, such that

$$\tilde{x}(t) = \Re\{x(t)e^{j2\pi f_c t}\}, \tag{3}$$

where $x(t)$ is the complex envelope of the signal.

In the absence of background noise, the received signal at the channel output is

$$\begin{aligned} \tilde{r}(t) &= \Re\{\sum_{n=1}^{N} \alpha_n(t)x(t-\tau_n(t))e^{j2\pi f_c(t-\tau_n(t))}\} \\ &= \Re\{r(t)e^{j2\pi f_c t}\}, \end{aligned}$$

where $r(t)$ is the complex envelope of the received signal and can be represented as

$$r(t) = \sum_{n=1}^{N} \alpha_n(t)e^{-j2\pi f_c \tau_n(t)}x(t-\tau_n(t)). \tag{4}$$

As the mobile user moves, $\alpha_n(t)$ and $\tau_n(t)$ are a function of $t \Longrightarrow$ the channel is linear time-variant.

$\Longrightarrow$ The channel impulse response depends on the instant that the impulse is applied to the channel.

Channel impulse response

Linear time-invariant (LTI) channel:

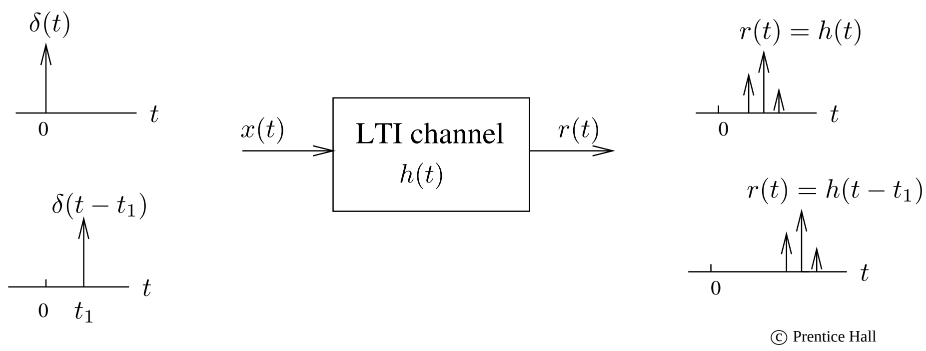


Figure 5: The linear time-invariant channel model

We can use $h(t)$ to describe the channel, where t is a variable describing propagation delay, assuming that the impulse is always applied to the channel at time zero.

Linear time-variant (LTV) channel:

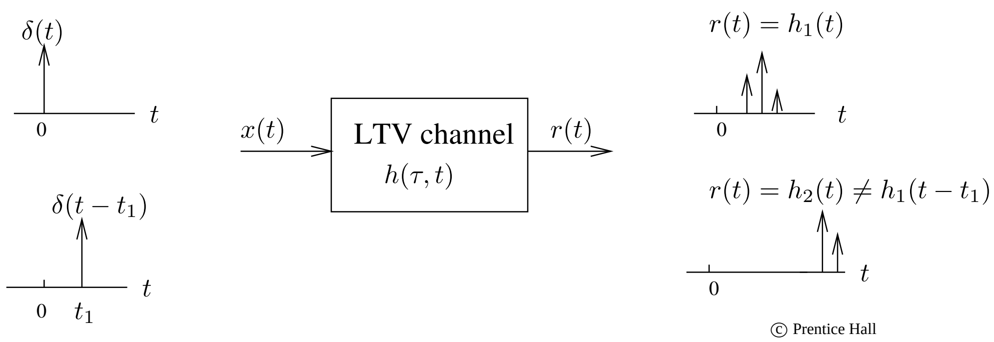


Figure 6: The linear time-variant channel model

The channel impulse response is a function of two variables: one describing when the impulse is applied to the channel, the other describing the moment of observing the channel output or the associated propagation delay.

Definition:

> The impulse response of an LTV channel, $h(\tau, t)$, is the channel output at t in response to an impulse applied to the channel at $t - \tau$.

The received signal is

$$r(t) = \int_{-\infty}^{\infty} h(\tau, t)x(t-\tau)d\tau. \tag{5}$$

The channel impulse response for the channel with N distinct scatterers is then

$$h(\tau, t) = \sum_{n=1}^{N} \alpha_n(t)e^{-j\theta_n(t)}\delta(\tau - \tau_n(t)), \tag{6}$$

where $\theta_n(t) = 2\pi f_c \tau_n(t)$ represents the carrier phase distortion introduced by the nth scatterer.

Note:
$\tau_n(t)$ changes by $1/f_c \longrightarrow \theta_n(t)$ changes by 2π
$\Longrightarrow$ a small change in the propagation delay $\longrightarrow$ a small change in $\alpha_n(t)$ and $\tau_n(t)$, but a significant change in $\theta_n(t)$

That is, the carrier phase distortion $\theta_n(t)$ is much more sensitive to user mobility than the amplitude fluctuation $\alpha_n(t)$.

Time-variant transfer function

Linear time-invariant channel:

$$H(f) = \mathcal{F}[h(t)] \tag{7}$$

The channel output in the frequency domain is

$$R(f) = H(f)X(f). \tag{8}$$

Linear time-variant channel:

Definition

> The time-variant transfer function of an LTV channel is the Fourier transform of the impulse response, $h(\tau, t)$, with respect to the delay variable τ.

$$\begin{cases} H(f,t) = \mathcal{F}_\tau[h(\tau,t)] = \int_{-\infty}^{\infty} h(\tau,t)e^{-j2\pi f\tau}d\tau \\ h(\tau,t) = \mathcal{F}_f^{-1}[H(f,t)] = \int_{-\infty}^{\infty} H(f,t)e^{+j2\pi f\tau}df \end{cases}$$

where the time variable t can be viewed as a parameter.

The received signal is

$$\begin{aligned} r(t) &= \int_{-\infty}^{\infty} h(\tau, t) x(t-\tau) d\tau \\ &= \int_{-\infty}^{\infty} x(t-\tau)[\int_{-\infty}^{\infty} H(f,t)\exp(j2\pi f\tau)df]d\tau \\ &= \int_{-\infty}^{\infty} H(f,t)\{\int_{-\infty}^{\infty} x(t-\tau)\exp[-j2\pi f(t-\tau)]d\tau\}\exp(j2\pi ft)df \\ &= \int_{-\infty}^{\infty} H(f,t)[\int_{\infty}^{-\infty} x(\xi)\exp(-j2\pi f\xi)(-d\xi)]\exp(j2\pi ft)df \\ &= \int_{-\infty}^{\infty} H(f,t)X(f)\exp(j2\pi ft)df \\ &\triangleq \int_{-\infty}^{\infty} R(f,t)\exp(j2\pi ft)df \end{aligned}$$

where

$$R(f,t) = H(f,t)X(f)$$

and

$$X(f) = \mathcal{F}[x(t)].$$

Summary:

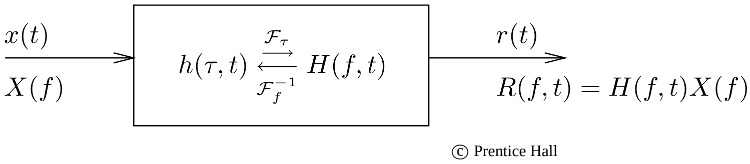


Figure 7: Frequency-time channel representation

where

$$
\begin{aligned}
r(t) &= \int_{-\infty}^{\infty} h(\tau,t)x(t-\tau)d\tau \\
R(f,t) &= H(f,t)X(f) \\
X(f) &= \int_{-\infty}^{\infty} x(t)\exp(-j2\pi ft)dt \\
r(t) &= \int_{-\infty}^{\infty} R(f,t)\exp(j2\pi ft)df.
\end{aligned}
$$

Example 2.1:
Consider an LTV channel with the impulse response given by

$$h(\tau, t) = 4\exp(-\tau/T)\cos(\Omega t), \quad \tau \geq 0,$$

where $T = 0.1$ ms and $\Omega = 10\pi$.

a. Find the channel time-variant transfer function $H(f, t)$.

b. Given that the transmitted signal is

$$x_1(t) = \begin{cases} 1, & |t| \leq T_0 \\ 0, & |t| > T_0 \end{cases},$$

where $T_0 = 0.025$ ms, find the received signal in the absence of background noise.

c. Repeat part (b) if the transmitted signal is

$$x_2(t) = x_1(t - T_1),$$

where $T_1 = 0.05$ ms.

d. What do you observe from the results of parts (b) and (c)?

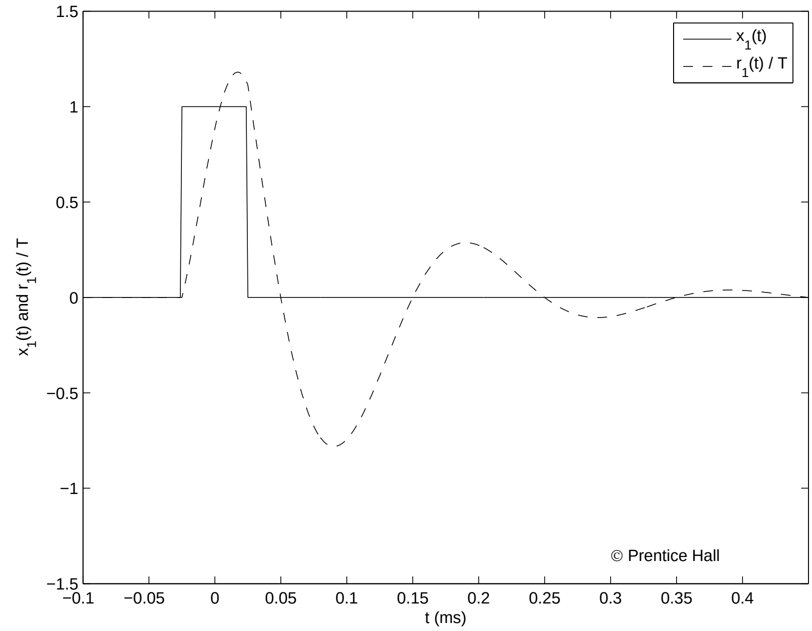
Solution:

a. The time-variant transfer function is

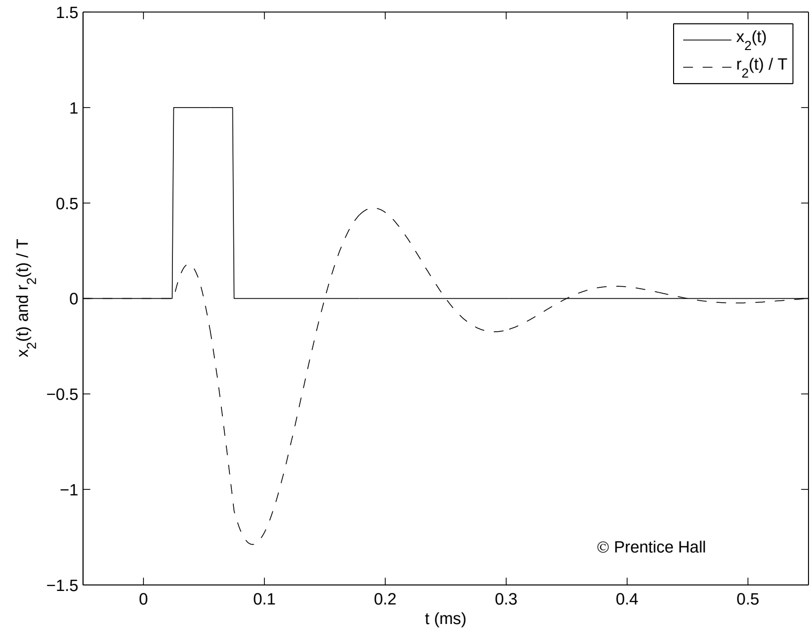
$$\begin{aligned} H(f,t) &= \mathcal{F}_\tau[h(\tau,t)] \\ &= \mathcal{F}_\tau[4\exp(-\tau/T)\cos(\Omega t)] \\ &= 4\cos(\Omega t)\mathcal{F}_\tau[\exp(-\tau/T)] \\ &= \frac{4T\cos(\Omega t)}{1 + j2\pi fT}. \end{aligned}$$

b. The received signal is calculated as follows:

$$\begin{aligned} r_1(t) &= \int_{-\infty}^{+\infty} h(\tau,t)x_1(t-\tau)d\tau \\ &= \int_0^{\infty} 4\exp(-\tau/T)\cos(\Omega t)x_1(t-\tau)d\tau \\ &= \begin{cases} 0, & t \leq -T_0 \\ 4T\cos(\Omega t)[1 - \exp(-\frac{t+T_0}{T})], & -T_0 < t < T_0 \\ 4T\cos(\Omega t)[\exp(-\frac{t-T_0}{T}) - \exp(-\frac{t+T_0}{T})], & t \geq T_0 \end{cases} \end{aligned}$$



(a)



(b)

Figure 8: The transmitted signals $x_i(t)$ and normalized received signals $r_i(t)/T$ (i = 1 and 2) in Example 2.1

The transmitted and received signals are plotted in Figure 8(a).

c. Similar to part (b), the received signal is calculated as follows:

$$
\begin{aligned}
r_2(t) &= \int_{-\infty}^{+\infty} h(\tau,t)x_2(t-\tau)d\tau \\
&= \int_0^{\infty} 4\exp(-\tau/T)\cos(\Omega t)x_1(t-T_1-\tau)d\tau \\
&= \begin{cases} 0, & t \le T_1 - T_0 \\ 4T\cos(\Omega t)[1-\exp(-\frac{t-T_1+T_0}{T})], & T_1 - T_0 < t < T_1 + T_0 \\ 4T\cos(\Omega t)[\exp(-\frac{t-T_1-T_0}{T}) - \exp(-\frac{t-T_1+T_0}{T})], & t \ge T_1 + T_0 \end{cases}
\end{aligned}
$$

The transmitted and received signals are plotted in Figure 8(b).

d. From Figure 8, it is observed that: (1) the received signals have a larger pulse width than the corresponding transmitted signals because the channel is time dispersive; and (2) even though the transmitted signal $x_2(t)$ is $x_1(t)$ delayed by T_1, the received signal $r_2(t)$ is not $r_1(t)$ delayed by T_1 because the channel is time varying.

□

2.3 Channel Correlation Functions

Assumptions:

(a) the channel impulse response $h(\tau, t)$ is a wide-sense stationary (WSS) process;

(b) the channel impulse responses at τ_1 and τ_2, $h(\tau_1, t)$ and $h(\tau_2, t)$, are uncorrelated if $\tau_1 \neq \tau_2$ for any t.

$\rightarrow$ wide-sense stationary uncorrelated scattering (WSSUS) channel

Delay power spectral density (psd)

The autocorrelation function of $h(\tau, t)$ is

$$\begin{aligned} \phi_h(\tau_1, \tau_2, \Delta t) &\stackrel{\triangle}{=} \frac{1}{2} E[h^*(\tau_1, t) h(\tau_2, t + \Delta t)] \\ &= \phi_h(\tau_1, \Delta t) \delta(\tau_1 - \tau_2) \end{aligned}$$

or equivalently,

$$\phi_h(\tau, \tau + \Delta\tau, \Delta t) = \phi_h(\tau, \Delta t) \delta(\Delta\tau), \tag{9}$$

where

$$\phi_h(\tau, \Delta t) = \int \phi_h(\tau, \tau + \Delta\tau, \Delta t) d\Delta\tau.$$

At $\Delta t = 0$, we define

$$\phi_h(\tau) \stackrel{\triangle}{=} \phi_h(\tau, 0). \tag{10}$$

From Eqs. (9) – (10), we have

$$\begin{aligned} \phi_h(\tau) &= \mathcal{F}_{\Delta\tau}[\phi_h(\tau, \tau + \Delta\tau, \Delta t)]|_{\Delta t = 0} \\ &= \mathcal{F}_{\Delta\tau}\{\frac{1}{2} E[h^*(\tau, t) h(\tau + \Delta\tau, t)]\}. \end{aligned}$$

$\phi_h(\tau)$ measures the average psd at the channel output as a function of the propagation delay, τ, and is therefore called the delay psd of the channel, also known as the multipath intensity profile. The nominal width of the delay psd pulse is called the multipath delay spread, denoted by T_m.

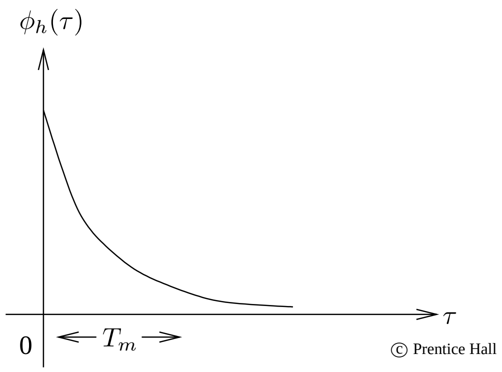


Figure 9: Delay power spectral density

The nth moment of the delays

$$\bar{\tau^n} = \frac{\int \tau^n \phi_h(\tau) d\tau}{\int \phi_h(\tau) d\tau}. \tag{11}$$

The mean propagation delay, or first moment, denoted by $\bar{\tau}$, is

$$\bar{\tau} = \frac{\int \tau \phi_h(\tau) d\tau}{\int \phi_h(\tau) d\tau} \tag{12}$$

and the rms (root-mean-square) delay spread, denoted by σ_τ, is

$$\sigma_\tau = \left[\frac{\int (\tau - \bar{\tau})^2 \phi_h(\tau) d\tau}{\int \phi_h(\tau) d\tau} \right]^{1/2}. \tag{13}$$

In calculating a value for the multipath delay spread, it is usually assumed that

$$T_m \approx \sigma_\tau.$$

Frequency and Time Correlation Functions

The autocorrelation function of $H(f,t)$is

$$
\begin{aligned}
\phi_H(f_1,f_2,t,\Delta t) &\triangleq \frac{1}{2}E[H^*(f_1,t)H(f_2,t+\Delta t)] \\
\stackrel{\text{WSS}}{\Longrightarrow}\ \phi_H(f_1,f_2,\Delta t) &= \frac{1}{2}E[H^*(f_1,t)H(f_2,t+\Delta t)] \\
\stackrel{\text{US}}{\Longrightarrow}\ \phi_H(f_1,f_2,\Delta t) &= \frac{1}{2}E\{[\int h(\tau_1,t)e^{-j2\pi f_1\tau_1}d\tau_1]^*[\int h(\tau_2,t+\Delta t)e^{-j2\pi f_2\tau_2}d\tau_2]\} \\
&= \int\int \frac{1}{2}E[h^*(\tau_1,t)h(\tau_2,t+\Delta t)]e^{-j2\pi(f_2\tau_2-f_1\tau_1)}d\tau_1 d\tau_2 \\
&\stackrel{\text{WSS}}{=} \int\int \phi_h(\tau_1,\tau_2,\Delta t)e^{-j2\pi(f_2\tau_2-f_1\tau_1)}d\tau_1 d\tau_2 \\
&\stackrel{\text{US}}{=} \int\int \phi_h(\tau_1,\Delta t)\delta(\tau_1-\tau_2)e^{-j2\pi(f_2\tau_2-f_1\tau_1)}d\tau_1 d\tau_2 \\
&= \int \phi_h(\tau_1,\Delta t)e^{-j2\pi(f_2-f_1)\tau_1}d\tau_1 \\
&= \int \phi_h(\tau,\Delta t)e^{-j2\pi(\Delta f)\tau}d\tau \quad (\Delta f = f_2 - f_1) \\
&\triangleq \phi_H(\Delta f,\Delta t) \quad \text{— time-frequency correlation function.}
\end{aligned}
$$

Frequency correlation function:

Let $\Delta t = 0$, we have

$$
\begin{aligned}
\phi_H(\Delta f) &\triangleq \frac{1}{2}E[H^*(f,t)H(f+\Delta f,t)] \\
&= \int \phi_h(\tau)e^{-j2\pi\Delta f\tau}d\tau.
\end{aligned}
$$

Note:

- $\phi_h(\tau)$ and $\phi_H(\Delta f)$ is a pair of Fourier transform.
- uncorrelated scattering $\Longrightarrow$WSS $H(f,t)$ w.r.t. f
- $\phi_H(\Delta f)$ characterizes the correlation of channel gains at f and $f+\Delta f$ for any t.

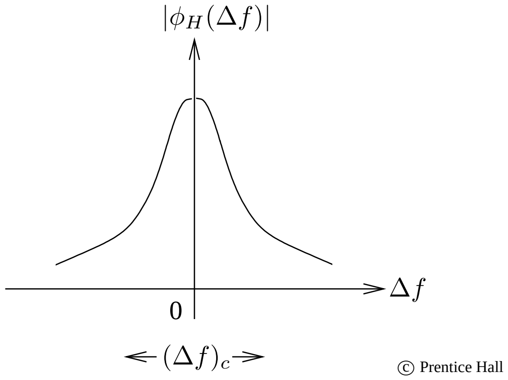


Figure 10: Frequency-correlation function and channel coherence bandwidth

$\phi_H(\Delta f)$ provides a measure of the "frequency coherence" of the channel.

- coherence bandwidth: recall the channel with 2 propagation paths

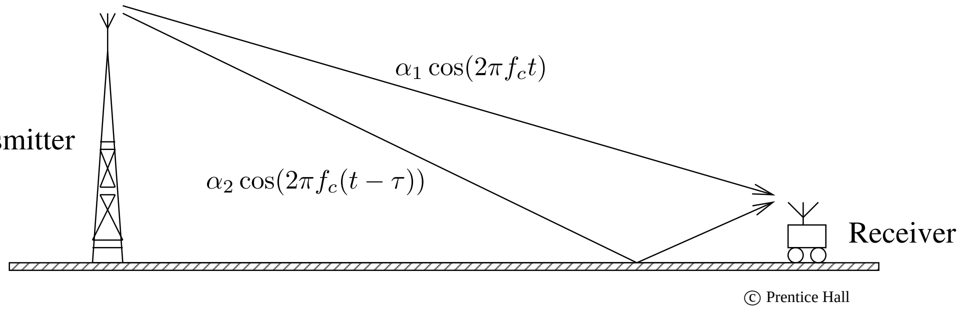


Figure 11: A channel with two propagation paths

The effect of the 2-path channel on the received signal depends on the signal frequency. With the same τ, two signals $A\cos(2\pi f_1 t)$ and $A\cos(2\pi f_2 t)$ will be affect differently.

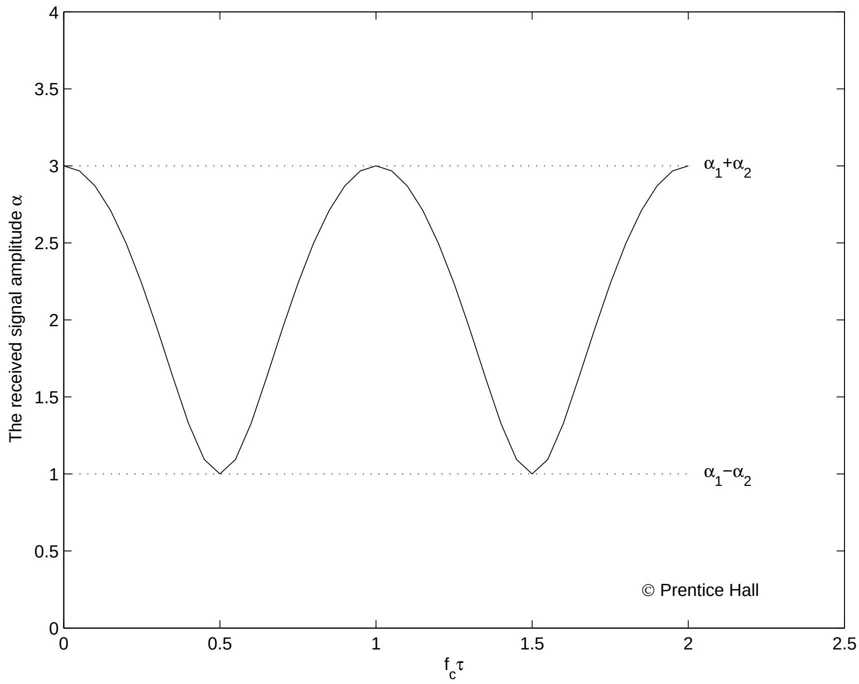


Figure 12: The amplitude fluctuation of the two-path channel with $\alpha_1 = 2$ and $\alpha_2 = 1$

$\Longrightarrow |f_1 - f_2| \uparrow \quad \Longrightarrow |r_1(t)|$ and $|r_2(t)|$ at any t will be uncorrelated.

Definition: The maximum frequency difference for which the signals are still strongly correlated is called the coherence bandwidth of the channel, denoted by $(\Delta f)_c$.

$\Longrightarrow$ Two sinusoids with frequency separation larger than $(\Delta f)_c$ are affected differently by the channel at any t.

- Let W_s denote the bandwidth of the transmitted signal.

  If $(\Delta f)_c < W_s$, the channel is said to exhibit frequency selective fading which introduces severe ISI to the received signal;

  If $(\Delta f)_c \gg W_s$, the channel is said to exhibit frequency nonselective fading or flat fading which introduces negligible ISI.

- $\phi_h(\tau) \leftrightarrow \phi_H(\Delta f) \quad \Longrightarrow \quad (\Delta f)_c \approx 1/T_m$.

Time correlation function $\phi_H(\Delta t)$:

Letting $\Delta f = 0$ in the time-frequency correlation function $\phi_H(\Delta f, \Delta t)$, we have

$$\phi_H(\Delta t) \stackrel{\triangle}{=} \phi_H(0, \Delta t) = \frac{1}{2} E[H^*(f,t)H(f,t+\Delta t)]. \tag{14}$$

- $\phi_H(\Delta t)$ characterizes, on average, how fast the channel transfer function changes with time at each frequency.
- The nominal width of $\phi_H(\Delta t)$, $(\Delta t)_c$, is called the coherence time of the fading channel.
- If the channel coherence time is much larger than the symbol interval of the transmitted signal, the channel exhibits slow fading.
- $\phi_H(\Delta t)$ is independent of f due to the US assumption $\Longrightarrow$ US in the time domain is equivalent to WSS in the frequency domain.

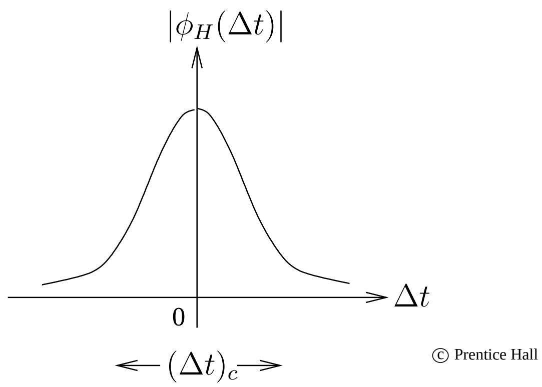


Figure 13: Time-correlation function and channel coherence time

Doppler Power Spectral Density

Doppler shifts:

An LTV channel introduces Doppler frequency shifts: Given the transmitted signal frequency f_c, the received signal frequency is $f_c+\nu(t)$, where $\nu(t)$ is the Doppler frequency shift and is given by

$$\nu(t) = \frac{V f_c}{c} \cos\theta(t).$$

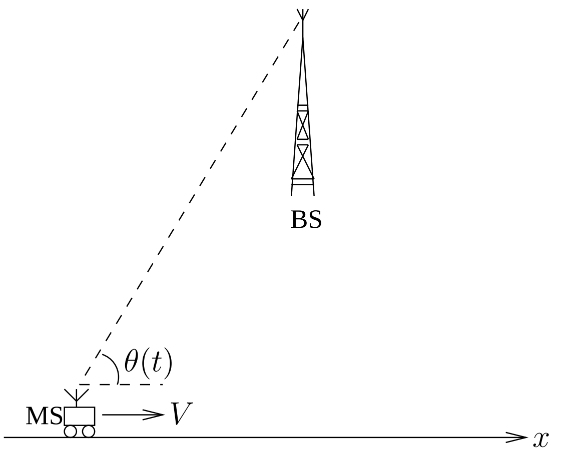


Figure 14: The Doppler effect

Consider the wireless channel with N distinct scatterers where the propagation delay can be approximated by its mean value $\bar{\tau}$.

$$h(\tau, t) \approx \sum_{n=1}^{N} \alpha_n(t) \exp[-j2\pi f_c \tau_n(t)]\delta(t-\bar{\tau}) \stackrel{\triangle}{=} Z(t)\delta(\tau-\bar{\tau}), \tag{15}$$

$$\begin{aligned} r(t) &= \int_{-\infty}^{+\infty} h(\tau, t)x(t-\tau)d\tau \\ &= \int_{-\infty}^{+\infty} [Z(t)\delta(\tau-\bar{\tau})]x(t-\tau)d\tau \\ &= Z(t)x(t-\bar{\tau}). \end{aligned}$$

In the frequency domain, the received signal is

$$\begin{aligned} R(f) &= \mathcal{F}[r(t)] \\ &= \mathcal{F}[Z(t)x(t-\bar{\tau})] \\ &= \mathcal{F}[Z(t)] \star \mathcal{F}[x(t-\bar{\tau})] \\ &= \mathcal{F}[Z(t)] \star [X(f)e^{-j2\pi f\bar{\tau}}], \end{aligned}$$

$\Longrightarrow$ The channel indeed broadens the transmitted signal spectrum by introducing new frequency components, a phenomenon referred to as frequency dispersion.

Doppler-spread function $H(f, \nu)$:

$H(f, \nu)$ is the channel gain associated with Doppler shift ν to the input signal component at frequency f.

$$R(f) = \int_{-\infty}^{+\infty} X(f - \nu)H(f - \nu, \nu)d\nu. \tag{16}$$

Relation between $H(f, t)$ and $H(f, \nu)$:

$$\begin{cases} H(f, \nu) = \mathcal{F}_t[H(f, t)] = \int_{-\infty}^{+\infty} H(f, t)e^{-j2\pi vt}dt \\ \\ H(f, t) = \mathcal{F}_\nu^{-1}[H(f, \nu)] = \int_{-\infty}^{+\infty} H(f, \nu)e^{+j2\pi\nu t}d\nu \end{cases}$$

$\Longrightarrow$ *being time-variant in the time domain* can be equivalently described by *having Doppler shifts in the frequency domain.*

Autocorrelation function of $H(f,\nu)$:

$$\begin{aligned}
\Phi_H(f_1,f_2,\nu_1,\nu_2) &\stackrel{\triangle}{=} \frac{1}{2}E[H^*(f_1,\nu_1)H(f_2,\nu_2)\\
&= \int\int \frac{1}{2}E[H^*(f_1,t_1)H(f_2,t_2)]e^{j2\pi\nu_1 t_1}e^{-j2\pi\nu_2 t_2}dt_1dt_2\\
&\stackrel{\text{WSSUS}}{=} \int\int \phi_H(\Delta f,\Delta t)e^{-j2\pi[\nu_2(t_1+\Delta t)-\nu_1 t_1]}d\Delta t dt_1\\
&\quad \text{(where } \Delta f = f_2 - f_1 \text{ and } \Delta t = t_2 - t_1\text{)}\\
&= \int \phi_H(\Delta f,\Delta t)e^{-j2\pi\nu_2\Delta t}d\Delta t \int e^{-j2\pi(\nu_2-\nu_1)t_1}dt_1\\
&= \Phi_H(\Delta f,\nu_2)\delta(\nu_2-\nu_1)
\end{aligned}$$

where

$$\Phi_H(\Delta f,\nu) = \int \phi_H(\Delta f,\Delta t)e^{-j2\pi\nu\Delta t}d\Delta t$$

is the Fourier transform of $\phi_H(\Delta f,\Delta t)$ w.r.t. Δt.

At $\Delta f = 0$,

$$\Phi_H(\nu) \stackrel{\triangle}{=} \Phi_H(0,\nu) = \int_{-\infty}^{\infty} \phi_H(\Delta t)e^{-j2\pi\nu\Delta t}d\Delta t.$$

• $\Phi_H(\nu)$ is the Fourier transform of the channel correlation function $\phi_H(\Delta t)$.

$\Longrightarrow$ $\Phi_H(\nu)$ is psd as a function of the Doppler shift ν.

$\Longrightarrow$ $\Phi_H(\nu)$ is called *Doppler power spectral density* function.

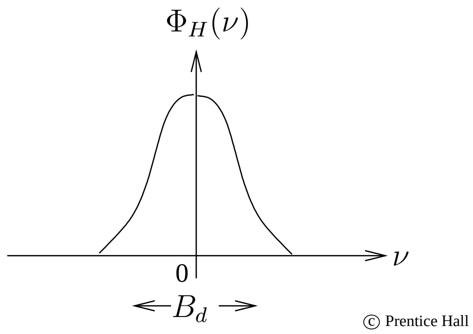


Figure 15: Doppler power spectral density and Doppler spread

- The nominal width of the Doppler psd, B_d, is called the Doppler spread.

- Since $\phi_H(\Delta t) \leftrightarrow \Phi_H(\nu)$,

$$(\Delta t)_c \approx \frac{1}{B_d}.$$

- The mean Doppler shift is

$$\bar{\nu} = \frac{\int \nu \Phi_H(\nu) d\nu}{\int \Phi_H(\nu) d\nu}$$

and the rms Doppler spread is

$$\sigma_\nu = [\frac{\int (\nu - \bar{\nu})^2 \Phi_H(\nu) d\nu}{\int \Phi_H(\nu) d\nu}]^{1/2}.$$

As an approximation, it is usually assumed that

$$B_d \approx \sigma_\nu.$$

Example 2.3 Delay psd and frequency-correlation function
Consider a WSSUS channel whose time-variant impulse response is given by

$$h(\tau, t) = \exp(-\tau/T)n(\tau)\cos(\Omega t + \Theta), \ \ \tau \geq 0,$$

where T and Ω are constants, Θ is a random variable uniformly distributed in $[-\pi, \ +\pi]$, and $n(\tau)$ is a random process independent of Θ, with $E[n(\tau)] = \mu_n$ and $E[n(\tau_1)n(\tau_2)] = \delta(\tau_1 - \tau_2)$.

a. Calculate the delay psd and the multipath delay spread.

b. Calculate the frequency correlation function and the channel coherence bandwidth.

c. Determine whether the channel exhibits frequency-selective fading for GSM systems with $T = 0.1$ ms.

Solution:

a. From Eq. (11), we have

$$\begin{aligned} \phi_h(\tau) &= \mathcal{F}_{\Delta\tau}\{\frac{1}{2}E[h^*(\tau,t)h(\tau+\Delta\tau,t]\} \\ &= \mathcal{F}_{\Delta\tau}\{\frac{1}{2}E[n(\tau)n(\tau+\Delta\tau)]E[e^{-(2\tau+\Delta\tau)/T}\cos^2(\Omega t+\Theta)]\}, \quad \tau \geq 0 \\ &= \mathcal{F}_{\Delta\tau}\{\frac{1}{4}e^{-(2\tau+\Delta\tau)/T}\delta(\Delta\tau)E[1+\cos(2\Omega t+2\Theta)]\}, \quad \tau \geq 0 \\ &= \frac{1}{4}e^{-2\tau/T}, \ \ \tau \geq 0, \end{aligned}$$

where

$$E[\cos(2\Omega t + 2\Theta)] = \int_{-\pi}^{+\pi} \cos(2\Omega t + 2\theta)\frac{1}{2\pi}d\theta = 0.$$

For $\tau < 0$, $\phi_h(\tau) = 0$.

The mean propagation delay is

$$\bar{\tau} = \frac{\int_0^\infty \tau \phi_h(\tau) d\tau}{\int_0^\infty \phi_h(\tau) d\tau} = \frac{\int_0^\infty \tau \frac{1}{4} e^{-2\tau/T} d\tau}{\int_0^\infty \frac{1}{4} e^{-2\tau/T} d\tau} = \frac{T}{2}$$

and the multipath delay spread is

$$T_m \approx \sigma_\tau = [\frac{\int (\tau - \bar{\tau})^2 \phi_h(\tau) d\tau}{\int \phi_h(\tau) d\tau}]^{1/2} = [\frac{\int \tau^2 \phi_h(\tau) d\tau}{\int \phi_h(\tau) d\tau} - \bar{\tau}^2]^{1/2} = \frac{T}{2}.$$

b. The frequency correlation function is

$$\begin{aligned} \phi_H(\Delta f) &= \mathcal{F}[\phi_h(\tau)] \\ &= \int_0^\infty \frac{1}{4} e^{-2\tau/T} e^{-j2\pi(\Delta f)\tau} d\tau \\ &= \frac{T}{8 + j8\pi T(\Delta f)}. \end{aligned}$$

The coherence bandwidth is

$$(\Delta f)_c \approx 1/T_m = \frac{2}{T}.$$

c. With $T = 0.1$ ms, we have $(\Delta f)_c = 20$ kHz. The GSM channels have a bandwidth of 200 kHz. Since $(\Delta f)_c \ll 200$ kHz, the channel fading is frequency selective.

$\square$

Example 2.4 Doppler psd
For the channel specified in Example 2.3 with $\Omega = 10\pi$, find

a. the Doppler psd,

b. the mean Doppler shift and the rms Doppler spread,

c. the channel coherence time, and

d. whether the channel exhibits slow fading for GSM systems.

Solution:

a. The Doppler psd can be calculated by taking the Fourier transform of the time correlation function $\phi_H(\Delta t)$. In this way, we need to calculate the correlation function $\phi_h(\tau, \Delta t)$ first. For the WSS channel, we have for $\tau \geq 0$

$$\begin{aligned}
& \phi_h(\tau, \Delta t) \\
= \ & \mathcal{F}_{\Delta\tau}\{\frac{1}{2}E[h^*(\tau, t)h(\tau + \Delta\tau, t + \Delta t)]\} \\
= \ & \mathcal{F}_{\Delta\tau}\{\frac{1}{2}E[e^{-\tau/T}n(\tau)\cos(\Omega t + \Theta) \cdot e^{-(\tau+\Delta\tau)/T}n(\tau + \Delta\tau)\cos(\Omega t + \Omega\Delta t + \Theta)]\} \\
= \ & \mathcal{F}_{\Delta\tau}\{\frac{1}{4}e^{-(2\tau+\Delta\tau)/T}E[n(\tau)n(\tau + \Delta\tau)]E[\cos(\Omega\Delta t) + \cos(2\Omega t + \Omega\Delta t + 2\Theta)]\} \\
= \ & \mathcal{F}_{\Delta\tau}\{\frac{1}{4}e^{-(2\tau+\Delta\tau)/T}\delta(\Delta\tau)\cos(\Omega\Delta t)\} \\
= \ & \frac{1}{4}e^{-2\tau/T}\cos(\Omega\Delta t), \ \ \tau \geq 0.
\end{aligned}$$

The time correlation function is then

$$\begin{aligned}
\phi_H(\Delta t) &= \phi_H(\Delta f, \Delta t)|_{\Delta f=0} \\
&= \int_{-\infty}^{+\infty} \phi_h(\tau, \Delta t) d\tau \\
&= \frac{1}{4}\cos(\Omega\Delta t)\int_0^{+\infty} e^{-2\tau/T} d\tau \\
&= \frac{T}{8}\cos(\Omega\Delta t).
\end{aligned}$$

The Doppler psd is

$$\begin{aligned}
\Phi_H(\nu) &= \mathcal{F}[\phi_H(\Delta t)] \\
&= \mathcal{F}[\frac{T}{8}\cos(\Omega\Delta t)] \\
&= \frac{T}{16}[\delta(2\pi\nu - \Omega) + \delta(2\pi\nu + \Omega)].
\end{aligned}$$

That is, the channel introduces two Doppler shifts, $\pm\Omega/2\pi = \pm 5$ Hz, with equal psd.

b. The mean Doppler shift is zero as the two Doppler shifts are negative of each other and have the same psd. The rms Doppler spread is

$$\begin{aligned}
\sigma_\nu &= \{\frac{\int_{-\infty}^{+\infty} \nu^2 \cdot \frac{T}{16}[\delta(2\pi\nu - \Omega) + \delta(2\pi\nu + \Omega)]d\nu}{\int_{-\infty}^{+\infty} \frac{T}{16}[\delta(2\pi\nu - \Omega) + \delta(2\pi\nu + \Omega)]d\nu}\}^{1/2} \\
&= \{\frac{\frac{T}{16}[(\frac{\Omega}{2\pi})^2 + (-\frac{\Omega}{2\pi})^2]}{\frac{T}{16}[1+1]}\}^{1/2} \\
&= \frac{\Omega}{2\pi},
\end{aligned}$$

which is 5 Hz.

c. The coherence time is

$$(\Delta t)_c \approx \frac{1}{\sigma_\nu} = 0.2 \text{ s}.$$

d. In GSM systems, the data rate $R_s = 270.833$ kbps, which corresponds to a symbol interval

$$T_s = \frac{1}{R_s} \approx 3.7 \times 10^{-6} \text{ s}.$$

Since $T_s \ll (\Delta t)_c$, the channel exhibits slow fading.

□

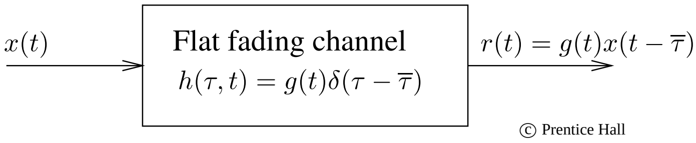
2.4 Large-Scale Path Loss and Shadowing

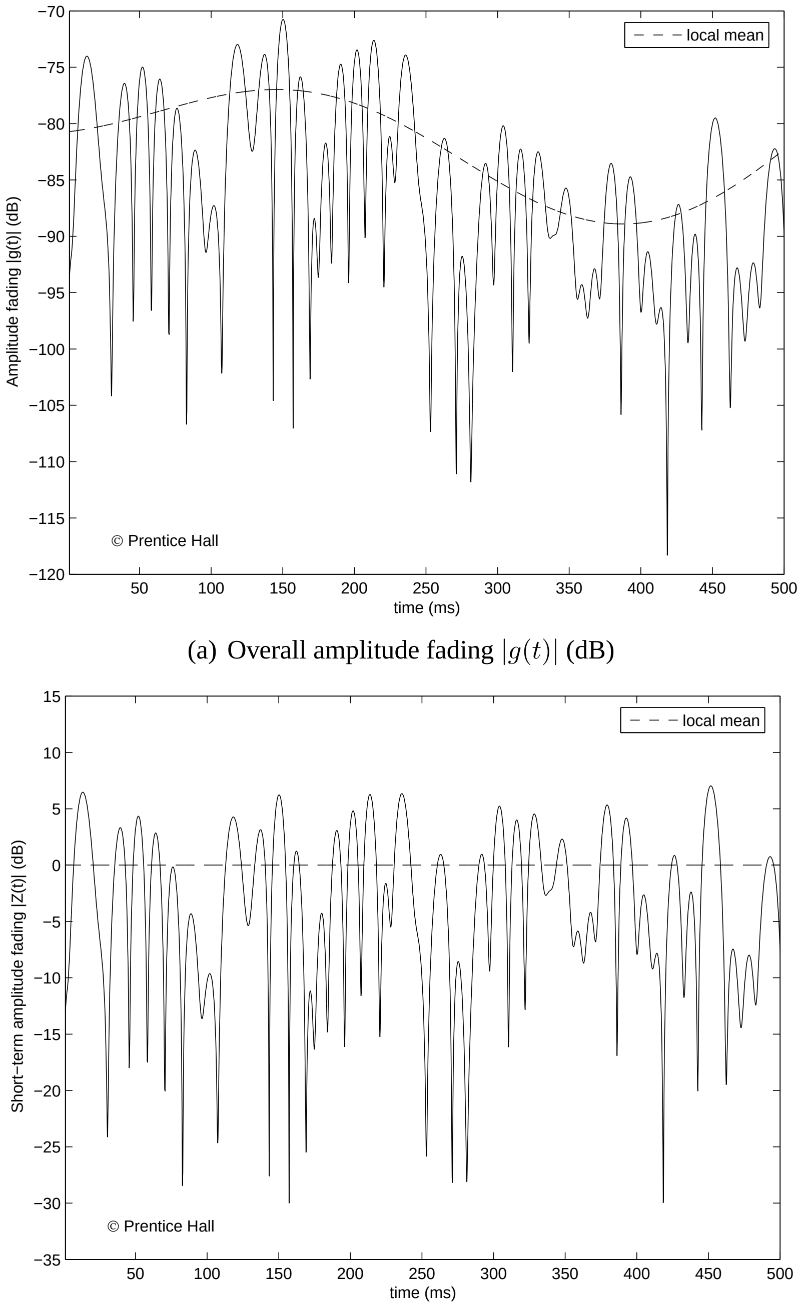
Consider a flat fading channel with the channel impulse response

$$h(\tau, t) \approx h(\bar{\tau}, t) \stackrel{\triangle}{=} g(t)\delta(\tau - \bar{\tau}).$$

The received signal is

$$r(t) = \int_{-\infty}^{+\infty} h(\tau, t)x(t - \tau)d\tau \approx g(t)x(t - \bar{\tau}).$$





(a) Overall amplitude fading $|g(t)|$ (dB)

(b) Short-term amplitude fading $|Z(t)|$ (dB)

Figure 16: Representation of long-term and short-term fading components

Free Space Propagation Model

When the distance between the transmitting antenna and receiving antenna is much larger than the wavelength of the transmitted wave and the largest physical linear dimension of the antennas, the power P_r at the output of the receiving antenna is given by

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

where

P_t	=	total power radiated by an isotropic source,
G_t	=	transmitting antenna gain,
G_r	=	receiving antenna gain,
d	=	distance between transmitting and receiving antennas,
λ	=	wavelength of the carrier signal = c/f_c,
c	=	3×10^8 m/s (velocity of light),
f_c	=	carrier frequency, and
$P_t G_t$	$\triangleq$	effective isotropically radiated power (EIRP).

The term $(4\pi d/\lambda)^2$ is known as the *free-space path loss* denoted by $L_p(d)$, which is

$$\begin{aligned} L_p(d) &= \frac{\text{EIRP} \times \text{Receiving antenna gain}}{\text{Received power}} \\ &= -10\log_{10}\left[\frac{\lambda^2}{(4\pi d)^2}\right] \quad \text{(dB)} \\ &= -20\log_{10}\left(\frac{c/f_c}{4\pi d}\right) \quad \text{(dB)}. \end{aligned}$$

In other words, the path loss is

$$L_p(d) = 20\log_{10} f_c + 20\log_{10} d - 147.56 \quad \text{(dB)}.$$

Note that the free-space path loss increases by 6 dB for every doubling of the distance and also for every doubling of the radio frequency.

Log-Distance Path Loss with Shadowing

Let $\bar{L}_p(d)$ denote the log-distance path loss. Then,

$$\bar{L}_p(d) \ \propto \ (\frac{d}{d_0})^{\kappa}, \quad d \geq d_0$$

or equivalently,

$$\bar{L}_p(d) = \bar{L}_p(d_0) + 10\kappa \log_{10}(\frac{d}{d_0}) \ \text{ dB}, \quad d \geq d_0$$

- d_0: 1 km for macrocells, 100 m for outdoor microcells, and 1 m for indoor picocells

Table 1: Path loss exponents for different environments

Environment	Path Loss Exponent, κ
free space	2
urban cellular radio	2.7 to 3.5
shadowed urban cellular radio	3 to 5
in building with line of sight	1.6 to 1.8
obstructed in building	4 to 6

• *Shadowing*: As the mobile moves in uneven terrain, it often travels into a propagation shadow behind a building or a hill or other obstacle much larger than the wavelength of the transmitted signal, and the associated received signal level is attenuated significantly.

• A log-normal distribution is a popular model for characterizing the shadowing process.

Let $\epsilon_{(\text{dB})}$ be a zero-mean Gaussian distributed random variable (in dB) with standard deviation σ_ϵ (in dB). The pdf of $\epsilon_{(\text{dB})}$ is given by

$$f_{\epsilon(\text{dB})}(x) = \frac{1}{\sqrt{2\pi}\sigma_\epsilon}\exp(-\frac{x^2}{2\sigma_\epsilon^2}).$$

Let $L_p(d)$ denote the overall long-term fading (in dB). Then,

$$\begin{aligned} L_p(d) &= \bar{L}_p(d) + \epsilon_{(\text{dB})} \\ &= \bar{L}_p(d_0) + 10\kappa\log_{10}(\tfrac{d}{d_0}) + \epsilon_{(\text{dB})} \quad \text{(dB)}. \end{aligned}$$

• $\epsilon_{(\text{dB})}$ follows the Gaussian (normal) distribution $\Longrightarrow$ ϵ in linear scale is said to follow a log-normal distribution with pdf given by

$$f_\epsilon(y) = \frac{20/\ln 10}{\sqrt{2\pi}y\sigma_\epsilon}\exp[-\frac{(20\log_{10} y)^2}{2\sigma_\epsilon^2}].$$

• σ_ϵ: 8 dB for an outdoor cellular system and 5 dB for an indoor environment.

2.5 Small-Scale Multipath Fading

Consider a flat fading channel with N distinct scatterers:

$$\begin{aligned} r(t) &= \sum_{n=1}^{N} \alpha_n(t) e^{-j2\pi f_c \tau_n(t)} x(t - \tau_n(t)) \\ &\approx [\sum_{n=1}^{N} \alpha_n(t) e^{-j2\pi f_c \tau_n(t)}] x(t - \bar{\tau}). \end{aligned}$$

The complex gain of the channel is

$$\begin{aligned} Z(t) &= \sum_{n=1}^{N} \alpha_n(t) e^{-j2\pi f_c \tau_n(t)} \\ &= Z_c(t) - jZ_s(t) \end{aligned}$$

where

$$\begin{aligned} Z_c(t) &= \sum_{n=1}^{N} \alpha_n(t) \cos\theta_n(t) \\ Z_s(t) &= \sum_{n=1}^{N} \alpha_n(t) \sin\theta_n(t) \end{aligned}$$

and $\theta_n(t) = 2\pi f_c \tau_n(t)$.

Also,

$$Z(t) = \alpha(t) \exp[j\theta(t)]$$

where

$$\alpha(t) = \sqrt{Z_c^2(t) + Z_s^2(t)}, \quad \theta(t) = \tan^{-1}[Z_s(t)/Z_c(t)].$$

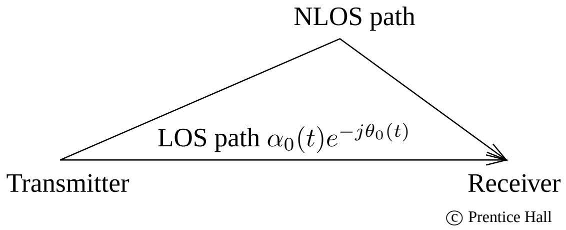


Figure 17: NLOS versus LOS scattering

Rayleigh fading (NLOS propagation)

$$E[Z_c(t)] = E[Z_s(t)] = 0. \tag{17}$$

Assume that, at any time t, for $n = 1, 2, \ldots, N$,

a. the values of $\theta_n(t)$ are statistically independent, each being uniformly distributed over $[0,\ 2\pi]$;

b. the values of $\alpha_n(t)$ are identically distributed random variables, independent of each other and of the $\theta_n(t)$'s.

$\Longrightarrow$

According to the central limit theorem, $Z_c(t)$ and $Z_s(t)$ are approximately Gaussian random variables at any time t if N is sufficiently large.

Z_c and Z_s are independent Gaussian random variables with zero mean and equal variance $\sigma_z^2 = \frac{1}{2}\sum_{n=1}^{N} E[\alpha_n^2]$.

$\Longrightarrow$

$$f_{Z_cZ_s}(x,y) = \frac{1}{2\pi\sigma_z^2}\exp[-\frac{x^2+y^2}{2\sigma_z^2}], \quad -\infty < x < \infty, \ -\infty < y < \infty.$$

$\Longrightarrow$ The amplitude fading, α, follows a Rayleigh distribution with parameter σ_z^2,

$$f_\alpha(x) = \begin{cases} \frac{x}{\sigma_z^2}\exp(-\frac{x^2}{2\sigma_z^2}), & x \geq 0 \\ 0, & x < 0 \end{cases}, \tag{18}$$

with $E[\alpha] = \sigma_z\sqrt{\pi/2}$ and $E(\alpha^2) = 2\sigma_z^2$;

The phase distortion follows the uniform distribution over $[0,\ 2\pi]$,

$$f_\theta(x) = \begin{cases} \frac{1}{2\pi}, & 0 \leq x \leq 2\pi \\ 0, & \text{elsewhere} \end{cases};$$

The amplitude fading α and the phase distortion θ are independent.

Rician Fading (LOS propagation)

$$Z(t) = Z_c(t) - jZ_s(t) + \Gamma(t),$$

where $\Gamma(t) = \alpha_0(t)e^{-j\theta_0(t)}$ is the deterministic LOS component.

$$E[Z(t)] = \Gamma(t) \neq 0.$$

The distribution of the envelope at any time t is given by the Rayleigh distribution modified by

a. a factor containing a non-centrality parameter, and

b. a zero-order modified Bessel function of the first kind.

The resultant pdf for the amplitude fading at any t, α, is known as the Rician distribution, given by

$$\begin{aligned} f_\alpha(x) &= \underbrace{\frac{x}{\sigma_z^2}\exp(-\frac{x^2}{2\sigma_z^2})}_{\text{Rayleigh}} \cdot \underbrace{\exp\{-\frac{\alpha_0^2}{2\sigma_z^2}\} \cdot I_0(\frac{\alpha_0 x}{\sigma_z^2})}_{\text{modifier}} \\ &= \frac{x}{\sigma_z^2}\exp(-\frac{x^2+\alpha_0^2}{2\sigma_z^2})I_0(\frac{\alpha_0 x}{\sigma_z^2}), \quad x \geq 0, \end{aligned}$$

where α_0 is $\alpha_0(t)$ at any t, α_0^2 is the power of the LOS component and is the non-centrality parameter, $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind and is given by

$$I_0(x) = \frac{1}{2\pi}\int_0^{2\pi} \exp(x\cos\theta)d\theta.$$

- The K factor:

$$K \triangleq \frac{\text{Power of the LOS component}}{\text{Total power of all other scatterers}} = \frac{\alpha_0^2}{2\sigma_z^2}.$$

$K \to 0 \Longrightarrow$ the Rician distribution $\to$ the Rayleigh distribution;

$K \to \infty \Longrightarrow$ the Rician fading channel $\to$ an AWGN channel.

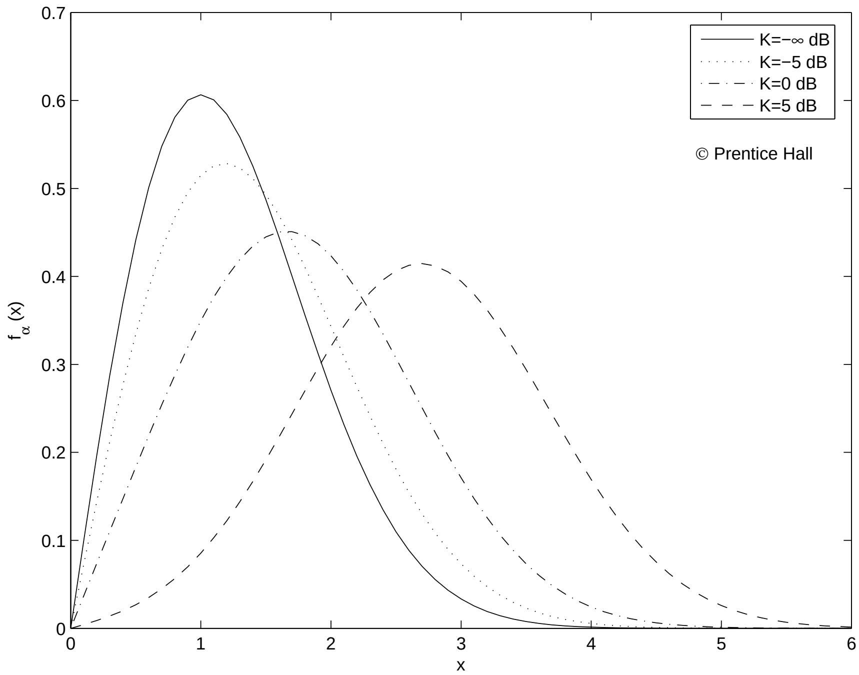


Figure 18: Rayleigh and Rician fading distributions with $\sigma_z = 1$

Second-Order Statistics - LCR and AFD

Level Crossing Rate (LCR):

The crossing rate at level R of a flat fading channel is the expected number of times that the channel amplitude fading level, $\alpha(t)$, crosses the specified level R, with a positive slope, divided by the observation time interval.

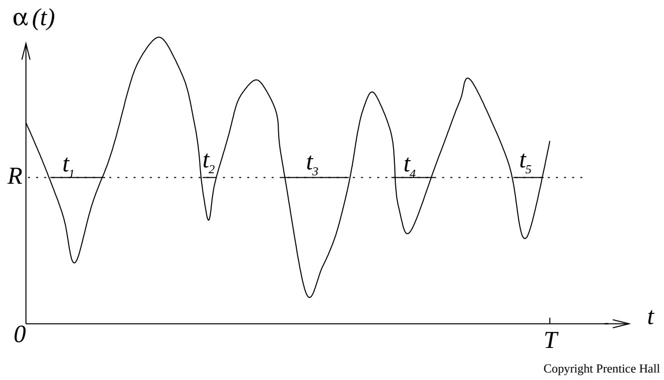


Figure 19: Level crossing rate and average duration of fades

$$N_R = \text{E[upward crossing rate at level } R].$$

Let $\dot{\alpha} = d\alpha(t)/dt$ denote the amplitude fading rate and let $f_{\alpha\dot{\alpha}}(x, y)$ denote the joint pdf of the amplitude fading $\alpha(t)$ and its derivative $\dot{\alpha}(t)$ at any time t. Then,

$$N_R = \int_0^\infty y f_{\alpha\dot{\alpha}}(x, y)|_{x=R} dy.$$

For the Rayleigh fading environment,

$$f_{\alpha\dot{\alpha}}(x, y) = \frac{x}{\sqrt{2\pi\sigma_{\dot{\alpha}}^2\sigma_z^2}} \exp[-\frac{1}{2}(\frac{x^2}{\sigma_z^2} + \frac{y^2}{\sigma_{\dot{\alpha}}^2})], \quad x \geq 0, \; -\infty < y < \infty,$$

where

$$\sigma_{\dot{\alpha}}^2 = \frac{1}{2}(2\pi\nu_m)^2\sigma_z^2$$

and ν_m is the maximum Doppler shift. The LCR is

$$\begin{aligned} N_R &= \int_0^\infty y \cdot \frac{R}{\sqrt{2\pi\sigma_{\dot{\alpha}}^2\sigma_z^2}} \exp[-\frac{1}{2}(\frac{R^2}{\sigma_z^2} + \frac{y^2}{\sigma_{\dot{\alpha}}^2})] dy \\ &= \sqrt{2\pi}\nu_m(\frac{R}{\sqrt{2}\sigma_z}) \exp(-\frac{R^2}{2\sigma_z^2}). \end{aligned}$$

Letting

$$\rho = \frac{R}{\sqrt{2}\sigma_z}$$

we have

$$N_R = \sqrt{2\pi}\nu_m\rho \exp(-\rho^2).$$

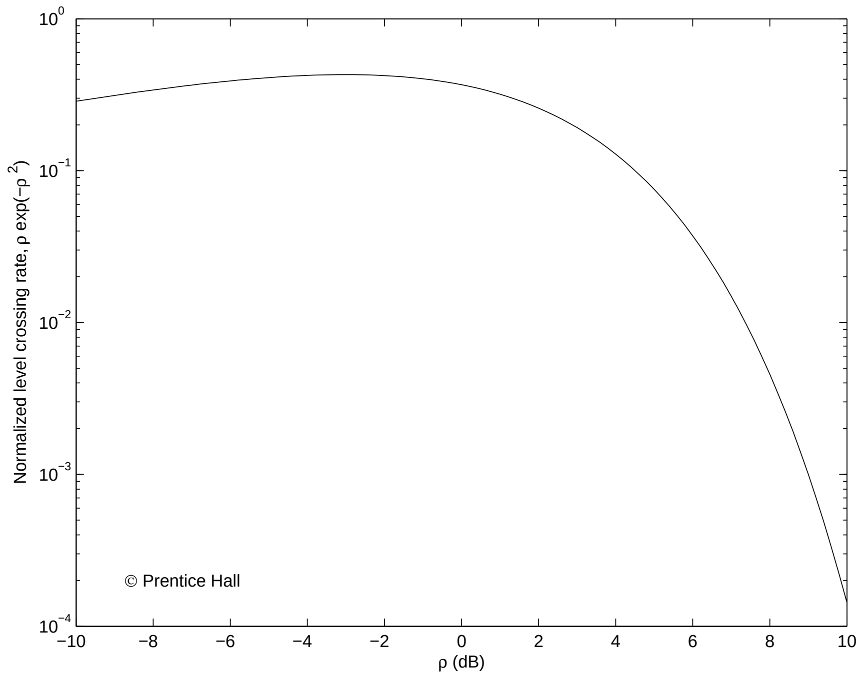


Figure 20: The normalized level crossing rate of the flat Rayleigh fading channel

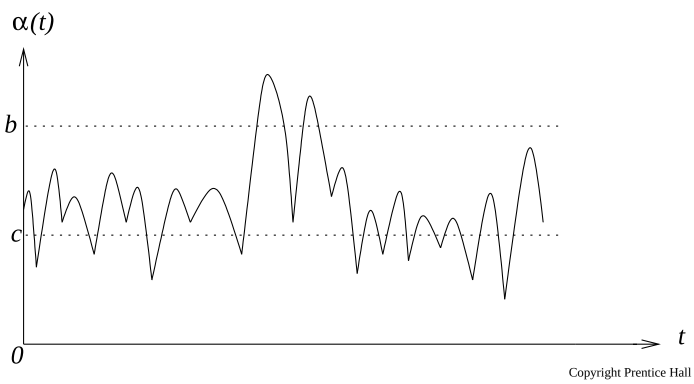


Figure 21: An example of amplitude fading level versus time

Average Fade Duration (AFD)

> The average fade duration at level R is the average period of time for which the channel amplitude fading level is below the specified threshold R during each fade period.

Let χ_R denote the AFD. It is a statistic closely related to the LCR. Mathematically, the AFD can be represented as

$$\begin{aligned}\chi_R \;=\; & \text{E[the period that the amplitude fading level stays below the threshold } R \\ & \quad \text{in each upward crossing].}\end{aligned}$$

$\Longrightarrow$

$$\begin{aligned}N_R \cdot \chi_R &= \lim_{T\to\infty} \frac{M_T}{T} \cdot \frac{\sum_{i=1}^{M_T} t_i}{M_T} \\ &= \lim_{T\to\infty} \frac{\sum_{i=1}^{M_T} t_i}{T} \\ &= P(\alpha \le R).\end{aligned}$$

For the Rayleigh fading environment, the cdf of α is

$$P(\alpha \le x) = \int_0^x f_\alpha(y)dy = 1 - \exp(-\frac{x^2}{2\sigma_z^2}).$$

The corresponding AFD is

$$\begin{aligned}\chi_R &= \frac{P(A \le R)}{N_R} \\ &= \frac{1 - \exp(-R^2/2\sigma_z^2)}{\sqrt{2\pi}\nu_m(R/\sqrt{2}\sigma_z)\exp(-R^2/2\sigma_z^2)} \\ &= \frac{\exp(\rho^2) - 1}{\sqrt{2\pi}\nu_m\rho}.\end{aligned}$$

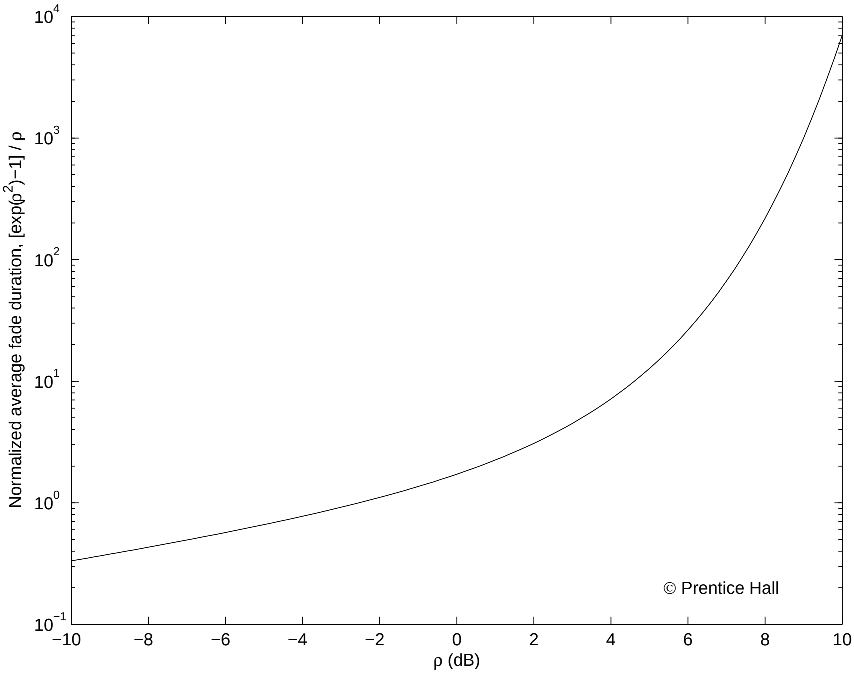


Figure 22: The normalized average fade duration of the flat Rayleigh fading channel

Example 2.8 The LCR N_R and AFD χ_R

Consider a mobile cellular system in which the carrier frequency is $f_c = 900$ MHz and the mobile travels at a speed of 24 km/h. Calculate the AFD and LCR at the normalized level $\rho = 0.294$.

Solution:

At $f_c = 900\ MHz$, the wavelength is $\lambda = \frac{c}{f_c} = \frac{3\times10^8}{900\times10^6} = \frac{1}{3}$ m. The velocity of the mobile is $V = 24$ km/h = 6.67 m/s. The maximum Doppler frequency is $\nu_m = V/\lambda = \frac{6.67}{1/3} = 20$ Hz. The average duration of fades below the normalized level $\rho = 0.294$ is

$$\chi_R = \frac{e^{\rho^2} - 1}{\sqrt{2\pi}\nu_m \rho} = \frac{e^{(0.294)^2} - 1}{\sqrt{2\pi} \times 20 \times 0.294} = 0.0061 \text{ s.}$$

The level crossing rate at $\rho = 0.294$ is

$$\begin{aligned} N_R &= \sqrt{2\pi}\nu_m \rho e^{-\rho^2} \\ &= \sqrt{2\pi} \times 20 \times 0.294 e^{-(0.294)^2} \\ &= 16 \text{ upcrossings/second.} \end{aligned}$$

□

Ch.3. Bandpass Transmission for Mobile Radio

3.1 Review of digital modulation

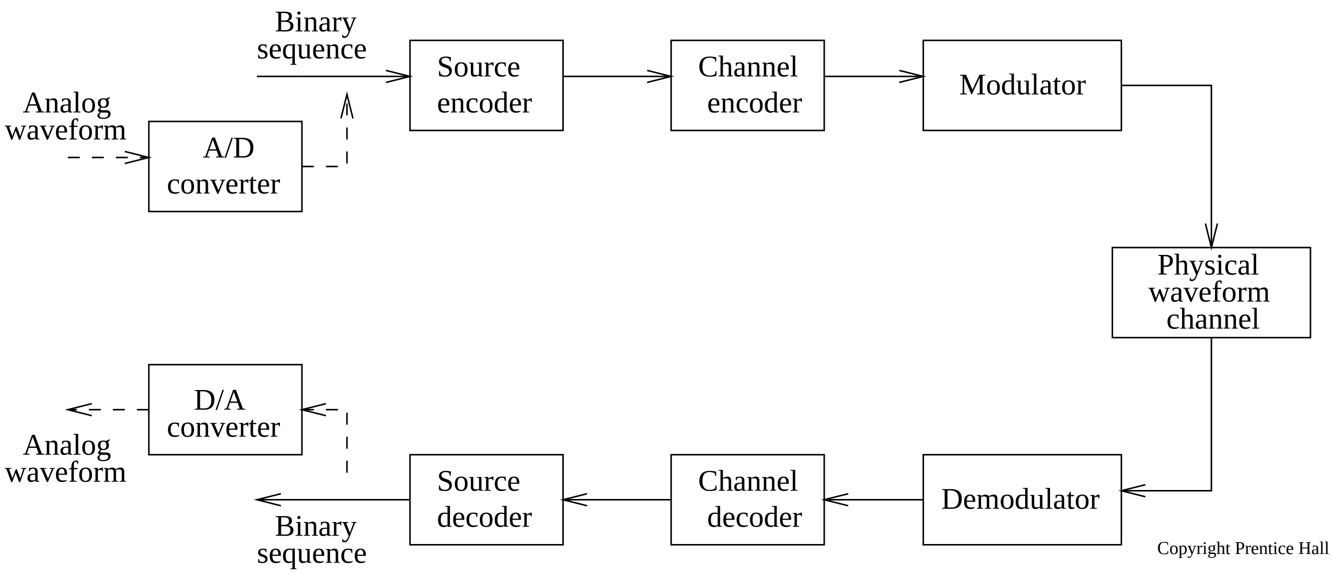


Figure 1: Functional block diagram of a binary digital communications system

Review:

- Bandpass digital transmission
- Signal space and decision regions
- Signal detection and optimal receiver (matched filter receiver)
- Power spectral density (psd) of the transmitted signal
- Probability of transmission error

Review: Quadrature phase shift keying (QPSK)

Modulation:

$$x_m(t) = \sqrt{\frac{2E_s}{T_s}}\cos(2\pi f_c t + \frac{2\pi(m-1)}{M} + \zeta_0), \quad 0 \le t \le T_s, \quad m = 1, 2, \cdots, M,$$

where $M = 4$, $\zeta_0 = 0$ or $\pi/4$, $E_s = \int_{-\infty}^{\infty} x_m^2(t)dt$; $E_s = lE_b$, $T_s = lT_b$.

For $\zeta_0 = \pi/4$, the modulated signal can be represented as

$$\begin{aligned} x_m(t) &= \sqrt{\frac{2E_s}{T_s}}\cos[(2m-1)\frac{\pi}{4}]\cos(2\pi f_c t) \\ &\quad - \sqrt{\frac{2E_s}{T_s}}\sin[(2m-1)\frac{\pi}{4}]\sin(2\pi f_c t), \quad 0 \le t \le T_s, \quad m = 1, 2, 3, 4. \end{aligned}$$

Letting

$$\begin{cases} \varphi_1(t) = \sqrt{\frac{2}{T_s}}\cos(2\pi f_c t), & 0 \le t \le T_s \\ \varphi_2(t) = \sqrt{\frac{2}{T_s}}\sin(2\pi f_c t), & 0 \le t \le T_s \end{cases},$$

we have

$$x_m(t) = \sqrt{E_s}\cos[(2m-1)\frac{\pi}{4}]\varphi_1(t) - \sqrt{E_s}\sin[(2m-1)\frac{\pi}{4}]\varphi_2(t).$$

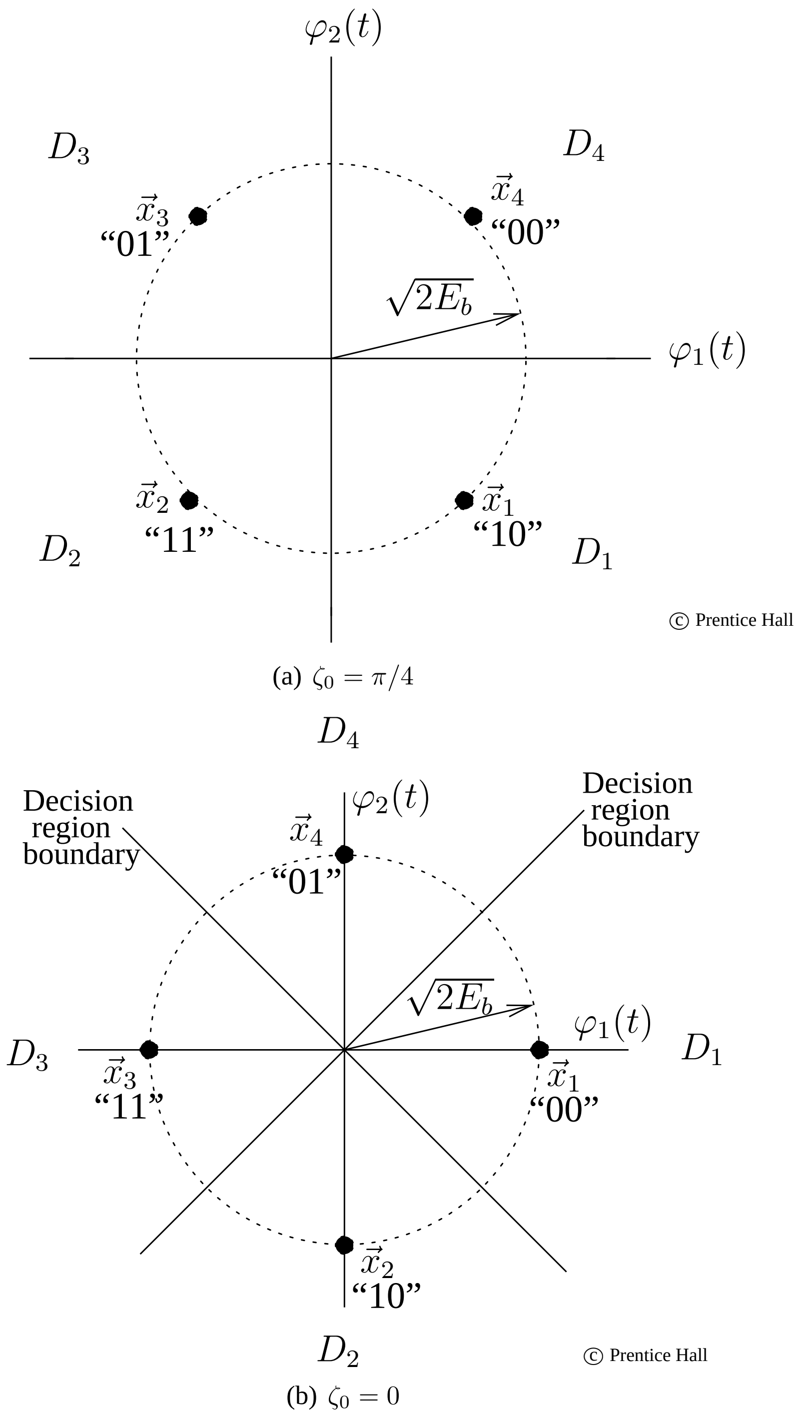
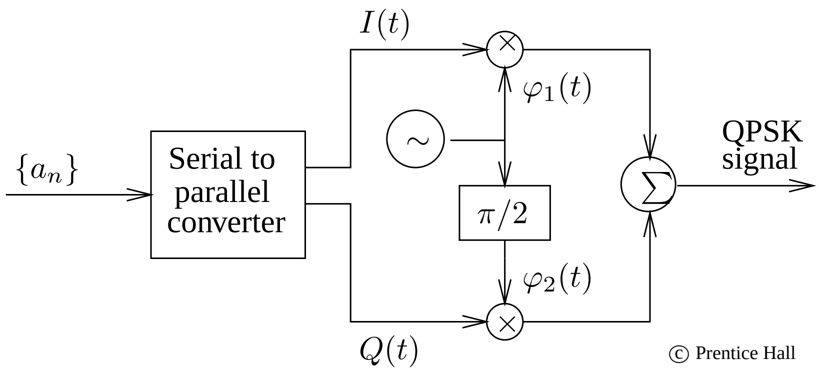
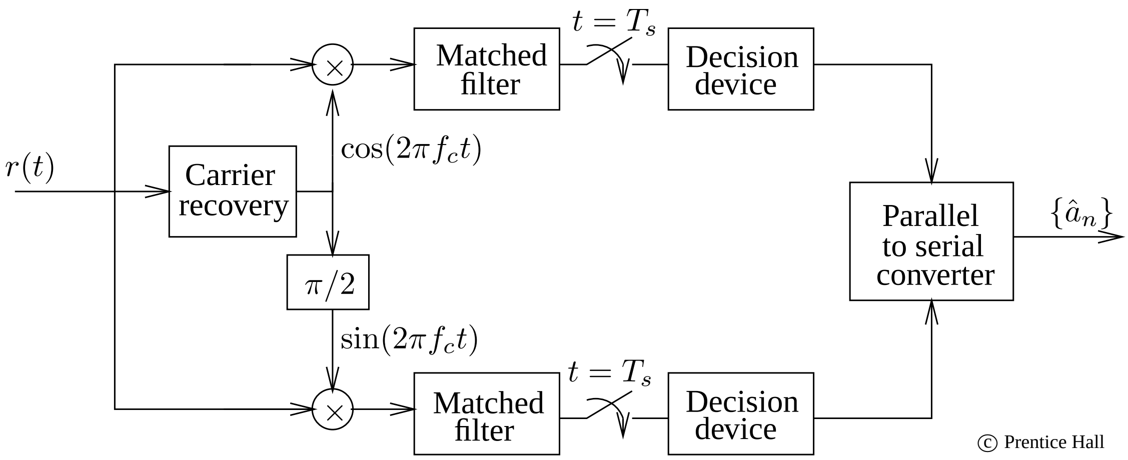


Figure 2: QPSK signal constellation, decision regions, and Gray encoding



(a) transmitter



(b) receiver (for $\zeta_0 = \pi/4$)

Figure 3: QPSK transmitter and receiver

Power spectral density (psd):

Consider a narrowband signal with carrier frequency f_c.

$$x(t) = \Re\{v(t)\exp[j2\pi f_c t]\},$$

where $v(t) = v_I(t) + jv_Q(t)$ is the complex envelope at baseband.

$$\begin{aligned}
\bar{R}_x(\tau) &= \frac{1}{T_c}\int_0^{T_c} E[x(t)x(t-\tau)]dt \\
&= \frac{1}{T_c}\int_0^{T_c} E\{\Re[v(t)e^{j2\pi f_c t}]\Re[v(t-\tau)e^{j2\pi f_c(t-\tau)}]\}dt \\
&= \frac{1}{2T_c}\int_0^{T_c} \{\Re[R_v(\tau)e^{j2\pi f_c \tau}] + \Re[E(v(t)v(t-\tau))e^{j2\pi f_c(2t-\tau)}]\}dt \\
&\approx \frac{1}{2}\Re[R_v(\tau)e^{j2\pi f_c \tau}],
\end{aligned}$$

where $R_v(\tau) = E[v(t)v^*(t-\tau)]$ is the correlation function of the baseband envelope $v(t)$, and

$$\int_0^{T_c} \Re[E(v(t)v(t-\tau))e^{j2\pi f_c(2t-\tau)}]\}dt \approx 0.$$

$\Longrightarrow$

$$\begin{aligned}
\Phi_x(f) &= \mathcal{F}[\bar{R}_x(\tau)] \\
&= \mathcal{F}\{\frac{1}{2}\Re[R_v(\tau)e^{j2\pi f_c \tau}]\} \\
&= \mathcal{F}\{\frac{1}{4}[R_v(\tau)e^{j2\pi f_c \tau} + R_v^*(\tau)e^{-j2\pi f_c \tau}]\} \\
&= \frac{1}{4}[\Phi_v(f-f_c) + \Phi_v(-f-f_c)],
\end{aligned}$$

where $\Phi_v(f) = \mathcal{F}[R_v(\tau)]$ is the psd of the baseband envelope $v(t)$.

If $\Phi_v(f)$ is an even function, then

$$\Phi_x(f) = \frac{1}{4}[\Phi_v(f - f_c) + \Phi_v(f + f_c)].$$

The psd of the baseband signal is

$$\Phi_v(f) = \frac{1}{T_s}|G(f)|^2 \sum_{k=-\infty}^{\infty} R_a(k) \exp(-j2\pi k f T_s),$$

where $G(f) = \mathcal{F}[g(t)]$ and $R_a(k) = E[a_n a^*_{n-k}]$.

For QPSK, $a_n = a_{1n} + j a_{2n}$ and

$$\begin{aligned} a_{1n} &= \begin{cases} +1, & \text{if the odd-numbered digit of the } n\text{th symbol is “1”} \\ -1, & \text{otherwise} \end{cases}, \\ a_{2n} &= \begin{cases} +1, & \text{if the even-numbered digit of the } n\text{th symbol is “1”} \\ -1, & \text{otherwise} \end{cases}, \end{aligned}$$

and the basic pulse is $g(t) = \sqrt{\frac{E_s}{T_s}}\Pi(t/T_s)$ with Fourier transform

$$G(f) = \mathcal{F}[g(t)] = \sqrt{\frac{E_s}{T_s}} \cdot T_s \text{sinc}(fT_s) \exp[-j2\pi f \frac{T_s}{2}].$$

The autocorrelation function of a_n is

$$R_a(k) = E\{[a_{1n} + j a_{2n}][a_{1(n-k)} - j a_{2(n-k)}]\} = 2\delta(k).$$

$\Longrightarrow$

$$\Phi_v(f) = 4E_b \text{sinc}^2(2fT_b).$$

$\Longrightarrow$ The psd of QPSK is then

$$\Phi_x(f) = E_b\{\text{sinc}^2[2(f - f_c)T_b] + \text{sinc}^2[2(f + f_c)T_b]\}.$$

3.2 Digital Modulation ($\pi/4$-DQPSK, MSK and GMSK)

$\pi/4$-**DQPSK**:

- Motivation:

  - more spectrally efficient than MSK (GMSK)
  - less envelope variation than QPSK $\rightarrow$ use of power efficient nonlinear amplifiers
  - simple receiver structure (e.g., using frequency discriminator)

- Superposition of two QPSK signal constellations offset by $\pi/4$ relative to each other

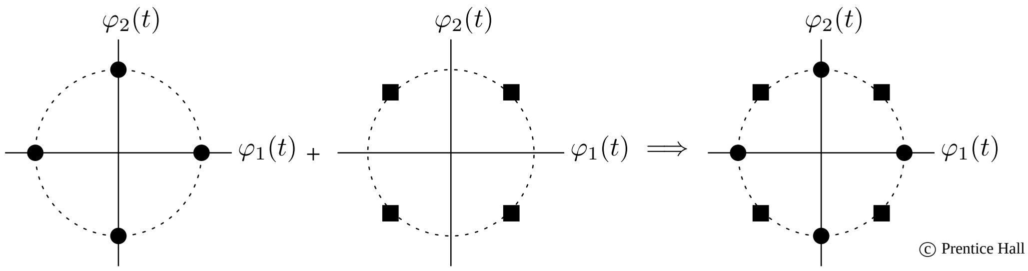


Figure 4: Signal constellation of $\pi/4$ shifted QPSK

  - The two signal constellations are used alternatively from symbol to symbol
  - There are 8 possible phases
  - The phase change between adjacent symbols $\in \{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\} \implies$
    (a) no phase change of $\pi \rightarrow$ amplitude fluctuation smaller than that of QPSK
    (b) no phase change of $0 \rightarrow$ symbol synchronization easier
    (c) noncoherent detection using a frequency discriminator

• $\pi/4$-DQPSK = $\pi/4$-shifted QPSK + differential encoding (to combat channel phase distortion)

Step 1: $\{a_n\}$ (at rate R_b) $\Longrightarrow \{a_{1n}\}$ and $\{a_{2n}\}$ (both at rate $R_b/2$)

Step 2: $(a_{1n}, a_{2n}) \Longrightarrow \phi_n$

Table 1: Mapping between the information symbol and phase difference in $\pi/4$-DQPSK

(a_{1n}, a_{2n})	ϕ_n
00	$\pi/4$
01	$3\pi/4$
11	$5\pi/4$
10	$7\pi/4$

Step 3: Generate the baseband in-phase and quadrature signal components as

$$\begin{cases} I_n = I_{n-1}\cos\phi_n - Q_{n-1}\sin\phi_n & \in \{0, \pm 1, \pm 1/\sqrt{2}\} \\ Q_n = I_{n-1}\sin\phi_n + Q_{n-1}\cos\phi_n & \in \{0, \pm 1, \pm 1/\sqrt{2}\} \end{cases}.$$

Also,

$$\begin{cases} I_n = \cos\Phi_n \\ Q_n = \sin\Phi_n \end{cases},$$

where Φ_n is the carrier phase for the nth symbol.

$\Longrightarrow$

$$\begin{cases} I_n = \cos\Phi_{n-1}\cos\phi_n - \sin\Phi_{n-1}\sin\phi_n = \cos(\Phi_{n-1} + \phi_n) \\ Q_n = \cos\Phi_{n-1}\sin\phi_n + \sin\Phi_{n-1}\cos\phi_n = \sin(\Phi_{n-1} + \phi_n) \end{cases}.$$

$\Longrightarrow$

$$\Phi_n = \Phi_{n-1} + \phi_n \quad \text{or} \quad \phi_n = \Phi_n - \Phi_{n-1}.$$

Step 4: Generate the transmitted signal over the current (nth) symbol interval as

$$\begin{aligned} x(t) &= \sqrt{\frac{2E_s}{T_s}}\cos(2\pi f_c t + \Phi_n) \\ &= \sqrt{E_s}\cos\Phi_n \underbrace{\sqrt{\frac{2}{T_s}}\cos(2\pi f_c t)}_{\varphi_1(t-nT_s)} - \sqrt{E_s}\sin\Phi_n \underbrace{\sqrt{\frac{2}{T_s}}\sin(2\pi f_c t)}_{\varphi_2(t-nT_s)}, \end{aligned}$$

where $f_c T_s$ is an integer and

$$\varphi_1(t) = \begin{cases} \sqrt{\frac{2}{T_s}}\cos(2\pi f_c t), & 0 \le t \le T_s \\ 0, & \text{otherwise} \end{cases}$$

$$\varphi_2(t) = \begin{cases} \sqrt{\frac{2}{T_s}}\sin(2\pi f_c t), & 0 \le t \le T_s \\ 0, & \text{otherwise} \end{cases}$$

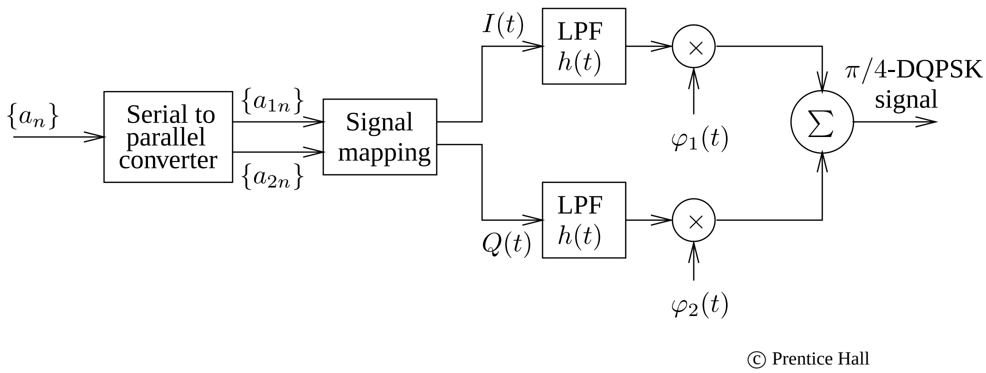


Figure 5: $\pi/4$-DQPSK transmitter

In the IS-54 standard, the LPFs shall have linear phase and square root raised cosine frequency response of the form

$$|H(f)| = \begin{cases} 1, & 0 \leq f \leq \frac{1-\alpha}{2T_s} \\ \sqrt{\frac{1}{2}\{1 - \sin[\frac{\pi(2fT_s-1)}{2\alpha}]\}}, & \frac{(1-\alpha)}{2T_s} \leq f \leq \frac{(1+\alpha)}{2T_s} \\ 0, & f > \frac{(1+\alpha)}{2T_s} \end{cases},$$

where $\alpha = 0.35$ is the roll-off factor.

The bandpass $\pi/4$-DQPSK signal in the nth symbol interval is given by

$$x_n(t) = \sqrt{E_s} I_n h(t - nT_s)\varphi_1(t - nT_s) - \sqrt{E_s} Q_n h(t - nT_s)\varphi_2(t - nT_s),$$

where $nT_s \leq t \leq (n+1)T_s$ and $\int_{-\infty}^{\infty} h^2(t)dt = 1$.

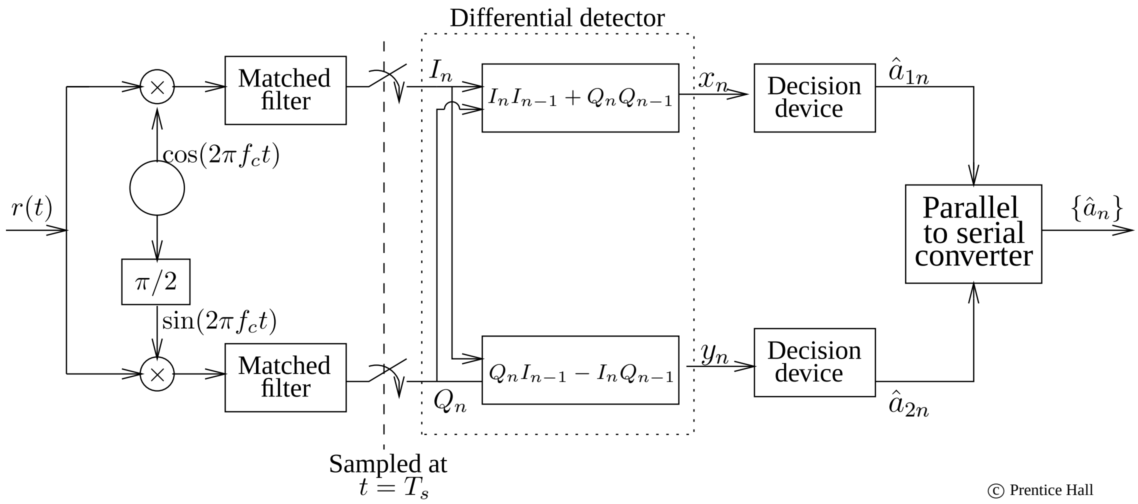


Figure 6: $\pi/4$-DQPSK receiver using a baseband differential detector

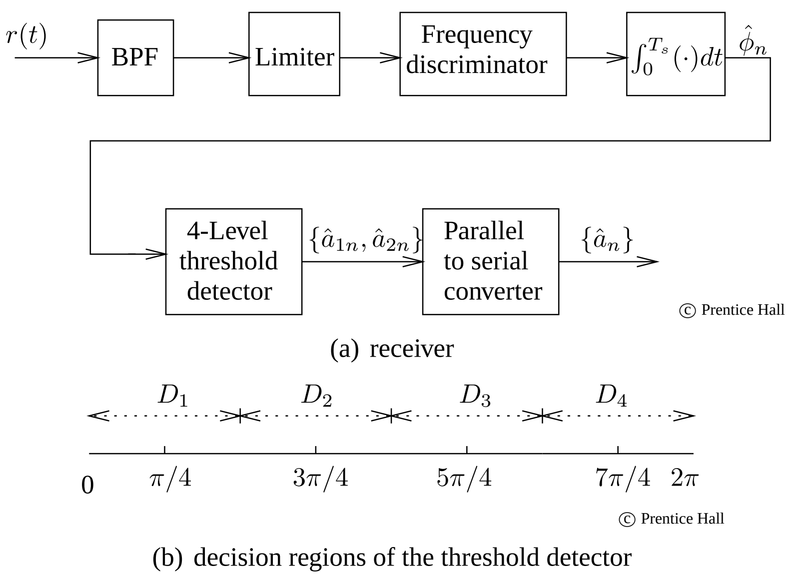


Figure 7: $\pi/4$-DQPSK noncoherent receiver using frequency discriminator

Minimum Shift Keying (MSK):

• a continuous phase frequency shift keying (CPFSK) scheme

• a binary FSK with modulation index $h = \Delta \cdot T_b = 1/2 \Longrightarrow$ minimum frequency separation that makes two FSK signals orthogonal $\rightarrow$ MSK

The frequency separation is

$$\Delta = f_2 - f_1 = 1/(2T_b).$$

$\Longrightarrow$

$$f_2 = f_c + \frac{1}{2}\Delta = f_c + \frac{1}{4T_b}, \quad \text{for symbol "1"}$$

and

$$f_1 = f_c - \frac{1}{2}\Delta = f_c - \frac{1}{4T_b}, \quad \text{for symbol "0"},$$

where $f_c = (f_1 + f_2)/2$ is the center of the two signal frequencies.

$\Longrightarrow$ With respect to the "carrier" phase $2\pi f_c t$, the phase change over each symbol (bit) interval is

$$\begin{cases} 2\pi(f_2 - f_c)T_b = \pi/2, & \text{for symbol "1"} \\ 2\pi(f_1 - f_c)T_b = -\pi/2, & \text{for symbol "0"} \end{cases}.$$

• For $t \in [0,\ T_b]$,

$$x(t) = \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_c t + \phi(t)],\ \ 0 \le t \le T_b,$$

where

$$\phi(t) = \phi(0) + 2\pi(\pm\frac{1}{2}\Delta)t = \phi(0) \pm \frac{\pi t}{2T_b},\quad 0 \le t \le T_b,$$

with "+" sign for symbol "1" and "$-$" sign for symbol "0", and where $\phi(0)$ is the phase at $t = 0$.

Given $\phi(0) = 0$, we have $\phi(t) = 0$ or π for $t = 2nT_b$, and $\phi(t) = \pi/2$ or $-\pi/2$ for $t = (2n+1)T_b$, where n is a positive integer. Within each symbol interval, the phase changes linearly with t.

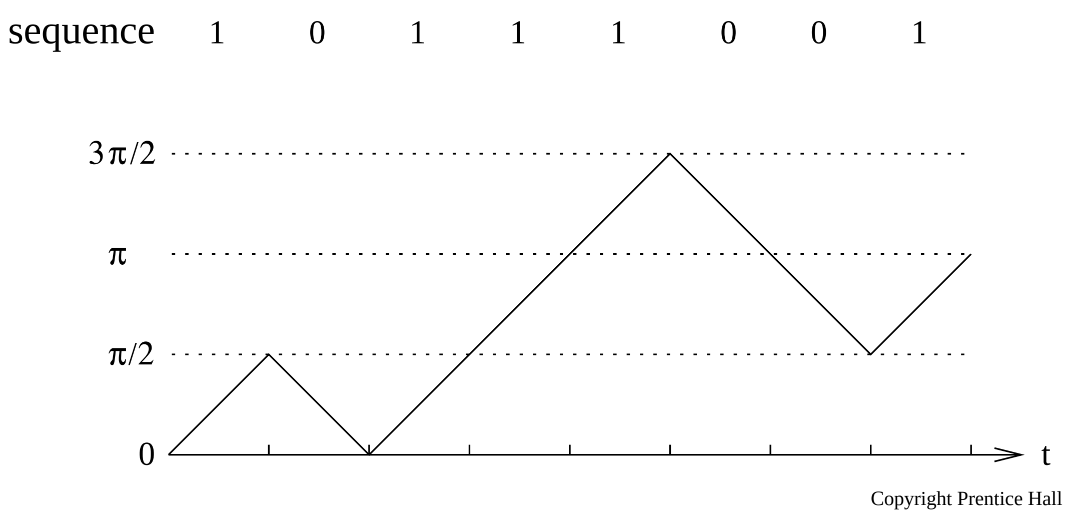


Figure 8: MSK phase trajectory for binary sequence "10111001" with $\phi(0) = 0$

• MSK signal representation:

For $t \in [0,\ T_b]$,

$$
\begin{aligned}
x(t) &= \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_c t + \phi(t)] \\
&= \sqrt{\frac{2E_b}{T_b}} \cos[\phi(t)] \cos(2\pi f_c t) - \sqrt{\frac{2E_b}{T_b}} \sin[\phi(t)] \sin(2\pi f_c t)
\end{aligned}
$$

where

$$
\begin{aligned}
\cos[\phi(t)] &= \cos[\phi(0)] \cos(\frac{\pi t}{2T_b}) \mp \sin[\phi(0)] \sin(\frac{\pi t}{2T_b}) \\
&= \cos[\phi(0)] \cos(\frac{\pi t}{2T_b}), \quad -T_b \le t \le T_b.
\end{aligned}
$$

The time interval can be extended to $[-T_b,\ T_b]$ as the cosine is an even function.

To find an expression of $\sin[\phi(t)]$, let $a_0 \in \{-1, +1\}$ and $a_1 \in \{-1, +1\}$ ($+1$ for symbol "1" and -1 for symbol "0") denote the information to be sent over the bit intervals $[0,\ T_b]$ and $[T_b,\ 2T_b]$ respectively.

For $t \in [0,\ T_b]$, we have

$$\phi(t) = \phi(0) + a_0 \frac{\pi t}{2T_b}.$$

$\Longrightarrow$

$$\phi(0) = \phi(T_b) - a_0 \frac{\pi T_b}{2T_b}.$$

$\Longrightarrow$

$$\begin{aligned}
\phi(t) &= \begin{cases} \phi(0) + a_0 \frac{\pi t}{2T_b}, & 0 \le t \le T_b \\ \phi(T_b) + a_1 \frac{\pi(t-T_b)}{2T_b}, & T_b \le t \le 2T_b \end{cases} \\
&= \begin{cases} \phi(T_b) + a_0 \frac{\pi(t-T_b)}{2T_b}, & 0 \le t \le T_b \\ \phi(T_b) + a_1 \frac{\pi(t-T_b)}{2T_b}, & T_b \le t \le 2T_b \end{cases}.
\end{aligned}$$

$\Longrightarrow$

$$\begin{aligned}
\sin[\phi(t)] &= \begin{cases} \sin[\phi(T_b)] \cos[a_0 \frac{\pi(t-T_b)}{2T_b}] + \cos[\phi(T_b)] \sin[a_0 \frac{\pi(t-T_b)}{2T_b}], & 0 \le t \le T_b \\ \sin[\phi(T_b)] \cos[a_1 \frac{\pi(t-T_b)}{2T_b}] + \cos[\phi(T_b)] \sin[a_1 \frac{\pi(t-T_b)}{2T_b}], & T_b \le t \le 2T_b. \end{cases} \\
&= \sin[\phi(T_b)] \cos[\frac{\pi(t-T_b)}{2T_b}], \quad 0 \le t \le 2T_b \\
&= \sin[\phi(T_b)] \sin(\frac{\pi t}{2T_b}), \quad 0 \le t \le 2T_b,
\end{aligned}$$

where $\cos[\phi(T_b)] = 0$ because $\phi(T_b) = \pm\frac{\pi}{2}$, and $\cos[a_0 \frac{\pi(t-T_b)}{2T_b}] = \cos[a_1 \frac{\pi(t-T_b)}{2T_b}] = \cos[\frac{\pi(t-T_b)}{2T_b}]$.

$\Longrightarrow$

$$x(t) = \sqrt{E_b} a_I \varphi_1(t) + \sqrt{E_b} a_Q \varphi_2(t), \quad 0 \le t \le T_b,$$

where

$$\begin{cases} a_I = \cos[\phi(0)] \\ a_Q = -\sin[\phi(T_b)] \end{cases}$$

and

$$\begin{cases} \varphi_1(t) &= \sqrt{\frac{2}{T_b}} \cos(\frac{\pi t}{2T_b}) \cos(2\pi f_c t), & -T_b \le t \le T_b \\ \varphi_2(t) &= \sqrt{\frac{2}{T_b}} \sin(\frac{\pi t}{2T_b}) \sin(2\pi f_c t), & 0 \le t \le 2T_b. \end{cases}$$

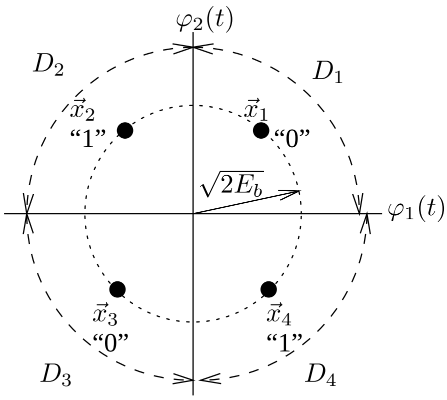
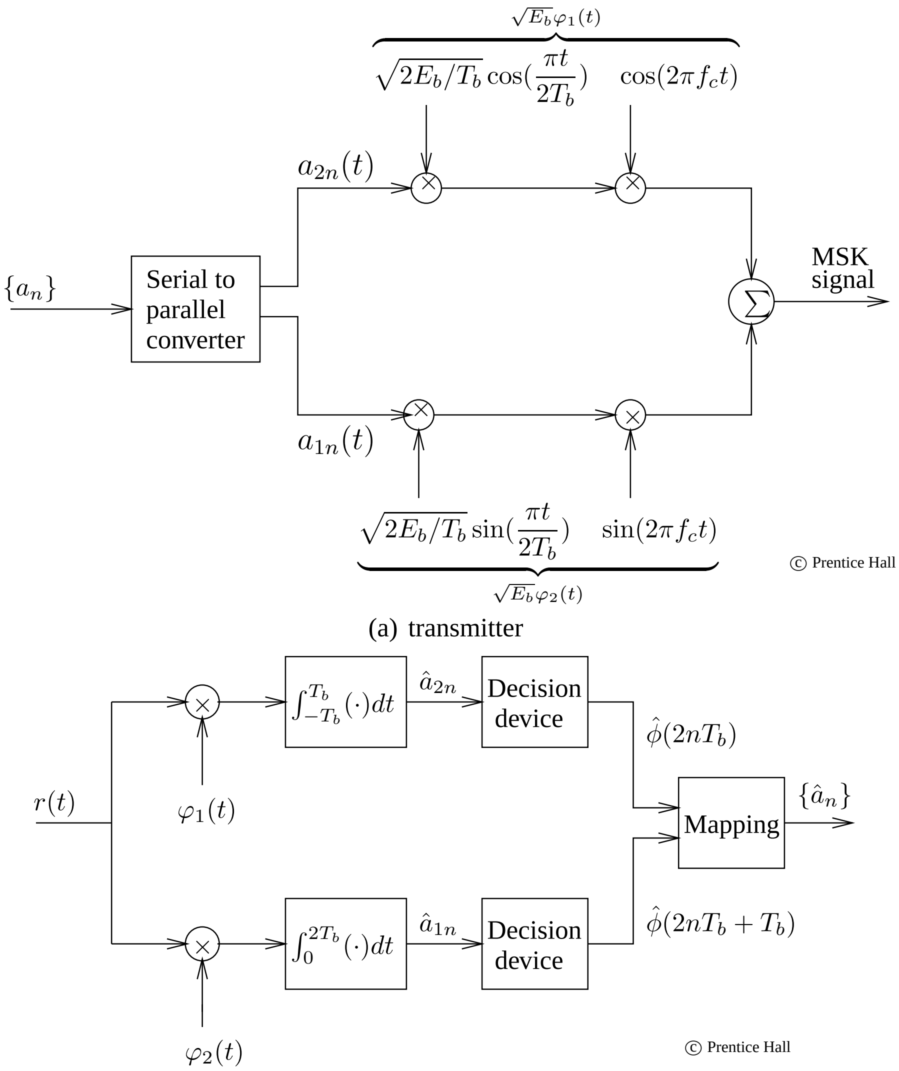


Figure 9: The signal space and decision regions of MSK

Table 2: Signal points in the MSK signal space

signal point	coordinates (a_I, a_Q)	$\phi(0)$	$\phi(T_b)$	phase change $\phi(T_b) - \phi(0)$	information symbol
$\vec{x}_1$	$(1, 1)$	0	$-\pi/2$	$-\pi/2$	“0”
$\vec{x}_2$	$(-1, 1)$	π	$-\pi/2$	$\pi/2$	“1”
$\vec{x}_3$	$(-1, -1)$	π	$\pi/2$	$-\pi/2$	“0”
$\vec{x}_4$	$(1, -1)$	0	$\pi/2$	$\pi/2$	“1”



(a) transmitter

(b) coherent receiver

Figure 10: MSK transmitter and receiver

Example 3.4 Orthogonality between the MSK signals

Verify that the minimum frequency separation between the two MSK signals is $1/2T_b$ in order to achieve orthogonal signaling.

Solution:

Without loss of generality, consider the transmission of a symbol over the time interval $[0, T_b]$. The MSK signal is

$$x_1(t) = A\cos[2\pi f_1 t + \phi(0)]$$

for symbol "0" and

$$x_2(t) = A\cos[2\pi f_2 t + \phi(0)]$$

for symbol "1", where the initial phase at $t = 0$, $\phi(0)$, is the same for both cases due to the constraint of continuous phase. For $x_1(t)$ and $x_2(t)$ to be orthogonal over the time interval $t \in [0, T_b]$, it is required that

$$\int_0^{T_b} x_1(t)x_2(t)dt = 0.$$

This requirement means that

$$\begin{aligned}
& \int_0^{T_b} \cos(2\pi f_1 t + \phi(0))\cos(2\pi f_2 t + \phi(0))dt = 0, \\
\Longrightarrow \quad & \int_0^{T_b} \cos[2\pi(f_1 + f_2)t + 2\phi(0)]dt + \int_0^{T_b} \cos[2\pi(f_2 - f_1)t]dt = 0, \\
\Longrightarrow \quad & \int_0^{T_b} \cos[2\pi(f_2 - f_1)t]dt = 0, \\
\Longrightarrow \quad & \sin[2\pi(f_2 - f_1)T_b] = 0, \\
\Longrightarrow \quad & \min\{(f_2 - f_1)\} = 1/2T_b,
\end{aligned}$$

where $\int_0^{T_b} \cos[2\pi(f_1 + f_2)t + 2\phi(0)]dt \approx 0$ because $(f_1 + f_2) \gg 1/T_b$. $\square$

Gaussian MSK (GMSK):

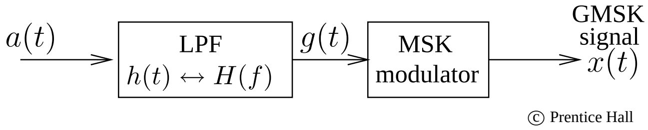


Figure 11: Gaussian MSK transmitter

$$H(f) \;=\; \exp[-\tfrac{\ln 2}{2}(\tfrac{f}{B})^2]$$

where B is the bandwidth of the filter.

$$h(t) = \mathcal{F}^{-1}[H(f)] = \sqrt{\frac{2\pi}{\ln 2}} B \exp[-\frac{2\pi^2 B^2 t^2}{\ln 2}].$$

$\Longrightarrow$

$$x(t) = \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_c t + \phi(t)].$$

To represent $\phi(t)$:

$$a(t) = \sum_{n=-\infty}^{\infty} a_n \Pi(\frac{t - nT_b}{T_b})$$

where

$$a_n = \begin{cases} 1, & \text{if the } n\text{th symbol is "1"} \\ -1, & \text{if the } n\text{th symbol is "0"} \end{cases}$$

and

$$\Pi(\frac{t}{T_b}) = \begin{cases} 1, & 0 \le t \le T_b \\ 0, & \text{otherwise} \end{cases} \quad .$$

With the impulse response $h(t)$, the output of the LPF is

$$g(t) = a(t) \star h(t) = \sum_{n=-\infty}^{\infty} a_n y(t - nT_b)$$

where

$$\begin{aligned} y(t) &= \Pi(\frac{t}{T_b}) \star h(t) \\ &= \int_{t-T_b}^{t} h(z)dz \\ &= \frac{1}{2}\{\text{erf}[\sqrt{\frac{2}{\ln 2}}\pi Bt] + \text{erf}[-\sqrt{\frac{2}{\ln 2}}\pi B(t - T_b)]\}, \end{aligned}$$

with $\text{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t \exp(-z^2)dz$.

The phase of the GMSK signal is then

$$\phi(t) = \frac{\pi}{2T_b} \int_{-\infty}^{t} g(z)dz.$$

Over the nth bit interval, $t \in [nT_b, (n+1)T_b]$, the phase is

$$\phi_n(t) = \phi(nT_b) \pm \frac{\pi}{2T_b} \int_{nT_b}^{t} y(z - nT_b)dz,$$

where the "+" sign is for $a_n = 1$ and the "$-$" sign is for $a_n = -1$. With the phase $\phi(t)$, the instantaneous frequency of the GMSK signal is

$$f(t) = f_c + \frac{1}{2\pi}\frac{d\phi(t)}{dt} = f_c \pm \frac{1}{4T_b} y(t - nT_b).$$

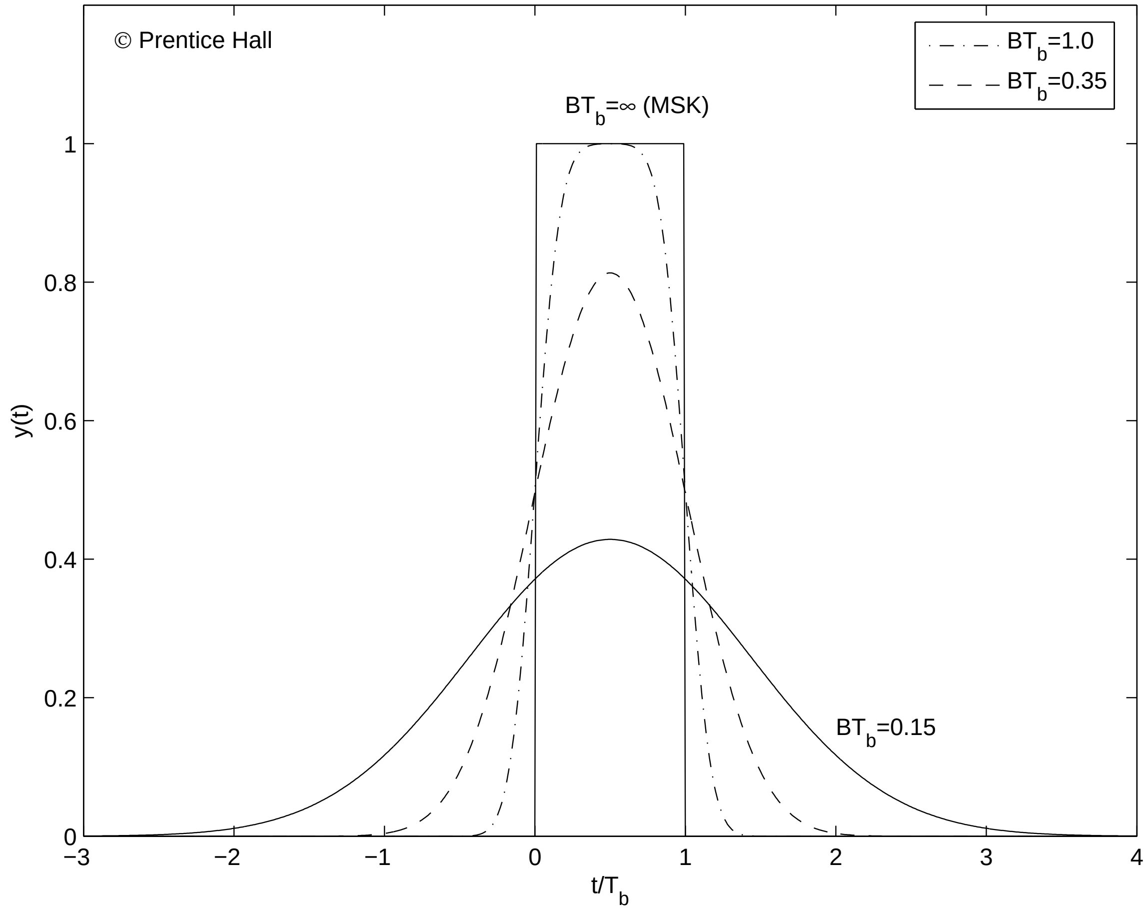


Figure 12: Characteristics of Gaussian shaping filter

Orthogonal Frequency Division Multiplexing (OFDM)

• OFDM is a block modulation scheme: a block of N serial symbols is converted into a block of N parallel modulated symbols, each of duration $T = NT_s$.

• If the rms delay spread of the channel is σ_τ, N is chosen so that $NT_s \gg \sigma_\tau$. $\Longrightarrow$ OFDM has the property of mitigating frequency-selective fading.

The complex envelope of OFDM is given by

$$\begin{aligned} v(t) &= \sqrt{\tfrac{2E_s}{T_s}} \textstyle\sum_{k=0}^{\infty} \sum_{n=0}^{N-1} a_{k,n} \tilde{\varphi}_n(t - kT) \\ &= \textstyle\sum_{n=0}^{N-1} v_n(t), \end{aligned}$$

where $a_{k,n}$ carries the information to be sent over the kth symbol interval and the nth subband.

$v_n(t)$ is the complex envelope of the signal transmitted in the nth subband

$$v_n(t) = \sqrt{\frac{2E_s}{T_s}} \sum_{k=0}^{\infty} a_{k,n} \tilde{\varphi}_n(t - kT)$$

where $\{\tilde{\varphi}_n(t)\}_{n=0}^{N-1}$ is a set of complex orthonormal waveforms and is given by

$$\tilde{\varphi}_n(t) = \begin{cases} \exp[j2\pi(n - \frac{N-1}{2})t/T], & t \in [0, T] \\ 0, & t \notin [0, T] \end{cases} .$$

$\Longrightarrow$ Each waveform in the set $\{v_n(t)\}_{n=0}^{N-1}$ corresponds to a distinct (nth) subcarrier with frequency $f_c + \frac{2n-(N-1)}{2T}$.

Consider $k = 0$.

$$v(t) = \sqrt{\frac{2E_s}{T_s}} \sum_{n=0}^{N-1} a_{0,n} \exp(j\frac{2\pi nt}{NT_s}) \exp[-j\frac{\pi(N-1)t}{NT_s}], \;\; 0 \le t \le NT_s.$$

$\Longrightarrow$

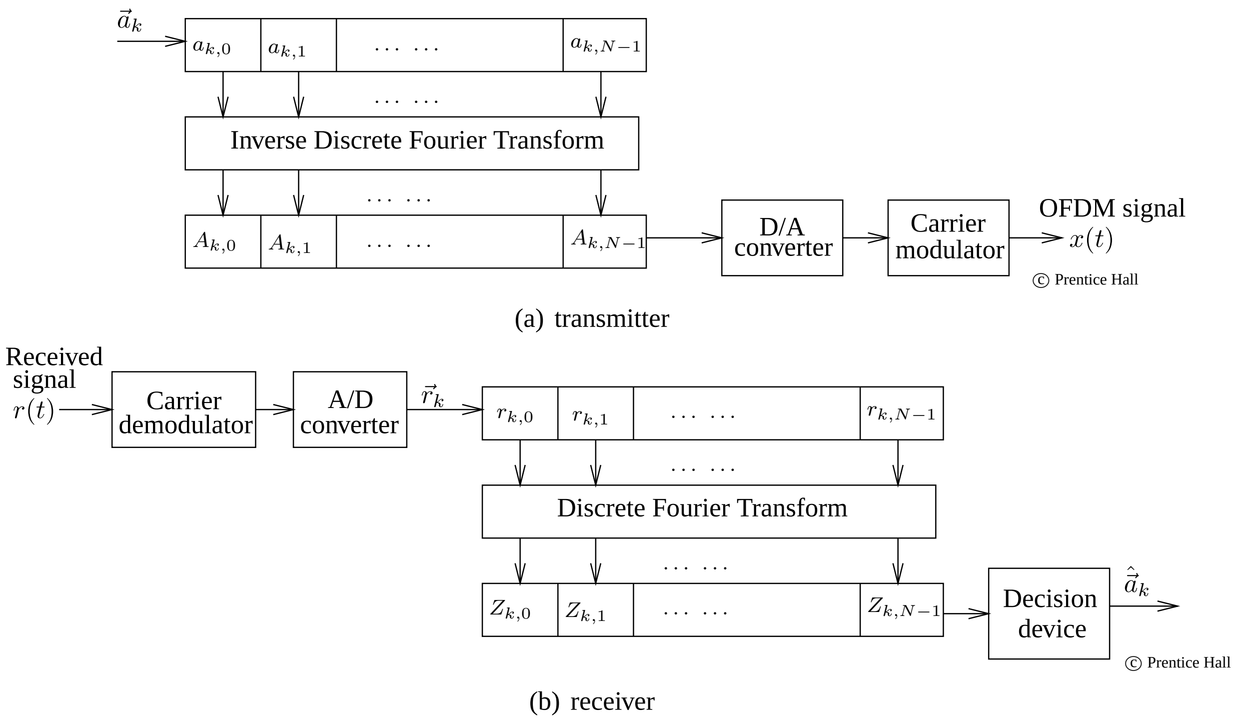
$$v(t) = \sqrt{\frac{2E_s}{T_s}} \sum_{n=0}^{N-1} a_{0,n} \exp(j\frac{2\pi nt}{NT_s}), \;\; 0 \le t \le NT_s,$$

with corresponding carrier $\exp[j2\pi(f_c - \frac{N-1}{2NT_s})t]$.

Sampling $v(t)$ at $t = \ell T_s$ yields

$$A_{0,\ell} \triangleq v(\ell T_s) = \sqrt{\frac{2E_s}{T_s}} \sum_{n=0}^{N-1} a_{0,n} \exp(j\frac{2\pi n\ell}{N}), \;\; \ell = 0, 1, 2, \cdots, N-1,$$

which is actually proportional to the inverse discrete Fourier transform (IDFT) of $\{a_{0,n}\}$.



(a) transmitter

(b) receiver

Figure 13: OFDM transmitter and receiver

3.3 Power Spectral Density (psd)

$$\Phi_x(f)\frac{1}{4}[\Phi_v(f-f_c)+\Phi_v(f+f_c)].$$

where

$$\Phi_v(f)=\frac{1}{T_s}|G(f)|^2\sum_{k=-\infty}^{\infty}R_a(k)\exp(-j2\pi kfT_s),$$

where $G(f)=\mathcal{F}[g(t)]$ and $R_a(k)=E[a_n a^*_{n-k}]$.

$\pi/4$-**DQPSK**:

• For $\pi/4$-DQPSK, $a_n=a_{1n}+ja_{2n}$. If $\{a_{1n}\}$ and $\{a_{2n}\}$ are independent,

$$\begin{aligned}
R_a(k) &= E[(a_{1n}+ja_{2n})(a_{1(n-k)}-ja_{2(n-k)})] \\
&= E[a_{1n}a_{1(n-k)}]+E[a_{2n}a_{2(n-k)}]+jE[a_{2n}]E[a_{1(n-k)}]-jE[a_{1n}]E[a_{2(n-k)}] \\
&= R_{a_1}(k)+R_{a_2}(k)
\end{aligned}$$

$\Longrightarrow$ The psd of $x(t)$ is the sum of the psd for the in-phase component and the psd for the quadrature component.

• Since the psd depends only on $|G(f)|$ not on $G(f)$, rotation of the signal constellation does not affect the psd;

In $\pi/4$-DQPSK, the two QPSK signal constellations have the same psd.

$\Longrightarrow$ $\pi/4$-DQPSK has the same psd as QPSK, which is given by

$$\Phi_x(f)=E_b\{\mathrm{sinc}^2[2(f-f_c)T_b]+\mathrm{sinc}^2[2(f+f_c)T_b]\}.$$

MSK and GMSK:

• MSK signal in $t \in [0,\ T_b]$ is

$$x(t) = \sqrt{\frac{2E_b}{T_b}}[a_I \cos\frac{\pi t}{2T_b}\cos(2\pi f_c t) + a_Q \sin\frac{\pi t}{2T_b}\sin(2\pi f_c t)], \quad 0 \le t \le T_b$$

$\Longrightarrow$

$$v(t) = \sum_{n=-\infty}^{\infty}[a_{I,n}g(t-nT_b) - ja_{Q,n}g(t-nT_b-T_b)],$$

where

$$g(t) = \sqrt{\frac{2E_b}{T_b}}\cos(\frac{\pi t}{2T_b}) \quad -T_b \le t \le T_b$$

which has Fourier transform

$$G(f) = \mathcal{F}[g(t)] = -\frac{1}{4\pi T_b}\sqrt{\frac{2E_b}{T_b}}\frac{\cos(2\pi f T_b)}{f^2 - \frac{1}{16T_b^2}}.$$

• $g(t-nT_b)$ and $g(t-nT_b-T_b)$ are orthogonal.

$\Longrightarrow$ The psd of $v(t)$ is the summation of the psd of $\sum_{n=-\infty}^{\infty} a_{I,n}g(t-nT_b)$ and psd of $-\sum_{n=-\infty}^{\infty} a_{Q,n}(t)g(t-nT_b-T_b)$.

$$\begin{aligned}
\Phi_v(f) &= \frac{1}{2T_b}E[(a_{I,n})^2]|\mathcal{F}[g(t)]|^2 + \frac{1}{2T_b}E[(a_{Q,n})^2]|\mathcal{F}[g(t-T_b)]|^2 \\
&= \frac{1}{2T_b}|G(f)|^2 + \frac{1}{2T_b}|G(f)\exp[-j2\pi f T_b]|^2 \\
&= \frac{32E_b}{\pi^2}[\frac{\cos(2\pi f T_b)}{1-(4fT_b)^2}]^2
\end{aligned}$$

$$\Phi_x(f) = \frac{8E_b}{\pi^2}\{\frac{\cos[2\pi(f-f_c)T_b]}{1-[4(f-f_c)T_b]^2}\}^2 + \frac{8E_b}{\pi^2}\{\frac{\cos[2\pi(f+f_c)T_b]}{1-[4(f+f_c)T_b]^2}\}^2.$$

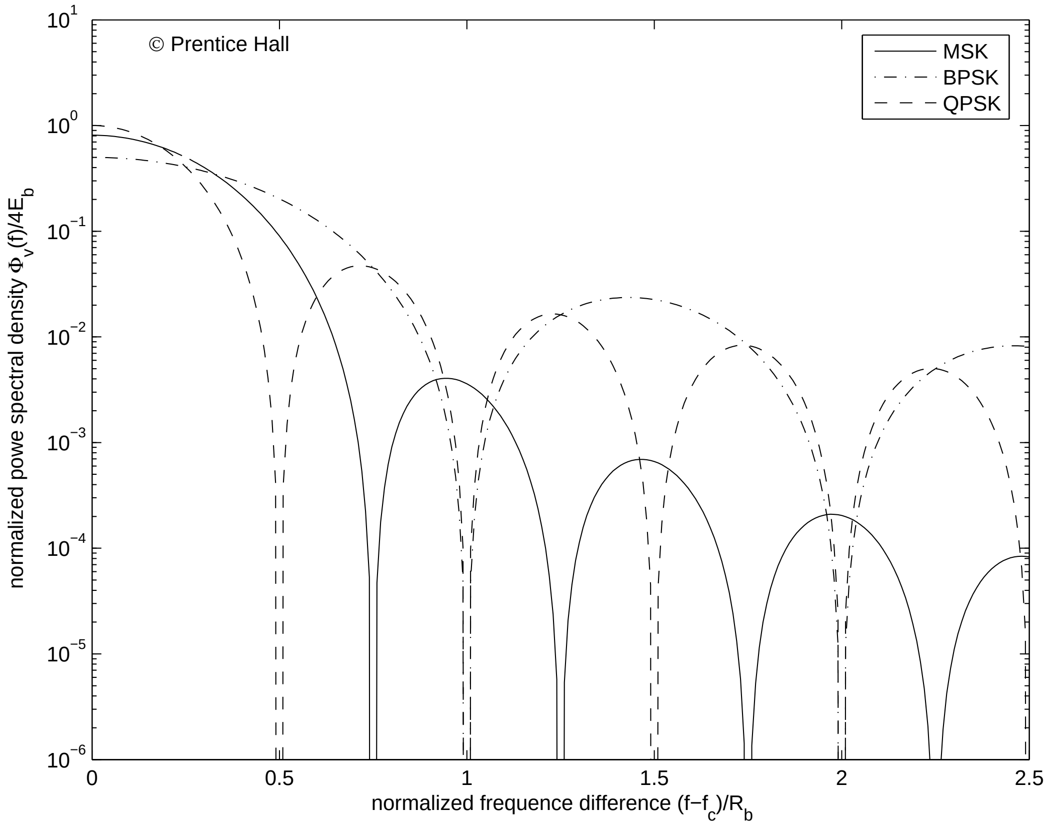


Figure 14: Normalized power spectral density of MSK, BPSK, and QPSK

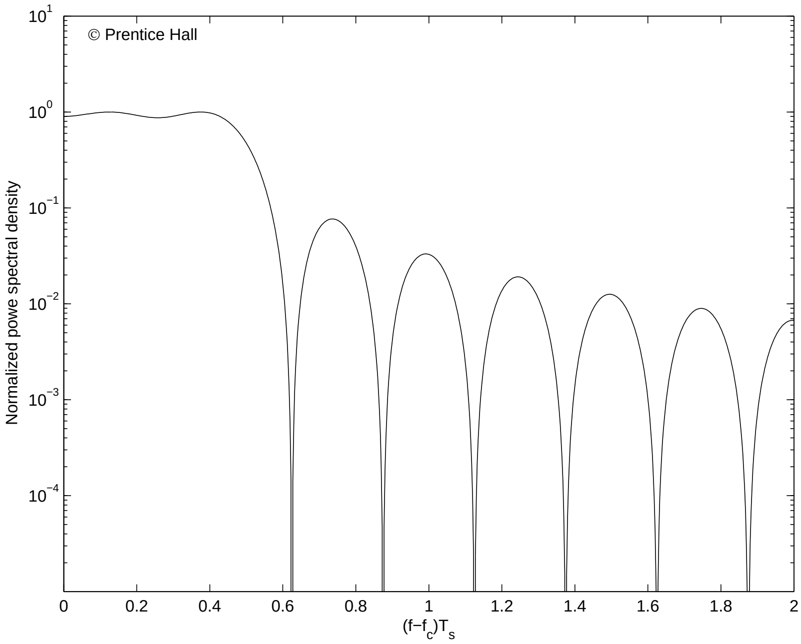
OFDM

OFDM can be viewed as independent modulation on orthogonal subcarriers separated by $\frac{1}{T} = \frac{1}{NT_s}$ Hz.

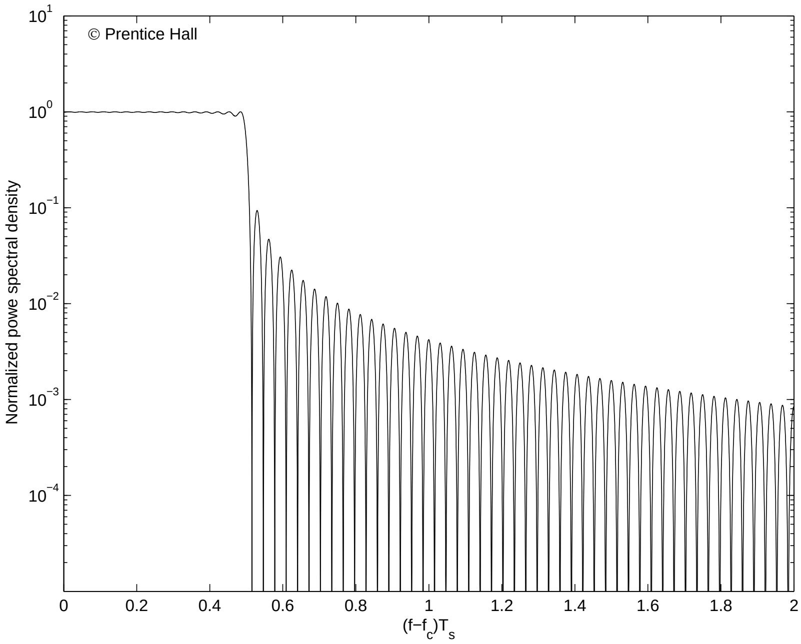
$\Longrightarrow$

$$\Phi_v(f) = \frac{1}{T}\sum_{n=0}^{N-1}\{|G(f - \frac{1}{T}(n - \frac{N-1}{2}))|^2 \sum_{l=-\infty}^{\infty} R_a(l,n)\exp(-j2\pi l f T)\},$$

where $R_a(l,n) = E[a_{k,n}a^*_{k-l,n}]$, $T = N/R_s$, and R_s is the information symbol rate.



(a) $N = 4$



(b) $N = 32$

Figure 15: Normalized OFDM psd as a function of the normalized frequency difference

Example 3.5 Calculation of psd

Determine an expression for the psd of the random pulse train

$$Y(t) = \sum_{n=-\infty}^{\infty} Y_n g(t - nT),$$

where $g(t)$ is a pulse shape defined on $[0, T]$, $Y_n = X_n + \frac{1}{2} X_{n-1}$, and the random variables X_n are independent and identically distributed, each following a Gaussian distribution with zero mean and variance σ^2.

Solution:

The psd of $Y(t)$ is given by

$$S_Y(f) = \frac{1}{T} |G(f)|^2 \sum_{k=-\infty}^{\infty} R_Y(k) \exp(-j2\pi k f T),$$

where $G(f) = \mathcal{F}[g(t)]$ and $R_Y(k) = E[Y_n Y_{n-k}]$.

The autocorrelation function of $\{X_n\}$ is

$$\begin{aligned} R_X(k) &= E[X_n X_{n-k}] \\ &= \begin{cases} E[X_n^2], & \text{if } k = 0 \\ E(X_n)E(X_{n-k}), & \text{if } k \neq 0 \end{cases} \\ &= \begin{cases} \sigma^2, & \text{if } k = 0 \\ 0, & \text{if } k \neq 0 \end{cases}, \end{aligned}$$

since $E(X_n) = E(X_{n-k}) = 0$ and $E(X_n^2) = Var(X_n) = \sigma^2$.

The autocorrelation function of $\{Y_n\}$ can be represented in terms of $R_X(k)$ as

$$\begin{aligned} R_Y(k) &= E[Y_n Y_{n-k}] \\ &= E[(X_n + 0.5X_{n-1})(X_{n-k} + 0.5X_{n-k-1})] \\ &= E[X_n X_{n-k} + 0.5X_n X_{n-k-1} + 0.5X_{n-1}X_{n-k} + 0.25X_{n-1}X_{n-k-1}] \\ &= R_X(k) + 0.5R_X(k+1) + 0.5R_X(k-1) + 0.25R_X(k) \\ &= \begin{cases} 1.25\sigma^2, & \text{if } k = 0 \\ 0.5\sigma^2, & \text{if } k = +1 \text{ or } k = -1 \\ 0, & \text{otherwise} \end{cases} . \end{aligned}$$

The psd of $Y(t)$ is therefore

$$\begin{aligned} S_Y(f) &= \frac{1}{T}|G(f)|^2[1.25\sigma^2 + 0.5\sigma^2(e^{-j2\pi fT} + e^{j2\pi fT})] \\ &= \frac{1}{T}|G(f)|^2[1.25\sigma^2 + \sigma^2\cos(2\pi fT)]. \end{aligned}$$

□

3.4 Probability of Transmission Error

Coherent BPSK in AWGN:

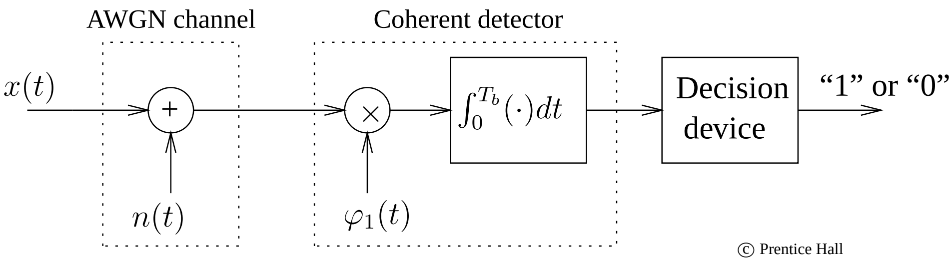


Figure 16: Coherent reception of BPSK in an AWGN channel

The transmitted signal is

$$x(t) = \begin{cases} \sqrt{E_b}\varphi_1(t), & \text{for symbol "1"} \\ -\sqrt{E_b}\varphi_1(t), & \text{for symbol "0"} \end{cases}$$

The received signal is

$$r(t) = x(t) + n(t),$$

where $n(t)$ represents the white Gaussian noise process with zero mean and two-sided psd $N_0/2$. The output of the correlator (or the matched filter at $t = T_b$)

$$r_1 = \int_0^{T_b} r(t)\varphi_1(t)dt \sim N(\mu_i, \sigma_N^2)$$

where $\mu_1 = \sqrt{E_b}$ given "1" being sent and $\mu_0 = -\sqrt{E_b}$ given "0" being sent, and $\sigma_N^2 = N_0/2$.

$\Longrightarrow$

$$f_{r_1}(x) = \frac{1}{\sqrt{2\pi}\sigma_N} \exp[-\frac{(x-\mu_i)^2}{2\sigma_N^2}], \quad i = 0, 1; -\infty < x < \infty.$$

The decision rule is:

$$\begin{cases} \text{if } r \geq 0, \text{ symbol "1" was sent;} \\ \text{if } r < 0, \text{ symbol "0" was sent.} \end{cases}$$

With equally likely symbols "1" and "0", the BER is

$$\begin{aligned} P_b &= P(r \geq 0|\text{symbol "0" was sent})P(\text{symbol "0" was sent}) \\ &\quad + P(r < 0|\text{symbol "1" was sent})P(\text{symbol "1" was sent}) \\ &= \frac{1}{2}\{\int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_N} \exp[-\frac{(x-\mu_0)^2}{2\sigma_N^2}]dx + \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma_N} \exp[-\frac{(x-\mu_1)^2}{2\sigma_N^2}]dx\} \\ &= Q(\sqrt{\frac{2E_b}{N_0}}) \end{aligned}$$

Coherent QPSK in AWGN

- QPSK=sum of two orthogonal BPSK;
- The signal detections; at the two branches are independent in an AWGN channel
- BER for the odd-numbered digits
  = BER for the even-numbered digits
  = BER for BPSK.

The probability of correct symbol reception for QPSK is

$$P_c = [1 - Q(\sqrt{\frac{2E_b}{N_0}})]^2.$$

The probability of symbol error is therefore given by

$$P_s = 1 - P_c = 2Q(\sqrt{\frac{2E_b}{N_0}}) - [Q(\sqrt{\frac{2E_b}{N_0}})]^2,$$

which can be approximated by $2Q(\sqrt{\frac{2E_b}{N_0}})$ if $2E_b/N_0 \gg 1$.

<u>Coherent MSK in AWGN</u>:

$$x(t) = \sqrt{E_b} a_I \varphi_1(t) + \sqrt{E_b} a_Q \varphi_2(t), \quad 0 \le t \le T_b,$$

where

$$\begin{cases} a_I = \cos[\phi(0)] \\ a_Q = -\sin[\phi(T_b)] \end{cases}$$

and

$$\begin{cases} \varphi_1(t) &= \sqrt{\frac{2}{T_b}} \cos(\frac{\pi t}{2T_b}) \cos(2\pi f_c t), & -T_b \le t \le T_b \\ \varphi_2(t) &= \sqrt{\frac{2}{T_b}} \sin(\frac{\pi t}{2T_b}) \sin(2\pi f_c t), & 0 \le t \le 2T_b. \end{cases}$$

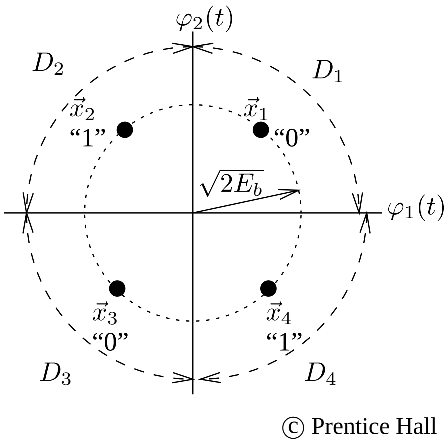


Figure 17: The signal space and decision regions of MSK

The signal constellation of MSK is similar to that of QPSK.

The differences between MSK and QPSK are:

- $$M = \begin{cases} 4 & \text{for QPSK} \rightarrow \text{1 symbol = 2 bits} \rightarrow \; E_s = 2E_b \\ 2 & \text{for MSK} \rightarrow \text{1 symbol = 1 bits} \rightarrow \; E_s = E_b \end{cases}$$

  $\Longrightarrow E_b$ in MSK corresponds to $E_s/2$ in QPSK;

- The orthonormal sets $\{varphi_1(t), \varphi_2(t)\}$ are different;

- QPSK has no memory, while MSK has a memory of one symbol due to the constraint of *continuous phase*.

In MSK, both decision regions D_1 and D_3 are for symbol "1", and both D_2 and D_4 are for "0". $\Longrightarrow$ A symbol error may not mean a detection error on the transmitted binary information bit.

In MSK the total decision region for "1" or "0" is half of the signal space, similar to that for the odd- or even-numbered digits in QPSK. $\Longrightarrow$ The probability of bit error for MSK is the same as the probability of bit error for QPSK.

$$P_b = Q(\sqrt{\frac{2E_b}{N_0}}).$$

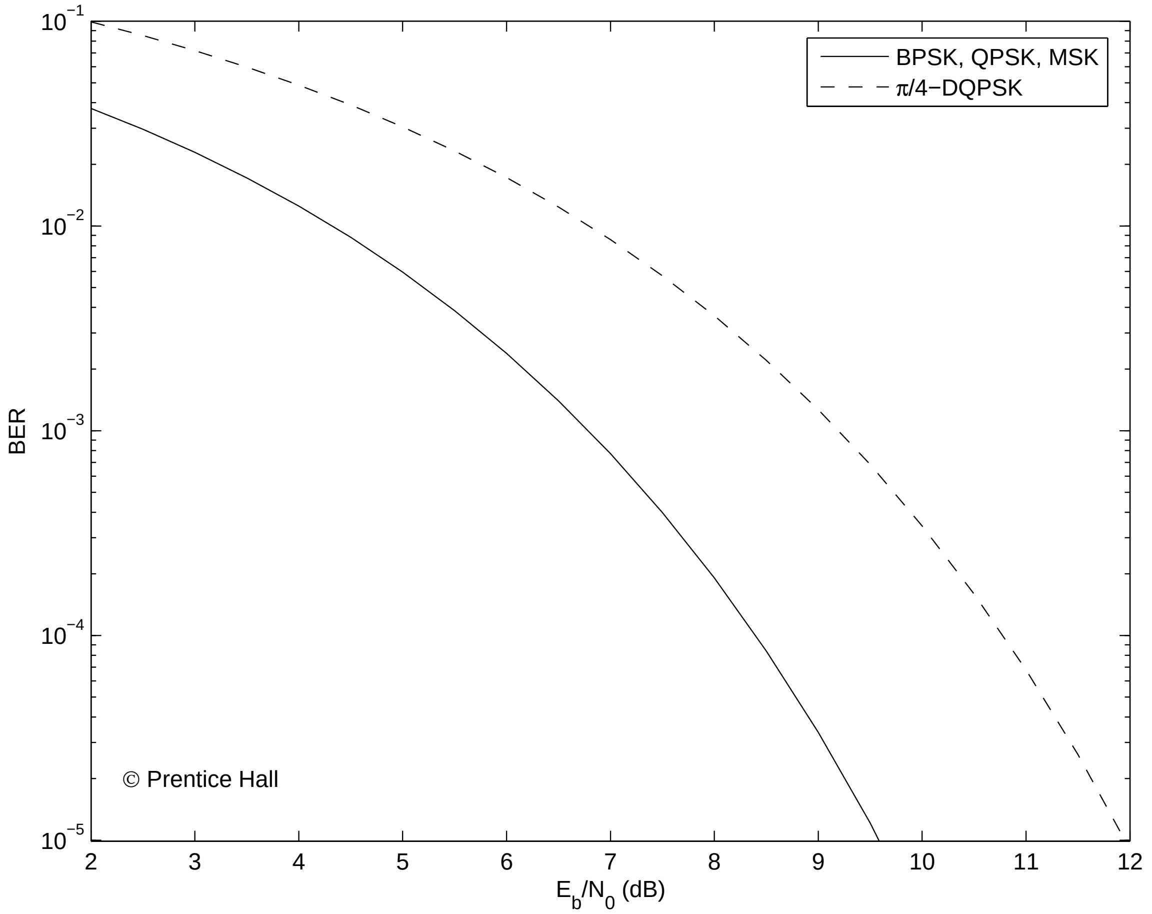


Figure 18: BER performance of coherently detected BPSK, QPSK and MSK, and differentially detected $\pi/4$-DQPSK in an AWGN channel

Example 3.6 Bit error rate for a binary system:

Consider a binary transmission system where, at the end of each symbol interval, the output of the demodulator is

$$r = \begin{cases} \alpha + n + Z, & \text{if "1" was sent} \\ -\alpha + n + Z, & \text{if "0" was sent} \end{cases},$$

where $\alpha\ (> 0)$ is a constant, n is a Gaussian random variable with zero mean and variance σ^2, Z is a discrete random variable with probability distribution $P(Z = z) = 0.25$, $P(Z = 0) = 0.5$, and $P(Z = -z) = 0.25$. Assuming that symbols "1" and "0" are equally likely, derive an expression for the probability of bit error in terms of α, σ, and z.

Solution:

The mean of random variable Z is

$$E[Z] = z \times P(Z = z) + 0 \times P(Z = 0) + (-z) \times P(Z = -z) = 0.$$

As a result, we have $E(n + Z) = E(n) + E(Z) = 0$ since n has a zero mean. Furthermore, as symbols "1" and "0" have equal probability, the optimal decision threshold of the decision device following the demodulator is 0. The decision rule is as follows

$$\begin{cases} \text{if } r \geq 0, \text{ then symbol "1" was sent;} \\ \text{if } r < 0, \text{ then symbol "0" was sent.} \end{cases}$$

Given $Z = 0$, the probability of bit error is

$$\begin{aligned} P_{b|Z=0} &= P(r < 0|\text{“1”}, Z = 0)P(\text{“1”}) + P(r \geq 0|\text{“0”}, Z = 0)P(\text{“0”}) \\ &= P(\alpha + n + Z < 0|Z = 0) \cdot \frac{1}{2} + P(-\alpha + n + Z \geq 0|Z = 0) \cdot \frac{1}{2} \\ &= \frac{1}{2}[P(n < -\alpha) + P(n \geq \alpha)] \\ &= Q(\frac{\alpha}{\sigma}). \end{aligned}$$

Similarly, given $Z = z$, we have

$$\begin{aligned} P_{b|Z=z} &= P(r < 0|\text{“1”}, Z = z)P(\text{“1”}) + P(r \geq 0|\text{“0”}, Z = z)P(\text{“0”}) \\ &= P(\alpha + n + Z < 0|Z = z) \cdot \frac{1}{2} + P(-\alpha + n + Z \geq 0|Z = z) \cdot \frac{1}{2} \\ &= \frac{1}{2}[P(n < -\alpha - z) + P(n \geq \alpha - z)] \\ &= \frac{1}{2}[Q(\frac{\alpha + z}{\sigma}) + Q(\frac{\alpha - z}{\sigma})] \end{aligned}$$

and, given $Z = -z$, we have

$$P_{b|Z=-z} = \frac{1}{2}[Q(\frac{\alpha - z}{\sigma}) + Q(\frac{\alpha + z}{\sigma})].$$

Hence, the probability of bit error is

$$\begin{aligned} P_b &= P_{b|Z=0}P(Z = 0) + P_{b|Z=z}P(Z = z) + P_{b|Z=-z}P(Z = -z) \\ &= \frac{1}{4}[2Q(\frac{\alpha}{\sigma}) + Q(\frac{\alpha + z}{\sigma}) + Q(\frac{\alpha - z}{\sigma})]. \end{aligned}$$

$\square$

BPSK in a slow Rayleigh fading channel:

Consider a flat Rayleigh fading channel with gain $Z(t) = \alpha(t)e^{j\theta(t)}$.

Assumption: $\theta(t)$ can be estimated accurately and then removed at the receiver.

Given the transmitted BPSK signal $x(t)$, $t \in [0, T]$, the received signal is

$$r(t) = \alpha x(t) + n(t).$$

If the symbol energy of $x(t)$ is E_b, then the symbol energy of $r(t)$ is $\alpha^2 E_b$.

$\Longrightarrow$ The received instantaneous SNR is $\alpha^2 E_b / N_0$.

$\Longrightarrow$ Given α, the BER is

$$P_{b|\alpha} = Q(\sqrt{\frac{2\alpha^2 E_b}{N_0}})$$

For a Rayleigh fading channel, α follows a Rayleigh distribution with pdf

$$f_\alpha(x) = \begin{cases} \frac{x}{\sigma_\alpha^2} \exp(-\frac{x^2}{2\sigma_\alpha^2}), & x \geq 0 \\ 0, & x < 0 \end{cases},$$

where $2\sigma_\alpha^2 = E(\alpha^2)$.

$\Longrightarrow$

$$
\begin{aligned}
P_b &= \int_0^\infty Q(\sqrt{2x^2\gamma_b})\frac{x}{\sigma_\alpha^2}\exp(-\frac{x^2}{2\sigma_\alpha^2})dx \\
&= \frac{1}{2\sigma_\alpha^2}\int_0^\infty Q(\sqrt{2y\gamma_b})\exp(-\frac{y}{2\sigma_\alpha^2})dy \qquad (\text{where } y = x^2) \\
&= \frac{1}{2\sigma_\alpha^2}\int_0^\infty [\frac{1}{\sqrt{2\pi}}\int_{\sqrt{2y\gamma_b}}^\infty \exp(-u^2/2)du]\exp(-\frac{y}{2\sigma_\alpha^2})dy \\
&= \frac{1}{2\sigma_\alpha^2}\frac{1}{\sqrt{2\pi}}\int_0^\infty \exp(-u^2/2)[\int_0^{u^2/2\gamma_b}\exp(-\frac{y}{2\sigma_\alpha^2})dy]du \\
&= \frac{1}{2}[1-\sqrt{\frac{\overline{\gamma}_b}{1+\overline{\gamma}_b}}].
\end{aligned}
$$

For $x \ll 1$, $\frac{1}{\sqrt{1+x}} \approx 1 - \frac{x}{2}$.
$\Longrightarrow$ For $\bar{\gamma}_b \gg 1$:

$$P_b \approx \frac{1}{4\bar{\gamma}_b}.$$

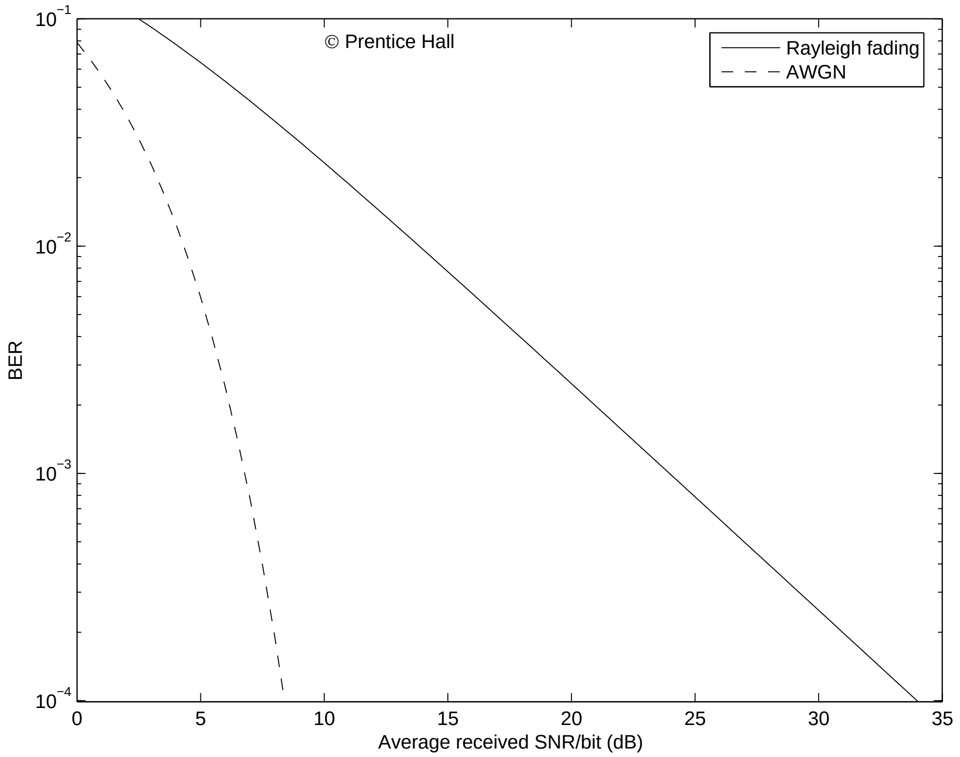


Figure 19: BER performance over flat Rayleigh fading channels

Example 3.8 BPSK in a fading channel:

Consider digital transmission via BPSK over a fading channel. At any instant, the channel provides a gain of 1.0 with a probability of 0.9 and a gain of 0.05 with a probability of 0.1. Derive the BER under the assumption of coherent detection.

Solution:

The BER performance of BPSK over the AWGN channel is

$$P_{b,AWGN} = Q(\sqrt{\frac{2E_b}{N_0}}).$$

Given that the channel gain is equal to 1.0, the conditional BER is $P_{b1} = P_{b,AWGN}$; Given that the channel gain is equal to 0.05, the conditional BER is $P_{b2} = Q(\sqrt{\frac{0.05^2 \times 2E_b}{N_0}})$. The overall BER is then

$$\begin{aligned} P_b &= P(\text{bit error}|\text{gain} = 1.0)P(\text{gain} = 1.0) + P(\text{bit error}|\text{gain} = 0.05)P(\text{gain} = 0.05) \\ &= 0.9P_{b1} + 0.1P_{b2} \\ &= 0.9Q(\sqrt{\frac{2E_b}{N_0}}) + 0.1Q(\sqrt{\frac{0.05^2 \times 2E_b}{N_0}}). \end{aligned}$$

Given that the transmitted bit energy is E_b, the average received bit energy is $\bar{E}_b = (1.0^2 E_b) \times 0.9 + (0.05^2 E_b) \times 0.1 = 0.90025 E_b$. The BER, in terms of the received SNR/bit, $\bar{\gamma}_b = \bar{E}_b / N_0$, is then

$$P_b = 0.9Q(\sqrt{2\bar{\gamma}_b/0.90025}) + 0.1Q(\sqrt{2\bar{\gamma}_b \times 0.0025/0.90025}).$$

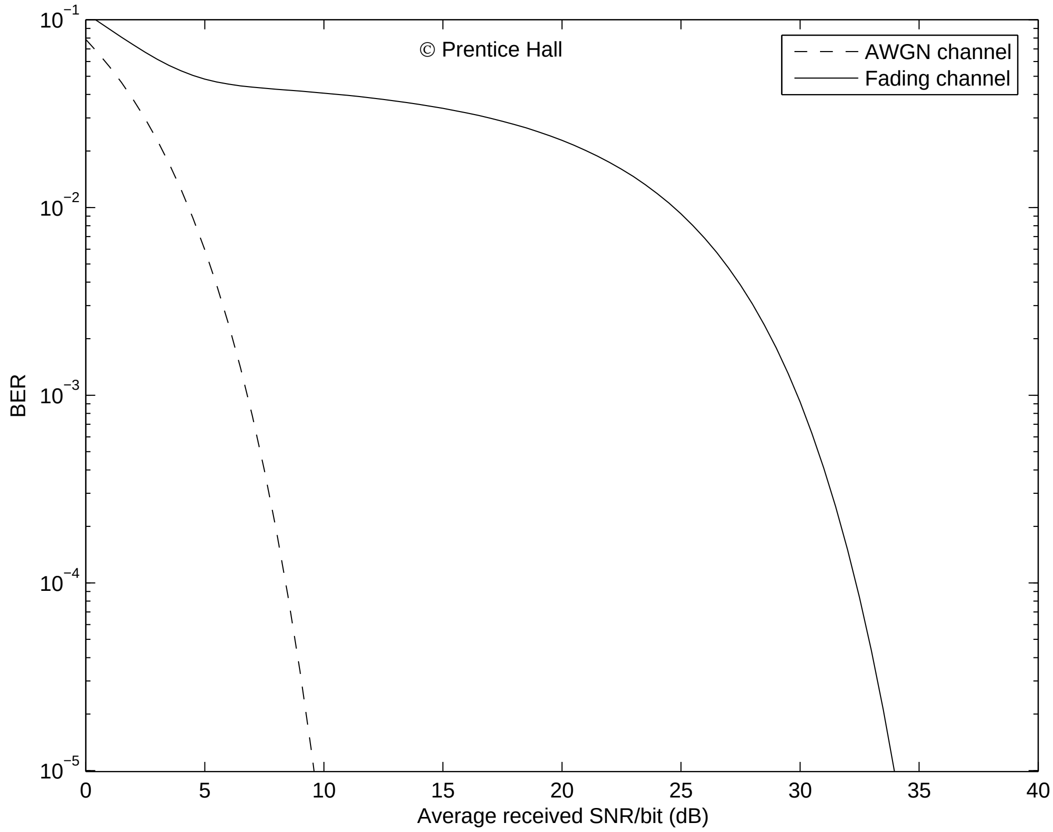


Figure 20: BER performance over the fading channel

□

Ch.4 Receiver Techniques for Fading Dispersive Channels

Techniques to mitigate the impairments due to a wireless fading dispersive channel:

- fading $\longrightarrow$ diversity;
- ISI (due to time dispersion) $\longrightarrow$ adaptive channel equalization in FDMA and TDMA and rake receiver in CDMA
- Power control to overcome the near-far effect
- Spread spectrum and Rake receiver
- Channel coding and interleaving
- Tellis-coded modulation and space-time coding

Principles of Diversity:

- more than one version of the transmitted signal
- uncorrelated channel fading in different diversity branches
- reduced chances of all the branches in deep fading simultaneously

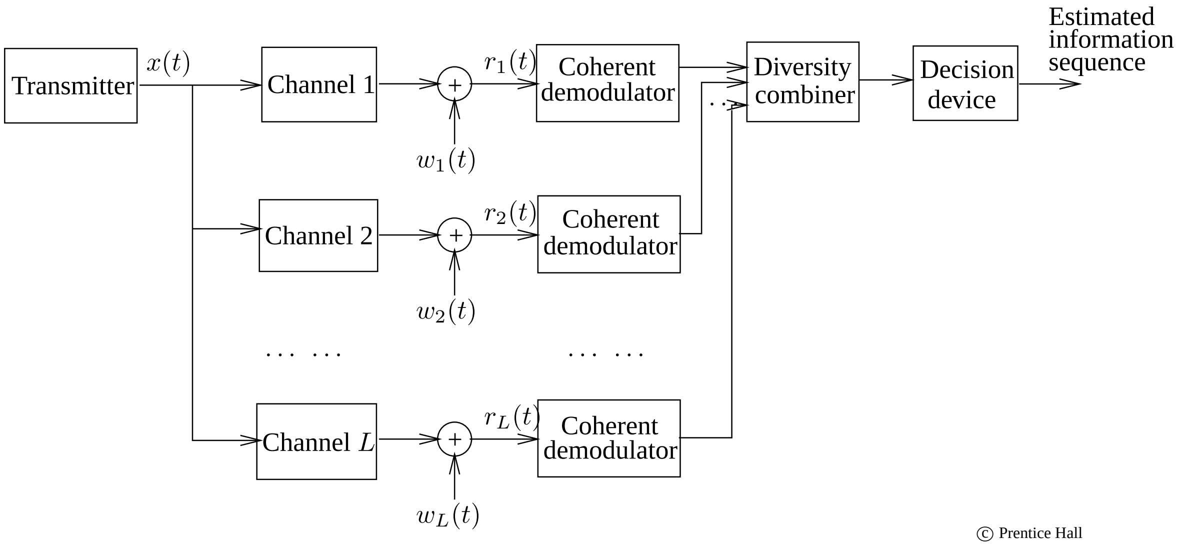


Figure 1: Illustration of diversity with coherent demodulation

Diversity Schemes

- **Frequency diversity**:
  - multiple transmission of the same information over different frequency slots
  - the separation between adjacent frequency slots should be larger than the channel coherence bandwidth $\Longrightarrow$ Channel fading over different slots is uncorrelated. If one frequency slot is in deep fading, then there is a large chance that other frequency slots are not in deep fading.

- **Time diversity**:
  - multiple transmission of the same information over different time slots
  - the separation between adjacent time slots should be larger than the channel coherence time

- **Space diversity**:
  - transmission using multiple transmitting and/or receiving antennas
  - the separation between adjacent antennas should be separated far enough to achieve uncorrelated channel gains. For example, for a Rayleigh fading channel, if the separation is $\lambda/2$ (where λ is the carrier wavelength), then the two channel gains are uncorrelated.

Combining Techniques

• **Maximal ratio combining**:

- Given the channel gains, $\alpha_l(t)\exp[j\theta_l(t)]$, $l = 1, 2, \ldots, L$, to remove the phase distortion introduced by the channel so that signals can be combined coherently, the weighting factors should have the component $\exp[-j\theta_l(t)]$;
- If the amplitude gain, $\alpha_l(t)$, of a channel is larger, then the received signal has a larger SNR, and more weight should be put on the received signal with better quality. $\Longrightarrow$ The weighting factor should be proportional to $\alpha_l(t)$.

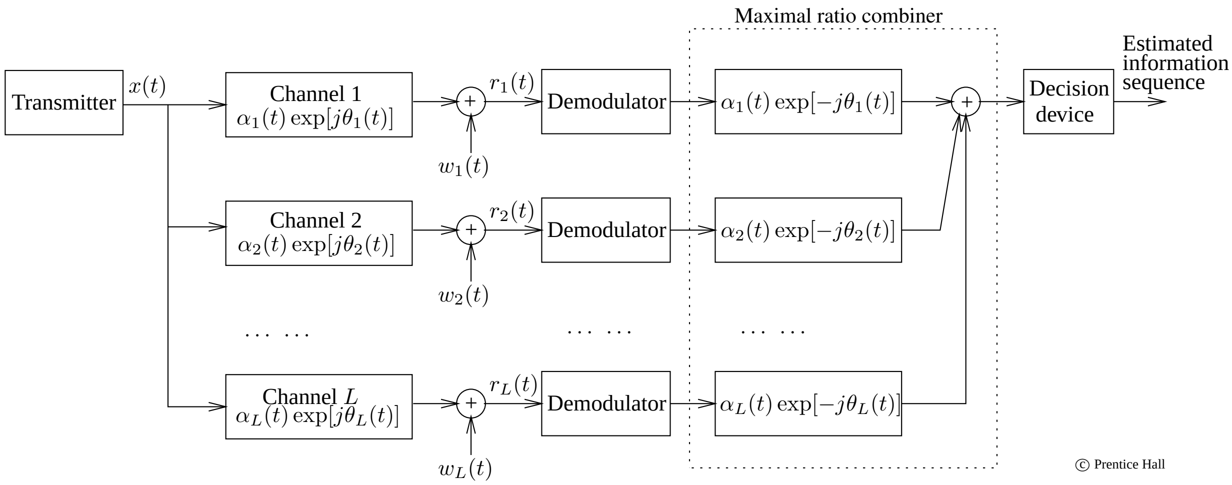


Figure 2: Diversity reception with maximal ratio combining

The maximal ratio combining achieves the best transmission performance at the cost of receiver complexity, as the receiver needs to estimate $\alpha_l(t)$ and $\theta_l(t)$.

- **Equal gain combining**:

  - The received signals from the L diversity branches are combined coherently and weighted equally $\Longrightarrow$ The weighting factors in the combiner are $\exp[-j\theta_l(t)]$, $l = 1, 2, \ldots, L$;
  - The receiver does not need the information of $\alpha_l(t)$;
  - The transmission performance is slightly degraded as compared with that of the maximal ratio combining.

- **Selective combining**:

  - The receiver monitors the SNR of the received signal from each diversity branch;
  - At any t, only the signal with the maximum SNR is used for detection;
  - It is easy to implement, especially for diversity over the reverse link where on mobile transmits and the received signals from several base stations need to be combined for detection;
  - The performance is slightly degraded from that with equal gain combining.

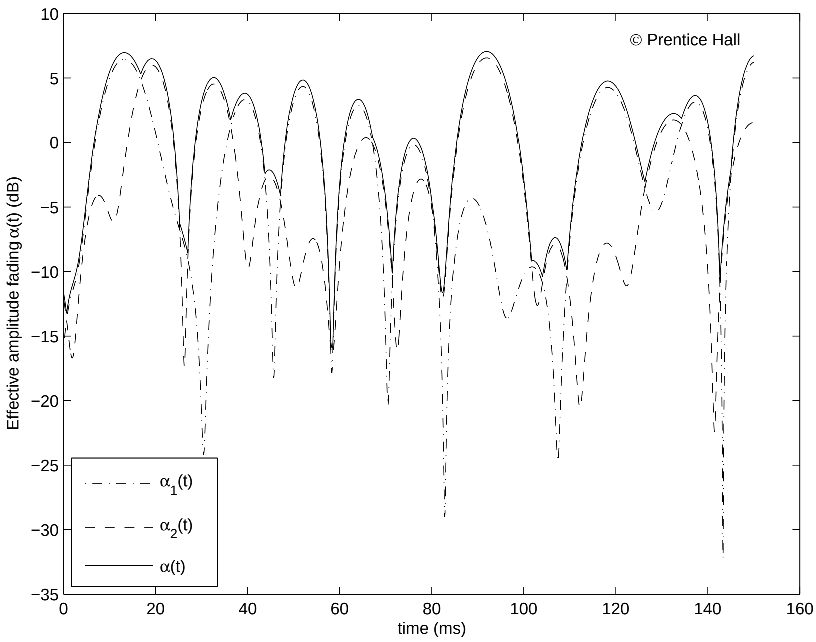


Figure 3: Illustration of amplitude fading with $L = 2$ selective diversity

Example 4.1 SNR improvement by selective diversity over Rayleigh fading
Consider Lth-order diversity with selective combining in a Rayleigh fading propagation environment. Each diversity channel exhibits independent and identically distributed (iid) fading. Find the pdf of the SNR in the diversity reception.

Solution:

At any instant, the amplitude fading of the lth channel, α_l, $l \in \{1, 2, \ldots, L\}$, has a pdf of

$$f_\alpha(x) = \begin{cases} \frac{x}{\sigma_\alpha^2} \exp(-\frac{x^2}{2\sigma_\alpha^2}), & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases},$$

where $\sigma_\alpha^2 = \frac{1}{2}E(\alpha^2)$. Let $\gamma_l = \alpha_l^2 E_b/N_0$ be the received SNR per bit of the lth channel at any instant, then the pdf of γ_l is

$$\begin{aligned} f_{\gamma_l}(y) &= f_{\alpha_l}(x)|dx/dy| \\ &= \frac{x}{\sigma_\alpha^2} \exp(-\frac{x^2}{2\sigma_\alpha^2})(2x\frac{E_b}{N_0})^{-1} \\ &= \frac{1}{2\sigma_\alpha^2(E_b/N_0)} \exp(-\frac{y}{2\sigma_\alpha^2(E_b/N_0)}) \\ &= \frac{1}{\Gamma_c} \exp(-\frac{y}{\Gamma_c}), \quad y \geq 0 \end{aligned}$$

which is an exponential distribution with mean $\Gamma_c = 2\sigma_\alpha^2(E_b/N_0) = E(\gamma_l)$ for $l \in \{1, 2, \ldots, L\}$. With selective diversity, at any instant, the effective received SNR per bit, γ, is the one from the strongest received signal,

$$\gamma = \max\{\gamma_1, \gamma_2, \ldots, \gamma_L\}.$$

For $x \geq 0$, γ has the following cdf

$$\begin{aligned} F_\gamma(x) &= P(\gamma \leq x) \\ &= P(\gamma_1 \leq x \ \cap \ \gamma_2 \leq x \ \cap \ \ldots \ \cap \ \gamma_L \leq x) \\ &= \prod_{l=1}^{L} P(\gamma_l \leq x) \quad \text{with independently faded channels} \\ &= [\int_0^x f_{\gamma_l}(z)dz]^L \quad \text{with identically distributed fading} \\ &= [1 - \exp(-\frac{x}{\Gamma_c})]^L, \end{aligned}$$

and for $x < 0$, the cdf of γ is $F_\gamma(x) = 0$. The pdf of γ is then

$$f_\gamma(x) = \frac{dF_\gamma(x)}{dx} = \frac{L}{\Gamma_c} \exp(-\frac{x}{\Gamma_c})[1 - \exp(-\frac{x}{\Gamma_c})]^{L-1}, \quad x \geq 0.$$

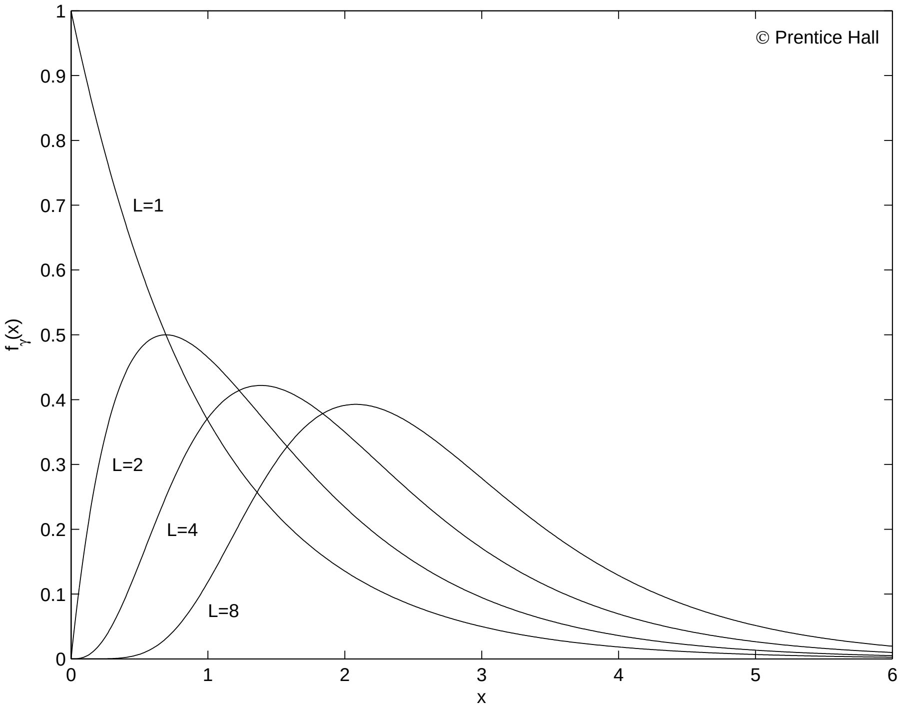


Figure 4: Probability density function of the received SNR per bit with selective diversity

Performance Improvement by Diversity

Assumptions:

- BPSK with coherent detection
- flat slow Rayleigh fading channel
- maximal ratio combining

The decision variable for the kth transmitted symbol can be represented by

$$\begin{aligned} r_k &= \sum_{l=1}^{L} [\alpha_{lk} \exp(-j\theta_{lk})][\alpha_{lk} \exp(j\theta_{lk}) x_k + n_{lk}] \\ &= [\sum_{l=1}^{L} \alpha_{lk}^2] x_k + [\sum_{l=1}^{L} \alpha_{lk} \exp(-j\theta_{lk}) n_{lk}] \\ &= g_k x_k + n_k, \end{aligned}$$

where where $x_k = \sqrt{E_b}$ for symbol "1" and $x_k = -\sqrt{E_b}$ for symbol "0",

$$g_k = \sum_{l=1}^{L} \alpha_{lk}^2, \quad n_k = \sum_{l=1}^{L} \alpha_{lk} \exp(-j\theta_{lk}) n_{lk}.$$

Therefore, given the weighting gain $\alpha_{lk} \exp(-j\theta_{lk})$, $l = 1, 2, \ldots, L$, the noise component n_k is a Gaussian random variable with zero mean and variance

$$\sigma_{k,n}^2 = \frac{N_0}{2} \sum_{l=1}^{L} \alpha_{lk}^2.$$

The SNR per bit at the output of the combiner for the kth symbol is then

$$\gamma_k = \frac{[g_k x_k]^2}{2\sigma_{k,n}^2} = \frac{[\sum_{l=1}^{L} \alpha_{lk}^2 x_k]^2}{N_0 \sum_{l=1}^{L} \alpha_{lk}^2} = \frac{E_b}{N_0} \sum_{l=1}^{L} \alpha_{lk}^2$$

where $\frac{E_b}{N_0}$ is the SNR value for the AWGN channel with $\alpha_{lk} = 1$ and $L = 1$.

In a Rayleigh fading environment, the α_{lk}'s are iid Rayleigh random variables with parameter σ_α^2. Therefore, γ_k follows a chi-square distribution with $2L$ degrees of freedom. Its pdf is given by

$$f_\gamma(x) = \frac{x^{L-1} \exp(-x/\Gamma_c)}{(L-1)!\Gamma_c^L}, \quad x \geq 0,$$

where $\Gamma_c = 2\sigma_\alpha^2 E_b/N_0$ is the average SNR per bit in each diversity channel. The mean SNR per bit after the combining is

$$\Gamma_b = E[\gamma_k] = L\Gamma_c,$$

which increases linearly with L.

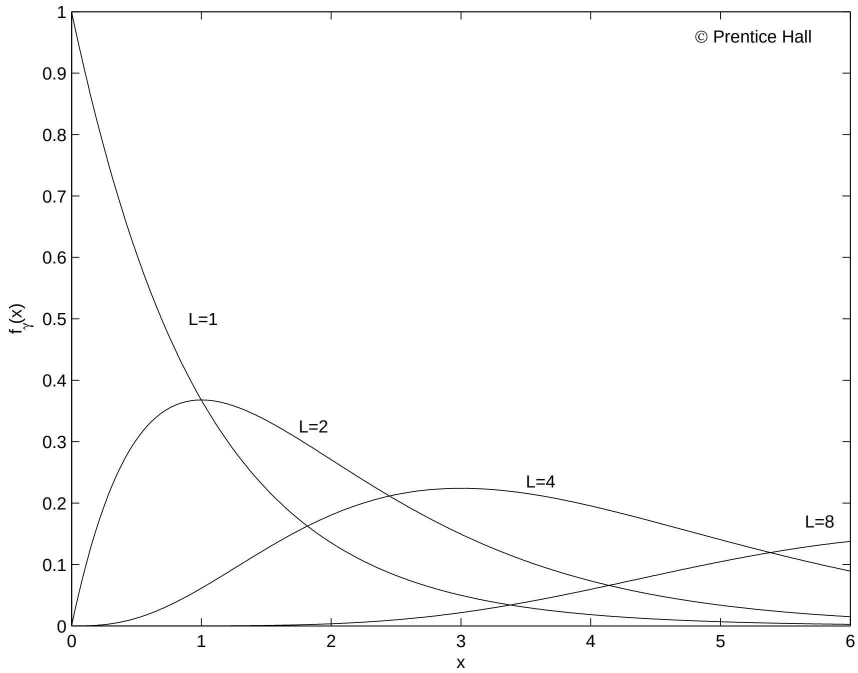


Figure 5: Probability density function of the received SNR per bit with maximal ratio combining

The analysis of the probability of bit error over a Rayleigh fading channel can be extended to the case with diversity, so that

$$\begin{aligned} P_b &= \int_0^\infty P_{e|\gamma}(x) f_\gamma(x) dx \\ &= \int_0^\infty Q(\sqrt{2x}) \frac{x^{L-1} \exp(-x/\Gamma_c)}{(L-1)!\Gamma_c^L} dx \\ &= [0.5(1-\mu)]^L \sum_{l=0}^{L-1} \binom{L-1+l}{l} [0.5(1+\mu)]^l \end{aligned}$$

where

$$\mu = \sqrt{\frac{\Gamma_c}{1+\Gamma_c}}.$$

For $\Gamma_c \gg 1$, we have $0.5(1+\mu) \approx 1$ and $0.5(1-\mu) \approx 1/4\Gamma_c$. Furthermore,

$$\sum_{l=0}^{L-1} \binom{L-1+l}{l} = \binom{2L-1}{L}.$$

As a result, for $\Gamma_c \gg 1$, the probability of error can be approximated by

$$P_b \approx (\frac{1}{4\Gamma_c})^L \binom{2L-1}{L}.$$

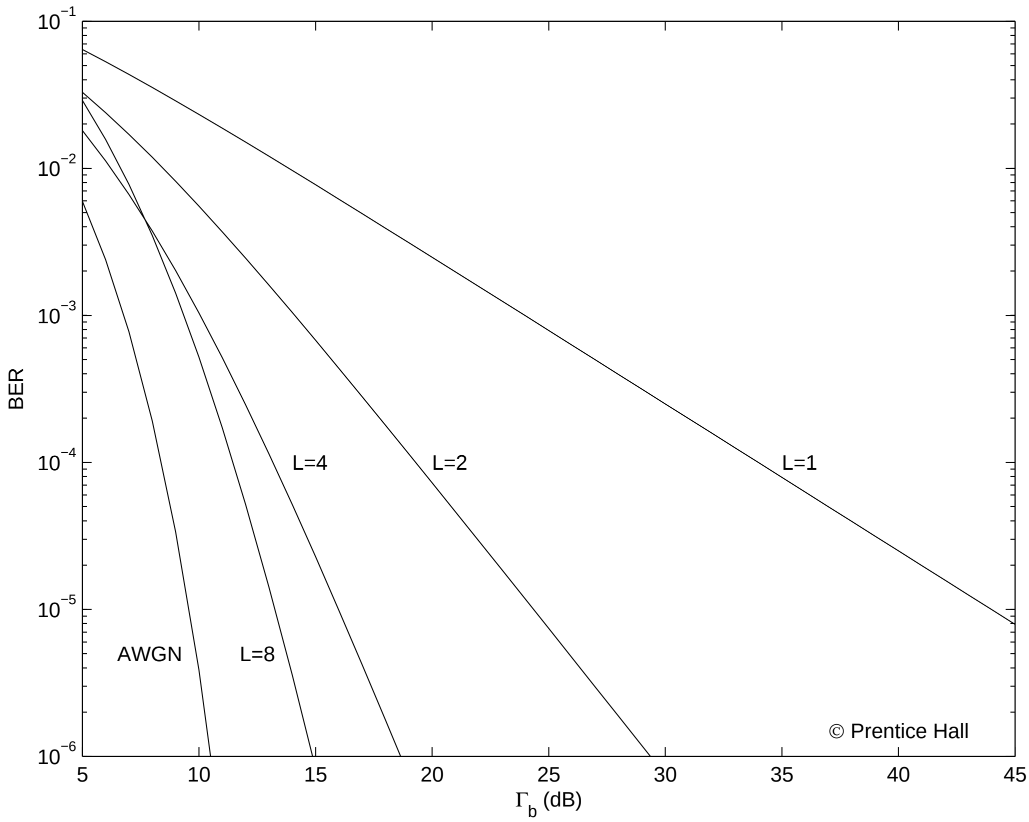


Figure 6: BER performance improvement using the diversity with maximal ratio combining

Example 4-2: BPSK with maximal-ratio diversity combining
Consider digital transmission via BPSK with second-order diversity. The channel gain in each branch takes on a value of 1.0, with probability 0.9, and a value of 0.05, with probability 0.1, over each symbol interval. The additive noise in each branch is white and Gaussian with two-sided psd equal to $N_0/2$. The two branches have independent channel gain and noise component. Derive the BER performance for maximal ratio combining.

Solution:

Let g_1 and g_2 denote the channel gains of the first and second branches, respectively, over each symbol interval. With maximal ratio combining, the effective channel gain over the symbol interval is

$$g = g_1^2 + g_1^2,$$

and the additive noise component in the decision device input is a Gaussian random variable with zero mean and variance

$$\sigma_N^2 = \frac{N_0}{2}(g_1^2 + g_2^2) = \frac{gN_0}{2}.$$

Let E_b denote the transmitted bit energy. The conditional probability of bit error is

$$P(e|g_1, g_2) = Q(\sqrt{\frac{2g^2E_b}{gN_0}}) = Q(\sqrt{\frac{2gE_b}{N_0}}).$$

The probability of bit error is then

$$\begin{aligned} P_b &= \sum_{x_i}\sum_{y_j} P(e|g_1 = x_i, g_2 = y_j)P(g_1 = x_i, g_2 = y_j) \\ &= 0.9^2 P(e|g_1 = 1.0, g_2 = 1.0) + 0.1^2 P(e|g_1 = 0.05, g_2 = 0.05) \\ &\quad + 0.9 \times 0.1 P(e|g_1 = 1.0, g_2 = 0.05) + 0.1 \times 0.9 P(e|g_1 = 0.05, g_2 = 1.0) \\ &= 0.81Q(\sqrt{\frac{4E_b}{N_0}}) + 0.01Q(\sqrt{\frac{0.01E_b}{N_0}}) + 0.18Q(\sqrt{\frac{2.005E_b}{N_0}}). \end{aligned}$$

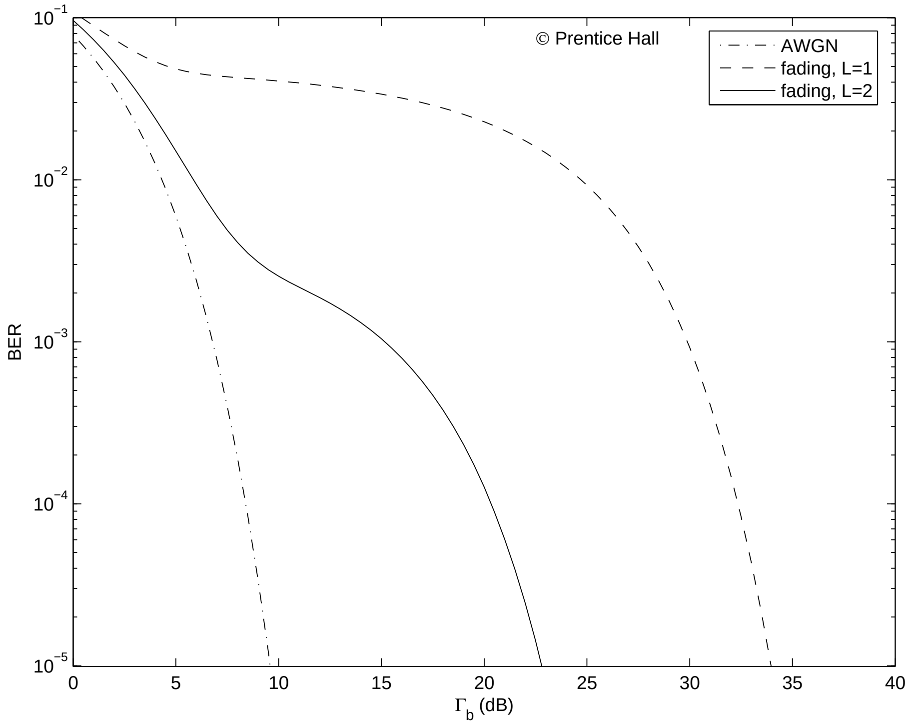


Figure 7: BER performance improvement via diversity

□

Channel Equalization

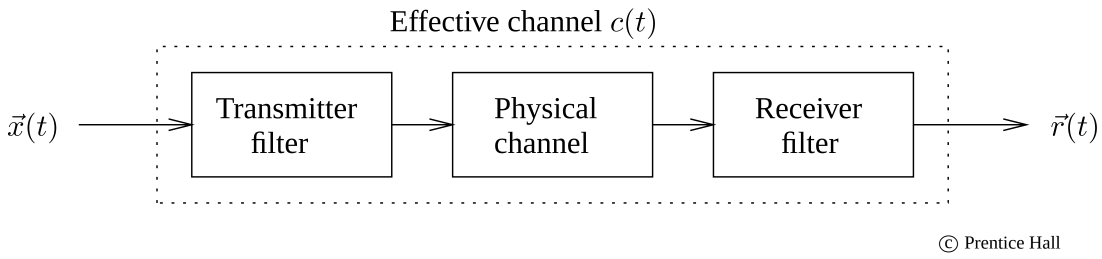


Figure 8: The effective channel with impulse response $c(t)$

The sampled received signal at $t = nT$ is

$$\begin{aligned}
\vec{r}(t)|_{t=nT} &= [\vec{x}(t) \star c(t) + \vec{n}(t)]|_{t=nT} \\
&= \{[\sum_{l=-\infty}^{\infty} \vec{x}_l \delta(t - lT)] \star c(t) + \vec{n}(t)\}|_{t=nT} \\
&= \{\sum_{l=-\infty}^{\infty} \vec{x}_l c(t - lT) + \vec{n}(t)\}|_{t=nT} \\
&= \sum_{l=-\infty}^{\infty} \vec{x}_l c(nT - lT) + \vec{n}(nT),
\end{aligned}$$

where $\vec{n}(t)$ is the filtered Gaussian noise vector.

In discrete-time sequence representation,

$$\begin{aligned}\vec{r}_n &= \sum_{l=-\infty}^{\infty} \vec{x}_l c_{n-l} + \vec{n}_n \\ &= \vec{x}_n \star c_n + \vec{n}_n \\ &= c_0 \vec{x}_n + \sum_{l=-\infty, l\neq n}^{\infty} \vec{x}_l c_{n-l} + \vec{n}_n.\end{aligned}$$

To minimize the probability of transmission error, the optimum receiver (in a mean-square error sense) consists of

- a matched filter, matched to the transmitter filter and the physical channel in tandem, to collect all the received signal energy
- an equalizer, to overcome the ISI
- a maximum likelihood decision device, to minimize the impact of the additive Gaussian noise

Example 4.5: Effect of ISI on transmission accuracy
In a digital transmission system using BPSK, the matched filter receiver designed for an AWGN channel is used. The physical channel introduces an additive Gaussian noise of zero mean and two-sided psd $N_0/2$. In addition, it introduces ISI. At the end of the nth symbol interval, the output of the demodulator is

$$r_n = x_n + 0.5x_{n-1} + n_n,$$

where the desired signal component $x_n = \sqrt{E_b}$ if symbol "1" was sent and $x_n = -\sqrt{E_b}$ if symbol "0" was sent, E_b is the transmitted signal symbol energy, and the noise component n_n is a Gaussian random variable with zero mean and variance $N_0/2$. Determine the probability of transmission error.

Solution:

Without ISI, the probability of transmission error is

$$P_b = Q(\sqrt{\frac{2E_b}{N_0}}).$$

With ISI, the effect of ISI depends on whether or not the previous and the current transmitted symbols are the same.

Table 1: The demodulator output at the end of current symbol in the absence of the noise component

previous symbol	current symbol	demodulator output without ISI	demodulator output with ISI
"0"	"0"	$-\sqrt{E_b}$	$-1.5\sqrt{E_b}$
"0"	"1"	$\sqrt{E_b}$	$0.5\sqrt{E_b}$
"1"	"0"	$-\sqrt{E_b}$	$-0.5\sqrt{E_b}$
"1"	"1"	$\sqrt{E_b}$	$1.5\sqrt{E_b}$

$$\begin{aligned}
P_b &= \sum P(\text{error}|\text{previous symbol, current symbol})P(\text{previous symbol, current symbol}) \\
&= \frac{1}{4}[P(\text{error}|\text{“0”},\text{“0”}) + P(\text{error}|\text{“0”},\text{“1”}) + P(\text{error}|\text{“1”},\text{“0”}) + P(\text{error}|\text{“1”},\text{“1”})] \\
&= \frac{1}{4}[P(-1.5\sqrt{E_b} + n_n > 0) + P(0.5\sqrt{E_b} + n_n < 0) \\
&\quad + P(-0.5\sqrt{E_b} + n_n > 0) + P(1.5\sqrt{E_b} + n_n < 0)] \\
&= \frac{1}{2}[Q(1.5\sqrt{\frac{2E_b}{N_0}}) + Q(0.5\sqrt{\frac{2E_b}{N_0}})].
\end{aligned}$$

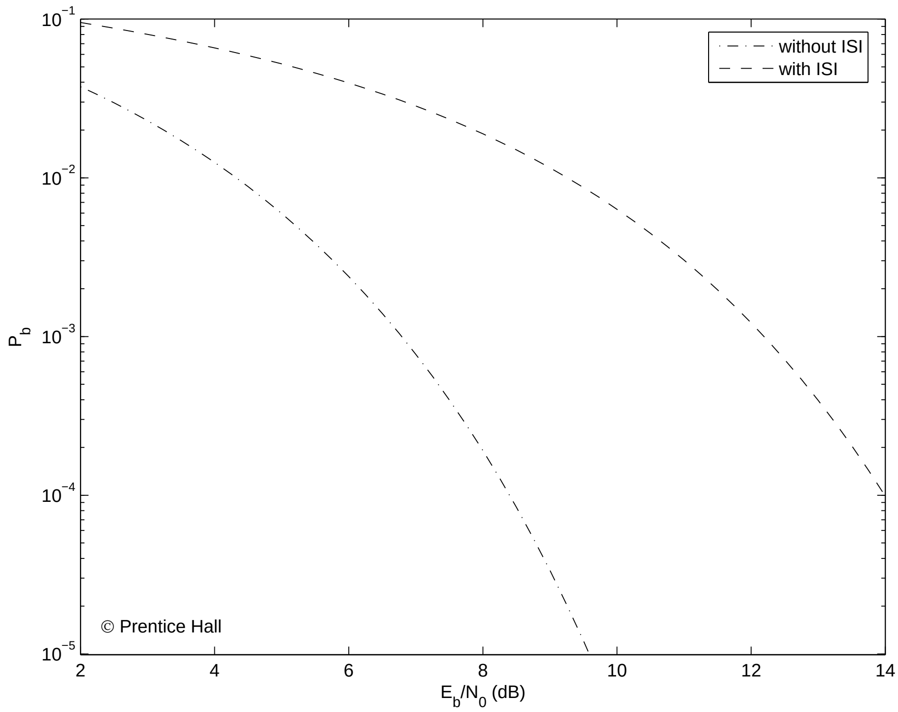


Figure 9: Comparison of the transmission accuracy with and without ISI

Linear Equalization

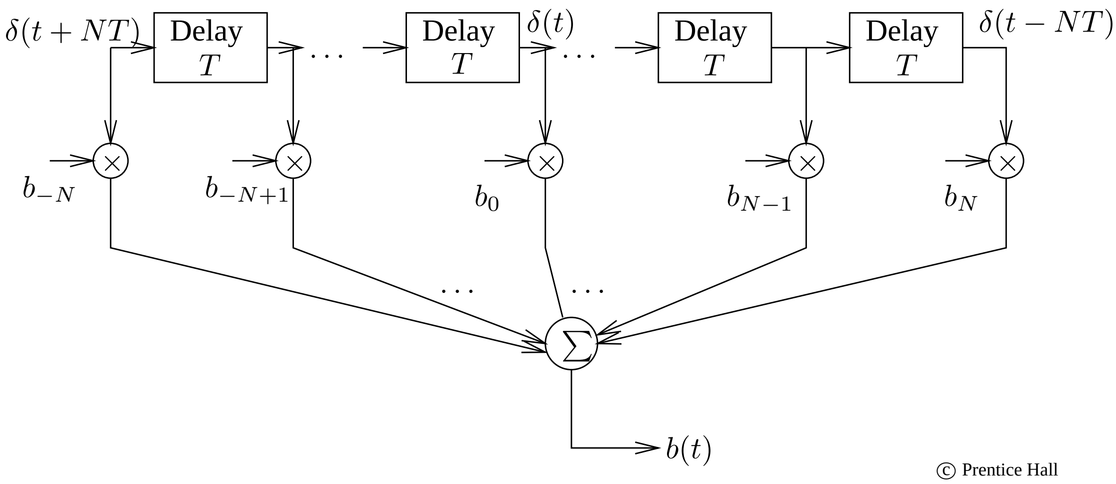


Figure 10: The tapped-delay-line linear equalizer

The impulse response of the equalizer is

$$b(t) = \sum_{k=-N}^{N} b_k \delta(t - kT).$$

The transfer function of the equalizer in the z domain is given by

$$B(z) = \sum_{k=-N}^{N} b_k z^{-k}.$$

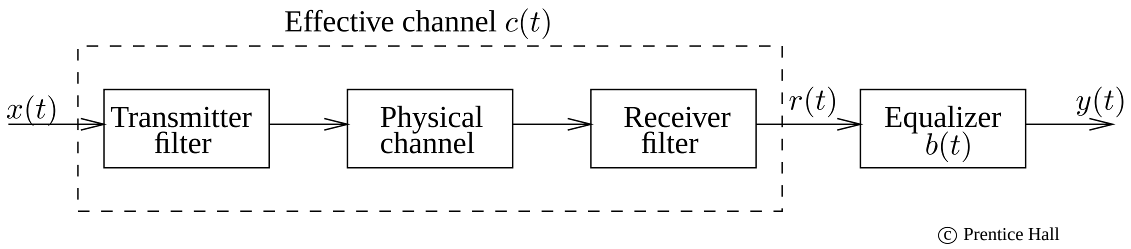


Figure 11: The equalized system

Let $R(z) = \sum_{n=-\infty}^{\infty} r_n z^{-n}$ denote the z transform of the discrete time sequence $\{r_n\}$.

The output of the equalizer in the z domain is then

$$\begin{aligned} Y(z) &= R(z)B(z) \\ &= \sum_{n=-\infty}^{\infty} r_n z^{-n} \sum_{k=-N}^{N} b_k z^{-k} \\ &= \sum_{n=-\infty}^{\infty} \sum_{k=-N}^{N} b_k r_n z^{-(k+n)}. \end{aligned}$$

Letting $l = k + n$, the above can be rearranged as

$$\begin{aligned} Y(z) &= \textstyle\sum_{l=-\infty}^{\infty} \sum_{k=-N}^{N} b_k r_{l-k} z^{-l} \\ &= \textstyle\sum_{l=-\infty}^{\infty} y_l z^{-l}, \end{aligned}$$

where

$$y_l = \sum_{k=-N}^{N} b_k r_{l-k}.$$

• Zero-forcing (ZF) linear equalizer

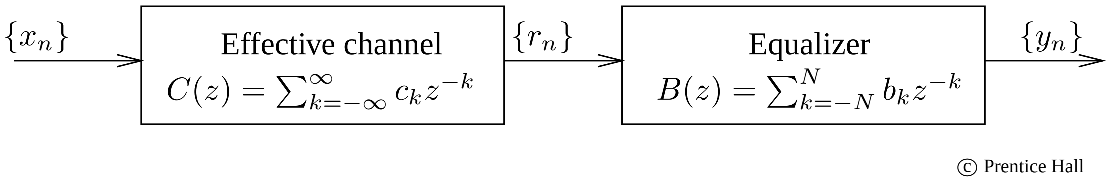


Figure 12: The equivalent discrete-time representation of the equalized system

The output of the effective channel at $t = nT$ is

$$r_n = x_n \star c_n.$$

$\Longrightarrow$

$$R(z) = X(z)C(z)$$

where

$$X(z) = \sum_{n=-\infty}^{\infty} x_n z^{-n}$$

$$C(z) = \sum_{n=-\infty}^{\infty} c_n z^{-n}$$

$\Longrightarrow$

$$Y(z) = R(z)B(z) = X(z)C(z)B(z).$$

The transfer function of the equalized system is

$$H(z) = \frac{Y(z)}{X(z)} = C(z)B(z),$$

$\Longrightarrow$

$$h_n = b_n \star c_n = \sum_{k=-N}^{N} b_k c_{n-k}.$$

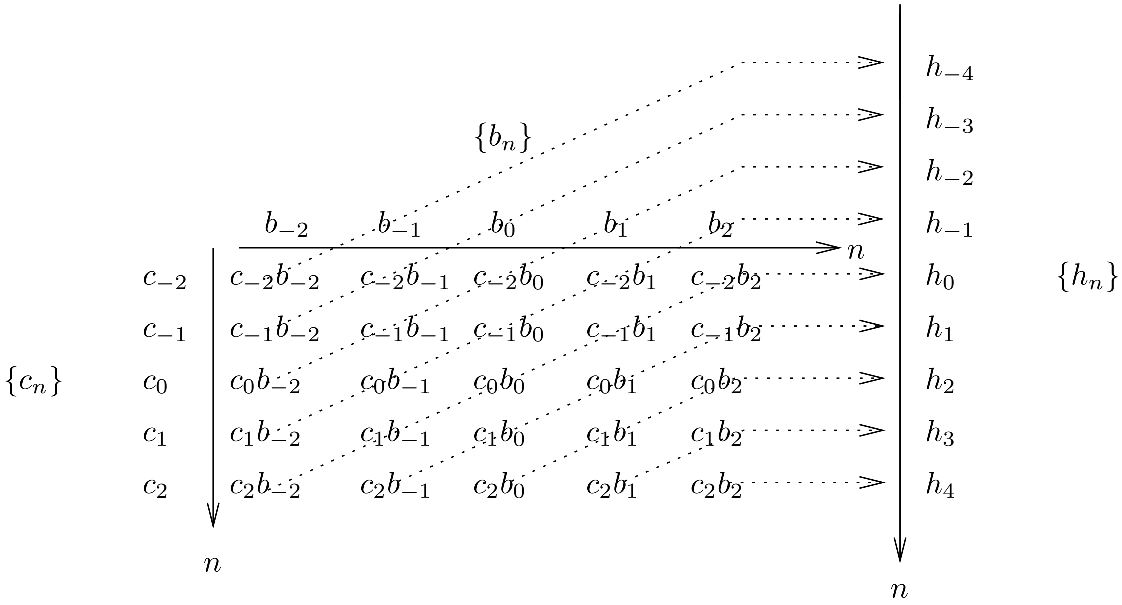


Figure 13: Convolution map for zero-forcing linear equalizer for $N = 2$

For ISI-free transmission,

$$H(z) = 1.$$

or,

$$h_n = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}.$$

Example 4.6: Zero-forcing linear equalizer with infinite taps
For the communication system in Example 4.5, design a linear equalizer to combat the ISI, and determine the probability of transmission error with the channel equalization.

Solution:

The effective channel has an impulse response (in the absence of additive noise) given by

$$c(t) = \delta(t) + 0.5\delta(t - T).$$

The channel transfer function is then

$$C(z) = \sum_{k=-\infty}^{\infty} c_k z^{-k} = 1 + 0.5z^{-1}.$$

To completely equalize the channel, the transfer function of the equalizer should be

$$\begin{aligned} B(z) &= \frac{1}{C(z)} \\ &= \frac{1}{1 + 0.5z^{-1}} \\ &= 1 - 0.5z^{-1} + 0.5^2 z^{-2} - \ldots + (-0.5z^{-1})^k + \ldots, \quad |0.5z^{-1}| < 1. \end{aligned}$$

The equalization function can be implemented by a tapped-delay-line linear filter with a large number of taps (to approximate the infinite number of taps).

In the presence of additive noise n_n at the end of the nth symbol interval at the effective channel output, the noise component at the equalizer output at the same instant is

$$\nu_n = n_n \star b_n = \sum_{k=0}^{\infty} (-0.5)^k n_{n-k}.$$

Since n_n is a zero-mean Gaussian random variable with variance $N_0/2$ and is independent from sample to sample, we have ν_n is also Gaussian with zero mean and variance equal to

$$\sigma_v^2 = \sum_{k=0}^{\infty} [(-0.5)^k]^2 (N_0/2) = 2N_0/3.$$

As a result, the probability of bit error with equalization is

$$P_b = Q(\sqrt{\frac{E_b}{2N_0/3}}) = Q(\sqrt{\frac{3E_b}{2N_0}}).$$

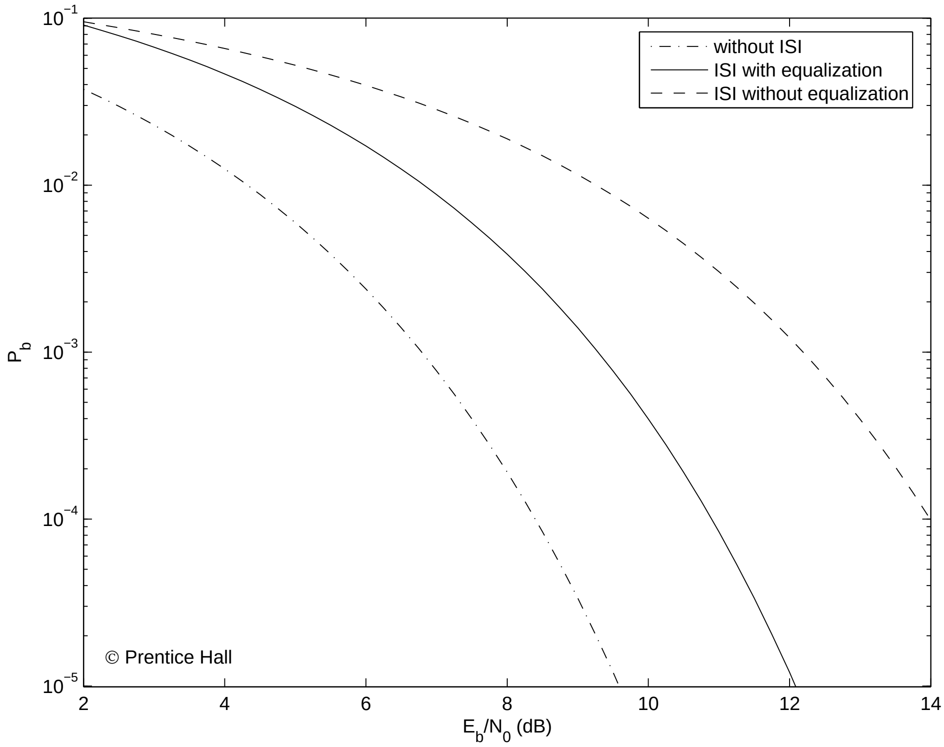


Figure 14: Comparison of the transmission accuracy with and without equalization

□

Example 4.7: Zero-forcing linear equalizer with finite taps
The impulse response of a dispersive effective channel is

$$c(t) = \exp(-\frac{|t|}{3T}), \quad -\infty < t < \infty,$$

where T is the transmitted symbol interval. Design a 3-tap zero-forcing linear equalizer for the channel.

Solution:

The discrete-time representation of the effective channel impulse response is $\{c_n\}$, where

$$c_n = c(t)|_{t=nT} = \exp(-\frac{|n|T}{3T}) = \exp(-\frac{|n|}{3}).$$

Let $\mathbf{b} = (b_{-1}, b_0, b_1)^T$ denote the tap coefficient vector of the 3-tap linear equalizer. The discrete-time representation of the equalized system impulse response is

$$h_n = c_n \star b_n = \sum_{k=-1}^{1} b_k c_{n-k}.$$

We want to determine $\mathbf{b}$ such that

$$h_n = \begin{cases} 1, & n = 0 \\ 0, & n = \pm 1 \end{cases}.$$

In matrix form, $\mathbf{b}$ can be computed by

$$\begin{bmatrix} 1 & \exp(-1/3) & \exp(-2/3) \\ \exp(-1/3) & 1 & \exp(-1/3) \\ \exp(-2/3) & \exp(-1/3) & 1 \end{bmatrix} \begin{bmatrix} b_{-1} \\ b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

This leads to

$$\begin{bmatrix} b_{-1} \\ b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 1 & \exp(-1/3) & \exp(-2/3) \\ \exp(-1/3) & 1 & \exp(-1/3) \\ \exp(-2/3) & \exp(-1/3) & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1.4726 \\ 3.1103 \\ -1.4726 \end{bmatrix}.$$

In fact, with the 3-tap equalizer, we have

$$h_n = \begin{cases} 1, & n = 0 \\ 0, & n = \pm 1, \pm 2, \pm 3, \pm 4, \pm 5, \pm 6, \pm 7, \pm 8 \end{cases}.$$

Therefore, the 3-tap linear equalizer is very effective in combating the ISI introduced by the effective channel. □

• Minimum mean-square error (MMSE) linear equalizer

A linear equalizer can significantly increase the impact of input additive noise on signal detection
$\Longrightarrow$ Choose the equalizer coefficients to minimize the mean squared error (MSE), which is the expected sum of the squares of all the ISI terms plus the noise power at the output of the equalizer
$\Longrightarrow$ An MMSE linear equalizer.

The output of the equalizer, sampled at $t = nT$, is

$$y_n = \sum_{k=-N}^{N} b_k r_{n-k}.$$

The desired output of the equalizer at $t = nT$ is x_n.

$\Longrightarrow$ The equalization error at $t = nT$ is

$$\epsilon_n = x_n - y_n.$$

The problem of choosing the equalizer tap coefficients is to solve the following minimization problem:

$$\min_{\mathbf{b}} E[\epsilon_n^2],$$

where $\mathbf{b} = (b_{-N}, \ldots, b_{-1}, b_0, b_1, \ldots, b_N)^T$ represents the linear equalizer of $2N+1$ taps, and the superscript T denotes matrix transposition.

The MSE of the equalization is

$$\begin{aligned}
\mathcal{E} &= E[\epsilon_n^2] \\
&= E[(x_n - \sum_{k=-N}^{N} b_k r_{n-k})^2] \\
&= \sum_{l=-N}^{N} \sum_{k=-N}^{N} b_k b_l R_r(k-l) - 2 \sum_{l=-N}^{N} b_l R_{xr}(k) + E(x_n^2),
\end{aligned}$$

where

$$\begin{aligned}
R_r(k-l) &= E[r_{n-l} r_{n-k}] \\
R_{xr}(k) &= E[x_n r_{n-k}]
\end{aligned}$$

and the expectation is taken with respect to the random information sequence corresponding to $\{x_n\}$ and the additive noise.

Solving for

$$\frac{\partial \mathcal{E}}{\partial b_k} = 0$$

leads to the necessary conditions for the minimum MSE as

$$\sum_{l=-N}^{N} b_l R_r(k-l) = R_{xr}(k), \quad k = 0, \pm 1, \pm 2, \ldots, \pm N.$$

• **Adaptive linear equalizer**

The characteristics of the wireless dispersive fading channel change randomly with time

$\Longrightarrow$ The equalizer coefficients should change according to the channel status so as to track the channel variations.

$\Longrightarrow$ An adaptive equalizer

A linear adaptive equalizer has the tap coefficients that minimize the MSE $\mathcal{E} = E[\epsilon_n^2]$.

$\Longrightarrow$

$$\frac{\partial \mathcal{E}}{\partial b_k} = E[2\epsilon_n \frac{\partial \epsilon_n}{\partial b_k}] = -2E[\epsilon_n r_{n-k}].$$

If the channel variation rate is much smaller than the symbol rate of the transmitted signal, the optimal tap coefficient vector $\mathbf{b}_{opt}$ can be determined iteratively under the assumption that $\mathbf{b}_{opt}$ does not change much over the period of the iterations.

$$\mathbf{b}(n+1) = \mathbf{b}(n) + \Delta \mathbf{b}(n),$$

$\Longrightarrow$

$$b_k(n+1) = b_k(n) + \alpha E[\epsilon_n r_{n-k}] \approx \alpha(\epsilon_n r_{n-k}), \quad k = -N, \cdots, -1, 0, 1, \cdots, N.$$

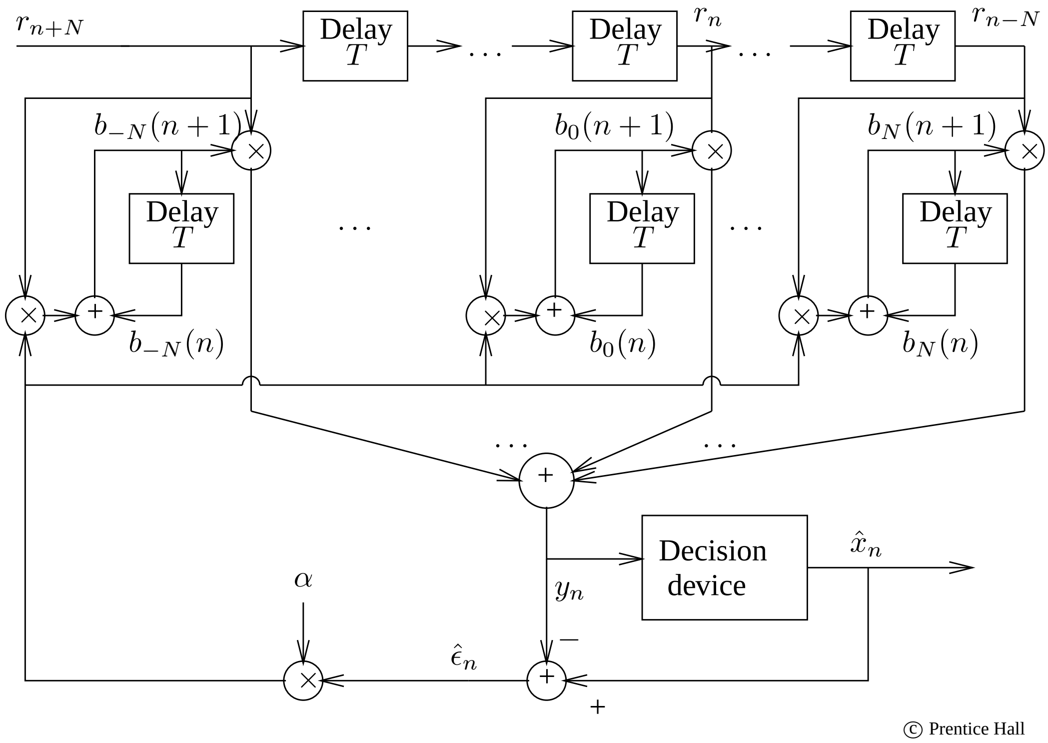


Figure 15: Decision-directed adaptive MSE linear equalizer

Decision Feedback Equalization

In the presence of channel noise, the effective channel output signal at $t = nT$ is

$$r_n = c_0 x_n + \sum_{l=-\infty}^{-1} c_l x_{n-l} + \sum_{l=1}^{\infty} c_l x_{n-l} + n_n.$$

- The first term represents the desired signal component with the channel gain c_0
- The second term is due to the precursors of the channel impulse response that occur before the main sample c_0 associated with the desired data symbol
- The third term is due to the postcursors of the channel impulse response that occur after the main sample c_0.

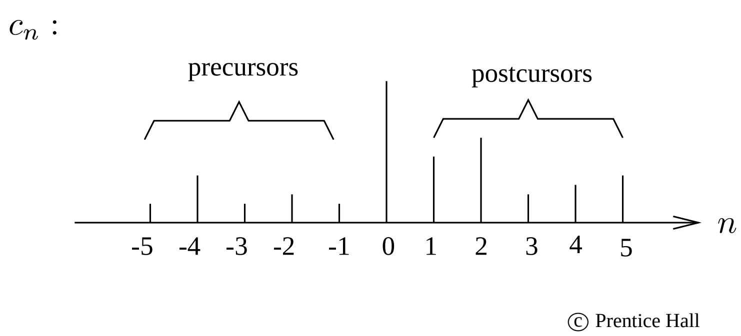


Figure 16: The discrete-time representation of the dispersive channel impulse response

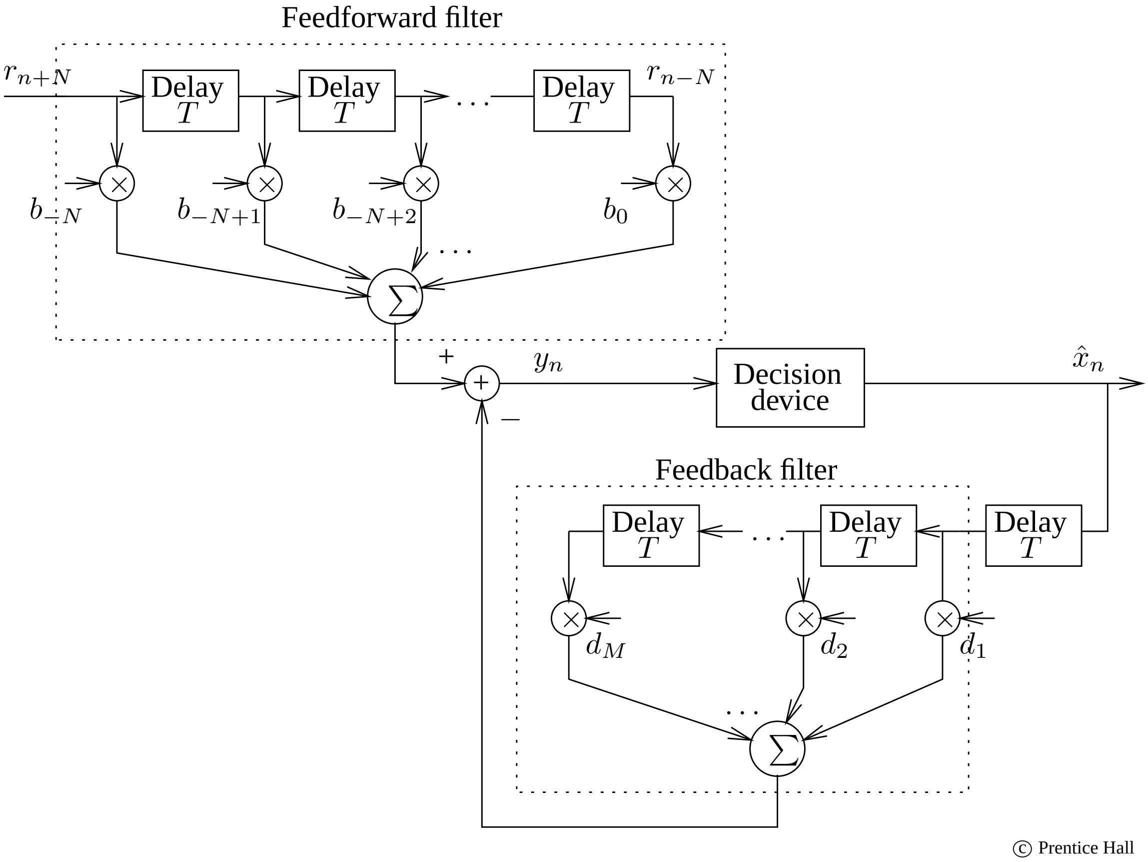


Figure 17: Structure of the DFE

• **Zero-forcing (ZF) DFE**

The transfer function of the channel is

$$C(z) = \sum_{k=-\infty}^{\infty} c_k z^{-k}.$$

The transfer function of the feedforward linear filter in the DFE is

$$B(z) = \sum_{k=-N}^{0} b_k z^{-k}.$$

The transfer function of the effective channel consisting of the dispersive channel (to be equalized) and the feedforward filter is then

$$C(z)B(z) = \sum_{n=-\infty}^{\infty} h_1(n) z^{-n},$$

where

$$h(n) = \sum_{k=-N}^{0} b_k c_{n-k}.$$

Under the constraint of the finite tap number $N+1$ $(> M)$, we can only force $h_1(n)$ to be zero for $n = -N, -N+1, \ldots, -2, -1$, by choosing the values of the tap coefficients according to the following equation

$$\begin{bmatrix} c_0 & c_{-1} & \ldots & c_{-N} \\ c_1 & c_0 & \ldots & c_{-N+1} \\ \vdots & \vdots & & \vdots \\ c_N & c_{N-1} & \ldots & c_0 \end{bmatrix} \begin{bmatrix} d_{-N} \\ d_{-N+1} \\ \vdots \\ d_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Also, in order to remove the postcursors by the feedback filter,

$$d_j = h_j, \quad j = 1, 2, \ldots, M.$$

$\Longrightarrow$ The equalizer output at $t = nT$ is

$$y_n = \sum_{k=-N}^{0} b_k r_{n-k} - \sum_{j=1}^{M} d_j \hat{x}_{n-j} = \sum_{k=-N}^{0} b_k r_{n-k} - \sum_{j=1}^{M} h_j \hat{x}_{n-j}.$$

Under the assumption that the residual precursors in the system impulse response after the feedforward filter are negligible, the output of the feedforward filter can be approximately given by

$$\sum_{k=-N}^{0} b_k r_{n-k} \approx x_n + \sum_{j=1}^{M} h_j x_{n-j} + v_n,$$

where $v_n = \sum_{k=-N}^{0} b_k n_{n-k}$ is due to the received noise components.

If we further assume $\hat{x}_{n-j} = x_{n-j}$, $j = 1, 2, \ldots, M$, then the DFE output is

$$\hat{x}_n = D[y_n] \approx D\left[(x_n + \sum_{j=1}^{M} h_j x_{n-j} + v_n) - (\sum_{j=1}^{M} h_j \hat{x}_{n-j})\right] = D[x_n + v_n]$$

where $D[\cdot]$ denotes the decision function.

If x_n is stronger than v_n, then $\hat{x}_n = x_n$.

Under the two assumptions, the equalization error is

$$\epsilon_n = x_n - y_n = v_n.$$

$\Longrightarrow$ The ISI due to the dispersive channel is completely eliminated.

• Minimum mean-square error (MMSE) DFE

Under the assumption of correct previously detected symbols, the output of the equalizer, sampled at $t = nT$, is

$$y_n = \sum_{k=-N}^{0} b_k r_{n-k} - \sum_{j=1}^{M} d_j x_{n-j}.$$

$\Longrightarrow$ The mean-square error is

$$\begin{aligned}\mathcal{E} &= E[\epsilon_n^2] \\ &= E[(x_n - y_n)^2] \\ &= E[(x_n - \sum_{k=-N}^{0} b_k r_{n-k} + \sum_{j=1}^{M} d_j x_{n-j})^2].\end{aligned}$$

To minimize the mean-square error $\mathcal{E}$, choose the tap coefficient b_k of the feedforward filter, $k = -N, -N+1, \ldots, 0$, so that

$$\begin{aligned}\frac{\partial \mathcal{E}}{\partial b_k} &= 2E[-r_{n-k}(x_n - \sum_{k=-N}^{0} b_k r_{n-k} - \sum_{j=1}^{M} d_j x_{n-j})] \\ &= -2E[r_{n-k}\epsilon_n] \\ &= 0,\end{aligned}$$

and choose the tap coefficient d_j of the feedback filter, $j = 1, 2, \ldots, M$, so that

$$\begin{aligned}\frac{\partial \mathcal{E}}{\partial b_{N+j}} &= 2E[x_{n-j}(x_n - \sum_{k=-N}^{0} b_k r_{n-k} - \sum_{j=1}^{M} d_j x_{n-j})] \\ &= 2E[x_{n-j}\epsilon_n] \\ &= 0.\end{aligned}$$

Under the assumption that $E(x_n x_{n-j}) = \sigma_x^2 \delta(j)$, where σ_x^2 is the transmitted symbol power and $\delta(j)$ is the Kronecker delta function.

$\Longrightarrow$

$$\sum_{k=-N}^{0} b_k R_c(k-l) + R_n(k-l)/\sigma_x^2 = \sum_{j=1}^{M} d_j c_{l-j} + c_l$$

for $l = -N, -N+1, \ldots, 0$ and

$$d_j = \sum_{k=-N}^{0} b_k c_{j-k}$$

for $j = 1, 2, \ldots, M$, where

$$R_n(k-l) = E[n_{n-l} n_{n-k}]$$

is the correlation function of the channel noise $\{n_n\}$ and

$$R_c(k-l) = \sum_{m=-\infty}^{\infty} c_{m-l} c_{m-k}$$

is the correlation function of the channel impulse response sample sequence $\{c_n\}$.

• Adaptive DFE

Letting

$$\begin{aligned}
\mathbf{b} &= (b_{-N}, b_{-N+1}, \ldots, b_{-1}, b_0)^T \\
\mathbf{d} &= (d_1, d_2, \ldots, d_M)^T \\
\mathbf{r}_n &= (r_{n+N}, r_{n+N-1}, \ldots, r_n)^T \text{ and} \\
\hat{\mathbf{x}}_{n-1} &= (\hat{x}_{n-1}, \hat{x}_{n-2}, \ldots, \hat{x}_{n-M})^T,
\end{aligned}$$

the equalization error at $t = nT$ can be represented as

$$\epsilon_n = x_n - (\mathbf{r}_n^T \mathbf{b} - \hat{\mathbf{x}}_{n-1}^T \mathbf{d}).$$

Choose the tap coefficient vector $(\mathbf{b}, \mathbf{d})^T$ in order to minimize the mean-square equalization error, $\mathcal{E} = E[\epsilon_n^2]$:

$$\begin{aligned}
\mathbf{b}(n+1) &= \mathbf{b}(n) + \alpha_1 \epsilon_n \mathbf{r}_n \\
\mathbf{d}(n+1) &= \mathbf{d}(n) - \alpha_2 \epsilon_n \hat{\mathbf{x}}_{n-1}.
\end{aligned}$$

Example 4.8: Performance of adaptive linear equalizer and adaptive DFE
Consider a 2-path wireless propagation channel, where the first path experiences Rician fading with a k-factor of 7.5 dB, and the second path experiences Rayleigh fading. The two paths have the same power of the diffusive signal component. Consider BPSK with coherent demodulation, and assume perfect carrier phase synchronization. The normalized fading rate is $\nu_m T = 0.0005$, where ν_m is the maximum Doppler shift and T is the BPSK symbol interval. The propagation delay difference between the two paths is equal to the symbol interval T. It is desired to evaluate the BER transmission performance when using (a) no equalizer, (b) a 7-tap adaptive linear equalizer, and (c) an adaptive DFE with a 3-tap feedforward filter and a 2-tap feedback filter. In the simulation, one sample is generated for each transmitted/received symbol.

Simulation results:

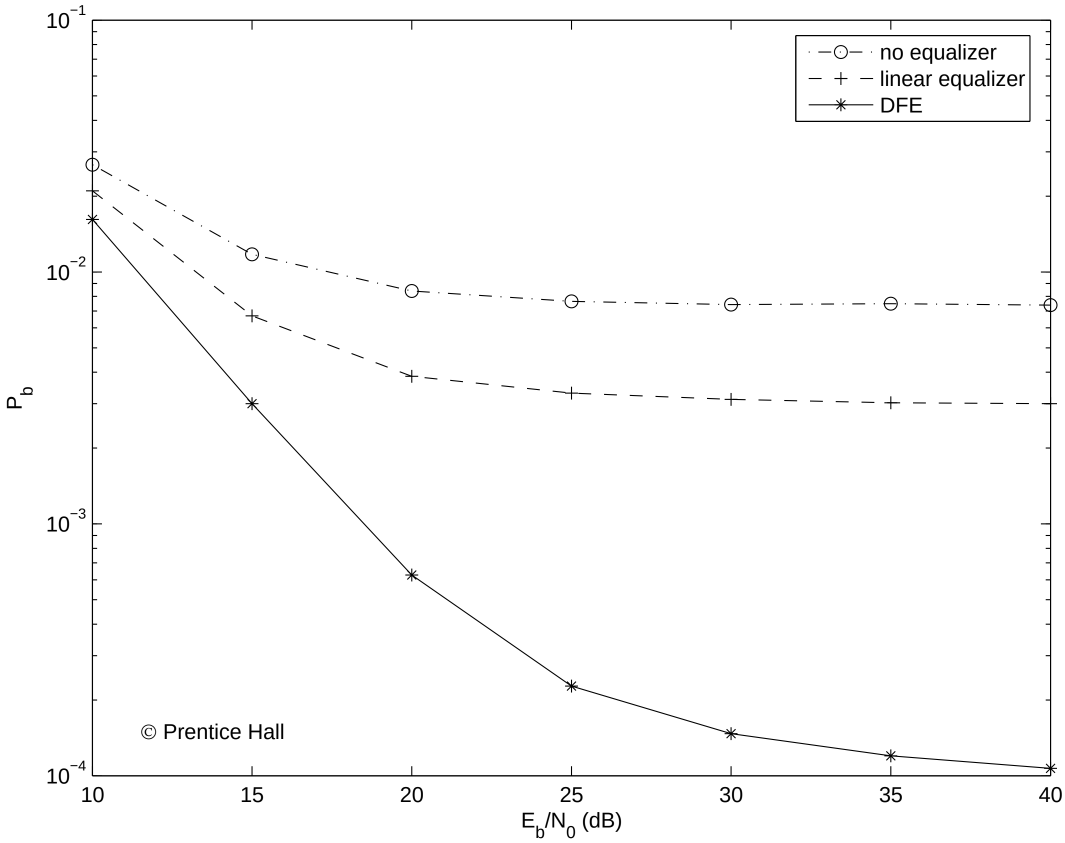


Figure 18: BER performance using accurate equalization error in the adaptive algorithms

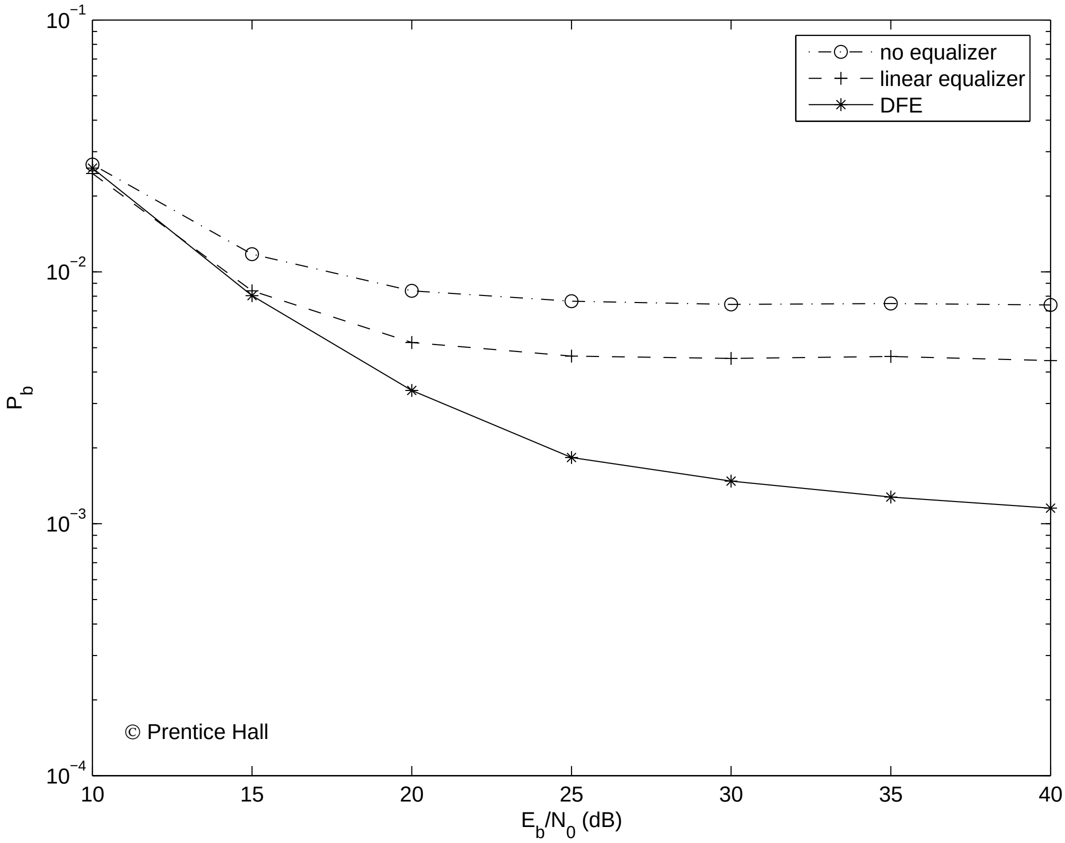


Figure 19: BER performance using estimated equalization error in the adaptive algorithms

□

Chapter 5 Fundamentals of Cellular Communications

- Cellular concept and frequency reuse
- Cochannel and adjacent channel interference
- Trunking and grade of service
- Mechanisms for capacity increase

5.1 Cellular Concept

- Given a propagation environment, increasing transmitted power will increase the service coverage area.

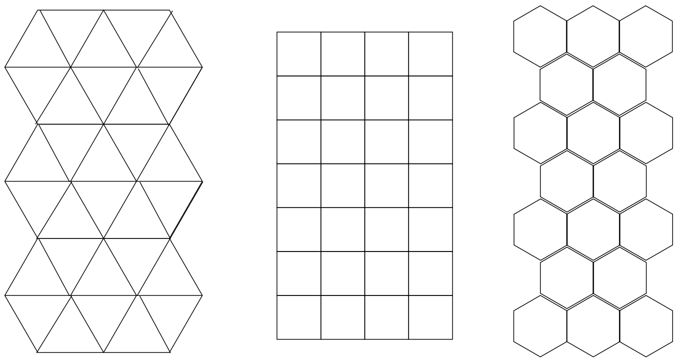
  $\Longrightarrow$ The coverage area can be controlled by using a proper transmitted power level.

- In cellular systems, the total service area is divided into a number of smaller areas, each of which is a radio cell.

  $\Longrightarrow$ Advantages:

  - Low transmitted power
  - Frequency reuse possible.

- Regular polygons may be used to represent the cell coverage.



(a) (b) (c)

- Hexagonal cells are popular because

  - closest to a circle
  - tight cellular packing
  - perfect partitioning of the service area.

- Frequency reuse is limited by co-channel interference. Cells which use the same frequency channels are called co-channel cells.

- Frequency is reused from cell cluster to cell cluster. No frequency channel is reused among cells in the same cell cluster.

  $\Longrightarrow$ Cells in each cell cluster use unique frequency channels.

Let

K - total number of channels in the system without frequency reuse

N - the number of cells in each cell cluster

J - total number of channels in each cell

then

$$K = JN \quad \text{or} \quad J = K/N.$$

Let

M - total number of cell clusters in the system

C - total number of channels in the system with frequency reuse

then

$$C = MK = MJN.$$

$\Longrightarrow$ cluster size $N \downarrow$

$\longrightarrow M \uparrow$ to cover the same area

$\longrightarrow C \uparrow$ for a given K value

That is, given K, C is maximized when N is minimized. However, the minimum N depends on the requirement on the co-channel interference level.

The cell cluster size N is also called the frequency reuse factor.

5.2 Frequency Reuse Factor

Consider hexagonal cells

- A hexagonal cells has exactly 6 equidistant neighbours.
- The lines joining the centers of any cell and each of its neighbours are separated by multiples of 60 degrees.

Nearest co-channel neighbours

To find the nearest co-channel neighbour of a particular cell, execute the folllowing two steps:

- move i cells along any chain of hexagons
- turn 60 degrees counterclockwise and move j cells

where the integers i and j are parameters for determining co-channel cells and for determining the size of the cell cluster (N)

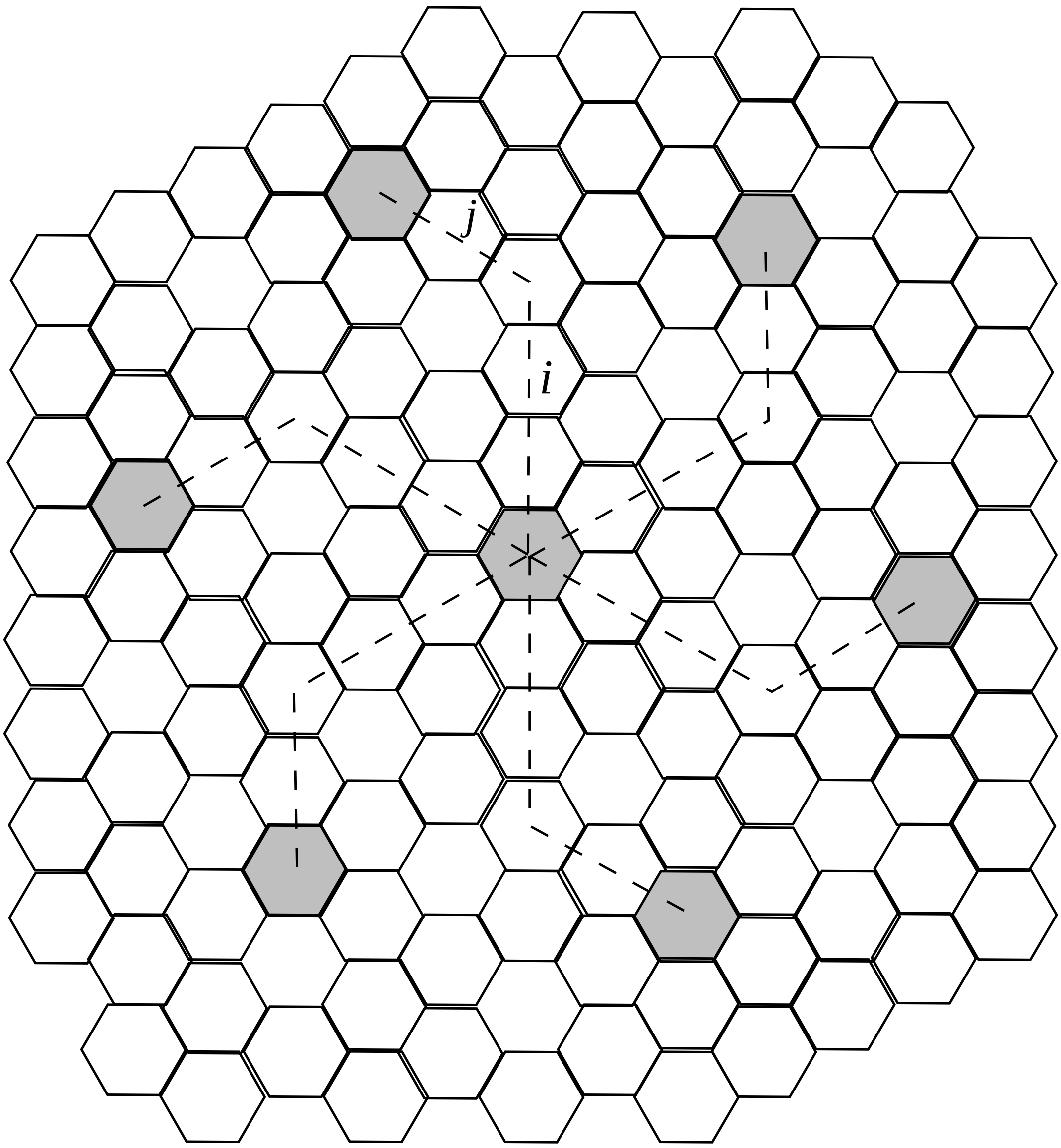
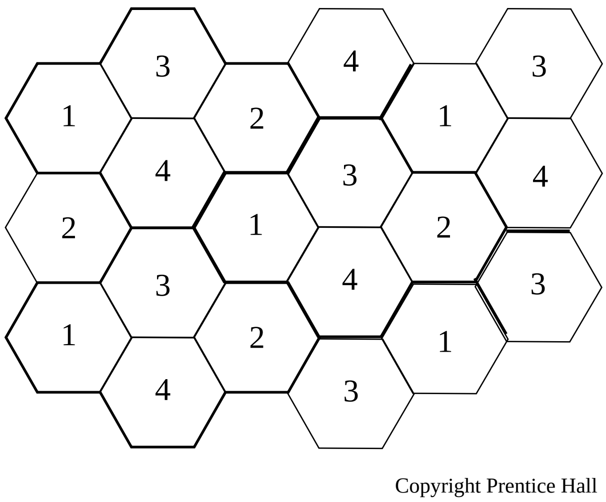
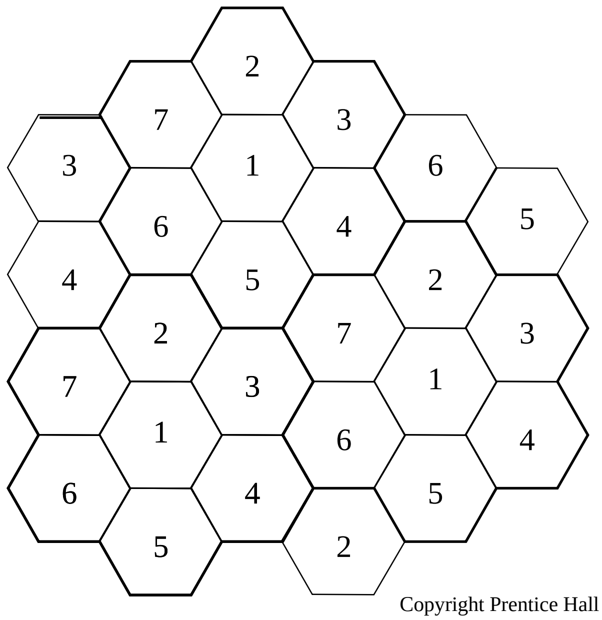


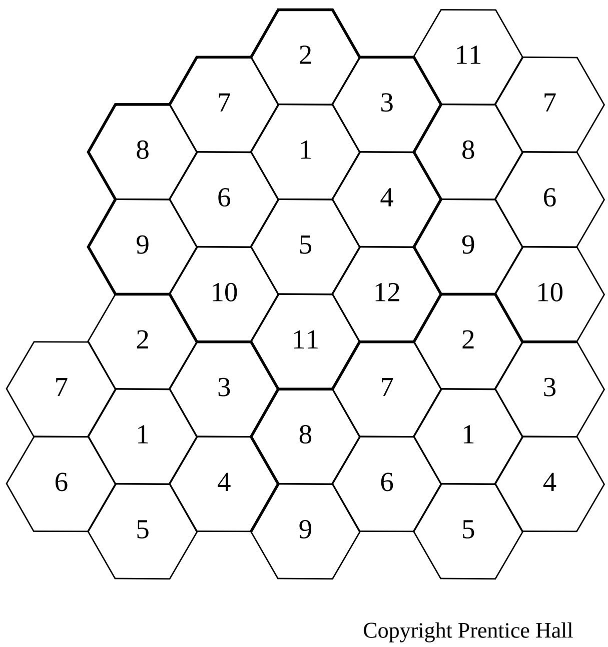
Figure 1: Locating cochannel cells in a cellular system



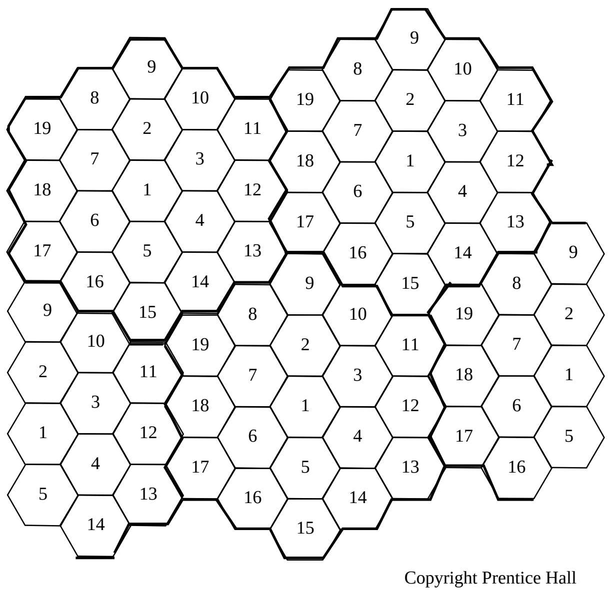
(a) $i = 2$ and $j = 0$



(b) $i = 1$ and $j = 2$



(c) $i = 2$ and $j = 2$



(d) $i = 2$ and $j = 3$

Figure 2: Cell clusters

Geometry of hexagonal cells

Let

R - radius of the cell (from center to vertex)

D - distance from the center of the candidate cell to the cell of the nearest co-channel cell

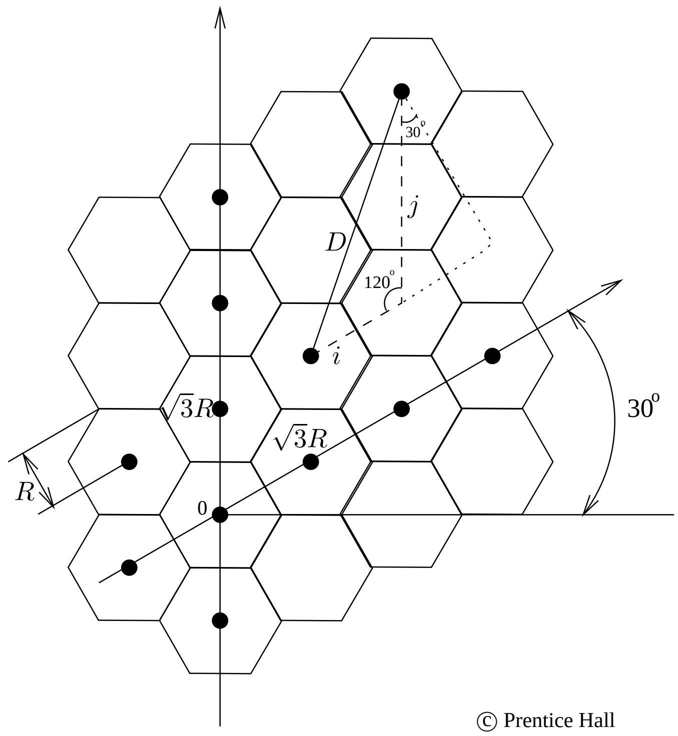


Figure 3: Distance between nearest cochannel cells

The actual distance between the centers of two adjacent cells is

$$2R\cos 30^0 = 2R\frac{\sqrt{3}}{2} = \sqrt{3}R.$$

$\Longrightarrow$

$$\begin{aligned} D^2 &= (j \cdot \sqrt{3}R\cos 30^0)^2 + (i \cdot \sqrt{3}R + j \cdot \sqrt{3}R \sin 30^0)^2 \\ &= (i^2 + j^2 + ij)3R^2 \end{aligned}$$

or,

$$\begin{aligned} D^2 &= (i \cdot \sqrt{3}R)^2 + (j \cdot \sqrt{3}R)^2 - 2(i \cdot \sqrt{3}R)(j \cdot \sqrt{3}R)\cos 120^0 \\ &= (i^2 + j^2 + ij)2R^2 \end{aligned}$$

It can be shown that the cell cluster size is given by

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

$\Longrightarrow$

$$D^2 = 3NR^2, \quad \text{or} \quad D = \sqrt{3N}R.$$

$$\Longrightarrow \quad D_{\text{norm}} \triangleq \frac{D}{\sqrt{3}R} = \sqrt{N}.$$

Frequency reuse ratio q

$$q \triangleq \frac{D}{R} = \frac{\sqrt{3N}R}{R} = \sqrt{3N}.$$

- The frequency reuse ratio q and the frequency reuse factor N carry the same information:

$$\begin{aligned} q \text{ (or } N) \uparrow \implies & \text{ cochannel interference} \downarrow \\ \implies & \text{ frequency reuse less often and system capacity} \downarrow \end{aligned}$$

- We should choose the minimum q (or N) subject to the constraint on the signal to cochannel interference ratio requirement.

Table 1: Frequency reuse ratio and cluster size

Frequency Reuse Pattern (i, j)	Cluster Size N	Frequency Reuse Ratio q
$(1, 1)$	3	3.00
$(2, 0)$	4	3.46
$(2, 1)$	7	4.58
$(3, 0)$	9	5.20
$(2, 2)$	12	6.00
$(3, 1)$	13	6.24
$(4, 0)$	16	6.93
$(3, 2)$	19	7.55
$(4, 1)$	21	7.94
$(3, 3)$	27	9.00
$(4, 2)$	28	9.17
$(4, 3)$	37	10.54

5.3 Cochannel and Adjacent Channel Interference

Cochannel interference

Let

N_I - the number of co-channel interfering cells

I_i - cochannel interference from the ith co-channel cell

S - the received power of the desired signal

The signal-to-cochannel interference ratio (S/I), also referred to as carrier-to-co-channel interference ratio (CIR), is

$$\frac{S}{I} = \frac{S}{\sum_{i=1}^{N_I} I_i}.$$

Consider only distance-dependent path loss. From chapter 2, we have

$$P_r(d) = P_0(d/d_0)^{-\kappa}$$

where

$P_r(d)$ – the received power at distance d ($\geq d_0$)

P_0 – the received power at distance d_0

κ - the path loss exponent

d - the distance between the transmitter and receiver.

Consider the forward link and assume that the transmitted power levels from all the BSs are the same, then

$$I_i \propto D_i^{-\kappa}$$

where D_i is the distance from the ith cochannel cell BS to the mobile.

When the mobile is at the cell boundary (the worst case),

$$S \propto R^{-\kappa}$$

$\Longrightarrow$

$$\frac{S}{I} = \frac{R^{-\kappa}}{\sum_{i=1}^{N_I} D_i^{-\kappa}}$$

which which is not a function of the transmitted power!

If we neglect cochannel interference from the second and other higher tiers, this means that $N_I = 6$. In the case that $r = R$ and using $D_i \approx D$ for $i = 1, 2, \ldots, N_I$,

$$\frac{S}{I} = \frac{(D/R)^{\kappa}}{N_I} = \frac{q^{\kappa}}{N_I} = \frac{(\sqrt{3N})^{\kappa}}{N_I}.$$

$\Longrightarrow$

$$q = \left(N_I \times \frac{S}{I}\right)^{1/\kappa} = \left(6 \times \frac{S}{I}\right)^{1/\kappa}.$$

Example 5.4 S/I ratio versus cluster size

Suppose the acceptable signal-to-cochannel interference ratio in a certain cellular communications situation is $S/I = 20$ dB or 100. Also, from measurements, it is determined that $\kappa = 4$. What is the minimum cluster size?

Solution:

The frequency reuse ratio can be calculated as

$$q \;=\; (6 \times 100)^{1/4} = 4.9492.$$

Then, the cluster size is given by

$$N \;=\; q^2/3 = 8.165 \simeq 9.$$

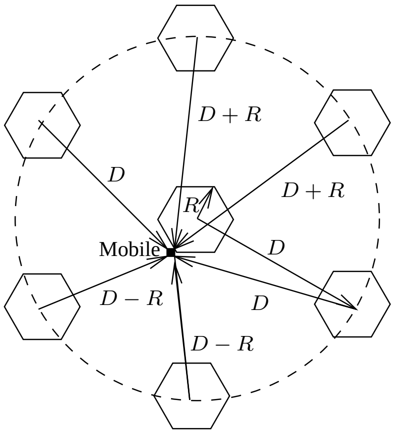
In this case, a 9-reuse pattern is needed for an S/I ratio of at least 20 dB. Since

$$q = D/R \quad \text{or} \quad D = qR,$$

D can be determined, given the cell radius R, and vice versa. Note that if N is less than 9, the S/I value would be below the acceptable level of 20 dB.

□

Consider a better approximation of the distances:



$$\frac{S}{I} \approx \frac{R^{-\kappa}}{2(D-R)^{-\kappa} + 2D^{-\kappa} + 2(D+R)^{-\kappa}}.$$

Substituting $D/R = q$, we have

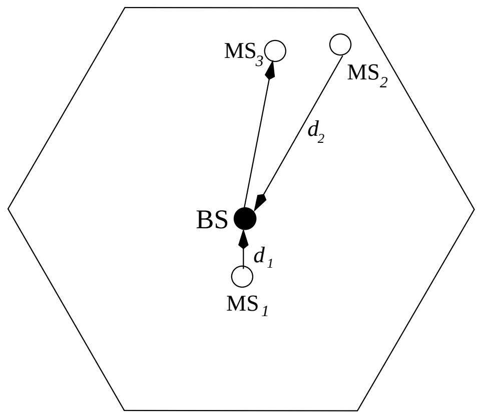
$$\frac{S}{I} = \frac{1}{2(q-1)^{-\kappa} + 2q^{-\kappa} + 2(q+1)^{-\kappa}}.$$

For $\kappa = 4$ and $N = 7$, $q = \sqrt{3N} = 4.6 \Longrightarrow S/I = 17.3$ dB

For $\kappa = 4$ and $N = 9$, $q = \sqrt{3N} = 5.2 \Longrightarrow S/I = 19.8$ dB

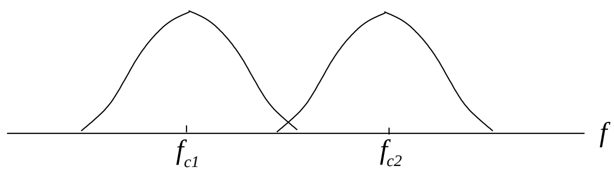
Design tradeoff: $N = 7$ or 9 if it requires $S/I = 18$ dB?

Adjacent channel interference

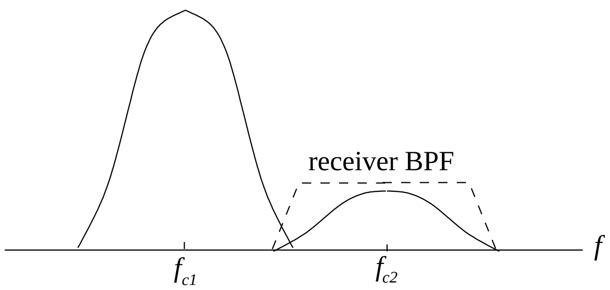


Near-far effect: $d_1 \ll d_2 \Longrightarrow P_{r1} \gg P_{r2}$ at the BS

psd: $P_{r1} = P_{r2}$



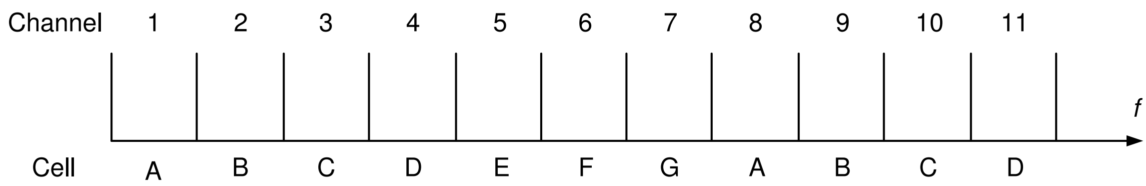
psd: $P_{r1} \gg P_{r2}$



For the signal from MS_2, the adjacent channel interference ↑↑ due to the near-far effect.

To reduce adjacent channel interference

- use modulation schemes which have small out-of-band radiation (e.g., MSK is better than QPSK)
- carefully design the receiver BPF
- use proper channel interleaving by assigning adjacent channels to different cells, e.g., for $N = 7$



- furthermore, do not use adjacent channels in adjacent cells, which is possible only when N is very large. For example, if N =7, adjacent channels must be used in adjacent cells
- use FDD or TDD to separate the forward link and reverse link.

5.4 Trunking and Grade of Service

- In cellular systems, a relatively small number of radio channels are used to serve a large population of mobile users, which is possible by frequency reuse and by trunking
- Trunking exploits the statistical behaviour of mobile users, so that a large number of mobile users can share the fixed radio channels in each cell on demand
- Based on traffic load, the number of radio channels in each cell should be determined in such a way that
    - all the channels are utilized efficienctly
    - call blocking rate is below a predetermined threshold
- Given a traffic load, number of channels $\uparrow \implies$ utilization efficiency $\downarrow$ and call blocking rate $\downarrow$
- The measure of traffic efficiency: 1 Erlang represents the amount of traffic intensity carried by a channel that is completely occupied, e.g., a radio channel that is occupied for 30 minutes during an hour carries 0.5 Erlangs of traffic
- The grade of service (GoS) is a measure of the ability of a user to access a trunked system during the busiest hour. GoS is typically given as
    - the likelihood that a call is blocked (for Erlang B systems), or
    - the likelihood that a call experiences a delay larger than a certain queueing delay (for Erlang C systems)

Definitions

- set-up time: the time required to allocate a trunked radio channel to a requesting user
- blocked call (lost call): call which cannot be completed at the time of request, due to congestion
- holding time (H): average duration of a typical call
- traffic intensity (ρ): measure of channel time utilization, which is the average channel occupancy measured in Erlangs
- load: traffic intensity across the entire trunked radio system, measured in Erlangs
- request rate (λ): the average number of call requests per unit time per user

Relations

- The traffic intensity offered by each user is (in Erlangs)

$$\rho_u = \lambda \cdot H$$

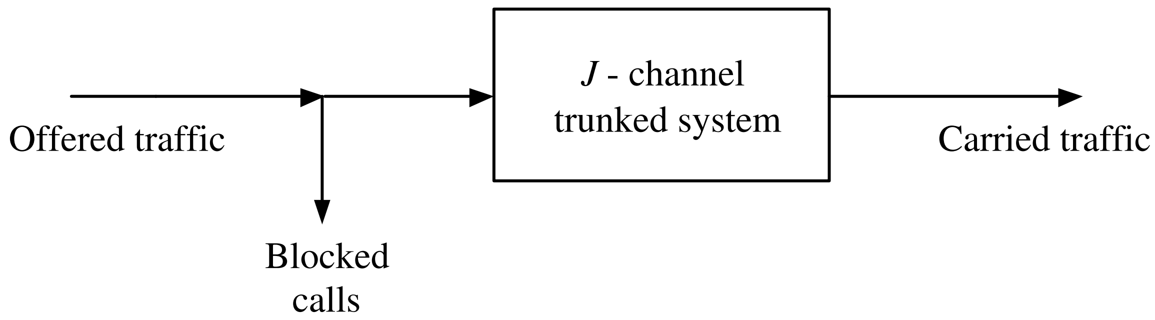
- For a system with u users and an unspecified number of channels, the total offered traffic intensity is (in Erlangs)

$$\rho = u \cdot \rho_u = u \cdot \lambda \cdot H$$

- In a Jchannel trunked system, if the traffic is equally distributed among the channels, then the traffic intensity per channel is

$$\rho_c = u \cdot \rho_u / J = u \cdot \lambda \cdot H / J \quad \text{(Erlangs)}$$

- Difference between offered traffic and carried traffic



where the intensity of the offered traffic is ρ, while the intensity of the carried traffic is $\rho(1 - P_B) \leq \rho$.

Types of trunked systems

Blocked calls cleared

- If no channels are available, the requesting user is blocked without access and is free to try again later;
- Call arrivals follows a Poisson distribution. Let X denote the number of calls arrivales per unit time, then

$$P(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}, \quad i = 0, 1, 2, \ldots$$

  where λ is the request rate (average number of call arrivals per unit time), and we have $E(X) = \lambda, V(X) = \lambda$;
- There are an infinite number of users;
- There are memoryless arrivals of calls, implying that all users, including blocked users, may request a channel at any time;
- Channel holding time, Y, follows an exponential distribution

$$f_Y(y) = \begin{cases} \alpha e^{-\alpha x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

  with $E(Y) = H = 1/\alpha$ and $V(Y) = 1/\alpha^2$;
- There are a finite number of channels available in the trunked pool.

$\Longrightarrow$ The system can be modeled by an M/M/C queue.

$\Longrightarrow$ The Erlang-B formula (the blocked calls cleared formula):

$$P(\text{blocking}) = \frac{\rho^J / J!}{\sum_{j=0}^{J} \rho^j / j!} \quad \text{– call blocking probability}$$

Note: The assumption that there are an infinite number of users in the system results in a conservative estimate of the GoS, as the blocking probability with a finite number of users is smaller than that obtained by the Erlang B formula.

Blocked calls delayed

- If a channel is not available immediately, the call request may be delayed until a channel becomes available;
- Other conditions (assumptions) are the same as those in the case of blocked calls cleared;
- The Erlang-C formula: The probability of a call not having immediate access to a channel is

$$P(\text{queueing}) = \frac{\frac{J\rho^J}{J!(J-\rho)}}{[\frac{J\rho^J}{J!(J-\rho)}] + \sum_{j=0}^{J-1}(\frac{\rho^j}{j!})}.$$

  Queueing gives rise to delay. The probability of nonzero delay is given by

$$P(\text{delay} > 0) = \frac{\rho^J}{\rho^J + J!(1-\frac{\rho}{J})\sum_{j=0}^{J-1}\frac{\rho^j}{j!}}.$$

- If no channels are immediately available the call is delayed, and the probability that the delayed call is forced to wait more than t seconds is

$$\begin{aligned} P(\text{delay} > t) &= P(\text{delay} > 0) \times P(\text{delay} > t|\text{delay} > 0) \\ &= P(\text{delay} > 0)\exp[-(J-\rho)\mu t]. \end{aligned}$$

- The average delay, $\bar{D}$, for all calls in the queueing system is given by

$$\bar{D} = P(\text{delay} > 0) \times \frac{1}{\mu(J-\rho)}.$$

- The average delay for those calls which are queued is $H/(J-\rho)$.

5.5 Capacity Enhancement in Cellular Systems

The capacity can be improved by

- cell splitting
- antenna sectoring
- dynamic channel assignment.

Cell splitting

- Subdivide a congested cell into smaller cells, each with its own base station and a corresponding reduction in antenna height and transmitted power.
- Cell splitting increases the capacity of a cellular system since it increases the number of times that channels are reused ($M \uparrow$).
- Reducing cell size increases handoffs, the number of base stations needed, and may result in a difficulty in finding a proper site for the base station.
- Old base station should be kept in some splitting cells.

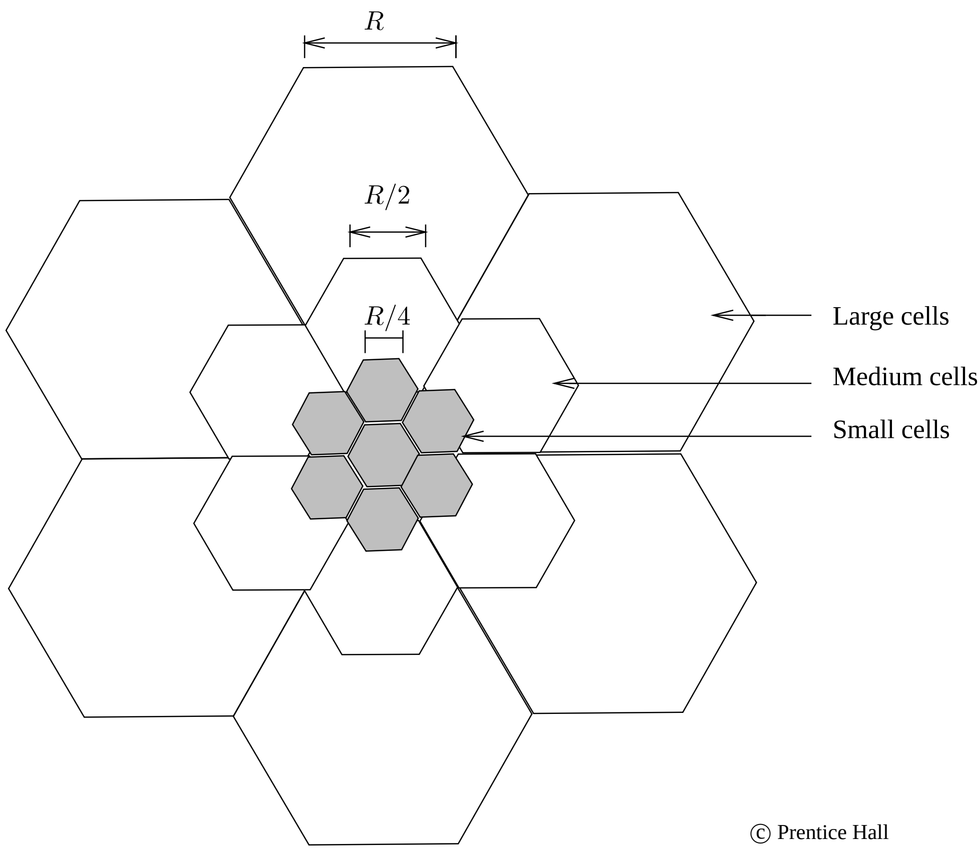


Figure 4: Illustrate of cell splitting from radius R to $R/2$ and to $R/4$

Let

P_{t1} the transmitted power of large cell BS

P_{t2} - the transmitted power of small cell BS

P_r - the received power at cell boundary

Then

P_r (large cell) $\propto P_{t1} \cdot R^{-\kappa}$

P_r (small cell) $\propto P_{t2} \cdot (R/2)^{-\kappa}$

On the basis of equal received power, we have

$$P_{t1} \cdot R^{-\kappa} = P_{t2} \cdot (R/2)^{-\kappa}$$

$\Longrightarrow$

$$P_{t1}/P_{t2} = 2^{\kappa}$$

$\Longrightarrow$

$$10 \log_{10}(P_{t1}/P_{t2}) = 10\kappa \log_{10} 2 \approx 3\kappa \text{ dB.}$$

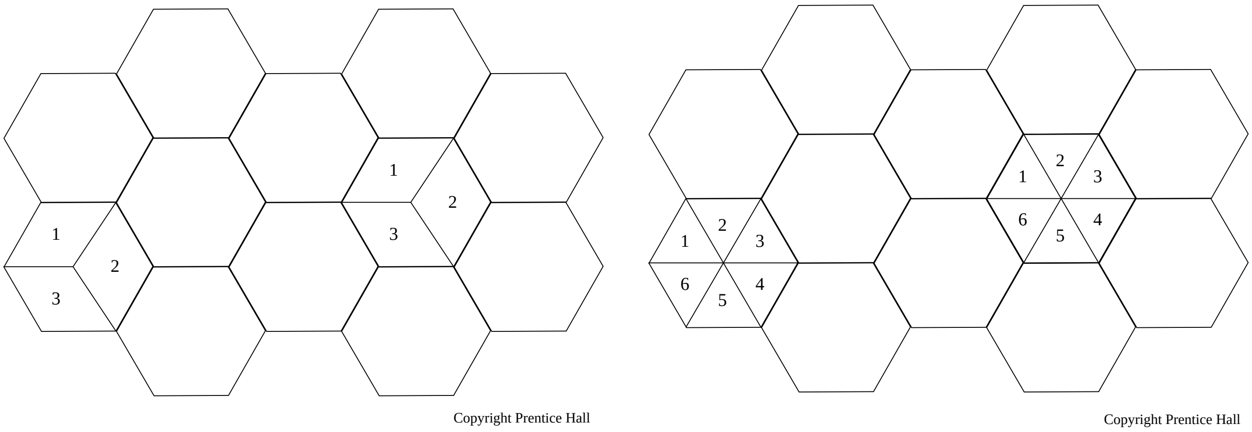
For $\kappa = 4, \;\; P_{t1}/P_{t2} = 12$ dB. In general

$R \longrightarrow R/2$ in cell splitting

$\Longrightarrow$ Cell area $\longrightarrow$(1/4) cell area

$\Longrightarrow$ Capacity is increased by 3 times (or 4 times in total).

Cell sectoring



(a) 3 sectors of $120°$ each (b) 6 sectors of $60°$ each

Assuming 7-cell reuse pattern, for the 3-sector case, the number of interferers in the first tier is reduced from 6 to 2.

The CIR is given by

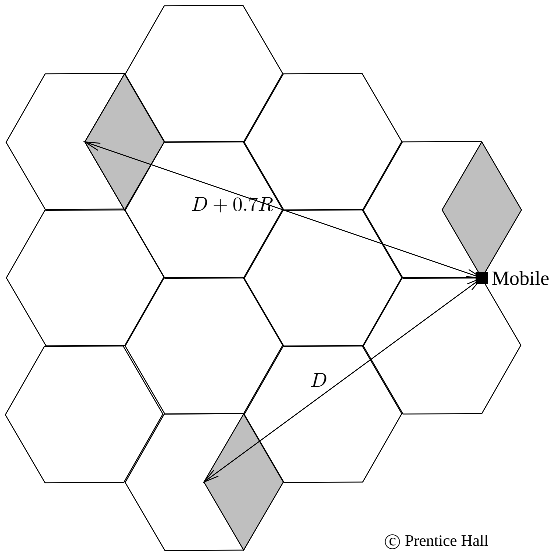
$$\frac{S}{I} = \frac{R^{-\kappa}}{\sum_{i=1}^{N_I}(D_i)^{-\kappa}} = \frac{R^{-\kappa}}{\sum_{i=1}^{2}(D_i)^{-\kappa}}$$

which is larger than the omnidirectional case where $N_I = 6$.

With $D_i \approx D$,

$$(\frac{S}{I})_{omni} = \frac{1}{6} \times q^{\kappa} \quad \text{and} \quad (\frac{S}{I})_{120°} = \frac{1}{2} \times q^{\kappa}.$$

Worst-case scenario in a 120° sectoring:



$$(\frac{S}{I})_{120^\circ} = \frac{R^{-\kappa}}{D^{-\kappa} + (D + 0.7R)^{-\kappa}} = \frac{1}{q^{-\kappa} + (q + 0.7)^{-\kappa}}.$$

For $\kappa = 4$ and $N = 7$ $(q = 4.6)$, we have $(S/I)_{120^\circ} = 24.5$ dB > 18 dB

$\Longrightarrow$ The 3-sector worst case for 7-cell reuse is acceptable for (S/I) requirement of 18 dB

Example 5.8 Cochannel interference with sectoring
In Example 5.6, it is shown that, with a frequency reuse factor of 7, base stations using omnidirectional antennas cannot satisfy the 18 dB signal-to-cochannel interference ratio requirement. Determine whether the use of 120° sectoring and 7-cell frequency reuse would satisfy the 18 dB requirement.

Solution:

For a 7-cell reuse, we have $q = \sqrt{3 \times 7} = 4.6 \Longrightarrow$

$$\left(\frac{S}{I}\right)_{120^\circ} = 285 \text{ or } 24.5\ dB.$$

Since this is greater than 18 dB, the 3-sector worst case for a 7-cell reuse is acceptable.

$\square$

Channel Assignment Strategies

Fixed channel assignment (FCA)

- 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 Channel allocation cannot adapt to traffic load dynamics.
- Borrowing option:
  - A cell is allowed to borrow channels from a neighbouring cell if all of its channels are already occupied.
  - Borrowing is supervised by the MSC to satisfy constraints on co-channel and adjacent channel interference.

Dynamic channel assignment (DCA)

- Voice channels are not allocated to different cells on a permanent basis.
- Each time a call request is made, the home BS requests a channel from the MSC.
- The MSC determines (dynamically) the availability of a channel and executes its allocation procedure accordingly.
- The MSC only allocates a given channel if the channel is not presently in use in the cell or any other cell which falls within the minimum restricted distance of the frequency reuse to avoid co-channel interference (co-channel reuse locking).
- DCA reduces the likelihood of call blocking, which increases the trunking capacity of the system, since all available channels under the control of the MSC are accessible to all the calls.
- DCA strategies require the MSC to collect real-time data on channel occupancy, traffic load distribution and radio signal strength indications of all the channel on a continuous basis.

Chapter 6 Multiple Access Techniques

Introduction

Radio cell: a geographical coverage area in which the services of mobile stations (MSs) are supported by a single base station (BS)

Forward link (downlink): BS $\rightarrow$ multiples MSs (one to many broadcasting)

Reverse link (uplink): MSs $\rightarrow$ BS (many to one multiple access)

Multiple access:

1. Multiple MSs want to access the common BS simultaneously
2. If two or more user signals arrive at the BS at the same time, there will be interferences, unless the signals are orthogonal
3. $x_i(t)$ and $x_j(t)$ are orthogonal in $[0, T]$ if

$$\int_0^T x_i(t)x_j(t)dt = 0 \qquad \text{for } i \neq j.$$

Question: How can we achieve the orthogonality?

1. Frequency division multiple access (FDMA)
2. Time-division multiple access (TDMA)
3. Code-division multiple access (CDMA)

6.1 Multiple Access Techniques for Mobile Communications

6.1.1 FDMA

1. The total bandwidth is divided into nonoverlapping frequency bands (channels)
2. Each user occupies a channel for the duration of the connection
   $\longrightarrow$ waste of resources
3. Narrowband transmission
4. Strict requirements on RF filters
5. Forward and reverse links use FDD

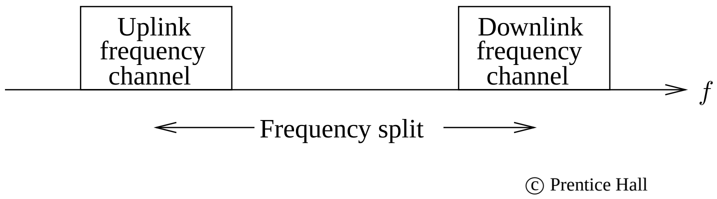


Figure 1: FDD for duplex operation

6.1.2 TDMA

1. Time is partitioned into frames
2. Each frame consists of N_{slot} data slots plus a header and a trailer
3. Each slot is for transmission of one information unit
4. A user continues to use the same slot in every frame during call connection $\longrightarrow$ waste of resources

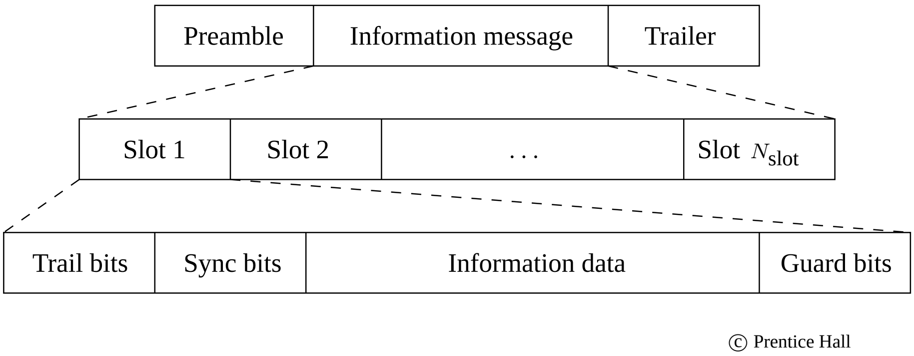
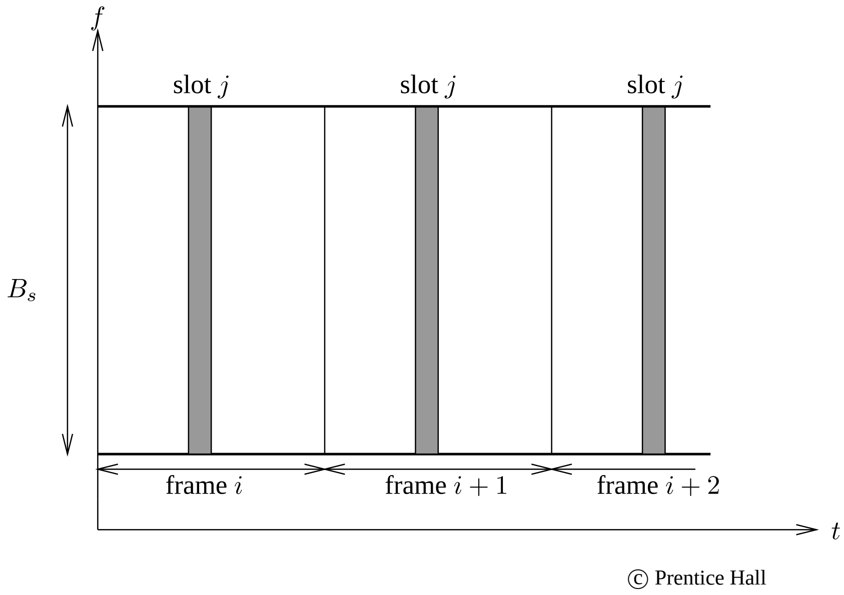


Figure 2: TDMA Frame Structure

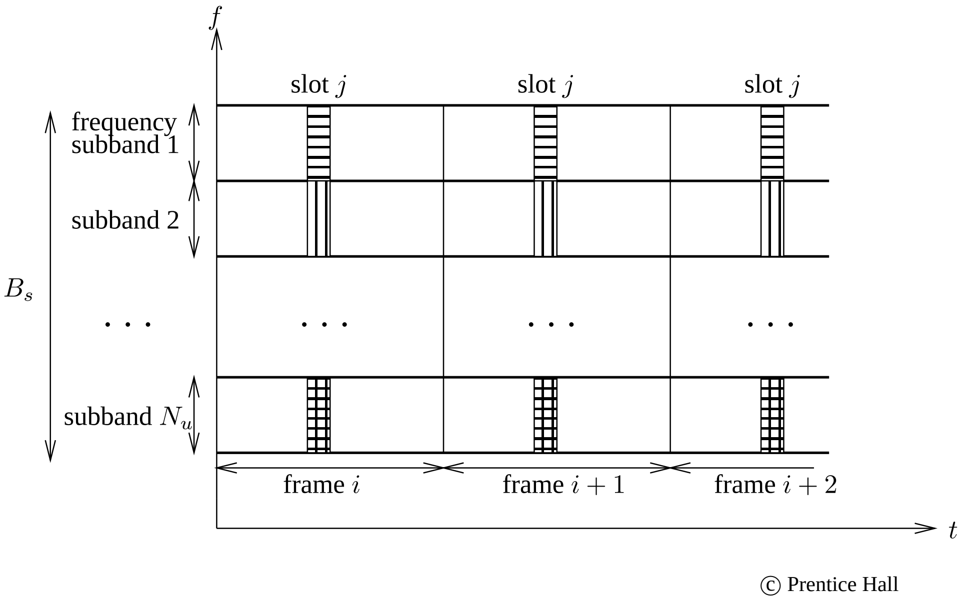
1. TDMA: wideband TDMA (W-TDMA) and narrowband TDMA (N-TDMA)

   W-TDMA: Each user occupies the total frequency bandwidth during its slots

   N-TDMA: The total frequency spectrum is divided into frequency subbands (channels); within each frequency channel, TDMA is used. $\longrightarrow$ Both time and frequency are partitioned.



(a) Wideband TDMA



(b) Narrowband TDMA

Figure 3: Wideband TDMA and narrowband TDMA

1. Forward and reverse links can use either FDD or TDD.

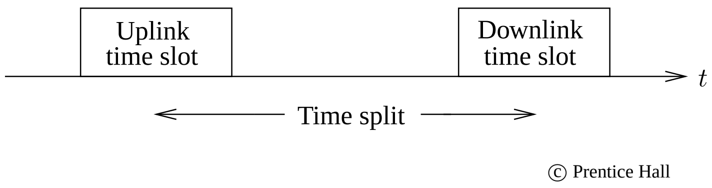


Figure 4: TDD for duplex operation

2. TDMA systems require strict time synchronization.

6.1.3 CDMA

Spread spectrum: the spectrum of the baseband message signal is spread by a significant order of magnitude larger than the minimum required bandwidth

Direct-sequence (DS) CDMA: The spread spectrum is achieved by directly multiplying the baseband signal with a high-rate pseudorandom noise (PN) sequence, called the spreading sequence

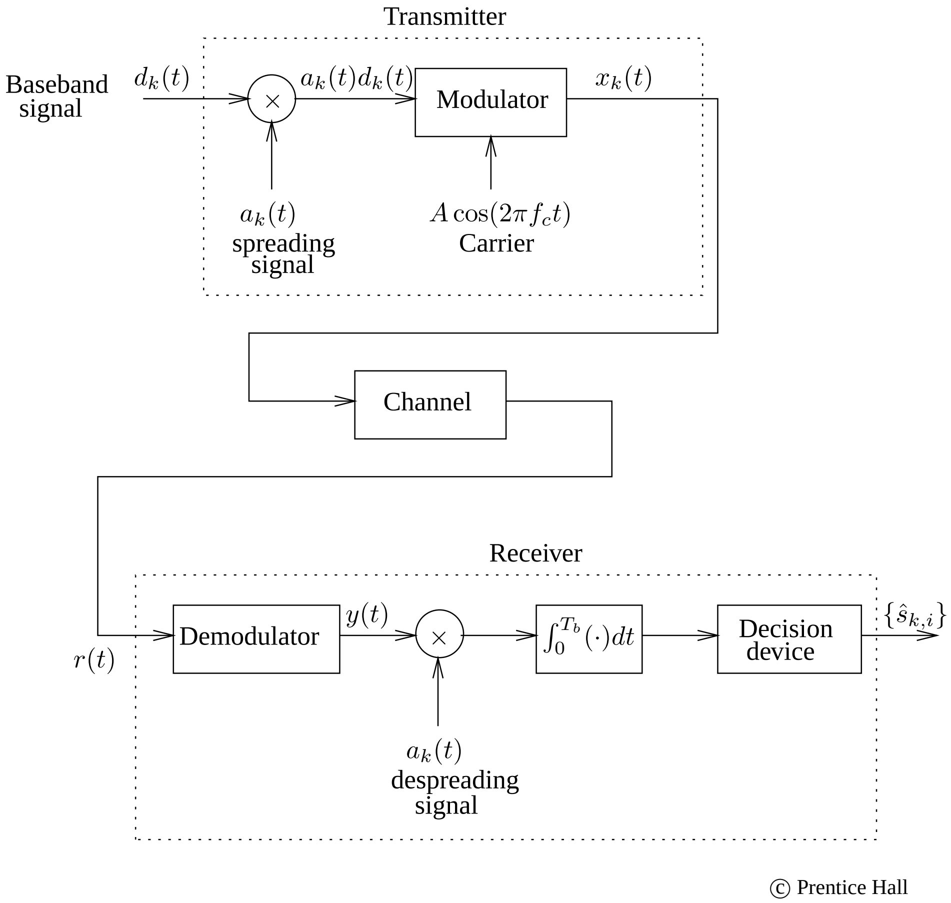


Figure 5: Function block diagram of user k transmitter and receiver in a DS-CDMA system

The information-carrying baseband signal $d_k(t)$ is

$$d_k(t) = \sum_i s_{k,i} \Pi(\frac{t - iT_b}{T_b})$$

where $s_{k,i} \in \{-1, +1\}$ is the ith binary information bit, T_b is the information bit interval, and $\Pi(t/T_b)$ is the rectangular pulse

$$\Pi(\frac{t}{T_b}) = \begin{cases} 1, & 0 \leq t \leq T_b \\ 0, & \text{otherwise.} \end{cases}$$

The spreading signal of the kth user is $a_k(t)$ and can be represented as

$$a_k(t) = \sum_l a_{k,l} P_{T_c}(t - lT_c)$$

where $a_{k,l} \in \{-1, +1\}$ is the lth chip of the binary PN sequence assigned to user k, $P_{T_c}(t)$ is the chip pulse waveform depending on baseband pulse shaping, and T_c is the chip interval.

For simplicity, we consider $P_{T_c}(t) = \Pi(t/T_c)$ and BPSK for the passband modulation.

Normally, $T_b/T_c = L$, where L is an integer.

Spreading: $d_k(t)a_k(t)$ with the same psd as that of $a_k(t) \Longrightarrow$ The bandwidth is increased by L times.

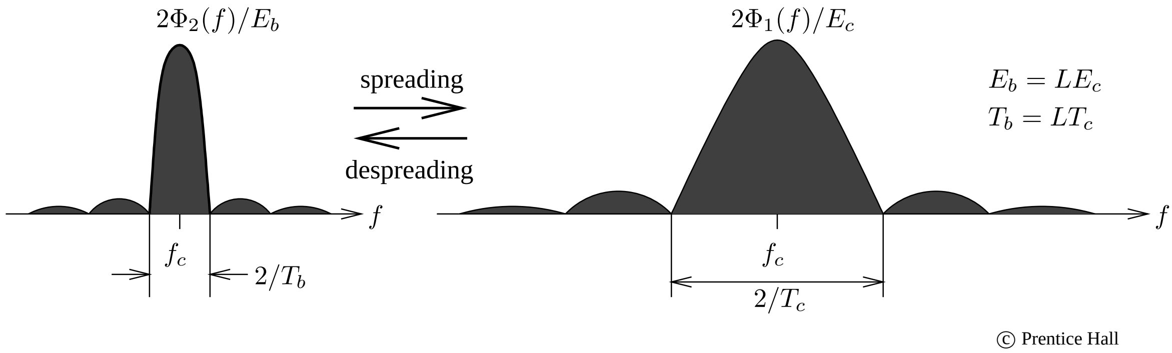


Figure 6: Illustration of signal power spectral density without and with spreading

Despreading: $[d_k(t)a_k(t)] \cdot a_k(t) = d_k(t) \cdot [a_k^2(t)] = d_k(t)$

Requirements:

1. The receiver should generate the same spreading waveform $a_k(t)$;
2. The locally generated $a_k(t)$ should accurately synchronized with that in the received signal $\longrightarrow$ PN code acquisition and tracking

Let K denote the number of mobile users in the cell. Assuming that all the K users in the cell are synchronized in time and have the same received signal power ($A_c^2/2$), $r(t)$ is given by

$$r(t) = \sum_{k=1}^{K} x_k(t) + I(t) + w(t)$$

where $I(t)$ is intercell interference and $w(t)$ is background white Gaussian noise with zero mean and two-sided psd $N_0/2$.

MAI (multiple access interference): intracell interference + intercell interference

$$\begin{aligned} r(t) &= A_c a_1(t) d_1(t) \cos(2\pi f_c t) \\ &\quad + \sum_{k=2}^{K} A_c a_k(t) d_k(t) \cos(2\pi f_c t) + I(t) + w(t). \end{aligned}$$

The demodulator translates the received signal centered at f_c to baseband centered at frequency zero. Its output, $y(t)$, can be written as

$$y(t) = \frac{A_c}{2} a_1(t) d_1(t) + \sum_{k=2}^{K} \frac{A_c}{2} a_k(t) d_k(t) + n(t)$$

where

$$n(t) \approx \frac{1}{T_c} \sum_{l} \{ \int_{lT_c}^{(l+1)T_c} [I(t) + w(t)] \cos(2\pi f_c t) dt \} \Pi(\frac{t - lT_c}{T_c})$$

is due to the intercell interference and additive background noise.

After despreading, we have

$$a_1(t) y(t) = \frac{A_c}{2} d_1(t) + \sum_{k=2}^{K} \frac{A_c}{2} a_1(t) a_k(t) d_k(t) + a_1(t) n(t).$$

To suppress the interference and noise, the next step in the receiver is to integrate the despread signal over each information symbol (bit) interval over which the desired signal component is a constant. The output of the integrator at the end of the ith symbol is

$$\int_{iT_b}^{(i+1)T_b} a_1(t)y(t)dt = \frac{A_cT_b}{2}\alpha_1 d_{1,i} + \frac{A_cT_b}{2}[\sum_{k=2}^{K} \alpha_k d_{k,i}] + n_i$$

where

$$\alpha_1 = \frac{1}{T_b}\int_{iT_b}^{(i+1)T_b} [a_1(t)]^2 dt = 1$$

is the autocorrelation of the spreading signal $a_1(t)$ over the symbol interval and

$$\alpha_k = \frac{1}{T_b}\int_{iT_b}^{(i+1)T_b} a_1(t)a_k(t)dt, \quad k = 2, 3, \ldots, K$$

is the crosscorelation between the spreading signals $a_1(t)$ and $a_k(t)$ over the symbol interval.

If all the spread signals $a_k(t)$, $k = 1, 2, \ldots, K$, are orthogonal in the symbol interval, then there is no intracell interference in the recovered baseband signal.

The effect of the intercell interference and background noise on the signal detection is given by the last term n_i, which is

$$n_i = \int_{iT_b}^{(i+1)T_b} a_1(t)n(t)dt.$$

In the frequency domain, the integrator is a lowpass filter with bandwidth approximately equal to $1/T_b$. The LPF lets the desired signal component $\frac{A_c}{2}d_1(t)$ go through without distortion and greatly reduces the interference and noise power. The despreading process significantly improves the signal-to-interference plus noise ratio (SINR) value.

The spread spectrum system performance is measured by the processing gain, G_p, defined as the SINR improvement achieved by despreading, i.e.,

$$G_p \triangleq \frac{\text{SINR after despreading}}{\text{SINR before despreading}} = \frac{W}{R_b} = \frac{T_b}{T_c} = L.$$

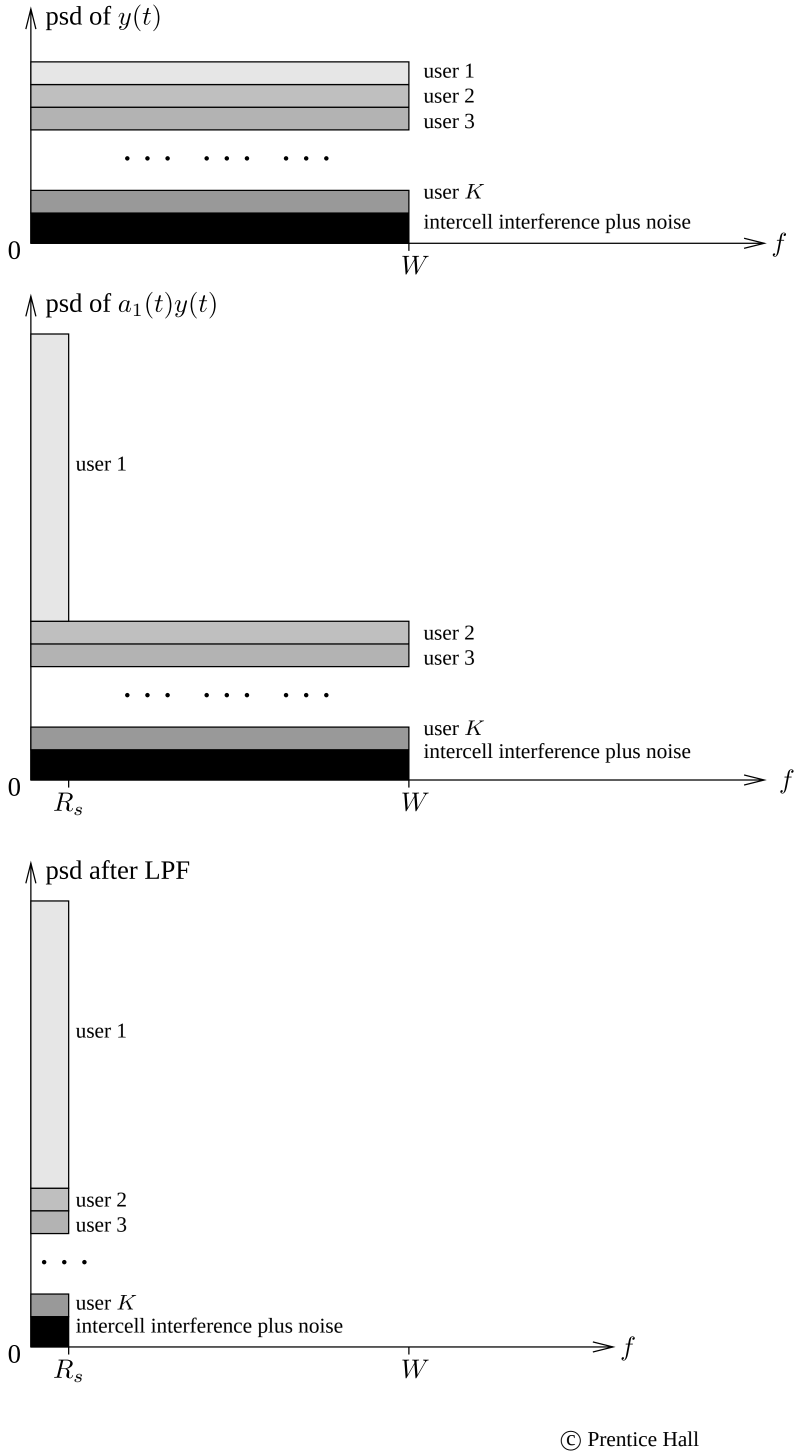


Figure 7: Illustration of signal power spectral density before and after despreading

Transmission performance:

• AWGN channel with two-sided noise psd $N_0/2$ in the absence of MAI

1. For the signal component $d_k(t)$, the effects of spreading and despreading cancel each other.
2. For the additive white Gaussian noise $n(t)$, because $a_k(t) = +1$ or -1 equally likely, $n(t)a_k(t)$ is still white Gaussian noise with the same statistics.

$\Longrightarrow$ The probability of transmission bit error is the same as that in the AWGN channel without spread spectrum. For example, if BPSK is used, the BER for the DS-CDMA user with coherent detection is

$$P_b = Q(\sqrt{\frac{2E_b}{N_0}}).$$

- AWGN channel with two-sided noise psd $N_0/2$ in the presence of MAI

  1. MAI will increase transmission error rate if the spreading sequences are not orthogonal.
  2. MAI can be approximated as a Gaussian process when the number of mobile users is large.
  3. With a large spread spectrum bandwidth W, the psd of the interference is approximately uniform over the bandwidth.

$\Longrightarrow$ The effect of the MAI on the transmission performance can be treated in the same way as the additive white Gaussian noise and the BER is

$$P_b = Q(\sqrt{\frac{2E_b}{n_0 + N_0}})$$

where n_0 is the two-sided psd of the MAI over the spread spectrum bandwidth.

Advantages of CDMA:

1. universal frequency reuse
2. diversity and the use of Rake receiver $\longrightarrow$ better transmission accuracy
3. soft handoff
4. soft capacity
5. utilization of source activity factor
6. facilitating multimedia traffic

Drawbacks of CDMA:

1. strict power control required
2. data rate limitation especially with a large processing gain
3. high complexity of transceivers (high chip rate, spread waveform synchronization, Rake receiver, power control, etc.)

6.2 Spectral Efficiency

The overall spectral efficiency depends on

1. channel spacing in kHz
2. cell size in km^2
3. frequency reuse factor
4. multiple access scheme

$\Longrightarrow$ The spectral efficiency $\eta = \eta_{\text{sys}} \times \eta_{\text{access}}$

$$\eta \triangleq \frac{\text{Total number of channels available for data in system}}{\text{(system bandwidth)(total coverage area)}} \quad \text{Channel/MHz}/\mathrm{km}^2$$

where the total number of channels for data = total number of channels - total number of control channels, taking into account frequency reuse. Or,

$$\eta \triangleq \frac{\text{Total traffic carried by the system}}{\text{(system bandwidth)(total coverage area)}} \quad \text{Erlang/MHz}/\mathrm{km}^2.$$

Techniques used to enhance the spectrum utilization:

1. source coding (data compression to reduce the source rate)
2. bandwidth reduction (modulation, detection, channel coding)
3. channel assignment
4. choice of multiple access method

6.2.1 FDMA Systems

Spectral efficiency of FDMA (η_{FDMA})

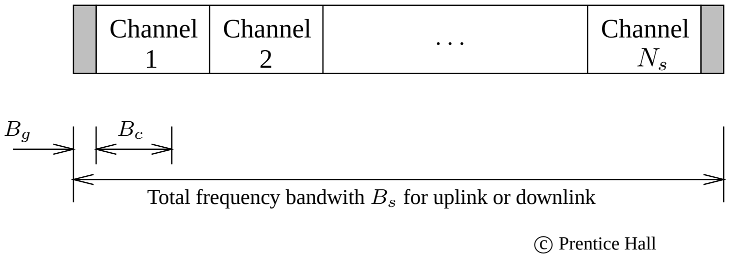


Figure 8: Channel spacing and guard bands in FDMA uplink or downlink

The number of channels, N_s, that can be simultaneously supported is

$$N_s = \frac{B_s - 2B_g}{B_c}.$$

Let N_{ctl} be the number of allocated control channels and N_{data} be the number of data channels in the system. Then the total number of available channels is

$$N_s = N_{data} + N_{ctl}.$$

The spectral efficiency of FDMA is defined as

$$\eta_{FDMA} = \frac{\text{bandwidth available for data}}{\text{system bandwidth}} = \frac{N_{data} B_c}{B_s} < 1.$$

Example 6.1 Spectral efficiency of FDMA
In the AMPS system, the system bandwidth is 12.5 MHz, the channel spacing is 30 kHz, and the edge guard spacing is 10 kHz. The number of channels allocated for control signaling is 21. Find

a) the number of channels available for message transmission, and

b) the spectral efficiency of FDMA.

Solution:
We have $B_s = 12.5$ MHz, $B_c = 30$ kHz, and $B_g = 10$ kHz. Therefore,

a) the number of available channels is

$$N_s = \frac{B_s - 2B_g}{B_c} = \frac{12.5 \times 1000 - 20}{30} = 416 \text{ channels, and}$$

b) the spectral efficiency of this FDMA system is

$$\eta_{FDMA} = \frac{30 \times (416 - 21)}{12.5 \times 1000} = 0.948.$$

□

System spectral efficiency η

The number of available channels per cluster is given by

$$N_{ch/cluster} = \frac{B_s - 2B_g}{B_c}.$$

The total number of channels available for data traffic per cluster is

$$N_{data/cluster} = N_{ch/cluster} - N_{ctl/cluster} = \frac{B_s - 2B_g}{B_c} - N_{ctl/cluster}$$

The total number of channels available for data traffic per cell is

$$N_{data/cell} = \frac{N_{data/cluster}}{N} = \frac{\frac{B_s - 2B_g}{B_c} - N_{ctl/cell}}{N}$$

which is also the cell capacity, N_c, defined as the maximum number of mobile stations that can be served at one time in each cell.

$\Longrightarrow$

$$\eta = \frac{\text{number of data channels per cluster}}{\text{system bandwidth times area of the cluster}} = \frac{N_{data/cluster}}{B_s \cdot (N \cdot A_{cell})}.$$

$\Longrightarrow$

$$\eta = \frac{1 - \frac{B_c}{B_s} \cdot (N_{ctl/cluster} + \frac{2B_g}{B_c})}{B_c \cdot N \cdot A_{cell}} \quad \text{channel/MHz/km}^2.$$

$\Longrightarrow$

$$\eta = \frac{1}{B_c \cdot N \cdot A_{cell}} - \frac{N_{ctl/cluster} + \frac{2B_g}{B_c}}{B_s \cdot N \cdot A_{cell}} \quad \text{channel/MHz/km}^2$$

where the second term on the right-hand side accounts for the overhead in FDMA.

Let η_t be the trunk (system) efficiency. The total traffic carried in a cluster, in Erlang, is $\eta_t \times N_{data/cluster}$.

$$\eta = \frac{\eta_t \cdot N_{data/cluster}}{B_s \cdot N \cdot A_{cell}} \quad \text{Erlang/MHz/km}^2.$$

$\Longrightarrow$

$$\eta = \frac{\eta_t}{B_c \times N \cdot A_{cell}} - \frac{\eta_t \cdot (N_{ctl/cluster} + \frac{2B_g}{B_c})}{B_s \times N \cdot A_{cell}} \quad \text{Erlang/MHz/km}^2$$

where the second term on the right-hand side is due to the overhead in FDMA.

Example 6.2 System spectral efficiency in channel/MHz/km^2
Suppose a cellular system in which the one-way bandwidth of the system is 12.5 MHz, the channel spacing is 30 kHz, and the guard band at each boundary of the spectrum is 10 kHz. If i) the cell area is 6 km^2, ii) the frequency reuse factor (cluster size) is 7, and iii) 21 of the available channels are used to handle control signaling, calculate the following.

a) The total number of available channels per cluster;

b) The number of available data channels per cluster;

c) The number of available data channels per cell;

d) The system spectral efficiency in units of channel/MHz/km^2.

Solution:

We allocate all of the available frequencies to one cluster and these frequencies or channels are distributed evenly among the N cells in the cluster.

a) The total number of available channels in the cluster is

$$N_{ch/cluster} = \frac{B_s - 2B_g}{B_c} = \frac{12.5 - 2 \times 0.01}{0.03} = 416.$$

b) The number of available data channels per cluster is

$$N_{data/cluster} = N_{ch/cluster} - N_{ctl/cluster} = 416 - 21 = 395.$$

c) The number of available data channels per cell is

$$N_{data/cell} = N_{data/cluster}/N = 395/7 = 56.$$

d) The overall spectral efficiency of the system is

$$\eta = \frac{N_{data/cell}}{B_s \cdot A_{cell}} = \frac{56}{12.5 \times 6} = 0.747 \quad \text{channel/MHz/km}^2.$$

□

6.2.2 TDMA Systems

W-TDMA:

Let

$$\begin{aligned} \tau_p &= \text{the time duration for the preamble} \\ \tau_t &= \text{the time duration for the trailer} \\ T_f &= \text{the frame duration} \\ L_d &= \text{the number of information data symbols in each slot} \\ L_s &= \text{the total number of symbols in each slot.} \end{aligned}$$

Then, we have

$$\eta_{W-TDMA} = \frac{T_f - \tau_p - \tau_t}{T_f} \times \frac{L_d}{L_s}.$$

N-TDMA:

$$N_u = \frac{B_s - 2B_g}{B_c}.$$

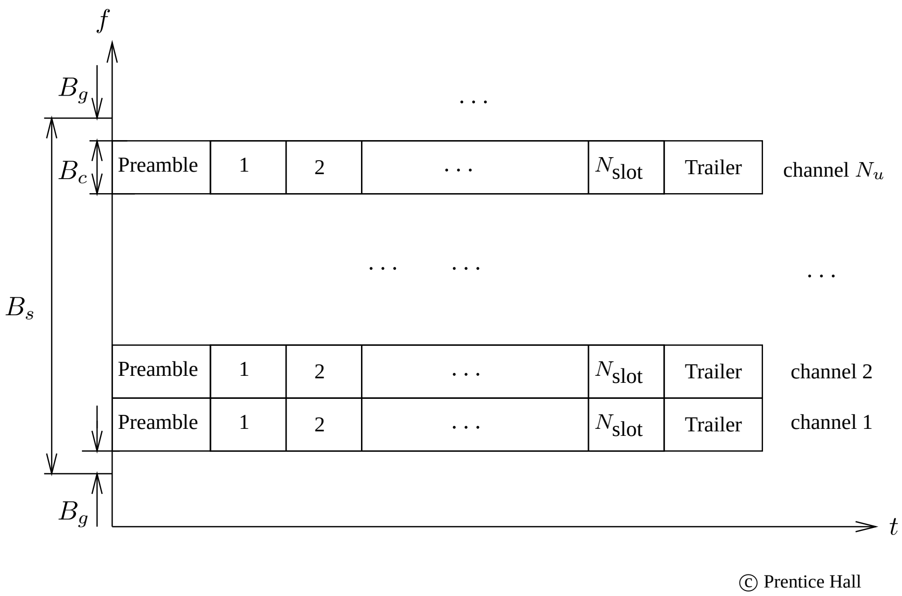


Figure 9: Narrowband TDMA format

The spectral efficiency of narrowband TDMA is then given by

$$\begin{aligned}\eta_{N-TDMA} &= \eta_{W-TDMA} \times \frac{B_c N_u}{B_s} \\ &= \frac{T_f - \tau_p - \tau_t}{T_f} \times \frac{L_d}{L_s} \times \frac{B_s - 2B_g}{B_s}.\end{aligned}$$

Cell capacity of TDMA Systems

The maximum number of simultaneous users that can be accommodated without frequency reuse is

$$N_s = N_u \cdot N_{\text{slot}}$$

where

$$N_u = \begin{cases} 1, & \text{for W-TDMA} \\ \frac{B_s - 2B_g}{B_c} > 1, & \text{for N-TDMA.} \end{cases}$$

With frequency reuse, all the N_s channels can be allocated to a single cell cluster. Then, the cell capacity, N_c, in a TDMA system with frequency reuse factor N is

$$N_c = \frac{N_u \cdot N_{\text{slot}}}{N}.$$

Taking into account of the source activity factor s_f, we have

$$N_c = \frac{N_u \cdot N_{\text{slot}}}{s_f \cdot N} \quad \text{user/cell}.$$

System spectral efficiency η

Let ϵ_{bw} denote the modulation efficiency, R_b the bit rate of each mobile source. The effective number of users, $N_{\text{effective}}$, that can be supported per cell is

$$N_{\text{effective}} = \epsilon_{bw} \times \frac{B_s}{R_b \cdot N}.$$

Taking into account of the overhead necessary for TDMA, the effective number should be modified to

$$N_{\text{effective}} = \epsilon_{bw} \times \frac{T_f - \tau_p - \tau_t}{T_f} \times \frac{L_d}{L_s} \times \frac{B_s - 2B_g}{R_b \cdot N}$$

The overall spectral efficiency of the system in bit/unit time/unit bandwidth/cell, η, can be expressed as

$$\begin{aligned} \eta &= N_{\text{effective}} \times \frac{R_b}{B_s} \\ &= \epsilon_{bw} \times \frac{T_f - \tau_p - \tau_t}{T_f} \times \frac{L_d}{L_s} \times \frac{B_s - 2B_g}{B_s} \times \frac{1}{N} \ \text{bit/s/Hz/cell}. \end{aligned}$$

Example 6.3 Spectral efficiency of the IS-54 system
Consider IS-54 (updated as IS-136), which is a synchronous N-TDMA/FDD system that uses a one way bandwidth of 25 MHz for the forward (or reverse) channel. The system bandwidth is divided into radio channels of 30 kHz, each supporting transmission at rate 16.2 kbps. Guard bands with $B_g = 20$ kHz are used. A single radio channel supports 3 speech channels. The frame duration is 40 ms, consisting of 6 time slots. The slot duration is then 40/6 $(\simeq 6.667)$ ms. Each slot consists of 324 bits, among which 260 bits are for information data and the remaining 64 bits are overhead for access control. The speech codec rate is 7.95 kbps, which corresponds to a gross bit rate of 13.0 kbps with channel encoding. If the frequency reuse factor is 7, find

a) the number of simultaneous users that can be accommodated in each cell cluster,

b) the cell capacity,

c) the spectral efficiency η_{N-TDMA} of TDMA, and

d) the overall spectral efficiency.

Solution:

Given:

$$\begin{aligned}
B_s &= 25\text{MHz} = 25000\text{kHz} \\
B_c &= 30\text{kHz} \\
B_g &= 20\text{kHz} \\
R_b &= 7.95\text{kbps} \\
N_{\text{slot}} &= 3 \\
T_f &= 40\text{ms} \\
\tau_s &= 0\text{ms} \\
\tau_t &= 0\text{ms} \\
L_d &= 260\text{bits} \\
L_s &= 324\text{bits} \\
\epsilon_{bw} &= 7.95\text{kbps} \times 3/30\text{kHz} = 0.265 \\
s_f &= 1 \\
N &= 7
\end{aligned}$$

a) The number of simultaneous users that can be accommodated is

$$N_u = \frac{B_s - 2B_g}{B_c} = \frac{25 \times 1000 - 2 \times 20}{30} = 832 \quad \text{user/cell cluster.}$$

b) The cell capacity of the synchronous TDMA system is

$$N_c \;=\; \frac{N_u N_{\text{slot}}}{s_f N} = \frac{832 \times 3}{1 \times 7} \simeq 356 \quad \text{user/cell.}$$

c) The spectral efficiency of the N-TDMA is

$$\eta_{N-TDMA} = \eta_{W-TDMA} \times \frac{B_c N_u}{B_s} = \frac{T_f - \tau_p - \tau_t}{T_f} \times \frac{L_d}{L_s} \times \frac{B_s - 2B_g}{B_s} \simeq 0.8.$$

d) The overall spectral efficiency is

$$\begin{aligned} \eta &= \epsilon_{bw} \times \frac{T_f - \tau_p - \tau_t}{T_f} \times \frac{L_d}{L_s} \times \frac{B_s - 2B_g}{B_s} \times \frac{1}{N} \\ &= 0.265 \times 3 \times \frac{40 - 0 - 0}{40} \times \frac{260}{324} \times \frac{25000 - 2 \times 20}{25000} \times \frac{1}{7} \\ &\simeq 0.0909 \quad \text{bit/s/Hz/cell.} \end{aligned}$$

□

6.2.3 DS-CDMA Systems

Cell capacity

Consider the uplink (from MSs to the BS) with accurate power control. The signal-to-interference ratio, S/I, is given by

$$\frac{S}{I} = \frac{R_b E_b}{B_s I_0} = \frac{1}{G_p} \frac{E_b}{I_0}.$$

$\Longrightarrow$

$$\frac{E_b}{I_0} = G_p \frac{S}{I}.$$

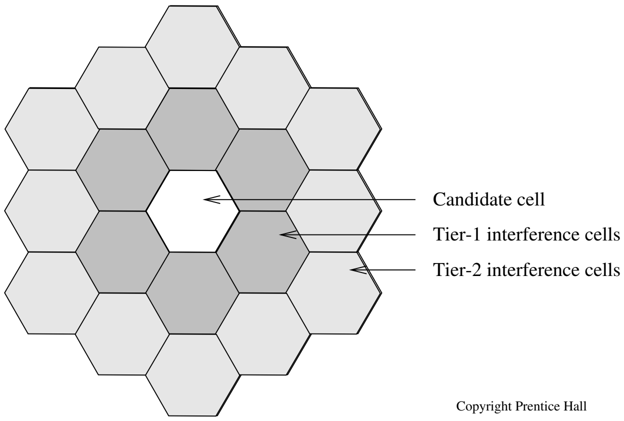


Figure 10: Interfering neighbors in hexagonal array

Let κ_f is the total intracell and intercell interference (on the average) normalized to the total intracell disturbance

$$\kappa_f = 1 + 6k_1 + 12k_2 + 18k_3 + \cdots .$$

Then the total interference power, including both intracell and intercell interference, is

$$I = [(N_{MS} - 1)S] \cdot \kappa_f.$$

The SINR can then be expressed as

$$\frac{S}{I} = \frac{1}{(N_{MS} - 1)\kappa_f + P_n/S}.$$

Define $\eta_f = 1/\kappa_f$ as the frequency reuse efficiency factor.

$$\frac{E_b}{I_0} = G_p \cdot \frac{S}{I} = \frac{G_p \cdot \eta_f}{(N_{MS} - 1) + \eta_f P_n/S}.$$

$\Longrightarrow$

$$N_{MS} = 1 + \eta_f [\frac{G_p}{E_b/I_0} - \frac{P_n}{S}].$$

Consider source activity factor (s_f), cell sectors (Q) with directive antennas, and imperfect power control (c_d):

$$\begin{aligned} I &= [(N_{MS} - 1)S] \cdot \kappa_f \\ \stackrel{s_f}{\Longrightarrow} \quad I' &= s_f[(N_{MS} - 1)S \cdot \kappa_f] + P_n \\ \stackrel{Q}{\Longrightarrow} \quad I'' &= \{s_f[(N_{MS} - 1)S \cdot \kappa_f] + P_n\}/Q \end{aligned}$$

$\Longrightarrow$ The SINR is now S/I'', from which we get

$$N_{MS} = 1 + \frac{\eta_f}{s_f}[\frac{QG_p}{E_b/I_0} - \frac{P_n}{S}].$$

With the capacity degradation factor c_d due to imperfect power control, we have

$$N_{MS} = 1 + \frac{c_d \eta_f}{s_f}[\frac{QG_p}{E_b/I_0} - \frac{P_n}{S}]$$

From the required BER, we can obtain the required E_b/I_0 value, $(E_b/I_0)^*$, then the cell capacity is

$$N_c = 1 + \frac{c_d \eta_f}{s_f}[\frac{QG_p}{(E_b/I_0)^*} - \frac{P_n}{S}].$$

System utilization

The system utilization is defined as the number of users, U_{MS}, that can be supported under the constraint that $E_b/I_0 \geq (E_b/I_0)^*$.

$$U_{MS} = 1 + \frac{c_d \eta_f}{s_f}[\frac{QG_p}{E_b/I_0} - \frac{P_n}{S}], \quad \text{with } E_b/I_0 \geq (E_b/I_0)^*.$$

Spectral efficiency of DS-CDMA systems

$$\eta = \frac{U_{MS}}{N_c} = \frac{(E_b/I_0)^*}{E_b/I_0}$$

or

$$\eta = U_{MS} \times \frac{s_f R_b}{B_s} \text{ bit/s/Hz}$$

where R_b is the constant bit rate (in units of bit/s/user) when a user is in an *on* state and B_s is the one-way system bandwidth in Hz.

Example 6.4 Capacity and utilization of CDMA system
Consider a DS-CDMA system where the one-way system bandwidth is 25 MHz. Suppose the data rate per user is 8 kbps. Assume perfect power control, one sector antenna, and persistent transmissions.

a) Calculate the E_b/I_0 specification required to support a maximum number of 250 users/cell if the signal-to-background noise ratio is 26 dB and the frequency reuse efficiency is 0.9.

b) If the actual operating value of E_b/I_0 is 12 dB, calculate the system utilization for the parameter values specified in part a).

Solution:

a) With the assumptions given in the problem, the E_b/I_0 specification can be written as

$$(E_b/I_0)^* = \frac{\eta_f G_p}{N_c - 1 + \eta_f P_n/S}.$$

It is given that S/P_n = 26 dB or 400, $\eta_f = 0.9$, $G_p = B_s/R_b = 25\times1000/8 = 3125$, and $N_c = 250$. Substituting these values in the above equation for $(E_b/I_0)^*$ yields

$$(E_b/I_0)^* = \frac{0.9 \times 3125}{250 - 1 + 0.9/400} \simeq 11.25 \text{ or } 10.5 \text{ dB}.$$

b) The utilization can be expressed as

$$U_{MS} = 1 + \eta_f \cdot [\frac{G_p}{E_b/I_0} - \frac{P_n}{S}], \quad \text{with } E_b/I_0 \geq 10.5 \text{ dB}.$$

If the actual operating value of E_b/I_0 is 12 dB, or 15.85, then the system utilization is

$$U_{MS} = 1 + 0.9 \times [\frac{3125}{15.85} - 1/400] \simeq 178.$$

□

Example 6.5 Spectral Efficiency of DS-CDMA
Consider the CDMA system in Example 6.4, with the same assumptions about parameter values.

a) If the actual operating value of E_b/I_0 is 12 dB, find the spectral efficiency for this CDMA system.

b) Suppose the actual operating value of E_b/I_0 is now 11 dB, find the spectral efficiency.

c) Discuss the impact of operating at a lower value of E_b/I_0 in terms of spectral efficiency and system utilization. What would be the benefit, if any, to the service provider?

Solution

a) For $E_b/I_0 = 12$ dB, the number of users that can be supported was 178 (Example 6.4), and the system capacity is 250. Therefore, the spectral efficiency is

$$\eta = 178/250 = 71.2\%$$

or

$$\eta = 178 \times \frac{8}{25 \times 1000} = 0.005696 \text{ bit/s/Hz.}$$

b) If the value of E_b/I_0 is now 11 dB or 12.59, the system utilization is

$$U_{MS} = 1 + 0.9 \times [\frac{3125}{12.59} - \frac{1}{400}] \simeq 224.$$

Therefore, the spectral efficiency is

$$\eta = 224/250 = 89.6\%$$

or

$$\eta = 224 \times \frac{8}{25 \times 1000} = 0.007168 \text{ bit/s/Hz.}$$

c) A reduction in the operating value of E_b/I_0 from 12 dB to 11 dB is accompanied by an increase in spectral efficiency from 71.2% to 89.6%. With a 1-dB decrease in E_b/I_0, the system utilization is increased from 178 to 224 users. This is significant from the revenue point of view for the service provider.

With the system operating at $E_b/I_0 = 11$ dB, which is 0.5 dB above the QoS requirement, the resultant BER would be acceptable for the system users.

□

Chapter 7. Mobility Management in Wireless Networks

- Wireless networks need a wireline backbone to extend the geographical coverage $\Rightarrow$ wireless/wireline interworking.

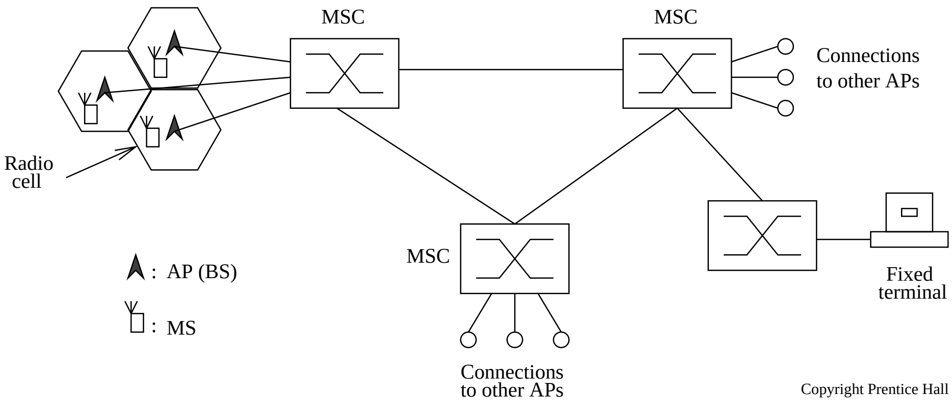


Figure 1: Wireless/wireline network

- Wireless networks have the flexibility to allow user roaming
  $\Rightarrow$ need for mobility management
- Mobility management $\Rightarrow$ handoff management and location management
- End-to-end information delivery requires quality of service (QoS) and grade of service (GoS) provisioning
- QoS parameters (at the packet level): packet loss, packet delay, delay jitter, throughput
- GoS parameters (at the call level): probability of blocking, probability of dropping, probability of forced termination
- Call admission control (CAC): needed to ensure satisfaction of GoS and QoS requirements.
- Maintaining ongoing sessions are more important than admitting new calls $\Rightarrow$ handoff calls should be given a higher access priority than new calls.

Handoff Management

Service continuity: When a user roams from one cell to another, its connection must be handed off from the currently serving Access Point (AP) to the new AP within a tolerable time interval (handoff completion time)

Handoff Strategies

- mobile controlled handoff (MCHO)
- network controlled handoff (NCHO)
- mobile assisted handoff (MAHO)

Handoff functions are performed at the mobile switching center (MSC)
MAHO reduces the burden on the network (used by GSM)

Handoff Types

- Hard handoff - break before make
- Soft handoff - make before break
- Backward handoff - handoff is predicted in advance and initiated via the existing radio link
- Forward handoff - handoff is initiated via the radio link associated with the candidate AP

Terminology

- **cluster**: A group of APs connected to the same MSC
- **intraswitch handoff**: handoffs within the same cluster
- **interswitch handoff**: handoffs between clusters
- **profile**: contains feedback information for decision-making in MAHO

Contents of Profile

- AP/MSC identity
- Signal strengths received from neighbouring APs

Question: Would it be necessary for the mobile to send signal strengths received from all APs in the cluster?

Answer: It is only necessary to send at most the signal strengths received from three adjacent APs.

Reason: The tagged MS may only receive strong signals from three adjacent APs (see Figs. **??** and **??**)

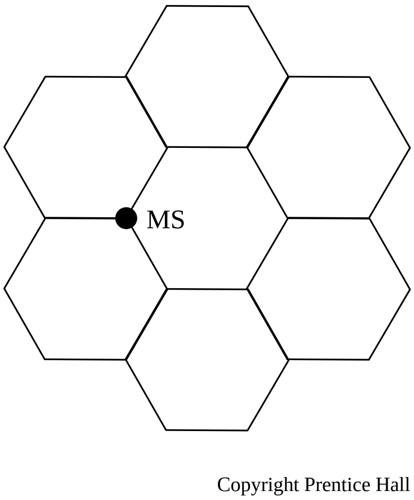


Figure 2: Mobile located at the vertex of three cells

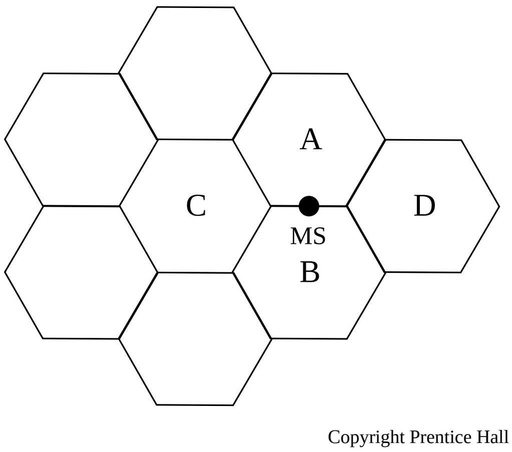


Figure 3: Mobile located at the boundary of two cells

The three adjacent cells may belong to a single cluster or at most three clusters relative to the serving MSC.

Exercise: Construct a Profile for a 2-tier cellular array by using the identification procedure laid out in Example 7.1, and using three strongest signal strengths.

Rate of Sending Profiles

- Let I_f be the interval for feeding back Profiles.
- The rate for sending Profiles is then $r_p = 1/I_f$.
- r_p is a function of the mean time, $E[t_s]$, for the mobile to move to the cell boundary, at which point a handoff has to take place. The rate of sending profiles is also related to the number of profiles, N_p, created during the cell dwelling time.
- $r_p \approx N_p/E[t_s]$

Question: Why does mobile movement have a direct effect on the rate of sending profiles?

Exercise: Assuming that the distance travelled by the mobile to reach the cell boundary is uniformly distributed between 0 and the maximum distance, and that the velocity is uniformly distributed between the minimum and maximum velocities, derive the expression for the probability density function of the the time taken to reach the cell boundary, as given by Eqn (7.3.1) in the text.

Handoff Scenarios

- User movements between Cells with APs directly connected to the same MSC constitute **intraswitch** handoff.
- User migration between cells belonging to different MSCs constitute **interswitch** handoff.
- Interswitch handoff requires the transfer of session information from the old serving MSC to the new MSC. hence takes more time to complete the handoff than intraswitch handoff.

Exercise: Study the solution for Example 7.2 by asking why each of the sequence of message exchanges is reasonable or necessary?

Location Management for Cellular Networks

- The geographical coverage of a cluster of APs connected to the same MSC is extended by interconnecting MSCs or MSCs to an Internet backbone network.
- A cellular network is thus an interconnection of subnetworks, each of which is overseen by an MSC.
- A given mobile is associated with a *home network*, one of the subnetworks in the cellular network.
- The mobile's permanent address resides in the home location register (HLR), co-located with the MSC of the home network.
- When the mobile moves away from its home network into a neighbouring subnetwork, or *foreign or visitor network*, it must update its registration with the HLR. This is done through the visitor location register (VLR) $\Rightarrow$ the VLR must establish an association with the HLR.
- Messages destined for the mobile are always delivered to its permanent address at the HLR, and then related to the mobile in its current location.
- In the visitor network, the mobile is assigned a care-of address, residing at the VLR, which is used to routing messages to the mobile in the visitor network.

Steps in Mobility Management

- **Handoff**:
  - When a mobile moves from one cell to another, an efficient handoff scheme must be in place to hand to mobile over to the new cell. - Under certain conditions, the mobile needs to update its registration with its HLR.

- **Authentication**:
  - While in the visitor network and attempting to register with its HLR through the VLR, the mobile's identity has to be authenticated to avoid the possibility of a third party masquerading that of the rightful owner.

- **Location Update**:
  - When the mobile moves within an RA, there is no need for the mobile to update its registration with its HLR. When the mobile moves into a new RA, it must update its registration with its HLR, i.e., to inform the HLR of its current location. This is done through its VLR.

- **Call Delivery**:
  - Messages destined for the mobile are always sent to the mobile's permanent address residing at its HLR. The HLR needs to know the current location of the mobile in order to route the messages to the mobile at its current location.

Two-Tiered Network Architecture

Both IS-41 and GSM' MAP (Mobile Applications Part) use a two-tiered network architecture.

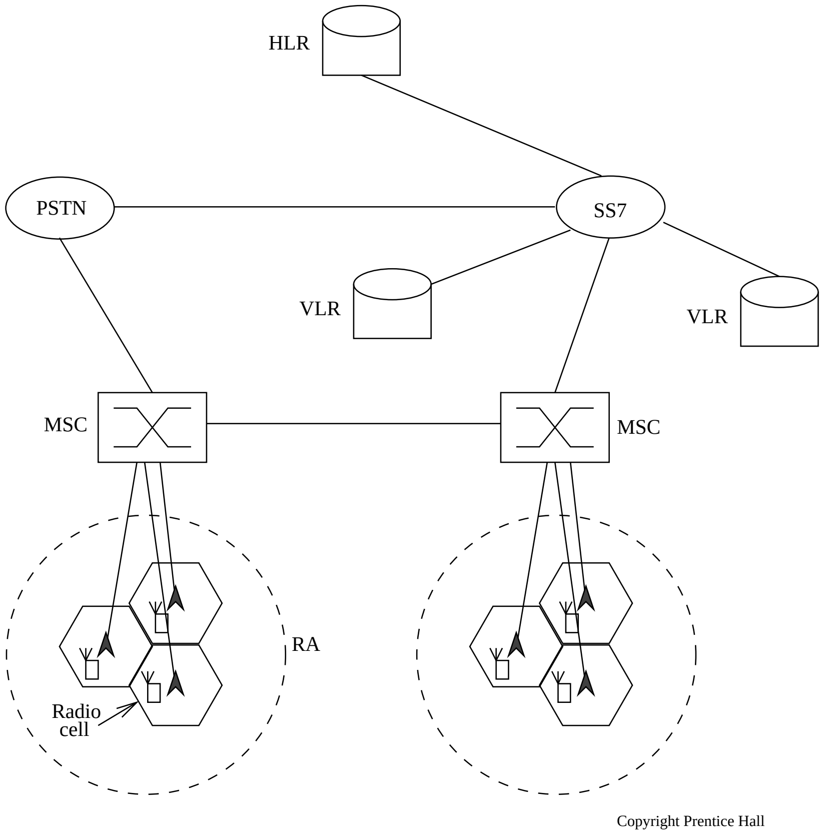


Figure 4: IS-41 network architecture

Both location update and call delivery involve signaling exchanges between various network entities. Signaling in IS-41 is handled by the SS7 signaling network.

- In the two-tiered network architecture, the only wireless links are those connecting MSs to BSs in the individual cells; the other subsystems are interconnected together via wireline links.
- A cluster of cells connected to the same MSC constitutes a registration area (RA).
- The MSC is the gateway that bridges the wireless and wireline domains, and provides typical network switching functions and coordinates the location registration and call setup.
- Migrations within the same RA does not require registration update with the HLR. Why?
- Migrations between RAs requires registration update with the HLR.
  Why?
- How does registration updates done when the MS moves to a new RA?
  - VLR $\rightarrow$ HLR
- From the mobile's perspective, how does the VLR differ from the HLR?

Signaling in the wired subnet of a Cellular Network

- All control signals transmission are in the wireline domain are handled by the SS7 signaling network.

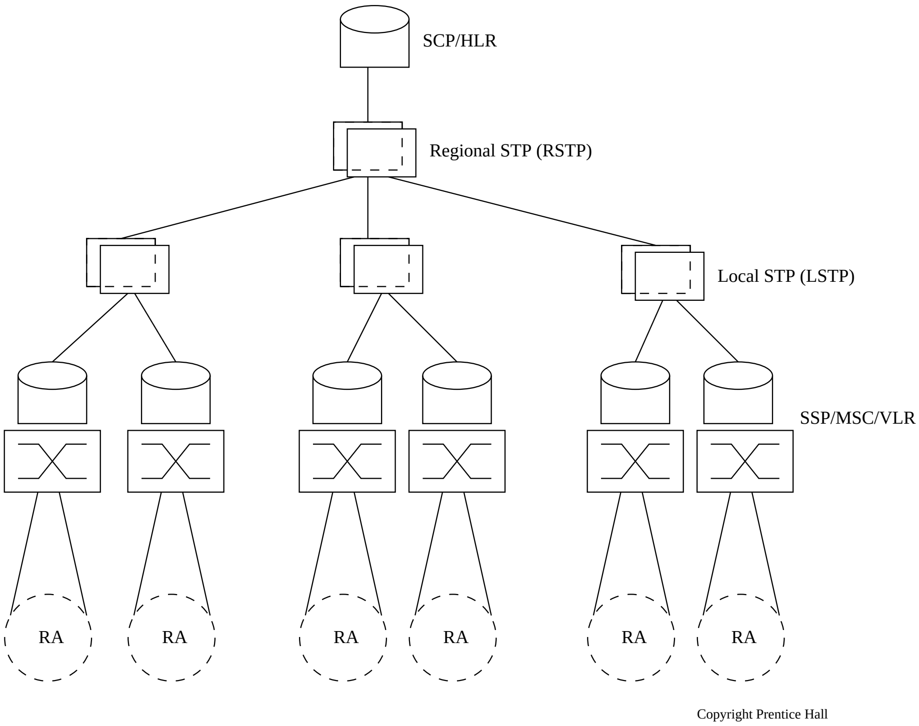


Figure 5: SS7 signaling network

Features of SS7 Signaling Network and Common Channel Signaling

- The SS7 network carries out network signaling exchange operations using common channel signaling (CCS).
- CCS is a digital communications technique that provides simultaneous transmission of user data, signaling data, and other related traffic throughout the network.
- CCS uses out-of-band signaling channels and is implemented is a time division multiplexing (TDM) format for serial data transmission.
- The SS7 network is organized in a tree structure. The HLR and SCP (service control point) are colocated at the root of the tree. The SCP contains the functionality associated with the HLR.
- RAs are located at the leaves of the tree.
- Signals from the root (SCP) are distributed to the subtrees by a RSTP (regional signal transfer point).
- Each subtree is supported by a local signal transfer point (LSTP).
- The SSP (service switching point), MSC and VLR are colocated.
- Each RA is served by one SSP., which is in turn connected to the LSTP.

Location Update

- When a mobile moves to a new RA, it has to identify its current location.
- Base stations in each RA periodically broadcast beacon signals; the mobile listens to the pilot signals from the base stations in the RA and uses these to identify its current RA.
- A mobile is identified by a mobile identification number (MIN); registration requests sent by the mobile should contain its MIN.

Call Setup

- When a mobile wants to send messages to a remote mobile, it send a request to its serving MSC in the registration area via its serving base station.
- The calling MSC locates the called MSC/VLR combination through its association with the HLR.
- The HLR routes the call request to the called MSC.
- The called MSC pages the called mobile to set up the call.

Paging by Polling

- Paging is used to locate a mobile within a registration area.
- *Polling*: The called MSC broadcasts a polling message containing the called mobile's MIN to all cells within the RA.
- The BS of each cell relays the paging message to the mobile terminals.
- Upon receiving the paging message, the called mobile responds to the called MSC, with the BS ID of its current cell.

Location Update Procedure

- Signal flow, based on IS-41 network architecture:

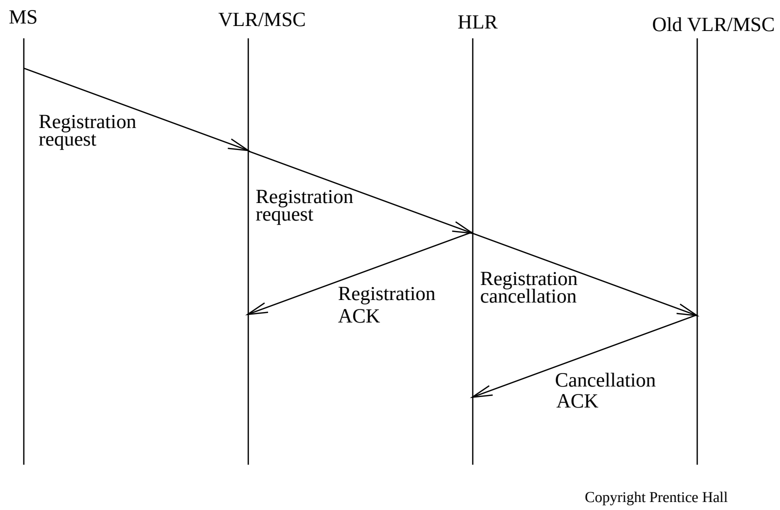


Figure 6: Location registration signaling flow diagram for IS-41

(1) The mobile determines in which RA it is located by monitoring the the beacon signals from the BSs in the RA.

(2) It sends a registration request, containing its MIN, to the MSC/VLR via its serving BS.

(3) The VLR updates its record and forwards the registration request to the HLR.

(4) The HLR determinees the mobile's identity from the MIN.

(5) The HLR performs the necessary authentication routine and updates the database entry for the mobile by recording the identity of the new serving VLR, and sends a registration acknowledgement message to the new VLR and a registration cancellation message to the old VLR.

(6) The old VLR returns a cancellation acknowledgement message to the HLR.

Call Setup Procedure

- Signal flow:

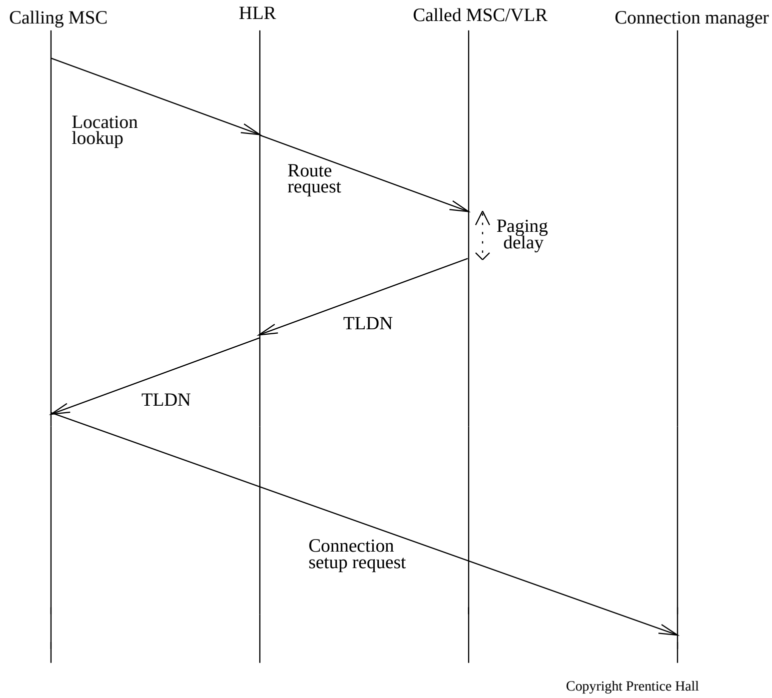


Figure 7: Call setup signaling flow diagram

(1) The calling BS receives a call request. The request should contain the MIN of the called mobile. The calling BS forwards the request to the VLR of its RA. The VLR then forwards the request to its corresponding MSC.

(2) The calling MSC sends a location lookup request to the HLR of the called mobile.

(3) The HLR of the called mobile determines the current VLR of the called mobile (from its database) and sends a routing request message to the VLR. The VLR then forwards the message to its corresponding MSC.

(4) The called MSC allocates a temporary local directory number to the called mobile and sends a reply to the HLR, together with the TLDN.

(5) The HLR forwards the TLDN to the MSC of the calling mobile.

(6) Using the TLDN, the MSC of the calling mobile initiates a connection request to the called MSC through the SS7 network.

Location Management for PCS Networks

- Future PCS networks, i.e., 3G and beyond, will provide higher transmission rates and operate at higher regions of the frequency spectrum (in the range of 1.85-1-99 GHz compared to 850 MHz range in 2G cellular systems).
- The two-tiered architecture of IS-41 and MAP is inadequate to support future PCS networks.
- To enlarge network capacity, future PCS networks will have smaller cells $\Rightarrow$ smaller registration areas and more frequent registration with the HLR.
- Capacity enlargement $\Rightarrow$ larger network $\Rightarrow$ HLR can be very far away from its access point.
- Higher capacity will support my users $\Rightarrow$ HLR can be a bottleneck to location management.
- Larger distances $\Rightarrow$ longer signaling delays.
- Possible Remedy: introduce methods to reduce the frequency of accessing the HLR.

Possible Methods

1. Pointer Forward - An Overlay Approach

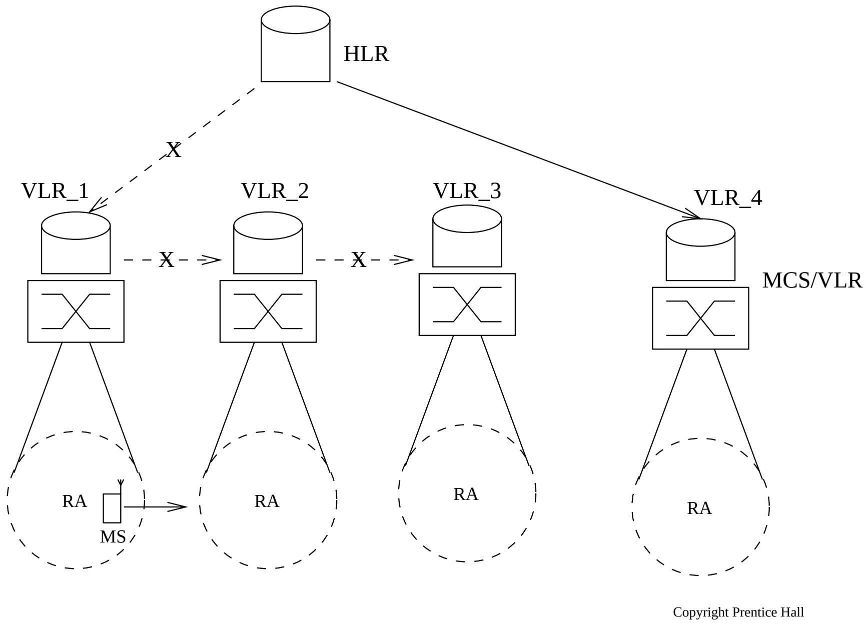


Figure 8: Pointer forwarding scheme

- n RAs, each associated with an MSC/V are link-chained together.
- The length of the link chain is $n - 1$.
- Only necessary to update registration with the HLR, when the mobile enters the first or the nth RA.
- If the delay of call setup in the entire link chain is too long, may have to force registration update with HLR after $K < (n - 1)$ moves, even if the mobile is still quite close to its last reported location.

2. Local Anchor Approach

- A VLR, colocated with an MSC, is selected as the local anchor (LA).
- The LA created for a given mobile acts as a virtual HLR as the mobile migrates from one RA to another RA.
- Thus, the LA oversees the registration for the mobile over several RAs.
- Different mobiles may have different LAs, and the LA for a mobile may change from time to time. Why?
- Once a VLR is selected as the LA for a particular mobile, an entry is set up in a table at this VLR (the LA) indicating the current serving MSC/VLR.

Signal Flow for Registration Update with the LA

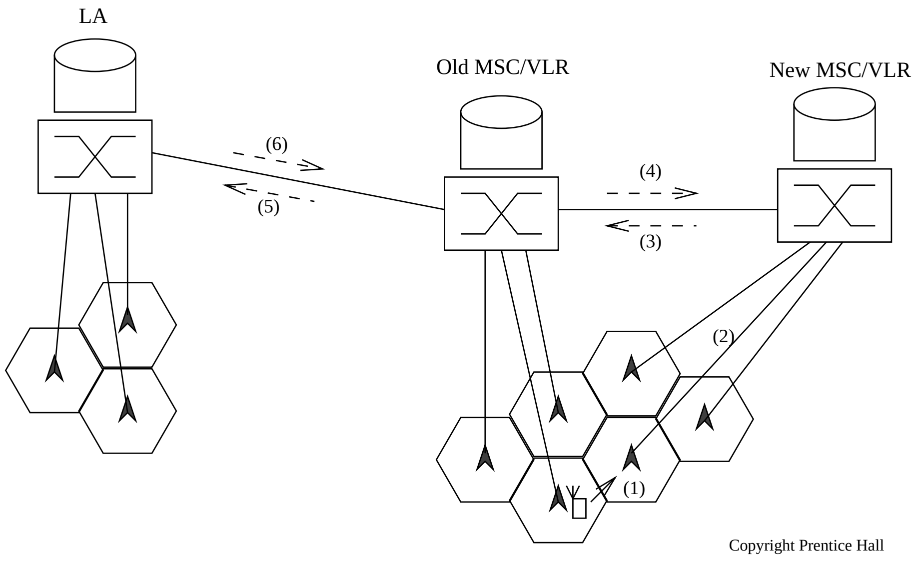


Figure 9: Signaling flow for reporting location change to the local anchor

- The numbered signal flows in the diagram illustrate the mobile migration, reporting to its new MSC/VLR and establishment of an association between the new MSC/VLR and the LA through the old MSC/VLR.

Call Delivery

- The MSC overseeing the RA where the calling mobile is located is called the calling MSC.
- The MSC associated with the RA where the called mobile resides is called the called MSC.
- When the calling mobile initiates a call, the calling MSC sends a location request message to the mobile's HLR.
- The numbered signal flows in the diagram below trace out the call delivery procedure.

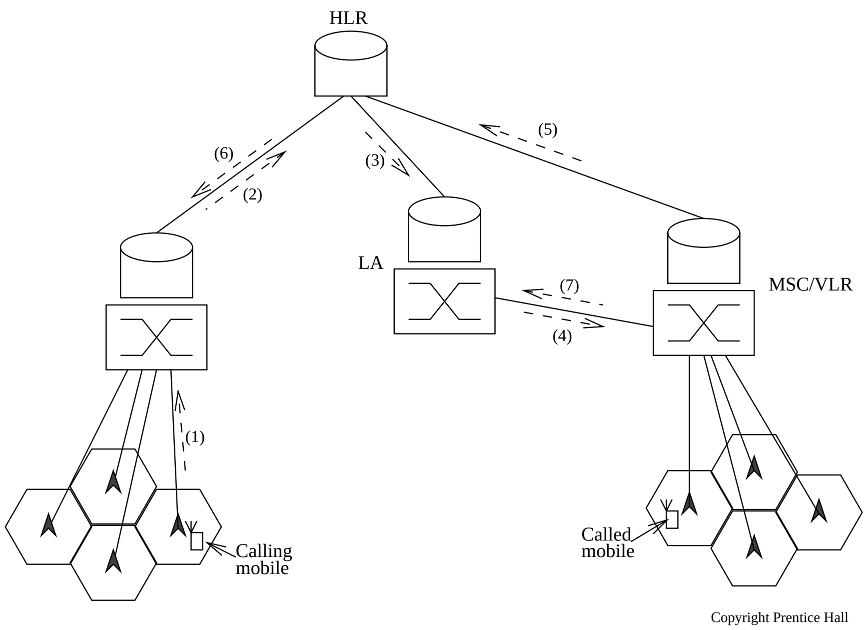


Figure 10: Signaling flow for call delivery in LA-based scheme

Performance Measure in Mobility Management

- The rate of boundary crossing is a tangible performance measure.
- Rate of boundary crossing = rate at which a randomly chosen mobile crosses the cell boundary $\times$ the number of mobiles per cell.
- Cells with BSs connected to the same MSC collectively form a *cluster*. When a mobile moves within the same cluster, it does not need to update registration with its HLR. Boundary crossings between cells within a cluster are referred to as intracluster crossings.
- Boundary crossings between clusters are referred to as intercluster crossings.
- The rate of boundary crossing in synonymous with the rate of handoff.
- Of interest are the intracluster handoff rate and intercluster handoff rate.
- Handoff rate calculations depend on system and traffic parameters, and mobility patterns. Mobility patterns can be obtained by measurement or analytical modeling. Analytical models are constructed based on known distributions, which do not precisely portray the mobility patterns, but yield sufficient guidance for the design of handoff and location update algorithms.
- System parameters include cell radius, R; number of cells in a cluster, N; cell perimeter, L_{cell}; cluster radius, $R_{cluster}$, cluster perimeter, $L_{cluster}$.
- Traffic parameters include population density of mobiles, ρ; mobile speed, V; traffic load per mobile station, λ_{ar} in Erlangs; percentage of powered mobile stations, δ; percentage of active mobiles among the powered stations, ϵ; number of crossings per cell per unit time, μ_{cell}; number of crossings per cluster per unit time, $\mu_{cluster}$.

Universal Mobility Model

a. Mobiles are uniformly distributed in the cell.

b. Direction of travel of all mobiles relative to the cell boundary is uniformly distributed on $[0,\ 2\pi]$.

c. After admitted into a cell, the mobile assumes a direction to travel at a constant speed V.

Handoff Rate

Cell topology: hexagonal

- Rate of mobiles departing a cell:

$$\mu_{cell} = \frac{\rho V L_{cell}}{\pi} = \frac{6\rho V R}{\pi}$$

- Cluster perimeter, $L_{cluster} \approx 6R_{cluster} \approx 6\sqrt{N}R$
- Rate of mobiles departing a cluster:

$$\mu_{cluster} \approx \frac{6\rho V\sqrt{N}R}{\pi}$$

Cell Splitting

- If a regular cell is split into K microcells, then the radius of the microcell is $R_{microcell} = R_{cell}/\sqrt{K}$.
- If a cell is split into K microcells, the total handoff rate in the cluster is increased by K-fold, while the intercluster handoff rate is increased by $\sqrt{K}$-fold.

Chapter 8. Wireless/Wireline Interworking

- Wireless/wireline interworking is a means to provide wider geographical coverage. The wireless front-end supports user roaming, while the wireline, e.g., the Internet, provides extended coverage.

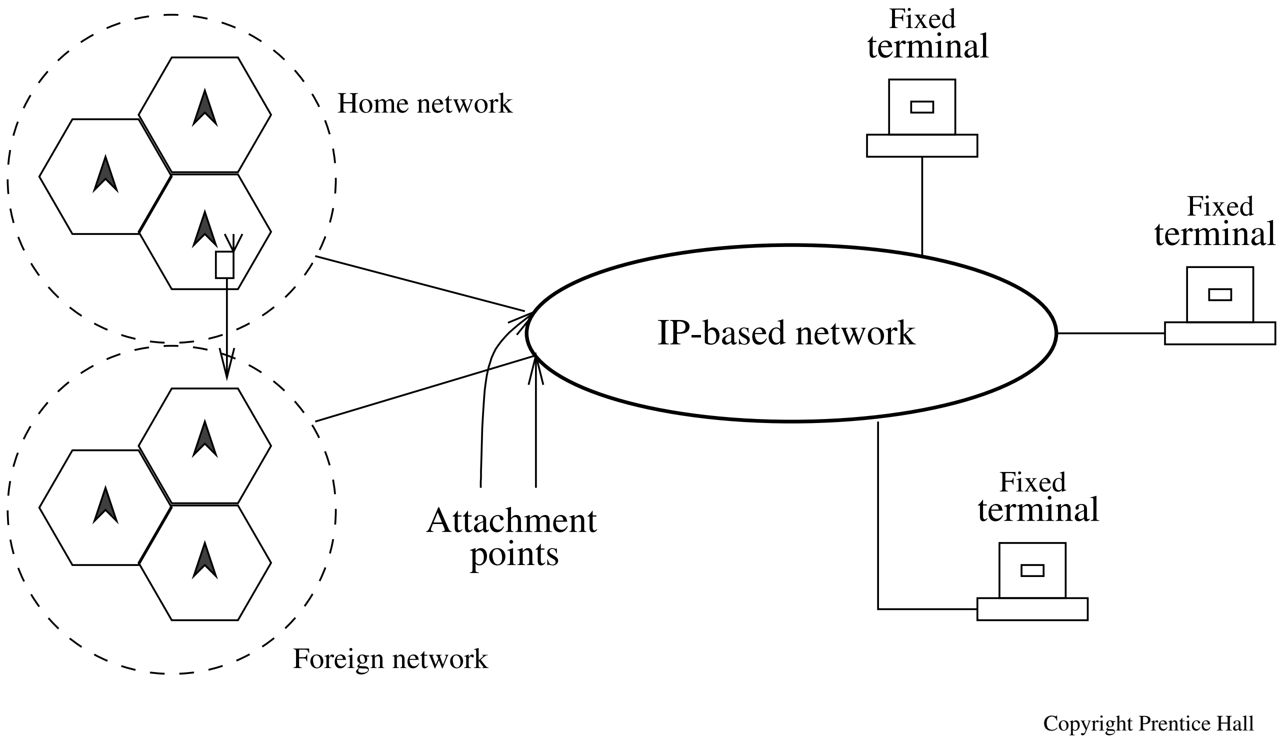


Figure 1: Hybrid wireless/IP-based network

- Internet Protocol (IP) is the main network layer packet transfer mechanism in Internet. IP only offers best effort service.
- Transmission Control Protocol (TCP) is a connection-oriented point-point transport layer protocol to provide end-to-end reliable and in-sequence data transfer over the best effort IP-based network.
- TCP uses a window mechanism to throttle the traffic at the edge router by shrinking the window size when necessary.
- Another transport layer protocol is User Datagram Protocol (UDP), which delivers packets from the source to destination without any reliability or in-sequence guarantee.
- Two of the most pressing problems in wireless/Internet interworking are
  - a) the delivery of messages from the Internet to the mobile at its current location, and - b) traffic control to protect network integrity and to satisfy end-to-end quality of service (QoS) requirements.

Mobile IP

Mobile IP (MIP) is an extension network proposed by the Internet Engineering Task Force (IETF) for delivery of messages from the Internet server to the mobile at its current location in a seamless manner.

Mobile IP uses an agent concept. The mobile has a *home agent* (HA) and a *foreign agent* (FA). The home agent maintains a database in which the mobile's home address resides.

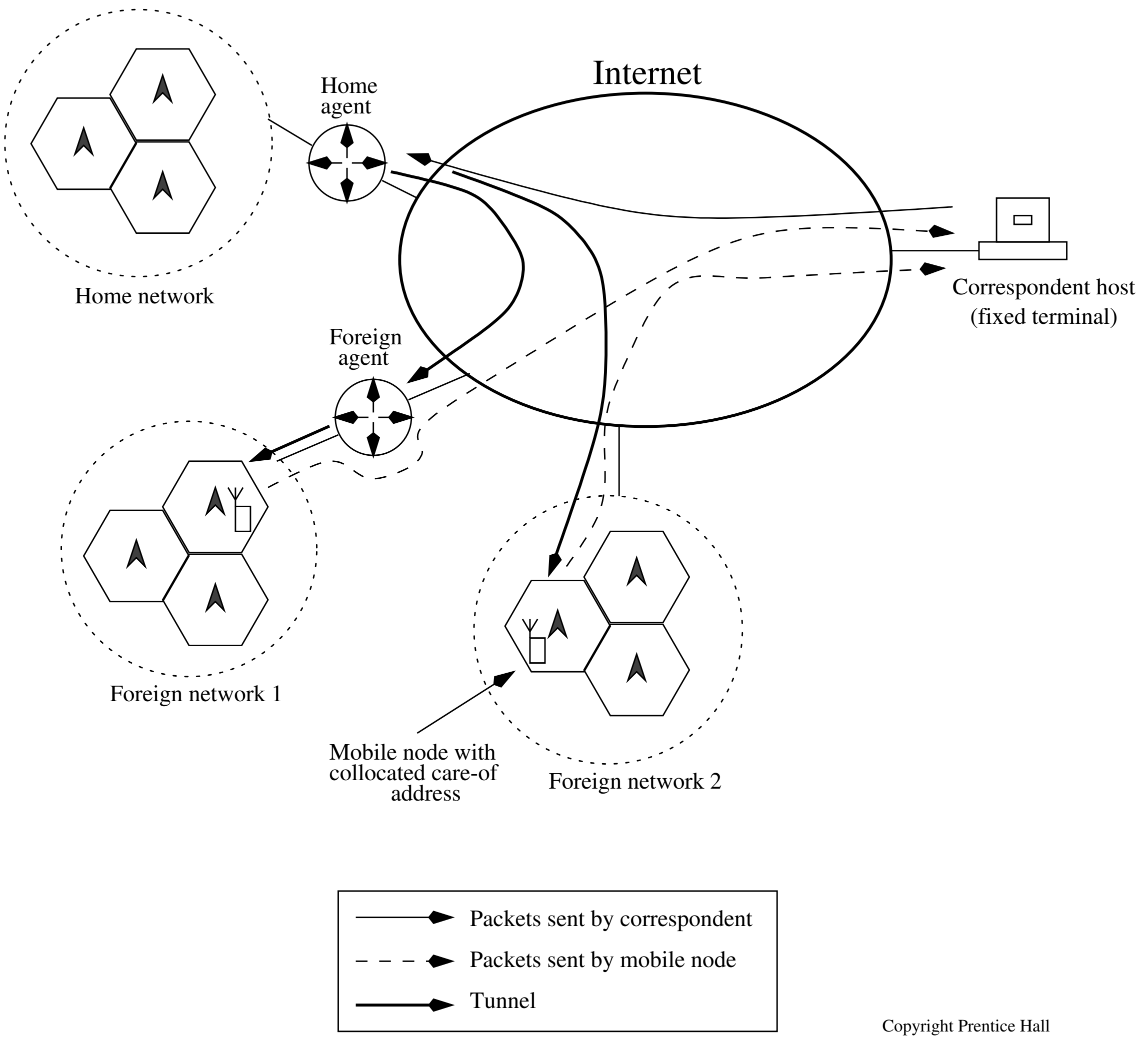
Messages from the **Correspondent Node** destined for the **Mobile Node** are delivered directly to the mobile's permanent address residing in its home network. Each mobile has two IP addresses: a fixed *home address* for identification and a *care-of address* for routing.

Entities which govern the operation of Mobile IP

- **Home address**: This is the permanent address assigned to mobile node.
- **Care-of address**: Datagrams sent to the mobile's home address and destined for the mobile have to be tunnelled to the foreign network where the mobile is currently visiting. The address in the foreign network to which the datagrams are tunnelled for delivery to the mobile is called the care-of address.
- **Collocated care-of address**: an externally obtained local IP address temporarily assigned to an interface of the mobile node.
- **Home agent**: a router with an interface on the mobile node's home network. The home agent intercepts packets destined to the mobile node's home adddress and tunnels them to the mobile node's current location via the care-of address.
- **Foreign agent**: a router with an interface on a mobile node's visiting network, which assists the mobile node in informing its home agent of its current care-of address.

- **Foreign agent care-of address**: an IP address of a foreign agent, which has an interface on the foreign network being visited by a mobile node. A foreign agent care-of address can be shared by many mobile nodes simultaneously.
- **Home network**: a network having a network prefix matching that of the mobile node's home address.
- **Foreign network**: a network other than a mobile node's home network to which the mobile node is currently connected.
- **Virtual network**: a network with no physical instantiation beyond its router. The router usually uses a conventional routing protocol to advertise how it can be reached to the virtual network.
- **Link**: a facility or medium over which nodes can communicate at the link layer.
- **Link-layer address**: an address that identifies the physical endpoint of a link. Usually, the link-layer address is the interface's Medium Access Control (MAC) address.
- **Mobile node's home link**: the link which has been assigned the same network-prefix as the network prefix of the mobile node's home address.
- **Mobile node's foreign link**: the link that the mobile node is visiting, which has been assigned the same network prefix as the network prefix of the mobile node's care-of address.
- **Agent advertisement**: the process in which foreign agents advertise their presence by using a special message.
- **Agent solicitation**: the message sent by a mobile node to request agent advertisement.
- **Tunnel**: the path followed by a datagram while it is encapsulated.
- **Binding entry**: an entry in the home agent's routing table. Mobile IP maps the mobile node's home address into its current care-of address.

Entities in Mobile IP



Operation of Mobile IP

The operation of Mobile IP is based on the cooperation of three major Processes:

Agent Discovery: From agent advertisements sent periodically by agents (home, foreign or both), agent discovery enables a mobile node to
- (a) determine whether it is connected to its home link or a foreign link,
- (b) detect whether it has changed its point of attachment, - (c) obtain a care-of address if it is connected to a foreign link.

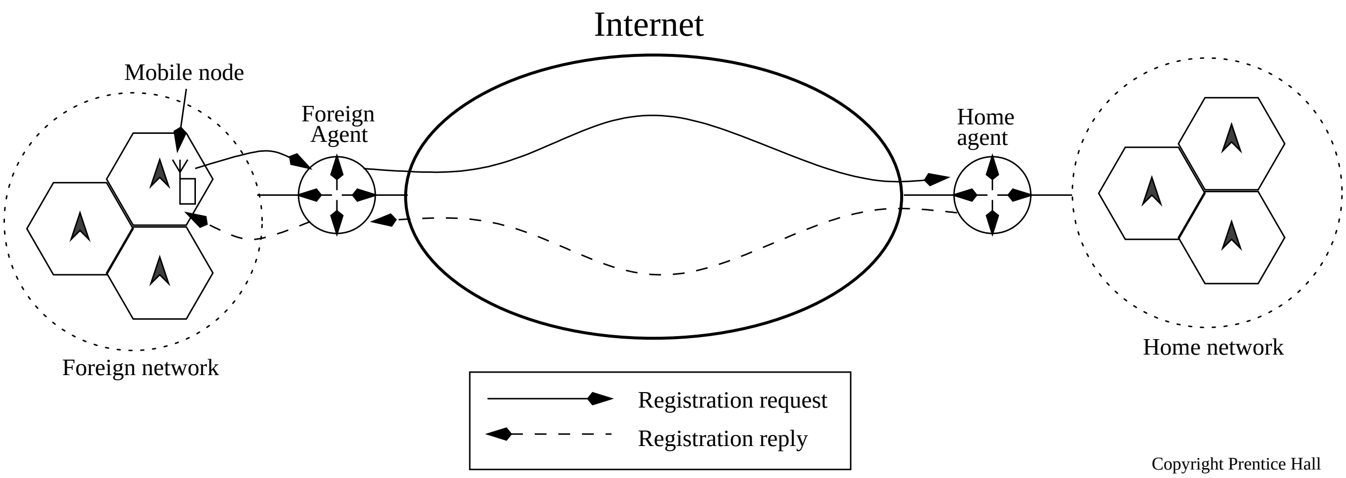
When a mobile node detects that it has moved, it acquires a care-of address by reading it directly from an agent advertisement, or contacting Dynamic Host Configuration Protocol (DHCP), or using the mannual configuration.

Registration: When a mobile moves to a foreign network, it requests the foreign agent to inform its home agent of the mobile's care-of address. Registration also involves reregistration upon expiration of a current registration and deregistration as the mobile node returns to its home link. Registration is in the form of request and reply between the mobile node and its home agent. Agent discovery message is carried by the Internet Control Message Protocol (ICMP) payload portion; registration message is carried by the User Datagram Protocol (UDP).

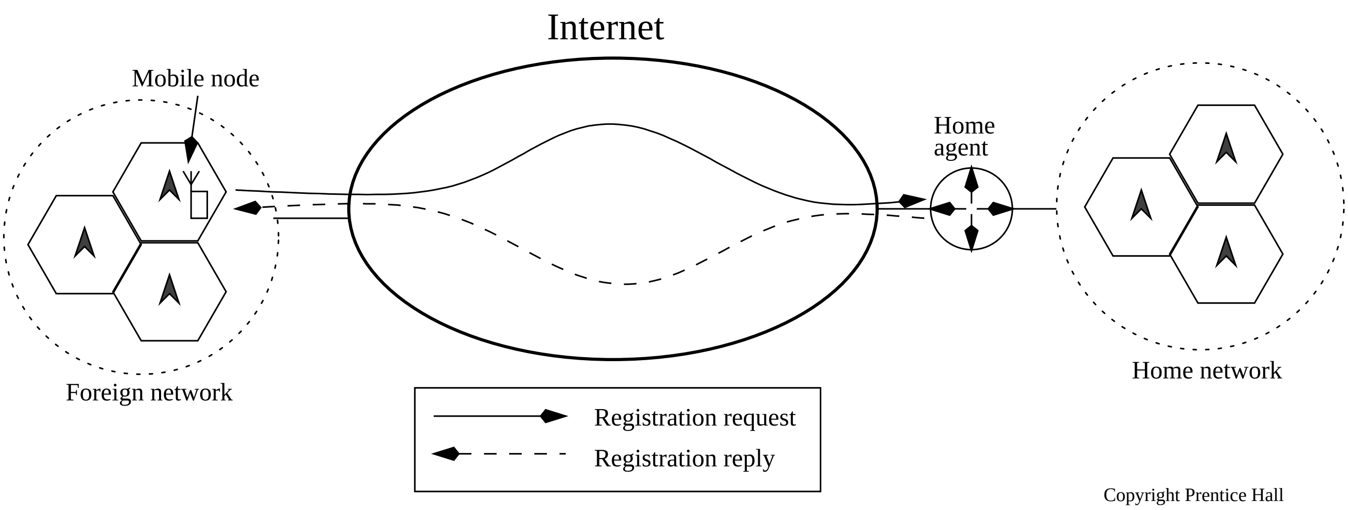
Tunnelling (routing): a process by which Mobile IP tunnels datagrams to the mobile node, whether it is or is not away from its home network.

Example of Registration and Deregistration

Registration with foreign agent's care-of address



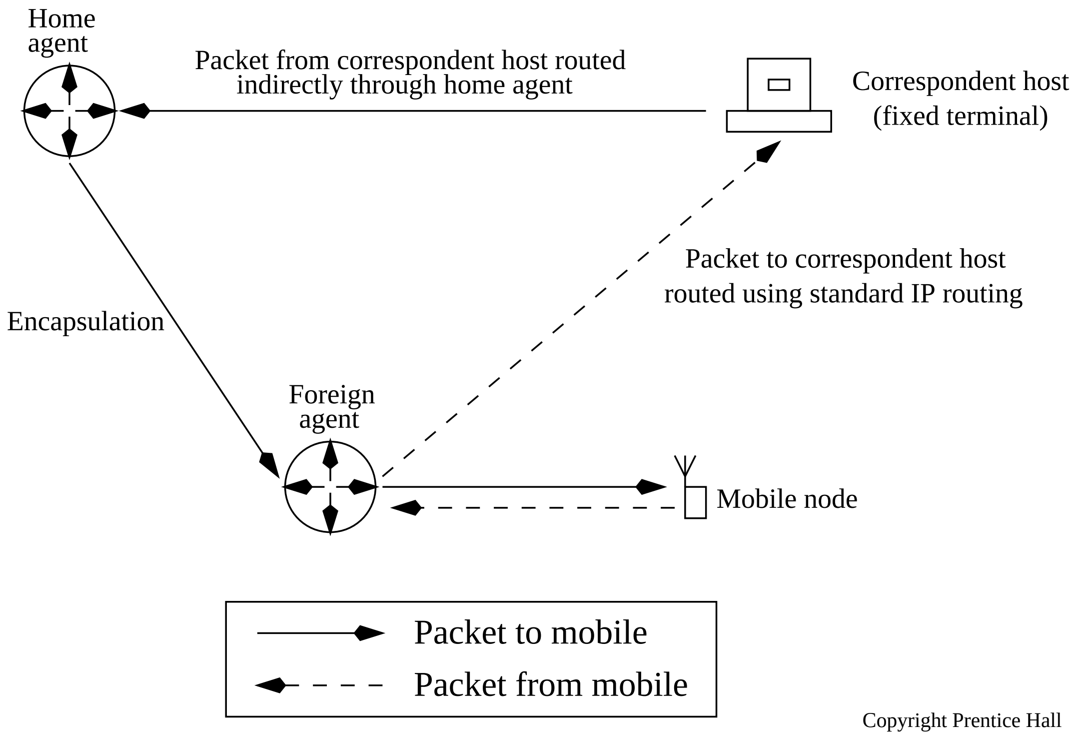
Deregistration with collocated care-of address



Routing in Mobile IP

Two routing approaches: Triangular routing and optimized routing

1. Triangle routing Datagram delivery steps



- A datagram from the correspondent for the mobile is sent to the mobile's home network using standard IP routing.
- Upon arrival at the home network, the datagram is intercepted by the home agent which, in turn, tunnels the datagram to the mobile's care-of address.
- At the foreign agent, the datagram is detunneled and delivered to the mobile.
- For datagrams sent by the mobile, standard IP routing is used to deliver each datagram to its destination, where the foreign agent is the default router.

What are the drawbacks of triangle routing?

2. Optimized routing

Salient feature

- Datagrams sent by the correspondent host are tunnelled directly to the mobile host, by passing the home agent.
- This direct tunnelling capability is facilitated by informing the correspondent host of the mobile host's care-of address when the mobile host moves to the foreign network and acquires a care-of address.
- The Mobile IP protocol with optimized routing allows every traffic source to cache and use binding copies.
- A moving host can always inform its previous foreign agent about the new care-of address, so that packets tunneled to the old location (because of out-of-date binding copy) can be forwarded to the current location.
- This should increase the overall quality of service in the case of high mobility.

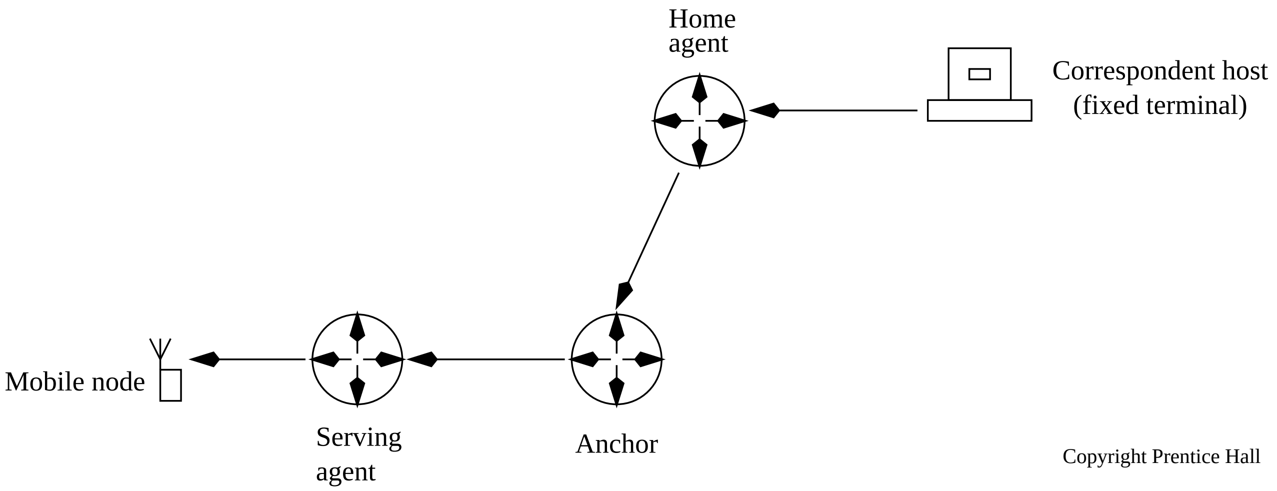
Disadvantages

- Optimized routing is much more complex than triangle routing. The complexity is mainly due to (i) large number of message exchanges, e.g., the correspondent host must be informed of the mobile host's care-of address, and (ii) security issues.
- The overhead incurred by message exchanges and processing (due to cache queries) can be critical.
- Cached bindings are possibly inconsistent since they are being kept in a distributed fashion.
- In a hostile environment, an intruder can easily cut off all communications to the mobile host by sending a bogus registration if it knows the mobile host's care-of address.
- This implies that authentication/security measures have to be incorporated in optimized routing $\Rightarrow$ increased complexity.

Methods to reduce registration costs

Two possible approaches: *Local anchor* and *Hierarchical routing*

1. Local Anchor Approach



Local Anchor Strategy

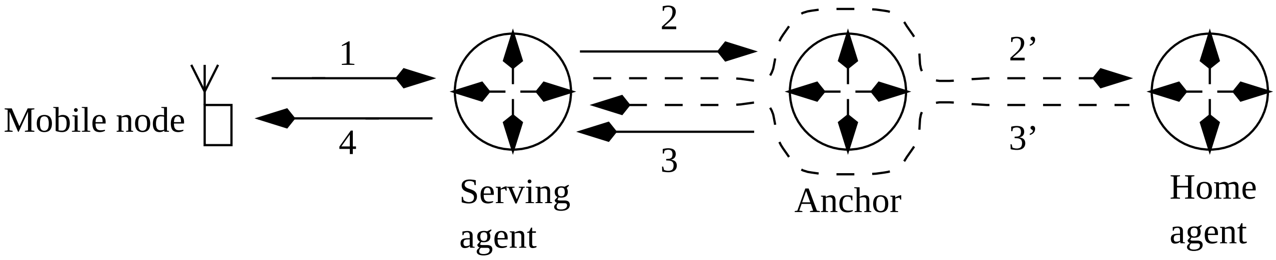
a. Choose one agent as the focus of an anchoring region and name this agent as an anchor.

b. When the mobile moves within the anchoring region, it does not need to register with its home agent; instead, it registers with the anchor,i.e., the local anchor acts as a virtual home agent.

c. When the mobile moves out of the anchoring region, it will register with its home agent and the new foreign agent will become the focus of the new anchoring region.

Question: Local Anchoring has the effect of reducing registration cost, but what effect does it have on packet delivery?

Example of Registration with Local Anchor Registration procedure



(1) The mobile sends the registration request indicating the current anchor agent and the home agent.

(2) There are two cases –
Case I: the new foreign agent decides that the mobile is still in its current anchoring region so it forwards the mobile's registration request to the anchor.
Case II: the new foreign agent decides that the mobile is out of its current anchoring region, so it forwards the mobile's registration request to the home agent.

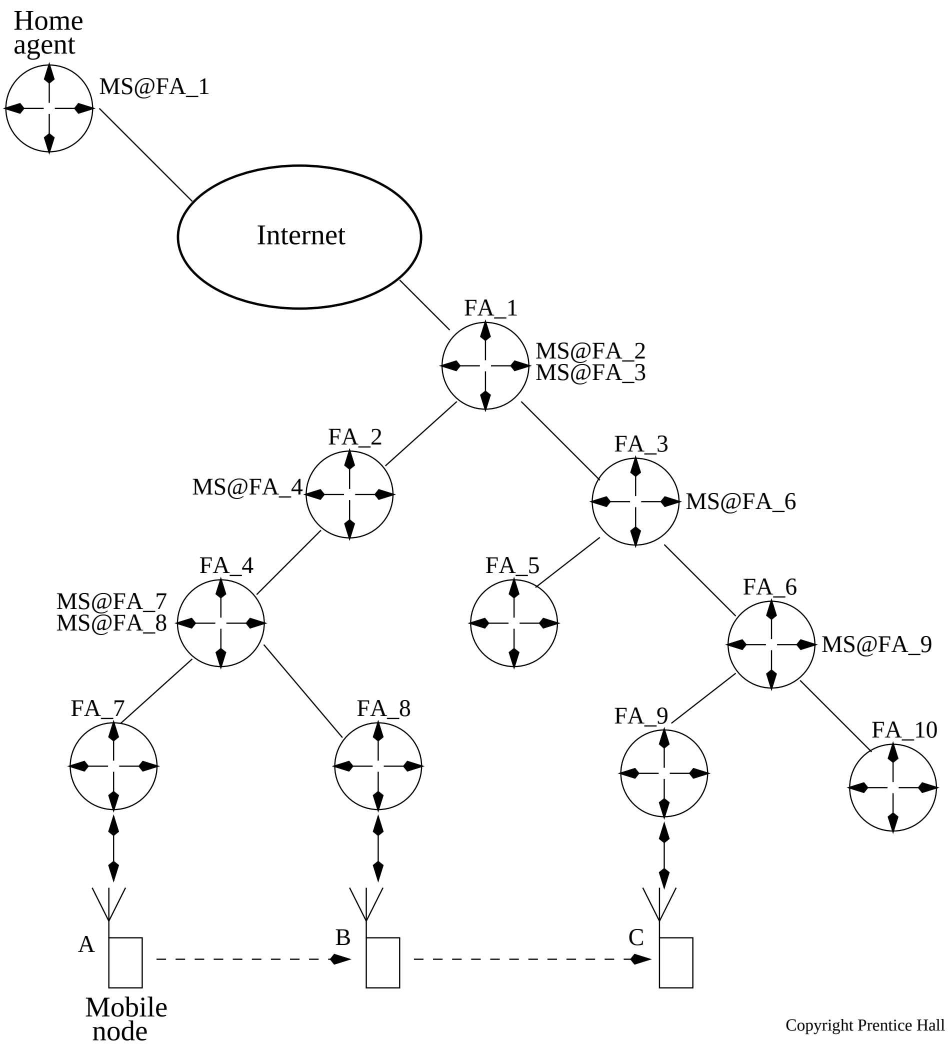
(3) The anchor agent or the home agent sends the registration reply back to the foreign agent.

(4) The foreign agent returns an acknowledgment to the mobile and indicates who, the anchor or the home agent, sends this registration reply. In Case I, the mobile knows that it has not moved out of the current anchoring region and the anchor does not change. In Case II, the foreign agent becomes the focus of the new anchoring region and the mobile will update its anchor agent's IP address for later use.

2. Hierarchical Routing

- Populate foreign agents in a tree structure.
- Agent advertisement is used as a means to advertise multiple foreign agents.
- Registration is localized to the foreign agent that is the lowest common ancestor of the care-of addresses at the multiple attachment points.

Example of binary tree structure



The home agent is at the root of the binary tree, and only directly sees the nearest foreign agent, i.e., FA_1, which is the common ancestor of two branches.

INTERNET PROTOCOL (IP)

- The Internet Engineering Task Force (IETF) was the international body that worked and evolved different version of the Internet Protocol. An early one was version 4 (IPv4), which was the protocol that supports Internet administration and operation.
- IPv4 was the basis for the original development of Mobile IP.
- The basic IPv4 header has a 32-bit source address and a 32-bit destination address.

IPv4 Header

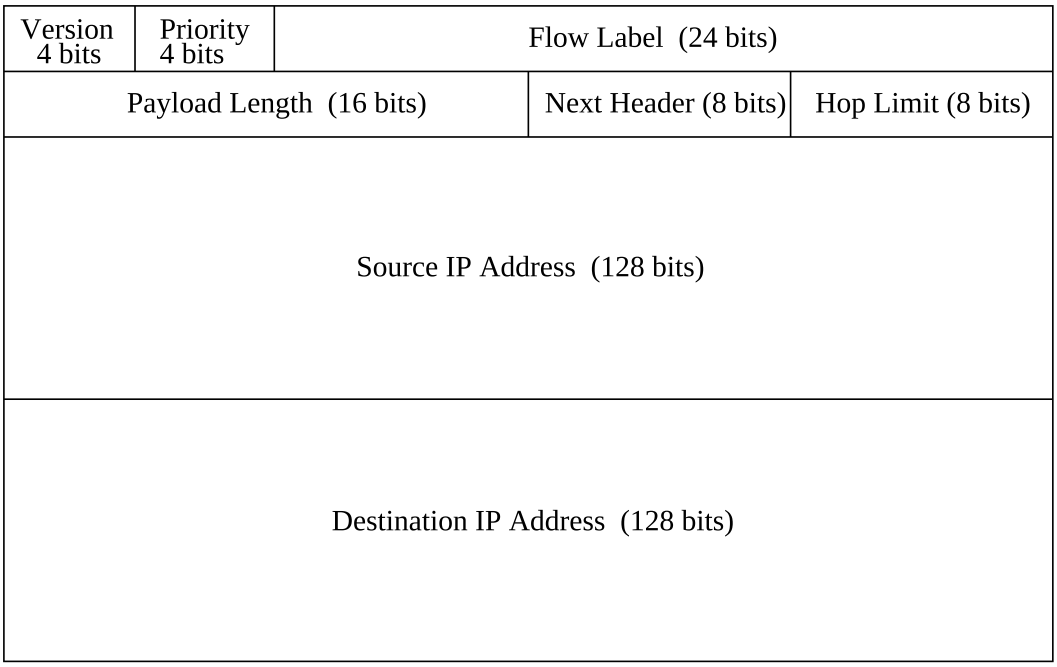
<table>
<tr><td>Version
4 bits</td><td>Header
Length</td><td>Type of Service
8 bits</td><td colspan="2">Total Length of Datagram
16 bits</td></tr>
<tr><td colspan="3">Datagram Identification (16 bits)</td><td>Flag
3 bits</td><td>Fragment Offset (13 bits)</td></tr>
<tr><td colspan="2">Time to Live
8 bits</td><td>Protocol
8 bits</td><td colspan="2">Header Checksum
16 bits</td></tr>
<tr><td colspan="5">Source IP Address (32 bits)</td></tr>
<tr><td colspan="5">Destination IP Address (32 bits)</td></tr>
<tr><td colspan="5">IP Options</td></tr>
<tr><td colspan="5">Data Portion of Datagram</td></tr>
</table>

- With only a 32-bit source and destination addresses, the address allocation scheme in IPv4 is insufficient to support the rapidly growing Internet subscriber population.

Internet Protocol version 6 (IPv6)

- The IETF introduced IPv6 to provide a remedy to the limitations inherent in IPv4, in terms of addressing, routing, mobility support, quality of service (QoS) provisioning, etc.
- The network entities of IPv6 are similar to those of IPv4, i.e., mobile nodes, home agent, home address, care-of address, etc., with the exception that IPv6 does not have the concept of foreign agent.
- The IPv6 header has a 128-bit source address and a 128-bit destination address.

IPv6 Header



- With a 128-bit address space, IPv6 has room to accommodate the Internet growth expansion.

IPv6 vs IPv4

- IPv4 has a two-level addressing scheme, with a fixed prefix and a host address. Once a network address is assigned to a particular, a block of IP addresses is assigned to that network, which may lead to a waste of address space.
- With a large address space, IPv6 allows the Internet to support more levels of addressing hierarchy, a much greater number of addressable nodes, and simpler auto-configuration of addresses, and supports unicast, anycast and multicast addressing.

Unicast address can be divided into two parts: a subnet prefix, which indicates the node's subnetwork; an interface ID.

Multicast address is divided into two groups: predefined groups (permanently assigned); transient groups (defined by specific organizations). Packets destined to a multicast address are sent to all nodes in that group.

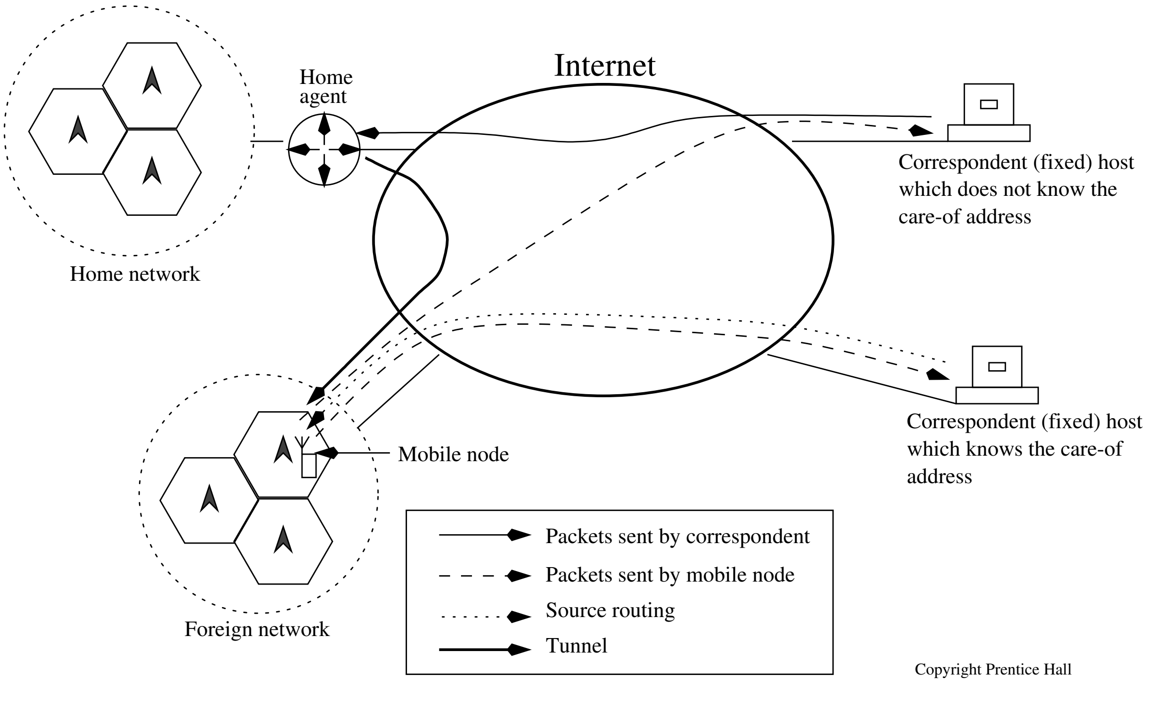
Anycast address is defined to identify sets of nodes such that a packet sent to an anycast address is delivered to any one of the nodes assigned that address. Packets destined to an anycast address are sent to only one node in that group.

The use of anycast address in IPv6 allows nodes to control the path along which their traffic flows.

Coexistence of IPv6 and IPv4 is facilitated by the use of IPv4-compatible-IPv6 addresses and IPv4-mapped-IPv6 addresses. IPv4-compatible-IPv6 addresses are assigned to those hosts and routers running IPv6, but must route traffic across IPv4 networks. IPv4-mapped-IPv6 addresses are assigned to those hosts running IPv4.

Mobile IPv6

- Mobile IPv6 uses the new and improved IPv6 *Routing Header* and the *Authentication Header*, and other pieces of IPv6 functionality.
- Different from Mobile IPv4, Mobile IPv6 has no foreign agent.
- The mobile node uses the *Address Autoconfiguration* procedure defined in IPv6 to acquire a collocated care-of address on a foreign link and report the care-of address to the home agent and selected correspondents.
- **Operation of Mobile IPv6**



- In Mobile IPv6, a correspondent node with knowledge of the care-of address can send packets directly to the mobile node using an IPv6 Routing Header. Those correspondent nodes with no knowledge of the care-of address have to send packets to the home agent for forwarding.

Mobile IPv6 has the following advantages over Mobile IPv4

a. The enormous address space in IPv6 allows very simple address autoconfiguration by means of Stateless Address Configuration (SAC). Because the mobile node can easily obtain a collocated care-of address by SAC, the foreign agent functionality is no longer needed. As a result, the foreign agent is eliminated from Mobile IPv6. This also implies that all Mobile IPv6 care-of addresses are collocated care-of addresses.

b. Mobile IPv6 uses the new IPv6 routing header to simplify routing to mobile nodes.

c. With the enhanced authentication header in IPv6 and the mandatory implementation of IP authentication header, Mobile IPv6 might adopt a wide scale of route optimization if a key management infrastructure becomes widely available on the Internet.

d. Mobile IPv6 uses both tunneling and source routing to deliver packets to mobile nodes. In the case of Mobile IPv4, tunneling is the only option.

Transmission Control Protocol (TCP)

- TCP is a transport layer window flow control strategy.
- The window size in TCP is regulated by a control signal fed back from the receiver to the transmitter.
- Flow control is based on acknowledgement (ACK) and timeout, and interlaced with slow start and fast retransmit phases.

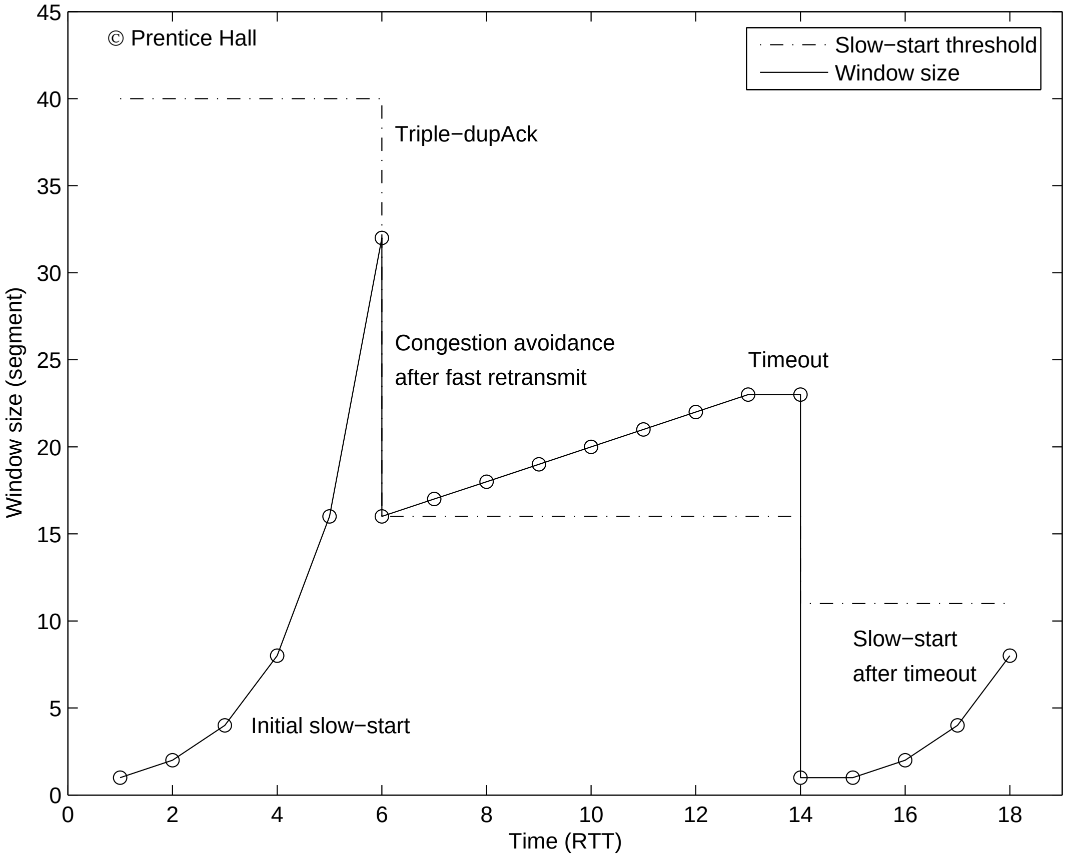


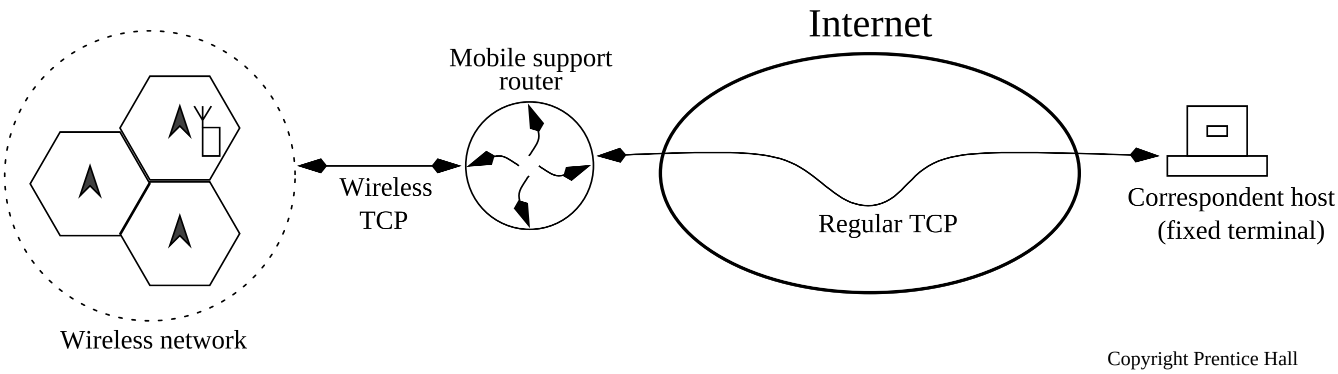
Figure 2: TCP window flow control strategy

- In the slow start phase, the window is small initially and increases until congestion is reached. Then the window is decreased by half and start fast retransmit to avoid congestion.
- When the sender fails to receive n acknowledgements , TCP enters a timeout phase, followed by slow start. 3 duplicate acknowledgements, i.e., triple-DupAck, is a normally accepted approach.
- TCP is thus an additive increase and multiplicative decrease window flow mechanism,i.e., increase by 1 and decrease by half the window size.

Modified TCP

- In TCP, packet loss indication via feedback is strictly based on network congestion. In a hybrid wireless/wireline network, packets might be already in error due to transmission errors and/or connection disruption because of handoff incompletion.
- If TCP can take transmission errors as well as congestion losses into consideration, then TCP can be a single connection across the hybrid network and provide end-to-end semantics.
- Conventional TCP is not aware of packet loss due to transmission errors. One way to modify TCP is to split the connection, and the resultant protocol is referred to as Indirect TCP (I-TCP) as shown in the diagram below.

Indirect Transmission Control Protocol (I-TCP)



- Here, TCP is split at the mobile support router, i.e., the base station, separating the wireless domain from the wireline domain.
- I-TCP is split into a regular TCP for the wireline network and wireless TCP for the wireless domain.

    a. the flow control and congestion control functionalities on the wireless link are separated from those on the wired link;

    b. a separate transport protocol for the wireless link can support notification of events such as disconnections, user movements and other features of the wireless link to the higher layers;

    c. a partition of the connection into two distinct parts allows the base station to manage much of the communication overhead for a mobile host.

Handling of Packet Loss at the Split Point

A. Berkeley Snoop Module

- Snoop, residing at the split point (base station) in I-TCP, is a solution for losses caused by high transmission errors.
- The **Berkeley Snoop Module** caches packets at the base station, inspects the TCP header of TCP data packets and acknowledgements that pass through, and buffers copies of the data packets.
- Using the information from the headers, the snoop module detects lost packets and performs local retransmissions across the wireless link to alleviate problems caused by a high BER.
- The Snoop module also implements its own retransmission timer and performs selective retransmissions when an acknowledgment is not received when no duplicate acknowledgements are received.
- Routing protocol is also modified to enable low-latency handoff to occur with negligible data losses.
- Experiment indicate that the Berkeley Snoop Module achieves a throughput up to 20 times that of regular TCP, and handoff latencies over 10 times shorter than those of other mobile routing protocols.
- The Snoop module does not perform as well in either the presence of lengthy disconnections or environments in which there are frequent handoffs.

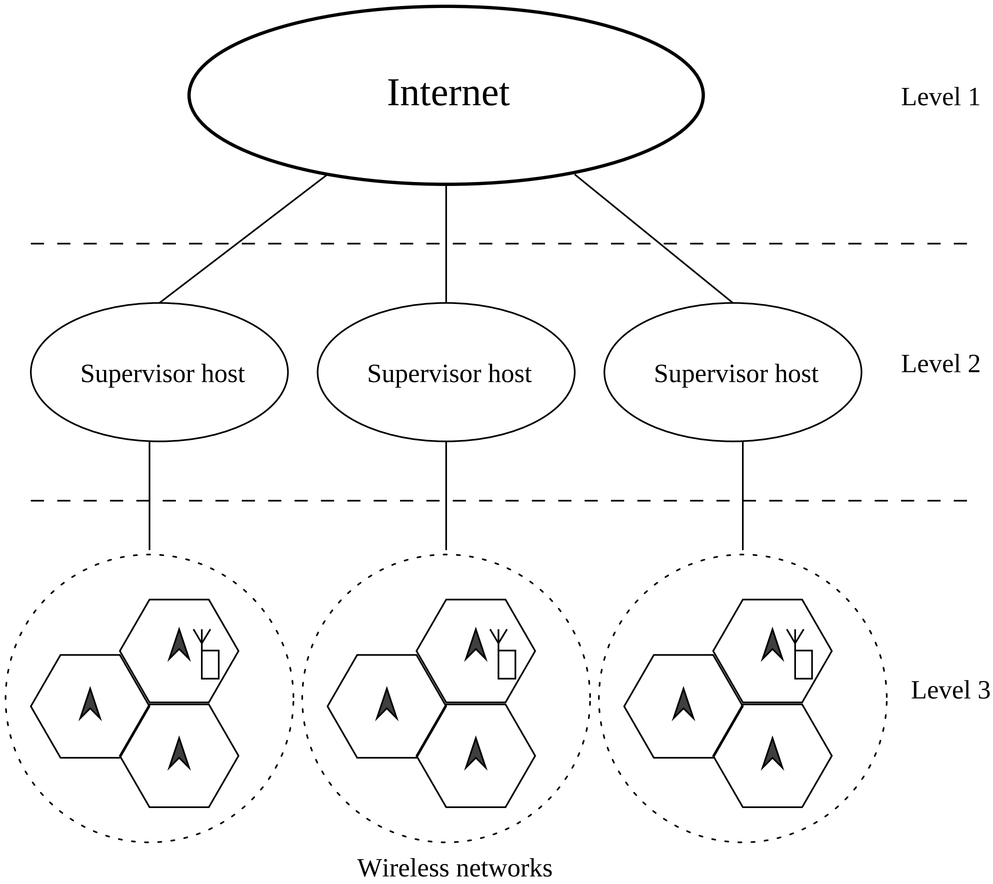
B. Fast Retransmission

- Fast retransmit is used to combat the effects of short disconnection on TCP throughput due to handoff disruptions.
- If the mobile host is disconnected for a lengthy period of time, the sender will automatically invoke congestion control because it does not receive acknowledgments for some segments.
- Fast Retransmit forces the mobile host to retransmit, in triplicate, the last old acknowledgment as soon as it finishes a handoff.
- This forces the sender to reduce the congestion window to one-half and retransmit one segment immediately.
- Fast Retransmit does not split the TCP connection.
- If the mobile host were disconnected for a long time, the sender would already have invoked congestion control and shrunk its window to one segment.
- Fast Retransmit will do little to improve the throughput because the sender's congestion window will repeatedly get shrunk to half of its previous size.

TCP for Mobile Cellular Networks (M-TCP)

- With the introduction of a *supervisor host* (SH) to interconnect a cluster of cell to the Internet, M-TCP is a three-level hierarchical architecture, as shown below.

M-TCP hierarchy



- Features of M-TCP

  a. When a mobile host roams from one cell to another, the two base stations do not need to transfer any state information if they are controlled by the same SH.

  b. The roaming mobile host remains within the domain of the same SH for long time periods because several base stations are controlled by the SH.

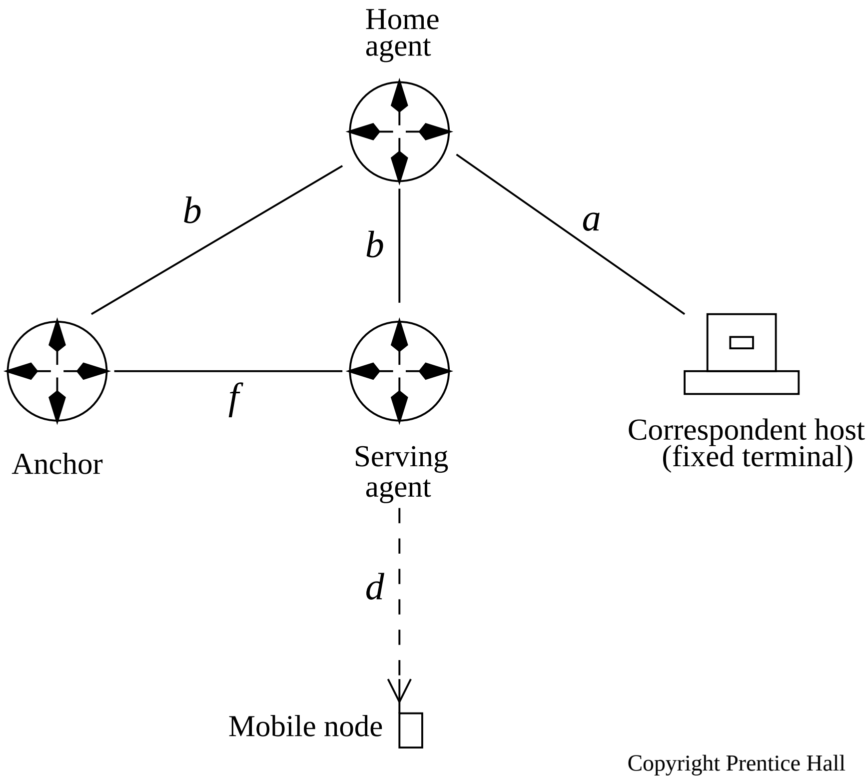
- By introducing the SH, M-TCP maintains end-to-end TCP semantics while it delivers excellent performance when mobile hosts encounter disconnections. This is done by splitting the TCP connection at the SH. As packets arrive from a sender on the Internet, an acknowledgment is sent back and the SH deals with ensuring the completion of delivery.

- The drawbacks of this scheme are complexity and possibility of buffer shortage at the SH.

NETWORK PERFORMANCE

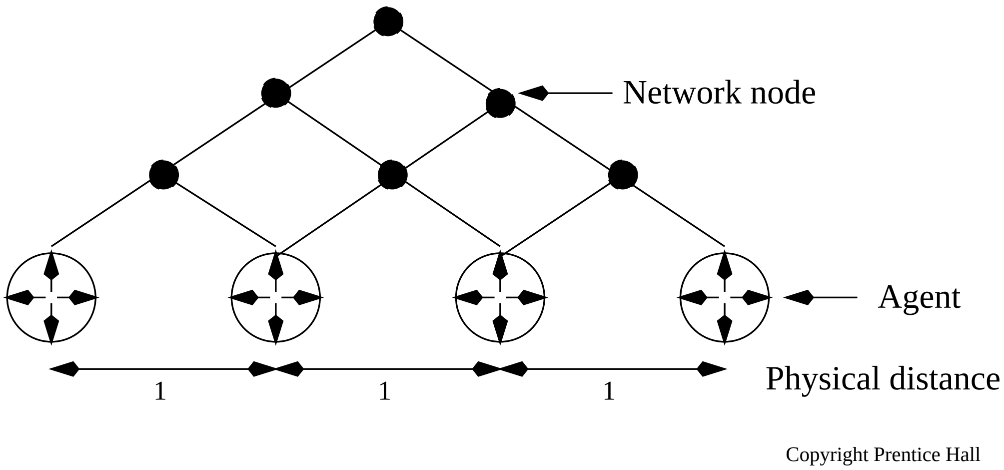
Approximate Performance Analysis of *Local Anchor Mobile IP*

- **Network model for cost (delay) analysis**



- The model shown above is for the case the mobile receives information from the correspondent host, in which the correspondent host generates data packets at a fixed rate λ.
- The parameters a, b, d, and f are the costs (delays) associated with the corresponding paths.

- Suppose that a, b, and d are fixed but f is a variable.
- We wish to quantify f.
- Assume the relationship between the physical distance and the network distance can be modeled by a bifork tree shown below:

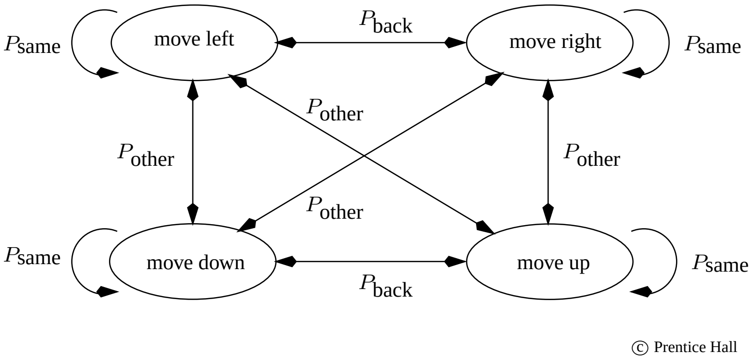


- In the bifork tree model, if any two agents are neighbors in the physical location, they will have a distance (and therefore a delay) of two hops between them.
- Let d_p be the physical distance between two agents and assume each hop has a unity delay. Then $f = 2 \times d_p$.

Two-dimensional Mobility Model

- A two-dimensional Markov mobility model is used to capture the effect of user movement on average handoff delay.
- The network is divided into service areas.
- User movements in the network are modeled by boundary crossings between serving areas.
- The residence time of the mobile in each serving area (the inter-handoff time) is assumed to be an exponentially distributed random variable with a constant mean value.
- We characterize the mobile's movement by a Markov state transition diagram based on the following assumptions:
    a. Calls to the mobile are modeled as a Poisson arrival process with a constant rate.
    b. Calls are assumed to be generated for randomly selected serving areas.
    c. The service areas are rectangle radio cell clusters of equal size.
    d. Thus, each cell cluster has four adjacent neighboring clusters; the mobile can move into one of four adjacent serving areas, with respect to the mobile's current serving area.
- Each cluster has a unique attachment point to the Internet.

- The direction of each movement is modeled as a Markov process, as shown below:

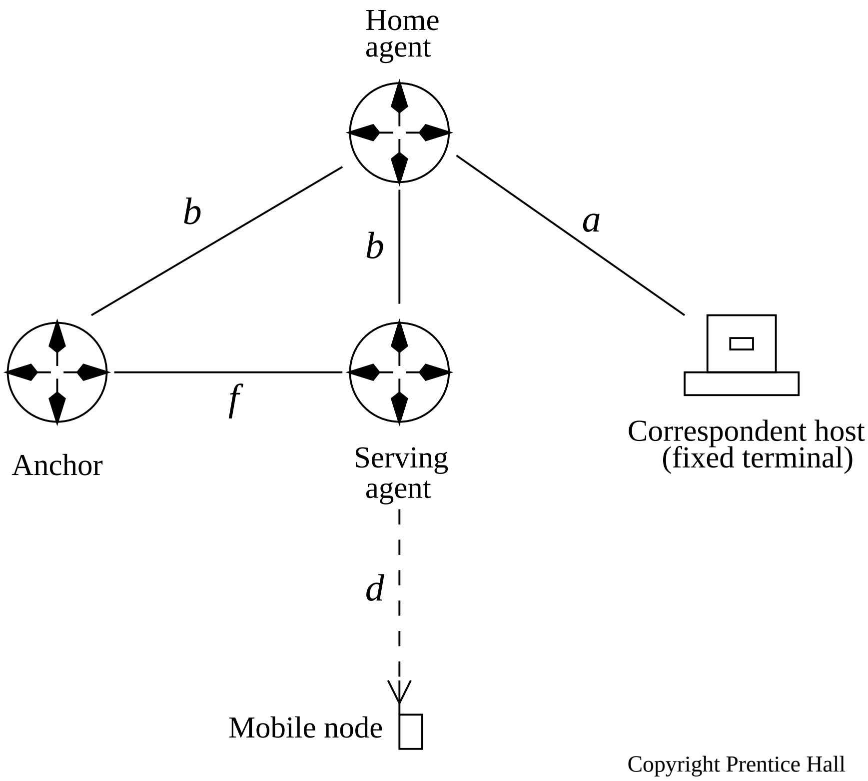


- P_{back} is the probability that the mobile will move back to its previous serving area.
- P_{same} is the probability that the next move will be in the same direction as the previous move.
- $P_{\text{other}} = (1 - P_{\text{same}} - P_{\text{back}})/2$ is the probability of the mobile moving in any other direction.

Handoff Delay with Local Anchor

Definition of handoff delay: The time interval from the instant when the mobile sends the registration request to the new foreign agent to the instant when the mobile is allowed to send/receive packets to/from the new foreign agent.

- Consider the local anchor-based network:



- Let r denote the fixed processing time (cost) at each agent.

Handoff delay within the anchoring region

- The registration request and reply only need to go through paths with delays d and f, incurring a registration delay of

$$t_{\text{h\_anchor}} = 2d + 2f + 3r. \tag{1}$$

Handoff outside the anchoring region

- The registration delay for an anchor-based indirect TCP (I-TCP) is given by

$$t_{\text{I-TCP}} = 2(b + d) + 3r. \tag{2}$$

- After a successful registration, the new foreign agent sends a message to the old foreign agent and requests TCP state transfer, incurring a transfer delay of

$$t_{\text{transfer}} = (n + 2)(f + r), \tag{3}$$

  where n is the TCP buffer size.

- When the mobile host has to register with its home agent, the total handoff delay, $t_{\text{h\_I-TCP}}$, for the anchor-based I-TCP scheme is

$$t_{\text{h\_I-TCP}} = 2(b + d) + 3r + (n + 2)(f + r). \tag{4}$$