無線通訊系統
(Wireless Communications Systems)

國立清華大學電機系暨通訊工程研究所
蔡育仁
台達館 821 室
Tel: 62210
E-mail: yrtsai@ee.nthu.edu.tw

Chapter 3
Physical Layer Technologies for Wireless Communication Systems

Modulation Schemes

Modulation

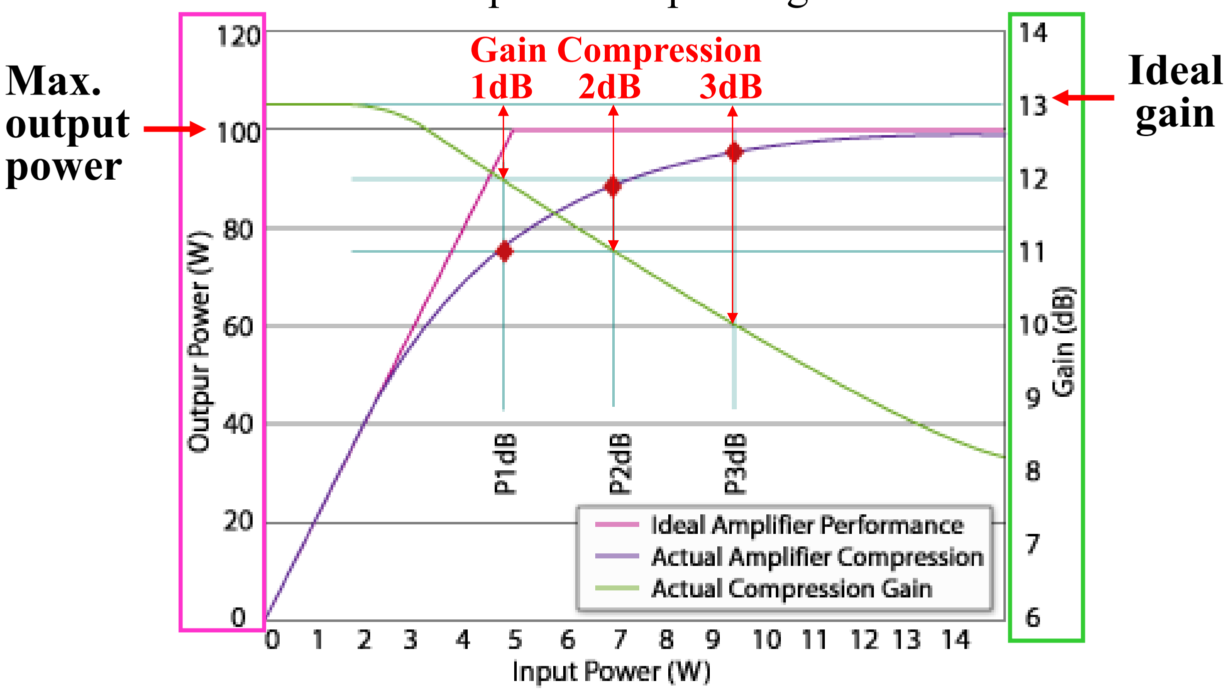
- Message information can be transmitted in the **amplitude**, **frequency** or **phase** of the carrier
- For mobile radio applications:
  - We shall use bandwidth and power resources most efficiently
- **Bandwidth efficiency:** measured in bits/sec/Hz
  - The modulation must have compact power density spectrum
- **Good BER performance:**
  - Good BER must be maintained in mobile radio environments
  - The modulation must be able to prevail over fading, ISI, Doppler spread, adjacent and co-channel interference, and thermal noise
- **Envelope properties:**
  - Portable transmitters normally use **nonlinear power amplifiers**
  - Modulation with a **relatively constant envelope** shall be used

Power Amplifier

- A RF **power amplifier** (PA) can amplify a small input RF signal of a certain shape to a larger one with the same shape at the output.
- The **gain** is defined as the ratio of output power to input power.
- Since the maximum output power is limited, a PA will definitely be **saturated** when input power gradually increases.
- The saturation of a PA is defined as the situation that the output signal **is not proportional to** the input signal.
  - **Distortion** occurs in the output signal
- It is highly desirable to operate a PA in the **linear region**
  - The input signal is faithfully re-produced at the output

Power Amplifier

- Theoretical and actual power-amplifier gain



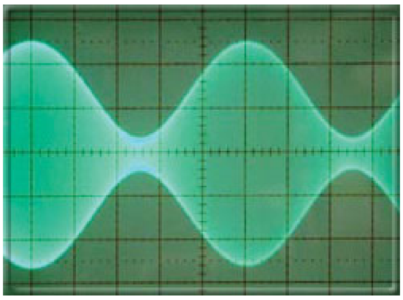
Power Amplifier

- The amount of **usable power** depends on the applications
  - It depends on how much **distortion** can be tolerated
- A commonly used figure-of-merit is the **1-dB compression point**. This is the output power level at which the gain is compressed by 1 dB.
  - However, the decreased gain affects only those parts of the signal with amplitudes **close to or greater than** the 1-dB compression point.

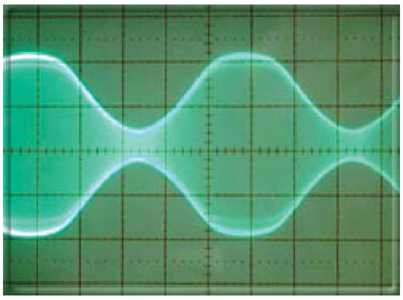
Power Amplifier

- For 1-dB compressed output waveform, the maximum input power is equal to the “**P1dB**” input power
  - **Smaller average output power with less distortion**
- For 3-dB compressed output waveform, the maximum input power is equal to the “**P3dB**” input power
  - **Larger average output power with more distortion**

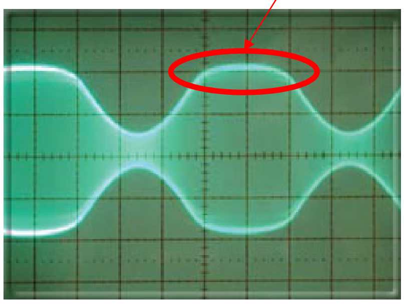
High frequency components exist



Input Waveform



1-dB Compressed Output Waveform



3-dB Compressed Output Waveform

Power Amplifier

- We may select the required amplifier power by **backing off** an acceptable amount from P1dB to prevent signal distortion.
- For example, if an application requires a maximum output power of **50 W**, the system designer may select a PA with a **P1dB of 100 W or higher** to ensure that the PA is within its **linear region** of operation.
  – It depends upon the **linearity requirement** of the system.
- For the same tolerable distortion, a signal with a high peak-to-average ratio will **degrade** the **average output signal power**, as well as the efficiency of the PA.
  – Or equivalently, it increases the **linearity** and **dynamic range** requirements of the PA

Modulation

- First generation systems:
  – **FM modulation** (AMPS, TACS, NMT)
- Second generation systems:
  – **π/4-DQPSK** (IS-136, PDC, PHS)
  – **GMSK** (GSM, DCS1800, DECT)
- Third generation systems:
  – CDMA – **QPSK／Offset QPSK** (cdma2000, WCDMA)
- Fourth generation systems:
  – OFDM – **BPSK, QPSK, 16-QAM, 64-QAM** (WiMAX, LTE)

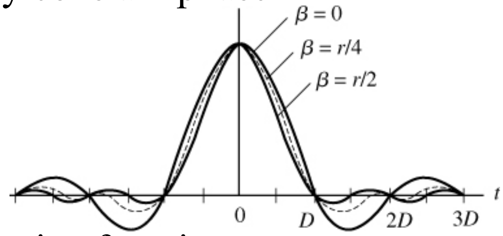
Phase Shift Keying (PSK)

- The M-PSK complex envelope can be expressed as

$$\tilde{s}(t) = A\sum_{n} b(t - nT, \mathbf{x}_n)$$

$$b(t, \mathbf{x}_n) = h_a(t)\exp\left(j\theta_n\right)$$

  - $\theta_n = \dfrac{2\pi}{M}x_n + \theta_0$ and θ_0 is an arbitrary constant phase
  - $x_n = n,\ n \in \{0, 1, \ldots, M - 1\}$
  - M is the alphabet size
  - $h_a(t)$ is the amplitude shaping pulse
    - Normally chosen to be a raised cosine function

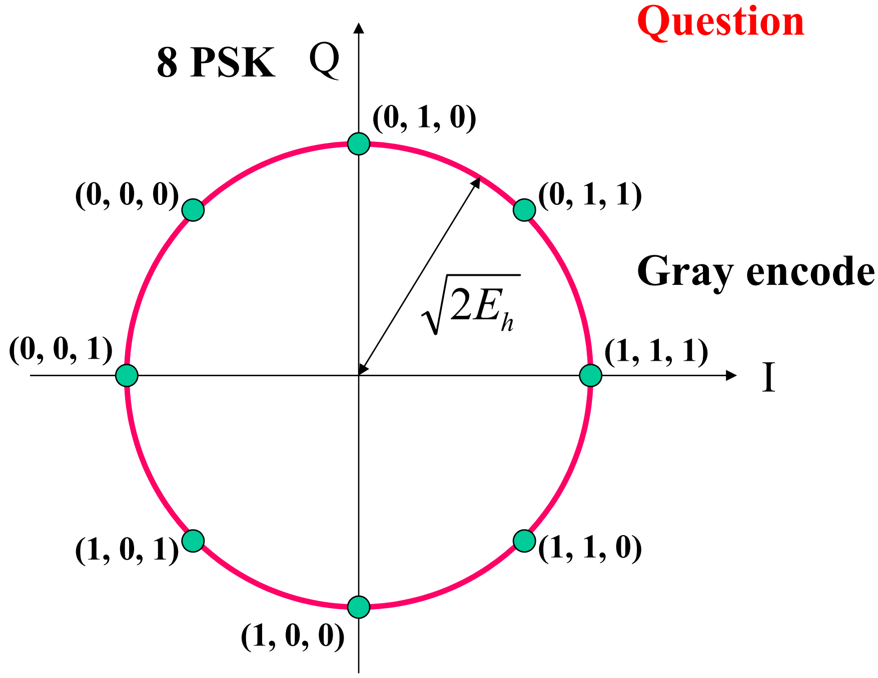


- The complex envelope is

$$\tilde{s}_n(t) = Ah_a(t)e^{j\theta_n}, \quad n = 0, 1, \cdots, M - 1$$

Phase Shift Keying (PSK)

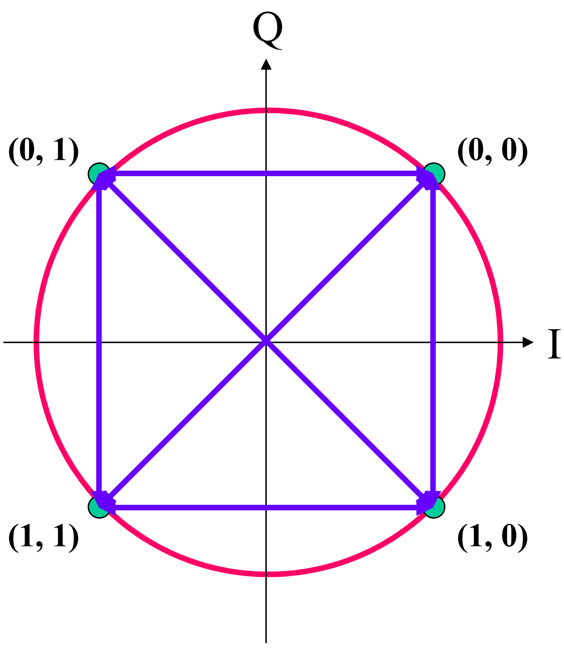
Question



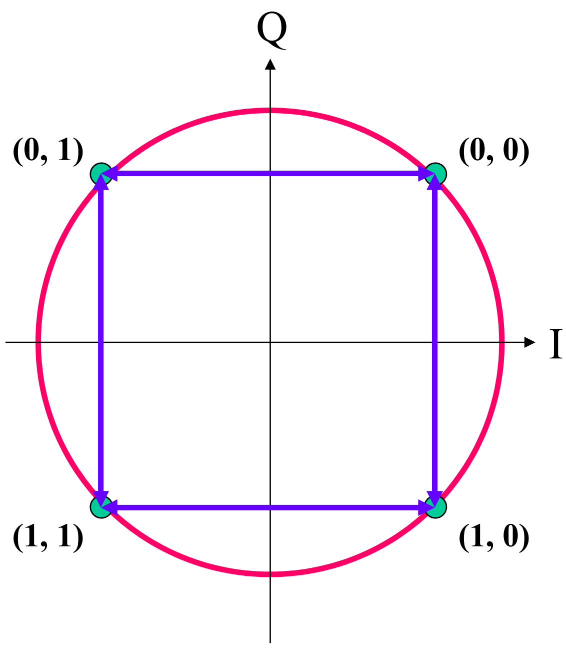
Offset QPSK (OQPSK)

- The QPSK signal can have either ± **90° or** ± **180°** phase shift from one baud interval to the next
- The OQPSK signal can change by only ± **90°**
- The phase trajectory does not pass through **the origin**
  - Reduce the peak-to-average ratio of the complex envelope
  - Make the OQPSK signal less sensitive to amplifier nonlinearity than the QPSK signal

Offset QPSK (OQPSK)

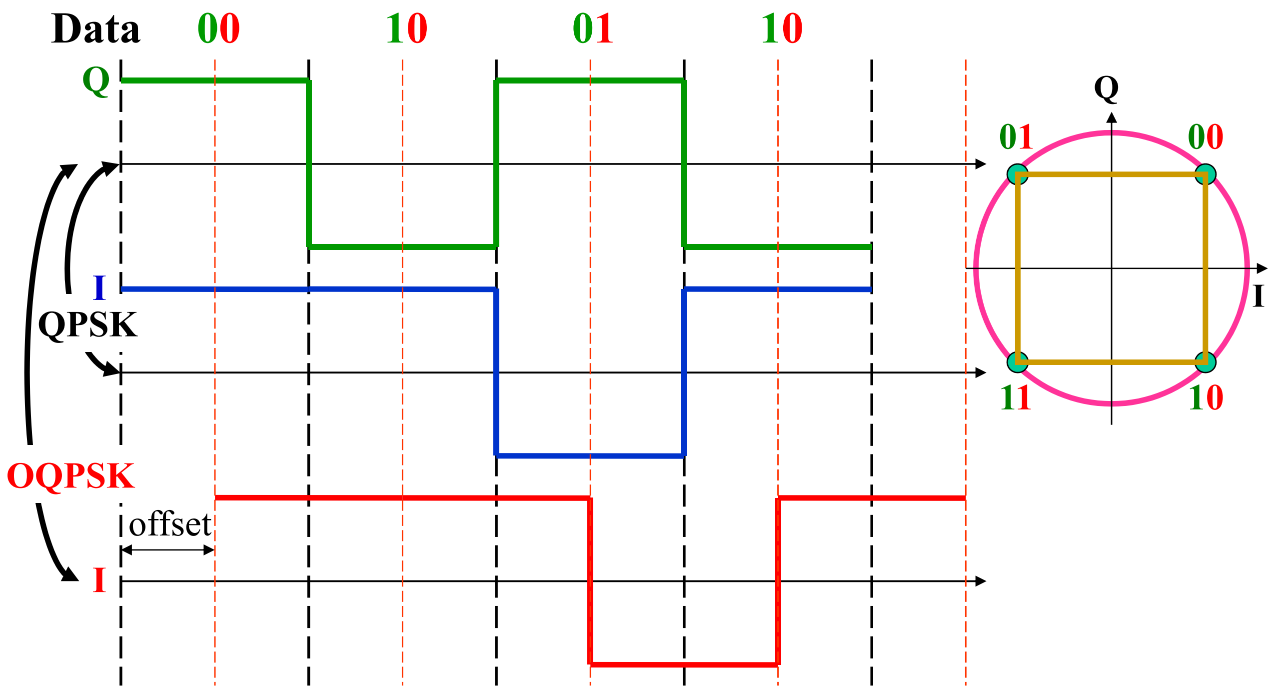


QPSK



OQPSK

Offset QPSK (OQPSK)



π/4-Differential QPSK (π/4-DQPSK)

- QPSK: **4** carrier phases (**Coherent detection is essential**)
- π/4-DQPSK: **8** carrier phases (**Non-coherent detection**)
- For quaternary source sequence $\{x_n\}$, $x_n \in \{\pm 1, \pm 3\}$

$$\Delta\theta_n = \theta_n - \theta_{n-1} = \begin{cases} -3\pi/4, & x_n = -3 \\ -\pi/4, & x_n = -1 \\ +\pi/4, & x_n = +1 \\ +3\pi/4, & x_n = +3 \end{cases}$$

  - The phase differences must be ± π/4 and ± 3π/4

$$\Delta\theta_n = x_n \frac{\pi}{4}$$

- The complex envelope is expressed as

$$\tilde{s}(t) = A \sum_n b(t - nT, \mathbf{x}_n)$$

π/4-Differential QPSK (π/4-DQPSK)

– where

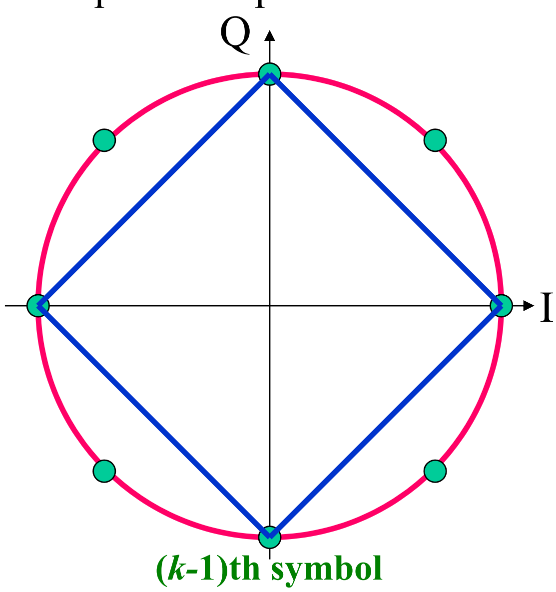
$$b(t,\mathbf{x}_n) = h_a(t)\exp\left\{ j\left(\theta_{n-1} + x_n\frac{\pi}{4} + \theta_0\right)\right\}$$
$$= h_a(t)\exp\left\{ j\frac{\pi}{4}\left(\sum_{k=-\infty}^{n-1} x_k + x_n\right) + j\theta_0\right\}$$

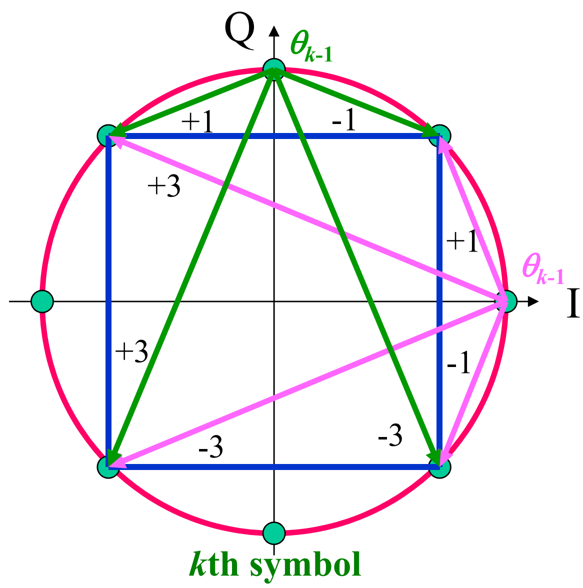
- Change the phase for every symbol: the **symbol synchronization** is much easier than QPSK
- **Non-coherent detection** can be applied in receivers
- There are two signal sets:

$$\{0, \pi/2, \pi, 3\pi/2\}$$
$$\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$$

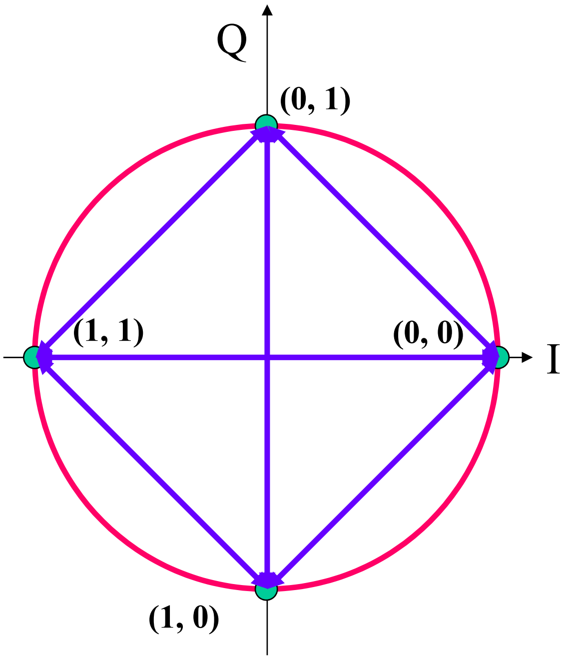
π/4-Differential QPSK (π/4-DQPSK)

- The phase trajectory does not pass through **the origin**
  – Reduce the linearity and dynamic range requirements of the power amplifier

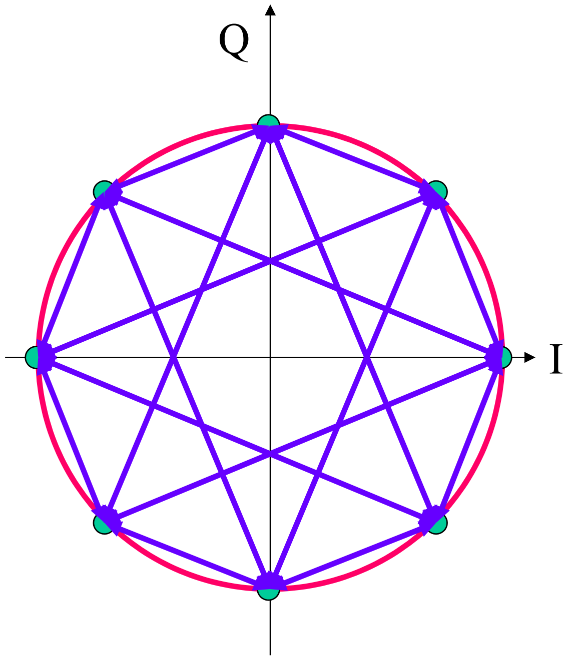




π/4-Differential QPSK (π/4-DQPSK)

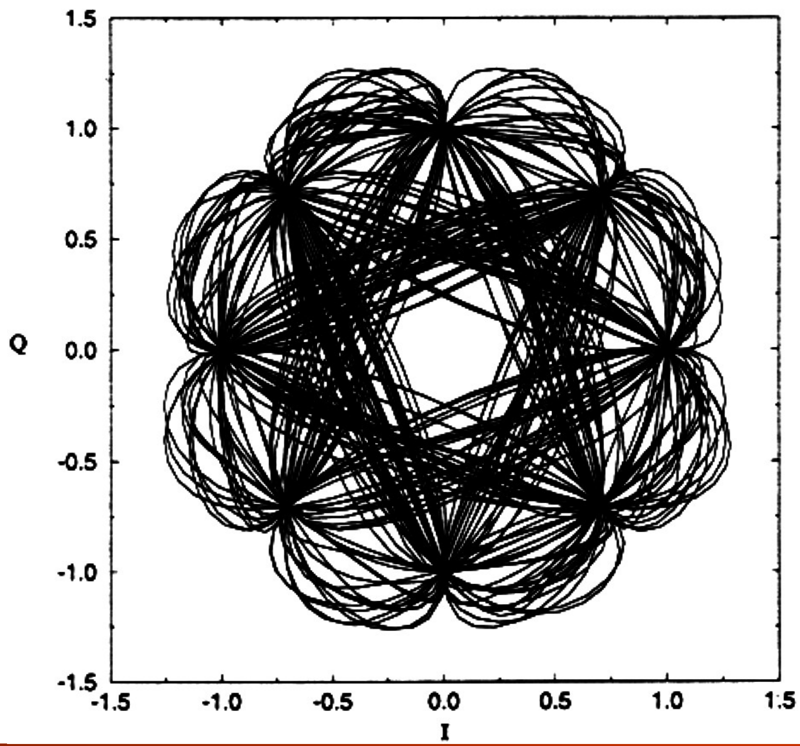


QPSK



π/4-DQPSK

π/4-Differential QPSK (π/4-DQPSK)



Continuous Phase Modulation (CPM)

- Continuous Phase Modulation: the carrier phase varies in a **continuous** manner
- CPM is attractive because it has **constant envelope** and **excellent spectral characteristics**
- The complex envelope is

$$\tilde{s}(t) = A\exp\left\{j\left(\phi(t)+\theta_0\right)\right\}$$

$$\phi(t) = 2\pi\int_0^t \sum_{k=0}^{\infty} h_k x_k h_f(\tau - kT)\, d\tau$$

  - $\phi(t)$: the **excess phase**: the phase difference to the zero phase
  - $\{x_k\}$: the data symbol sequence
  - $\{h_k\}$: the sequence of modulation indices
  - $h_f(t)$: the frequency shaping function

Continuous Phase Modulation (CPM)

- $h_f(t)$ is zero for $t < 0$ and $t > LT$
  - Full response CPM: $L = 1$
  - Partial response CPM: $L > 1$
- The phase shaping function:

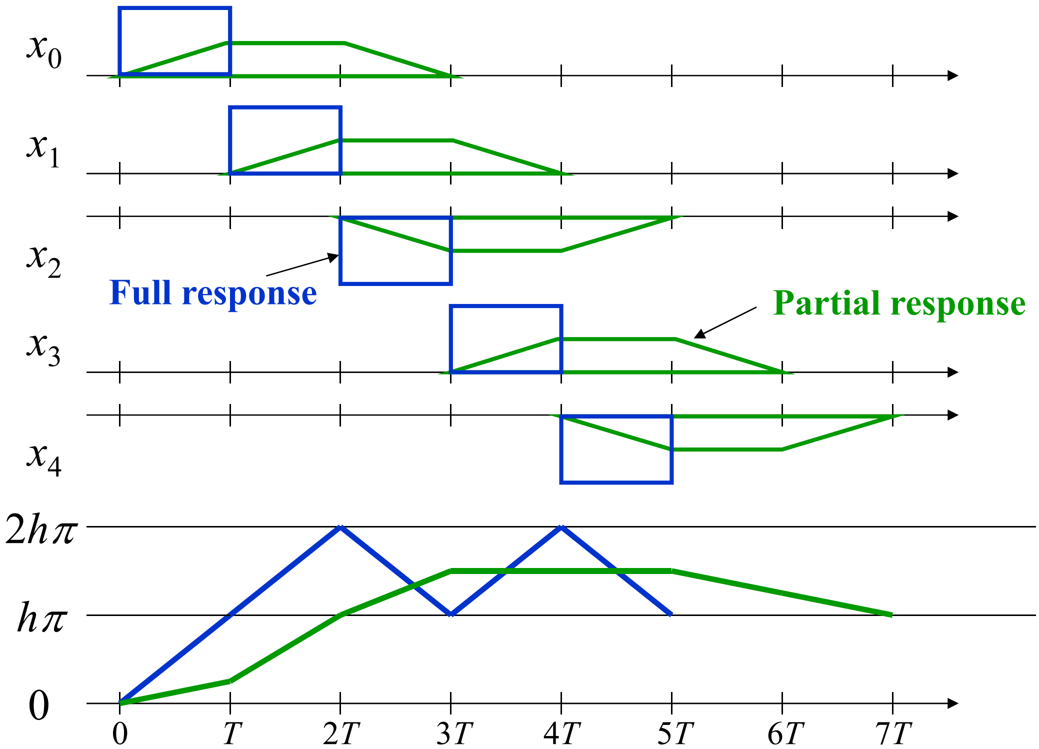
$$\beta(t) = \begin{cases} 0, & t < 0 \\ \int_0^t h_f(\tau)d\tau, & 0 \le t \le LT \\ 1/2, & t \ge LT \end{cases}$$

- For a full response CPM and $h_k = h$, the excess phase is

$$\phi(t) = 2\pi h\left[\underbrace{\int_0^{nT}\sum_{k=0}^{n-1} x_k h_f(\tau - kT)\, d\tau}_{\text{Response of the previous symbols}} + \underbrace{\int_{nT}^{t} x_n h_f(\tau - nT)\, d\tau}_{\text{Response of the current symbol}}\right]$$

$$= \pi h\sum_{k=0}^{n-1} x_k + 2\pi h x_n \beta(t - nT)$$

Continuous Phase Modulation (CPM)



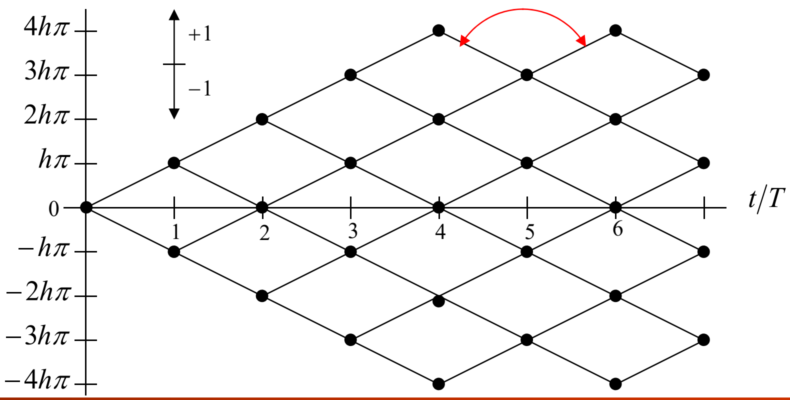
Continuous Phase Frequency Shift Keying

- Continuous phase frequency shift keying (CPFSK):
  - A special type of full response CPM
  - By using the **rectangular** shaping function $h_f(t) = u_T(t)$

$$\beta(t) = \begin{cases} 0, & t < 0 \\ t/2T, & 0 \le t \le T \\ 1/2, & t \ge T \end{cases}$$

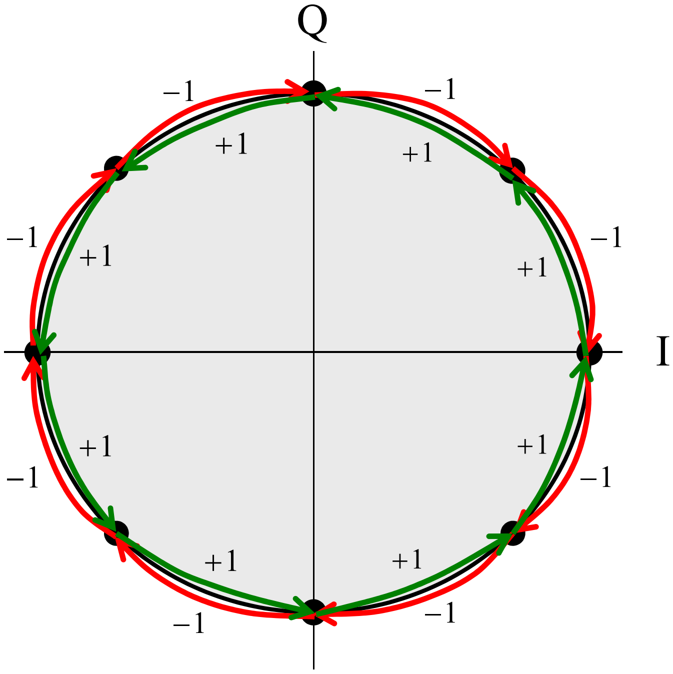
Two slopes
⇒Two frequencies
Representing '+1' & '–1'

- The phase tree:



Continuous Phase Frequency Shift Keying

- Binary CPM signal with $h = 1/4$



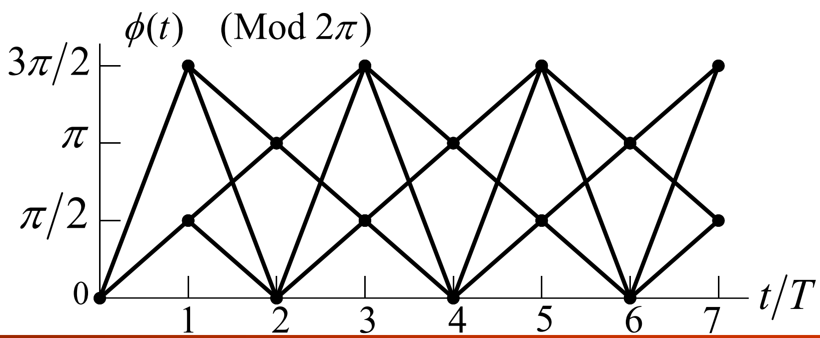
Minimum Shift Keying (MSK)

- Minimum Shift Keying is a special case of binary CPFSK
  - With modulation index $h = 1/2$
- **Carrier phase:** **Zero phase**

$$\phi_c(t) = 2\pi f_c t + \frac{\pi}{2}\sum_{k=0}^{n-1} x_k + \frac{\pi}{2} x_n \frac{t-nT}{T}$$

$$= \left(2\pi f_c + \frac{\pi x_n}{2T}\right)t + \frac{\pi}{2}\sum_{k=0}^{n-1} x_k - \frac{\pi n}{2} x_n$$

- Excess phase:



Minimum Shift Keying (MSK)

- The MSK band-pass waveform is

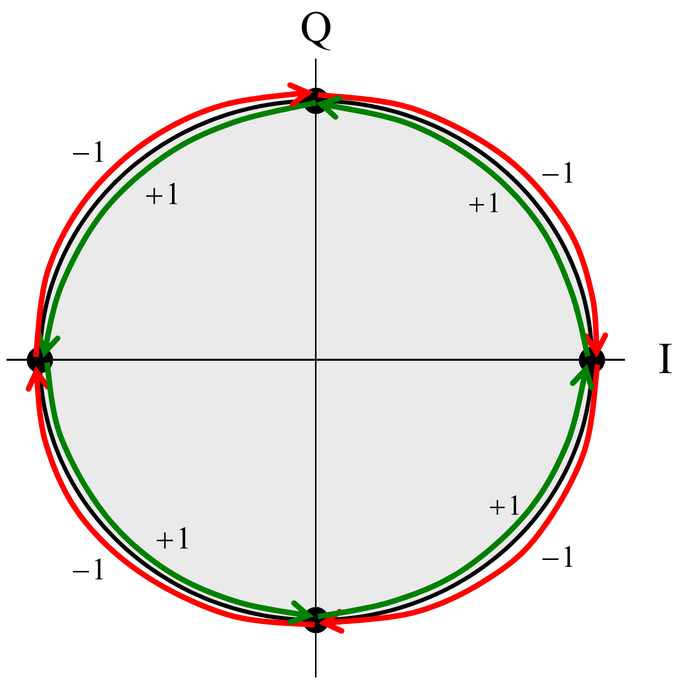
$$s(t) = A\cos\left(2\pi\left(f_c + \frac{x_n}{4T}\right)t + \frac{\pi}{2}\sum_{k=0}^{n-1} x_k - \frac{\pi n}{2} x_n\right), \quad nT \le t \le (n+1)T$$

- Note that an MSK signal has one of two possible frequencies:

$$f_L = f_c - \tfrac{1}{4T}, \quad \text{and } f_U = f_c + \tfrac{1}{4T}$$

- The difference $\Delta f = f_U - f_L = \frac{1}{2T}$ is the minimum frequency separation to ensure **orthogonality** between two sinusoid of duration T with coherent demodulation

Minimum Shift Keying (MSK)

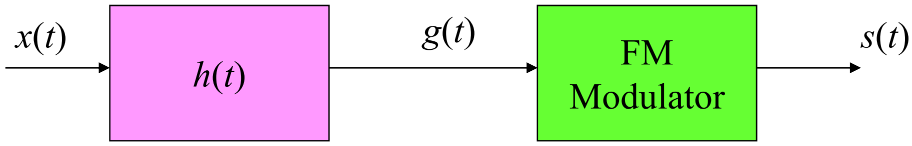


Gaussian Minimum Shift Keying (GMSK)

- **Gaussian Minimum Shift Keying** is achieved by **low-pass filtering** the modulating signal prior to modulation
  - The filtering removes the higher frequency components

  $\Rightarrow$ It results in a more **compact spectrum**

  $\Rightarrow$ It introduces the **inter-symbol interference (ISI)**



- The transfer function of the pre-modulation filter (low-pass) is

$$H(f) = \exp\left\{ -\left(\frac{f}{B}\right)^2 \frac{\ln 2}{2} \right\}$$

  - where B is the 3 dB bandwidth of the filter

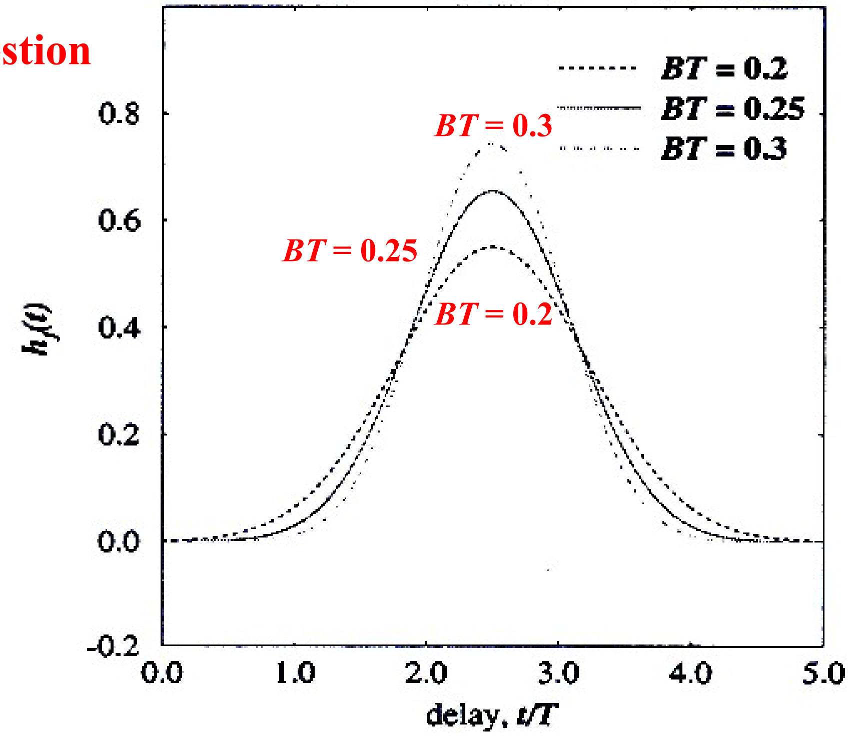
Gaussian Minimum Shift Keying (GMSK)

- The GMSK frequency shaping pulse filter:
  - The normalized filter bandwidth is BT
  - The shaping pulse duration $> T \Rightarrow$ ISI
  - BT decreases $\Rightarrow$ more compact in power spectral density $\Rightarrow$ ISI increases $\Rightarrow$ the bit error rate performance is degraded
- In GSM systems:

$$x_n \in \{-1, +1\}$$
$$h = 1/2$$
$$BT = 0.3$$

Gaussian Minimum Shift Keying (GMSK)

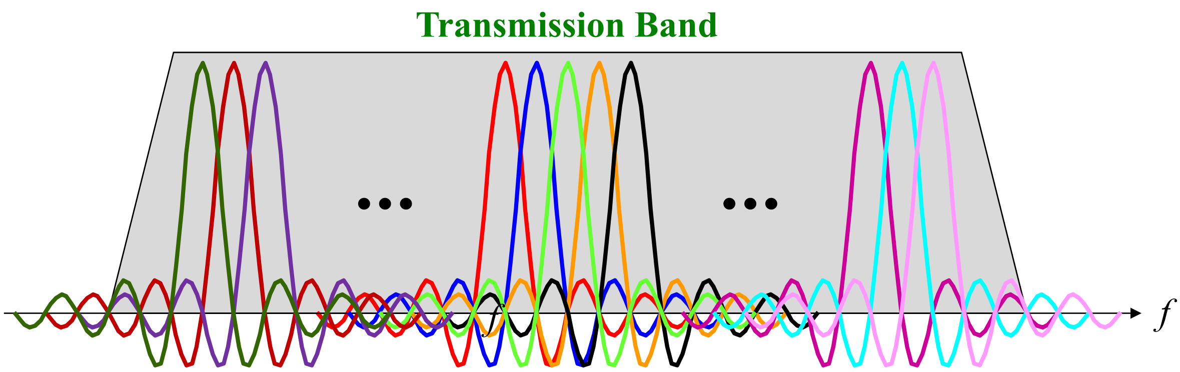
Question



OFDM

Orthogonal Frequency Division Multiplexing

- **Orthogonal frequency division multiplexing (OFDM)** is a promising technique because of its
  - High bandwidth efficiency and
  - Resistance to multipath fading
- OFDM transmission can be regarded as a kind of multi-carrier transmission.

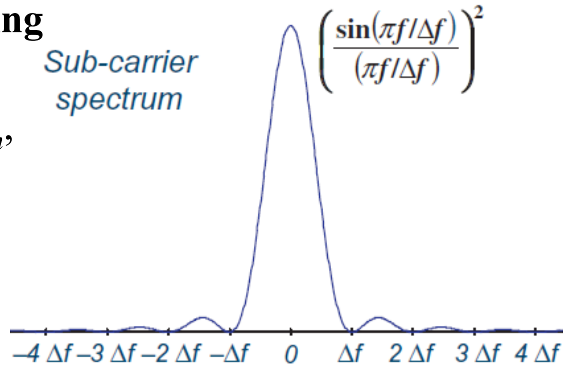


Orthogonal Frequency Division Multiplexing

- However, the following characteristics distinguish OFDM from a straightforward multi-carrier extension:
  - The use of a typically **very large number** of relatively narrowband subcarriers (e.g., several hundred subcarriers)
  - Simple **rectangular pulse shaping** is used

  $\Rightarrow$ A sinc-square-shaped per-subcarrier spectrum
  - **Tight frequency-domain packing** of the subcarriers

  $\Rightarrow$ A subcarrier spacing $\Delta f = 1/T_{sym}$, T_{sym} is the per-subcarrier modulation-symbol duration



Orthogonal Frequency Division Multiplexing

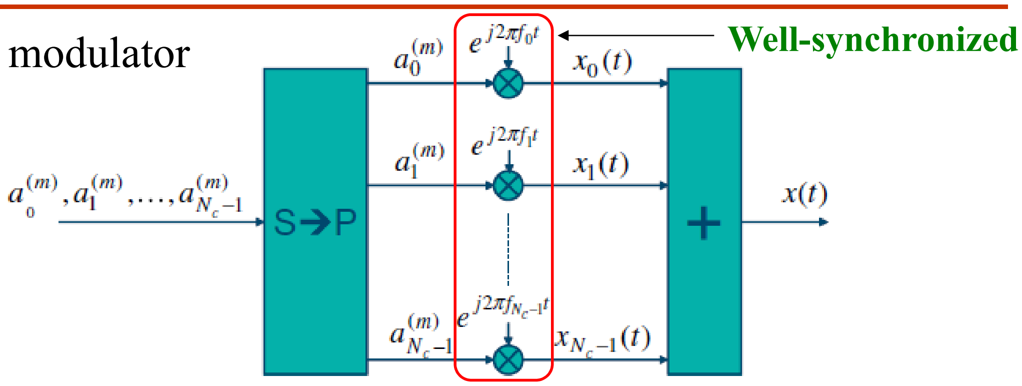
- **OFDM** can be regarded as a **frequency-division multiplexing (FDM)** scheme.
- A large number of closely-spaced **orthogonal subcarriers** are used to carry data.
- In complex baseband notation, a basic OFDM signal $x(t)$ during the time interval $mT_{sym} \le t \le (m+1)T_{sym}$ is expressed as:

$$x(t) = \sum_{k=0}^{N_c-1} x_k(t) = \sum_{k=0}^{N_c-1} a_k^{(m)} e^{j2\pi k \Delta f t}$$

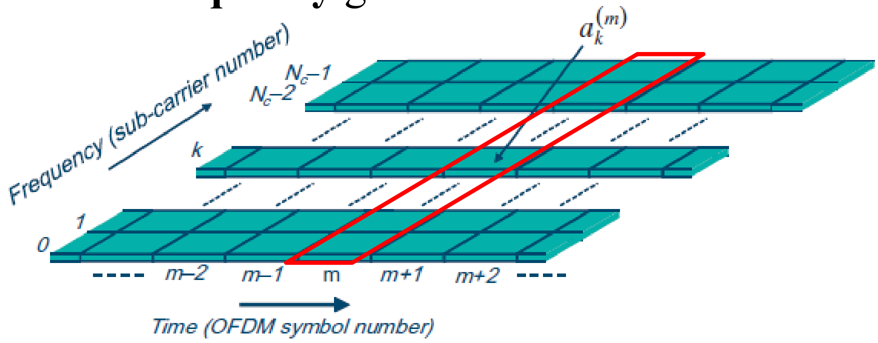
  – $x_k(t)$ is the k-th modulated subcarrier with $f_k = k\Delta f$
  – $a_k^{(m)}$ is the modulation symbol applied to the k-th subcarrier during the m-th OFDM symbol interval

OFDM Modulation

- OFDM modulator

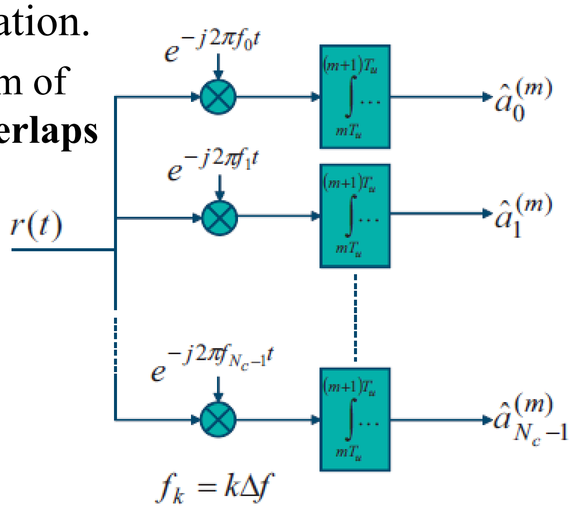


- OFDM **time–frequency** grid



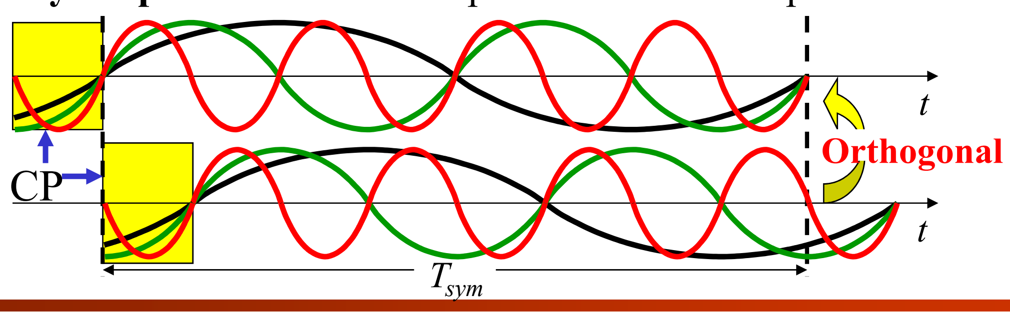
OFDM Demodulation

- The basic principle of OFDM demodulation consists of **a bank of correlators**, each of which corresponds to a subcarrier.
- Because of the orthogonality between subcarriers, in the ideal case, two OFDM subcarriers **do not** cause any interference to each other after demodulation.
  – Despite that the spectrum of neighbor subcarriers **overlaps**
- OFDM demodulator



OFDM Orthogonality

- The **subcarrier orthogonality** is due to
  – The specific **frequency-domain structure** of each subcarrier
  – The specific choice of a **subcarrier spacing** Δf equal to the **per-subcarrier modulation-symbol rate** $1/T_{sym}$
- Any **corruption** of the frequency-domain structure may lead to destruction in orthogonality ⇒ causing **mutual interference**
  – Time dispersion, **Doppler spread**, …
  – **Cyclic prefix** is inserted to prevent ISI in multipath channels



OFDM Implementation

- OFDM allows for low-complexity implementation by means of computationally efficient **Fast Fourier Transform** (FFT) processing.
  - The **sampling rate** f_s is a multiple of the subcarrier spacing Δf

$$f_s = 1/T_s = N \times \Delta f$$

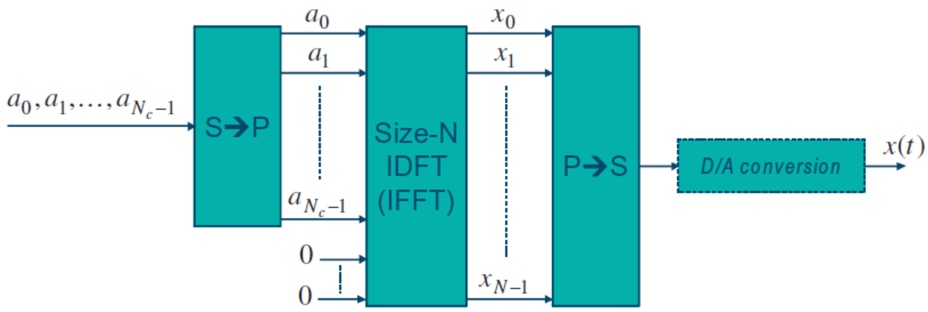
  - The FFT size N should exceed N_c with a sufficient margin, $N > N_c$
- The time-discrete OFDM signal can be expressed as
  - The sequence x_n is the Inverse Discrete Fourier Transform (IDFT) of the symbols $a_0, a_1, \ldots, a_{Nc-1}$, extended with zeros to length N

$$x_n = x(nT_s) = \sum_{k=0}^{N_c-1} a_k e^{j2\pi k\Delta f nT_s} = \sum_{k=0}^{N_c-1} a_k e^{j2\pi kn/N} = \sum_{k=0}^{N_c-1} a'_k e^{j2\pi kn/N}$$

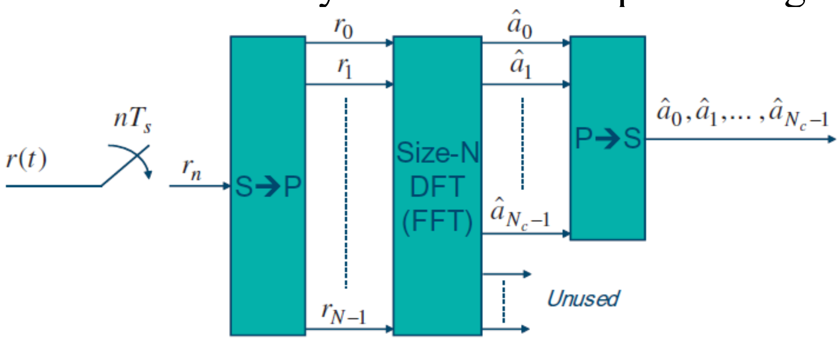
$$a'_k = \begin{cases} a_k & 0 \le k < N_c \\ 0 & N_c \le k < N \end{cases}$$

OFDM Implementation

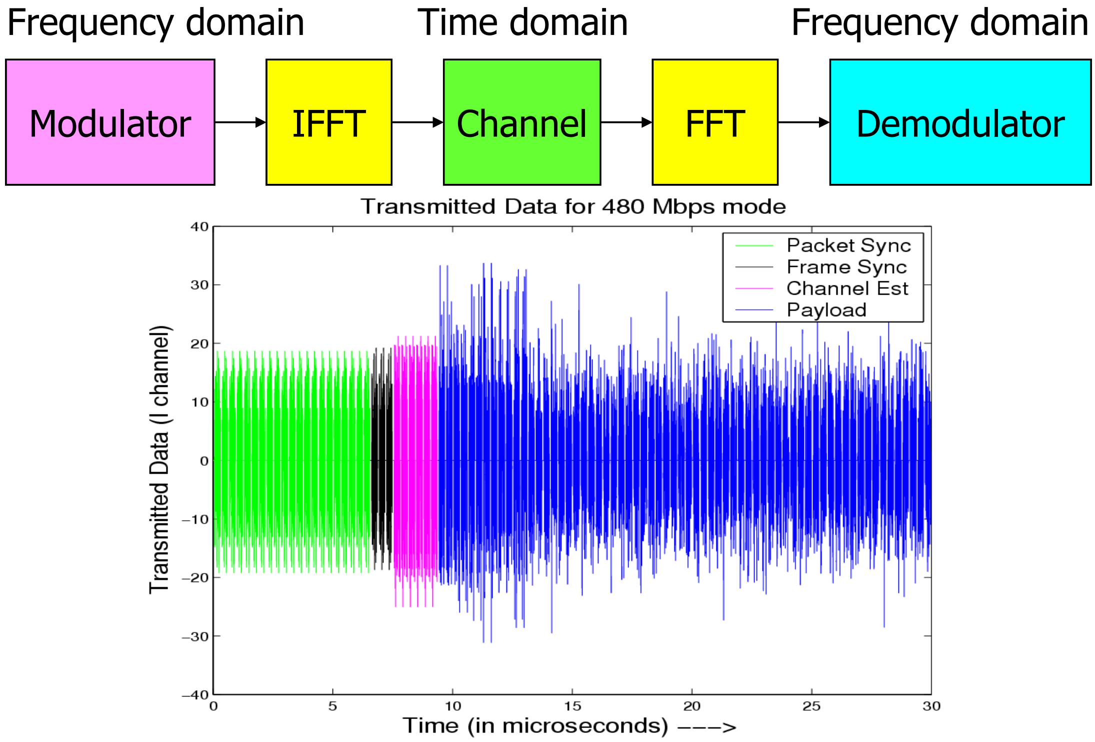
- OFDM **modulation** by means of **IFFT** processing



- OFDM **demodulation** by means of **FFT** processing



OFDM Transmission



OFDM Transmission

Question

- The data are divided into several **parallel data streams** or channels, one for each subcarrier.
- Each subcarrier applies a **conventional modulation scheme** (PSK, QAM, …) at a low symbol rate.
- OFDM may be regarded as using many **slowly-modulated narrowband signals** rather than one rapidly-modulated wideband signal.
  - The channel equalization is simplified
- OFDM requires very accurate **frequency synchronization** between the receiver and the transmitter
  - Including the **carrier** frequency and **sampling** frequency
  - Otherwise **inter-carrier interference (ICI)** is introduced

Digital Signaling on Flat Fading Channels

BER-BPSK

- The probability of bit error can be expressed as

$$P_b(\gamma_b) = Q(\sqrt{2\gamma_b})$$
$$\gamma_b = \alpha^2 E_b / N_0$$

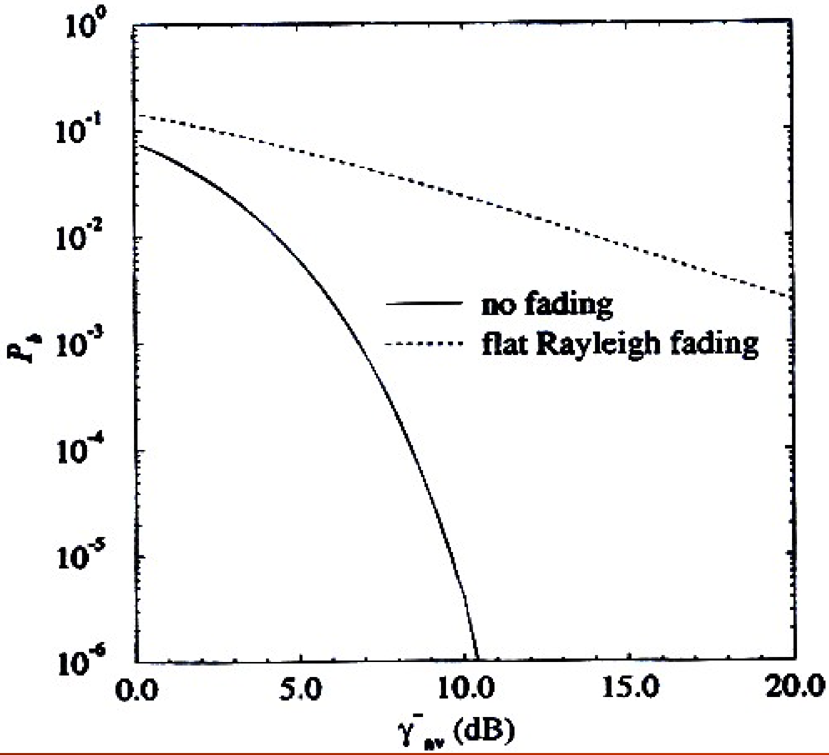
- The pdf of energy-to-noise ratio for flat Rayleigh fading channels is

$$p_{\gamma_b}(x) = (1/\bar{\gamma}_b)e^{-x/\bar{\gamma}_b}$$

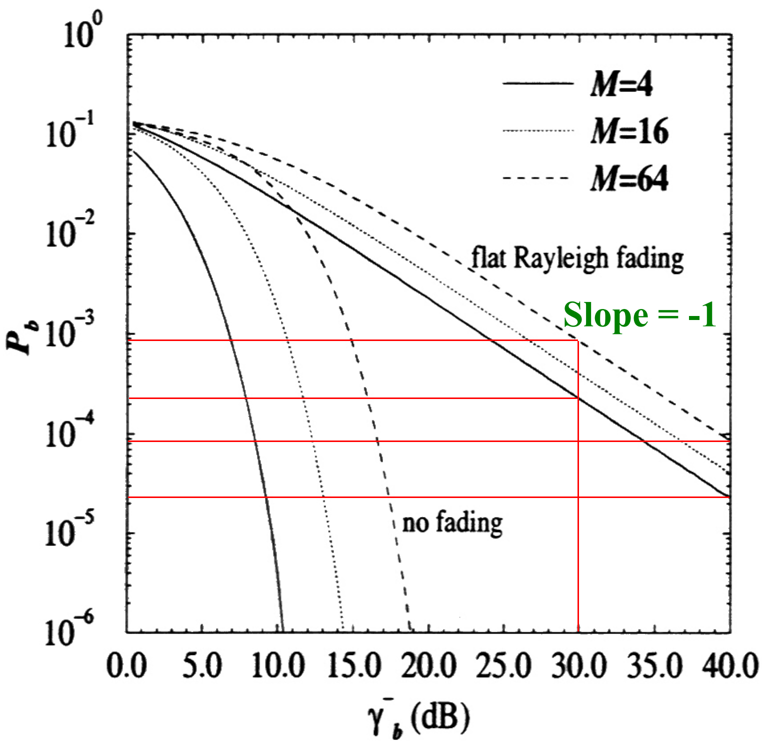
- The average probability of bit error becomes

$$\begin{aligned} P_b &= \int_0^\infty P_b(x) p_{\gamma_b}(x)\, dx \\ &= \int_0^\infty Q(\sqrt{2x}) p_{\gamma_b}(x)\, dx \\ &= \frac{1}{2}\left[1 - \sqrt{\frac{\bar{\gamma}_b}{1+\bar{\gamma}_b}}\right] \approx \frac{1}{4\bar{\gamma}_b} \quad \text{for } \bar{\gamma}_b >> 1 \end{aligned}$$

BER-BPSK and QPSK



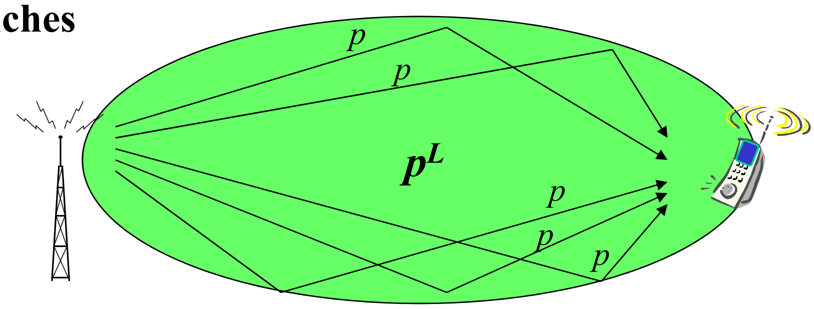
BER-M-QAM



Diversity Techniques

Diversity Techniques

- **Diversity** is a very effective solution for combating fading:
  - It can provide the receiver with multiple faded replicas of the same information signal
  - Let p denote the probability that the instantaneous signal-to-noise ratio is **below** some critical threshold **on a particular diversity branch**
  - p^L is the probability that all the instantaneous signal-to-noise ratios are **below** the same critical threshold **on all L diversity branches**

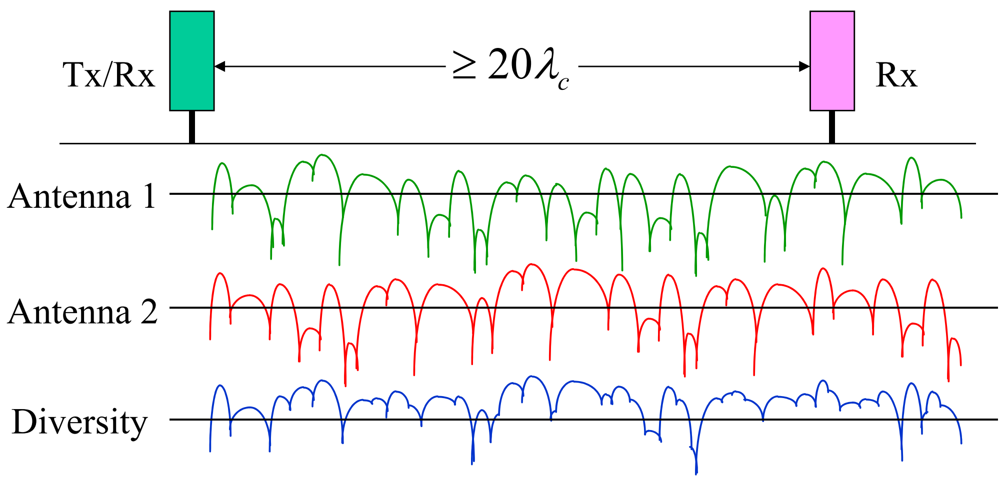


Diversity Techniques

- Diversity techniques:
  - Space
  - Angle
  - Polarization
  - Field
  - Frequency
  - Multipath
  - Time

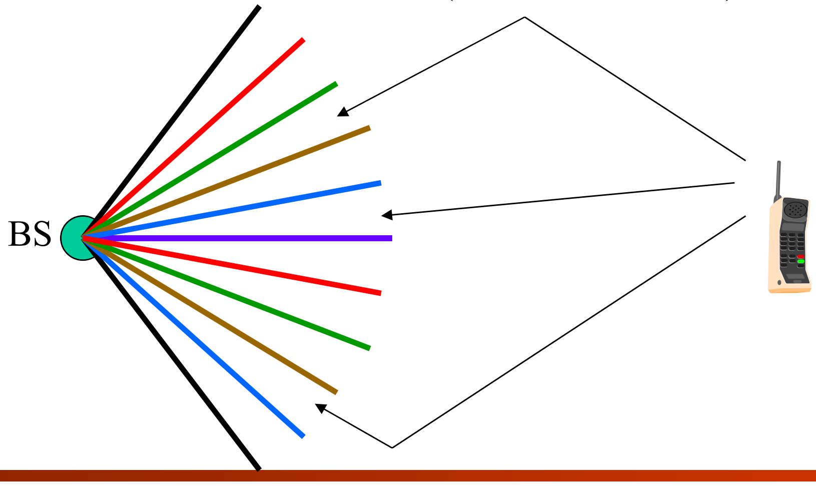
Space Diversity

- Space diversity (generally applied at BSs):
  - Achieved by using **multiple receive antennas**
  - The **spatial separation** between the antennas is chosen so that the diversity branches experience **uncorrelated fading**



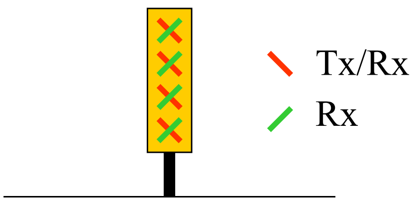
Angle Diversity

- Angle diversity:
  - A number of **directional antennas** are required
  - Each antenna selects plane waves arriving from a narrow range of angles
  - All branches are uncorrelated (WSSUS channel)

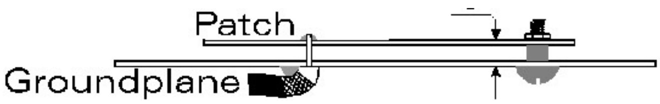


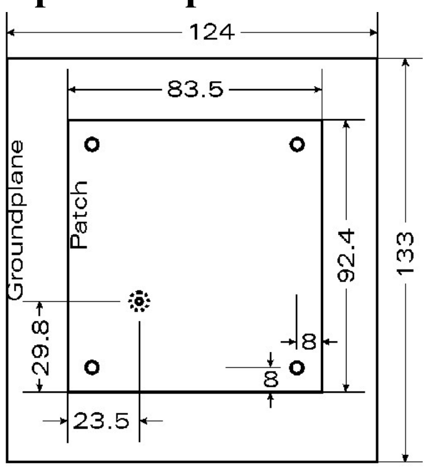
Polarization Diversity

- Polarization diversity: **Question**
  - The scattering environment tends to **depolarize** a signal
  - Receive antennas having different polarizations can be used to obtain diversity
  - At subscriber units: combination of **monopole** and **patch** antennas
  - At base stations: **slant 45° monopole**



Polarization Antenna





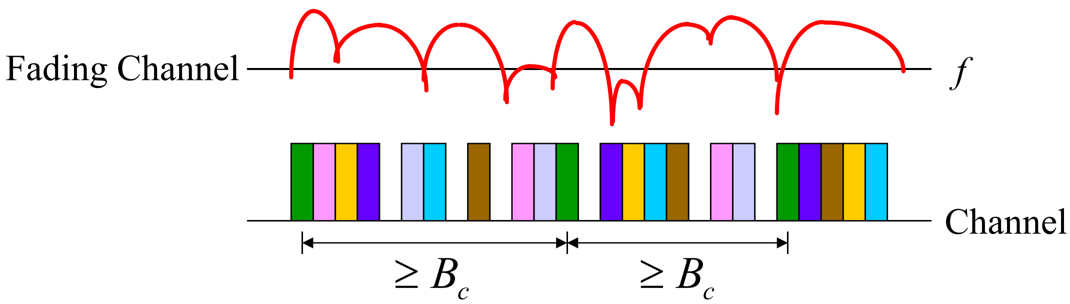
Field Diversity

- Field diversity:
  - The **electric** field and **magnetic** field components at any point are uncorrelated

Frequency Diversity

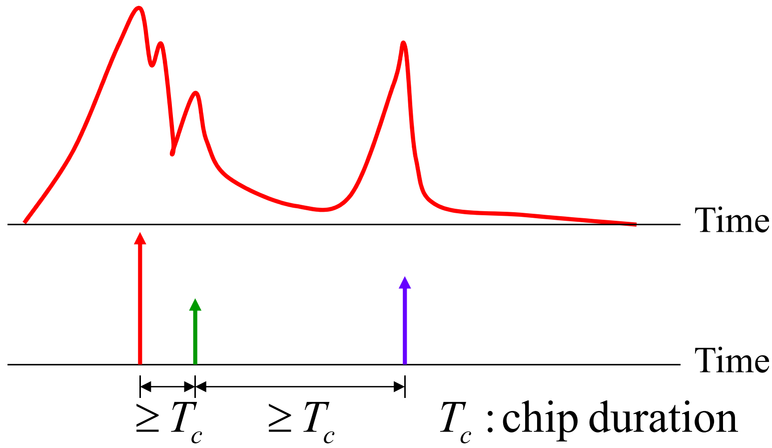
- Frequency diversity:
  - Using multiple frequency channels
  - Each channel is separated by at least the **coherence bandwidth** of the channel
  - **Frequency hopping spread spectrum (FHSS)** systems provide frequency diversity through the fast frequency hopping

    ⇒ Each symbol is transmitted on multiple hops (carriers)



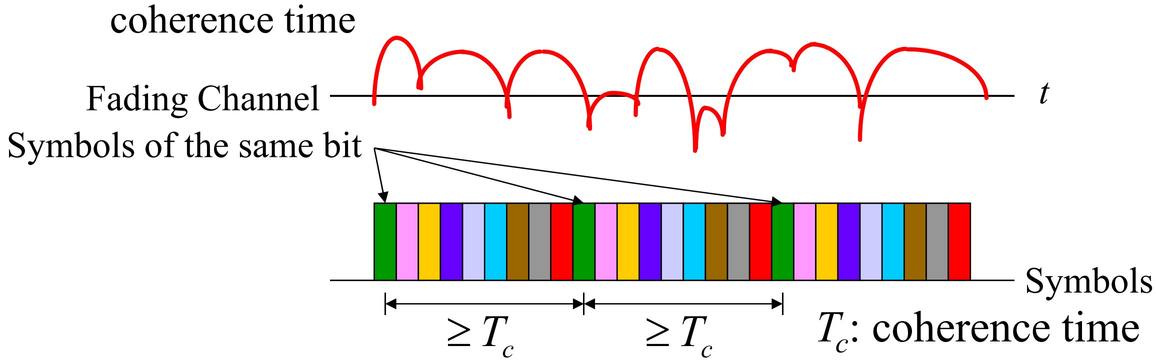
Multipath Diversity

- Multipath diversity:
  - Resolving multipath components at different delays
  - Time resolution must be high enough ⇒ **Wideband signals**
  - Using **direct sequence spread spectrum (DSSS)** signaling along with a **RAKE receiver** can achieve multipath diversity



Time Diversity

- Time diversity:
  - Using multiple time slots that are separated by at least the **coherence time** of the channel
  - **Error correction coding** with **interleaving** is an efficient method that provides time diversity
  - The coherence time of the channel depends on the **velocity** of the MS, and slow moving MSs have the channel with a large coherence time



Forward Error Correction Coding and Interleaving

Forward Error Correction Coding

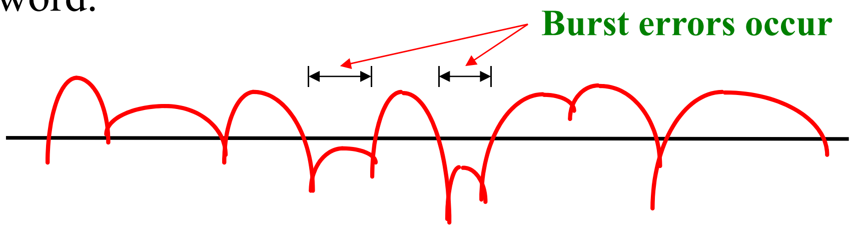
- **Forward error correction (FEC)** coding (or known as channel coding) is a technique used for controlling errors in data transmission
  - Over **noisy** and/or **unreliable (fading)** communication channels
- By introducing **redundancy** into an encoded message, the receiver can detect a limited number of errors that may occur anywhere in the message.
  - The **bandwidth efficiency** (throughput) is reduced, depending on the **code rate** ($0 < r \leq 1$)
  - For a large (small) code rate, the bandwidth efficiency is high (low), but the error-correction capability is low (high)
- FEC is also applied to **mass storage devices** to enable recovery of corrupted data and to improve the device reliability.

Forward Error Correction Coding

- There are two major categories of FEC codes:
  - **Block codes:** work on a fixed, predetermined number of bits or symbols (a fixed block size), such as Hamming codes, Reed-Solomon codes, BCH codes, LDPC codes, …
    - The error-correction capability depends on both the **code rate** and the **block size**.
  - **Convolutional codes:** work on a bit or symbol stream of **arbitrary length**
    - The most often used decoding approach is soft decoding with the **Viterbi algorithm**.
    - The error-correction capability depends on both the **code rate** and the **constraint length** (the memory size in the encoder).
- FEC codes generally expects errors to be **uniformly** distributed.

Interleaving

- **Interleaving** is frequently used in digital communications and storage systems to improve the performance of FEC codes.
- Many communication channels suffer **burst errors**, not random (independent) errors.
  - For an **AWGN channel**, the error pattern is random errors
  - For a **fading channel**, the error pattern is burst errors
- If the number of errors within a code word exceeds the error-correction capability, the decoder fails to recover the original code word.



Interleaving

- Interleaving mitigates the impact of burst errors by shuffling contiguous coded symbols across time, frequency, or other domains
  - To make the burst errors become uniformly distributed errors.
- Considering a sequence of coded symbols with a block size of 8 symbols, each row of symbols corresponds to a codeword
  - Without interleaving, the transmission order is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, …

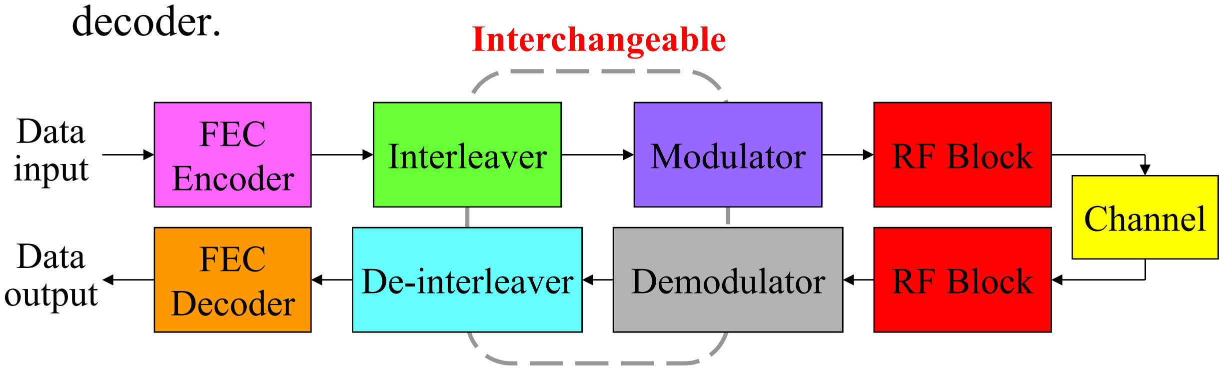
Without interleaving →

With interleaving ↓

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
449	450	451	452	453	454	455	456

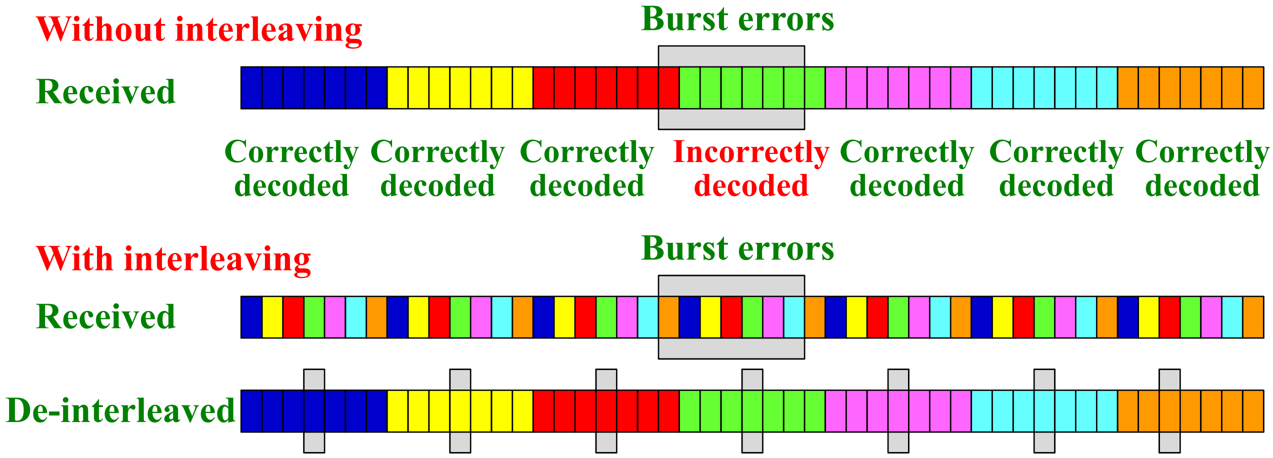
Interleaving

- With interleaving, the transmission order becomes 1, 9, 17, …, 449, 2, 10, 18, …
  - The adjacent symbols belong to different codewords.
  - The symbols of the same codeword are interleaved across time.
- At the transmitter, the interleaver is set after the FEC encoder.
- At the receiver, the de-interleaver is set in front of the FEC decoder.



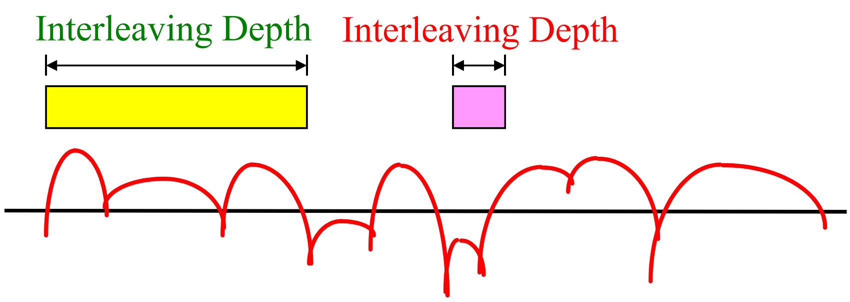
Interleaving

- When passing through the propagation channel, burst errors still occur in the received bit/symbol stream
  - Without interleaving, the receiver fails to decode some blocks
  - With interleaving, the burst errors becomes almost uniformly distributed errors and all blocks can be decoded correctly.



Interleaving

- The design of the **interleaving depth** relies on the **average envelope fade duration**.
- If the interleaving depth is **smaller** than the average envelope fade duration, all the symbols of a codeword may experience **the same fade duration** $\Rightarrow$ burst errors occur in a codeword
- The interleaving depth must be larger than the average envelope fade duration.



Diversity Combining

Diversity Combining

- Diversity combining refers to the method by which the signals from the diversity branches are combined
  - **Predetection** combining: diversity combining takes place **before** detection (takes place at RF)
  - **Postdetection** combining: diversity combining takes place **after** detection (takes place at baseband)
- For ideal coherent detection:
  - There is no difference in performance between predetection and postdetection combining

Diversity Combining

- Diversity combining methods:
  - **Selective Combining (SC)**
  - **Maximal Ratio Combining (MRC)**
  - **Equal Gain Combining (EGC)**
  - **Switched Combining (SW)**
- The received complex envelopes of the diversity branches:

$$\tilde{r}_k(t) = g_k \tilde{s}(t) + \tilde{n}_k(t), \quad k = 1, \cdots, L$$

  - $g_k = \alpha_k e^{j\phi_k}$ is the complex fading gain associated with the k^{th} branch
  - $\tilde{n}_k(t)$ is assumed to be independent from branch to branch
- The received signal vectors are

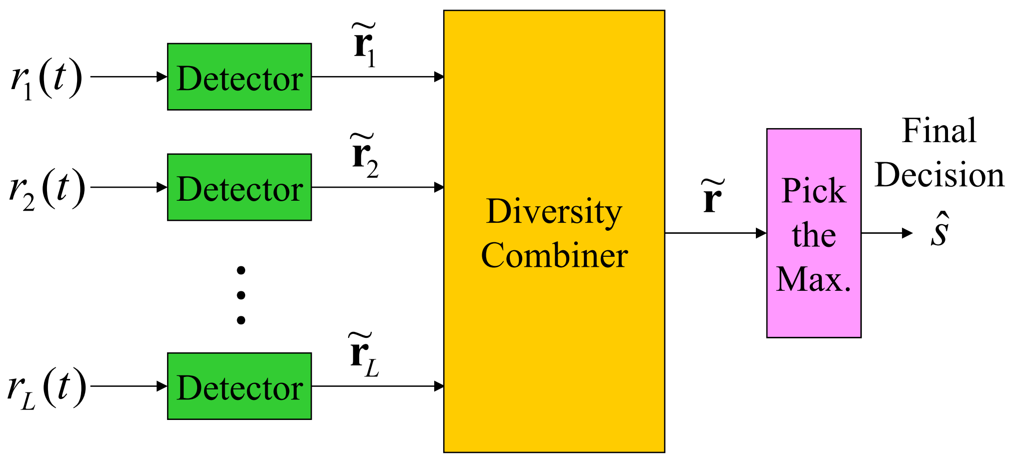
$$\tilde{\mathbf{r}}_k = g_k \tilde{\mathbf{s}} + \tilde{\mathbf{n}}_k, \quad k = 1, \cdots, L$$

Diversity Combining

- where $\tilde{\mathbf{r}}_k = \left[\tilde{r}_{k_1}, \tilde{r}_{k_2}, \cdots, \tilde{r}_{k_M}\right]$ is the received signal vector corresponding to **the *M* possible symbols**

$$\tilde{r}_{k_m} = g_k \tilde{s}_m + \tilde{n}_{k_m}, \quad m = 1, \cdots, M$$

Postdetection combining



Fading Gains

- The fading gains of the various diversity branches typically have **some degree of correlation**, which depends on
  - The type of diversity being used
  - The propagation environment
- For analytical purposes, the diversity branches are generally **assumed to be uncorrelated**
  - Branch correlation reduces the achievable diversity gain
  - The uncorrelated branch assumption gives **optimistic results**

Fade Distribution

- The **fade distribution** will affect the diversity gain
  - The relative advantage of diversity is **greater for Rayleigh fading** than Ricean fading
  - Ricean factor K increases $\Rightarrow$ less difference between the instantaneous received SNR on the various diversity branches
  - The performance will always **be better with Ricean than Rayleigh fading**

Predetection Selective Combining (SC)

- The branch with the highest signal-to-noise ratio is selected at any instant
- With Rayleigh fading, the instantaneous received bit energy-to-noise ratio on *k*-th diversity branch has the exponential pdf

$$p_{\gamma_k}(x) = \frac{1}{\bar{\gamma}_c} e^{-x/\bar{\gamma}_c}$$

  - $\bar{\gamma}_c$ is the average received bit energy-to-noise ratio for each diversity branch
- With SC, the branch with the **largest** bit energy-to-noise ratio is always selected

Predetection Selective Combining (SC)

- The effective instantaneous bit energy-to-noise ratio is

$$\gamma_s^S = \max\{\gamma_1, \gamma_2, \cdots, \gamma_L\}$$

  - L is the number of diversity branches
- If the diversity branches are independently faded, the cumulative distribution function (cdf) of γ_s^S is

$$F_{\gamma_k}(x) = \int_0^x \tfrac{1}{\bar{\gamma}_c} \exp(-\tfrac{y}{\bar{\gamma}_c})\, dy = -\exp(-\tfrac{y}{\bar{\gamma}_c}) \Big|_0^x = 1 - e^{-x/\bar{\gamma}_c}$$

$$F_{\gamma_s^S}(x) = P_r\left[\gamma_1 \le x, \gamma_2 \le x, \cdots, \gamma_L \le x\right] = \left[1 - e^{-x/\bar{\gamma}_c}\right]^L$$

- Differentiating the cdf gives the pdf

$$p_{\gamma_s^S}(x) = \frac{L}{\bar{\gamma}_c}\left[1 - e^{-x/\bar{\gamma}_c}\right]^{L-1} e^{-x/\bar{\gamma}_c}$$

Predetection Selective Combining (SC)

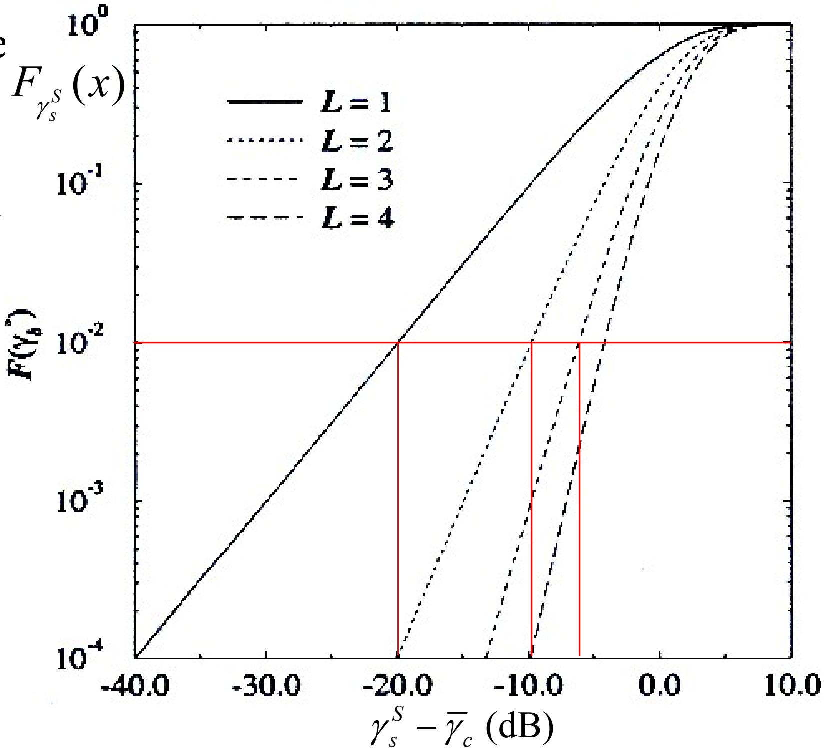
- The average bit energy-to-noise ratio with SC is

$$\bar{\gamma}_s^S = \int_0^\infty x\, p_{\gamma_s^S}(x)\, dx$$
$$= \int_0^\infty \frac{L\,x}{\bar{\gamma}_c}\left[1 - e^{-x/\bar{\gamma}_c}\right]^{L-1} e^{-x/\bar{\gamma}_c}\, dx$$
$$= \bar{\gamma}_c \sum_{k=1}^{L} \frac{1}{k}$$

- The largest diversity gain is obtained by using **2-branch diversity**:
  - $L = 1 \rightarrow 2 \Rightarrow 1.5/1 \Rightarrow 1.761$ dB
  - $L = 2 \rightarrow 3 \Rightarrow 1.83/1.5 \Rightarrow 0.8715$ dB

Predetection Selective Combining (SC)

- Assume the desire SNR is γ and the acceptable outage probability is 0.01
- For $L = 1$
  $\gamma - \bar{\gamma}_c = -20$ (dB)
  $\bar{\gamma}_c = \gamma + 20$ (dB)
- For $L = 2$
  $\gamma - \bar{\gamma}_c = -10$ (dB)
  $\bar{\gamma}_c = \gamma + 10$ (dB)
- For $L = 3$
  $\gamma - \bar{\gamma}_c = -6$ (dB)
  $\bar{\gamma}_c = \gamma + 6$ (dB)



Predetection Selective Combining (SC)

- The DPSK bit error probability with differential detection is:

$$P_b(\gamma_s) = \tfrac{1}{2} e^{-\gamma_s}$$

- With SC, the bit error probability for DPSK is:

Binomial expansion

$$(1-x)^{L-1} = \sum_{n=0}^{L-1} C_n^{L-1} (-1)^n x^n$$

$$\begin{aligned}
P_b &= \int_0^\infty P_b(x) p_{\gamma_s^S}(x)\,dx \\
&= \int_0^\infty \tfrac{1}{2}\exp(-x) \tfrac{L}{\bar{\gamma}_c}[1-\exp(-\tfrac{x}{\bar{\gamma}_c})]^{L-1} \exp(-\tfrac{x}{\bar{\gamma}_c})\,dx \\
&= \int_0^\infty \tfrac{L}{2\bar{\gamma}_c} \exp[-(1+\tfrac{1}{\bar{\gamma}_c})x][1-\exp(-\tfrac{x}{\bar{\gamma}_c})]^{L-1}\,dx \\
&= \frac{L}{2\bar{\gamma}_c} \sum_{n=0}^{L-1} C_n^{L-1} (-1)^n \int_0^\infty \exp[-(1+\tfrac{1}{\bar{\gamma}_c}+\tfrac{n}{\bar{\gamma}_c})x]\,dx \\
&= \frac{L}{2\bar{\gamma}_c} \sum_{n=0}^{L-1} C_n^{L-1} (-1)^n \frac{\bar{\gamma}_c}{\bar{\gamma}_c+n+1} = \frac{L}{2} \sum_{n=0}^{L-1} \frac{C_n^{L-1}(-1)^n}{\bar{\gamma}_c+n+1}
\end{aligned}$$

Predetection Selective Combining (SC)

- Diversity offers a very large improvement in the BER performance
- The bit error probability is now proportional to $\bar{\gamma}_c^{-L}$, not proportional to $\bar{\gamma}_c^{-1}$
- The largest diversity gain is achieved with 2-branch diversity

Predetection Selective Combining (SC)

- BER of DPSK

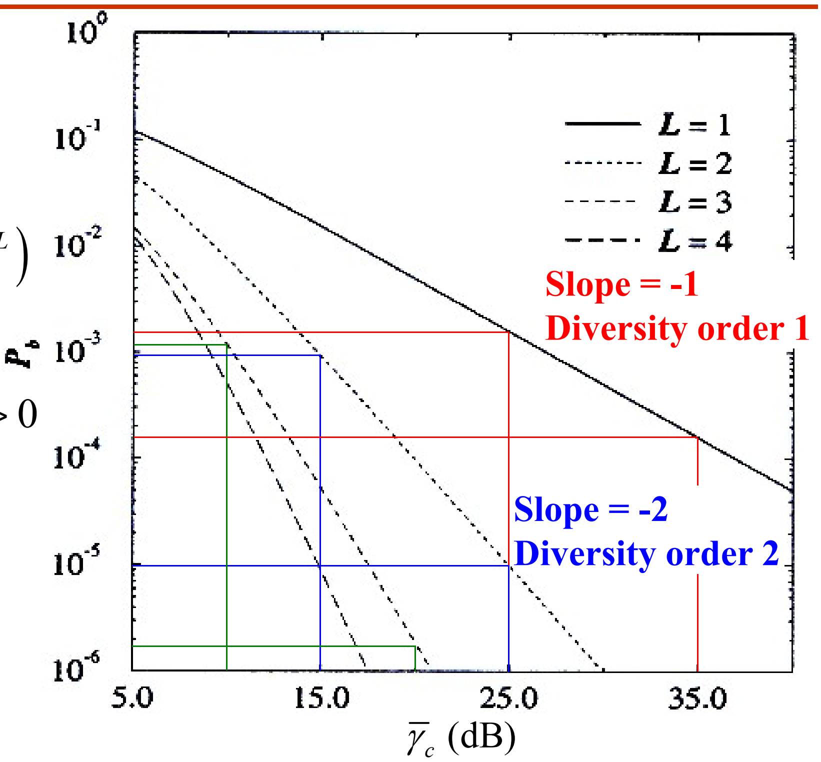
Assume $P_b = \beta \overline{\gamma}_c^{-L}$

$$\log_{10}(P_b) = \log_{10}\left(\beta \overline{\gamma}_c^{-L}\right)$$

$$= \log_{10}\beta - L\log_{10}(\overline{\gamma}_c)$$

$$\frac{\log_{10}(P_b)}{\log_{10}(\overline{\gamma}_c)} \approx -L, \quad \overline{\gamma}_c \gg 0$$

Diversity order



Maximal Ratio Combining (MRC)

- For Maximal Ratio Combining (MRC), the diversity branches must be **weighted** by their respective **complex fading gains** and then combined
- MRC results in a **maximum likelihood** (ML) receiver giving the **best possible performance** among the diversity combining techniques
- The ML receiver must have **complete knowledge** of the channel gain vector $\{\boldsymbol{g}\}$
- Postdetection combining:
  - The weighting and combining is performed after integration
- Predetection combining:
  - The weighting and combining is performed before integration

Maximal Ratio Combining (MRC)

- The received signal vector:

$$\tilde{\mathbf{r}} \triangleq (\tilde{r}_1, \tilde{r}_2, \cdots, \tilde{r}_L); \quad \tilde{r}_k(t) \triangleq g_k \tilde{s}(t) + \tilde{n}_k(t)$$

- The ML receiver chooses the symbol $\tilde{s}_m$ that maximizes the metric

ML criterion

$$\mu_1(\tilde{s}_m) = -\sum_{k=1}^{L} \left|\tilde{r}_k - g_k \tilde{s}_m\right|^2 = -\left(\sum_{k=1}^{L} \left|\tilde{r}_k\right|^2 \underbrace{-2\,\mathrm{Re}\left\{\left(\tilde{r}_k, g_k \tilde{s}_m\right)\right\} + \left|g_k \tilde{s}_m\right|^2}_{2\mu_2(\tilde{s}_m)}\right)$$

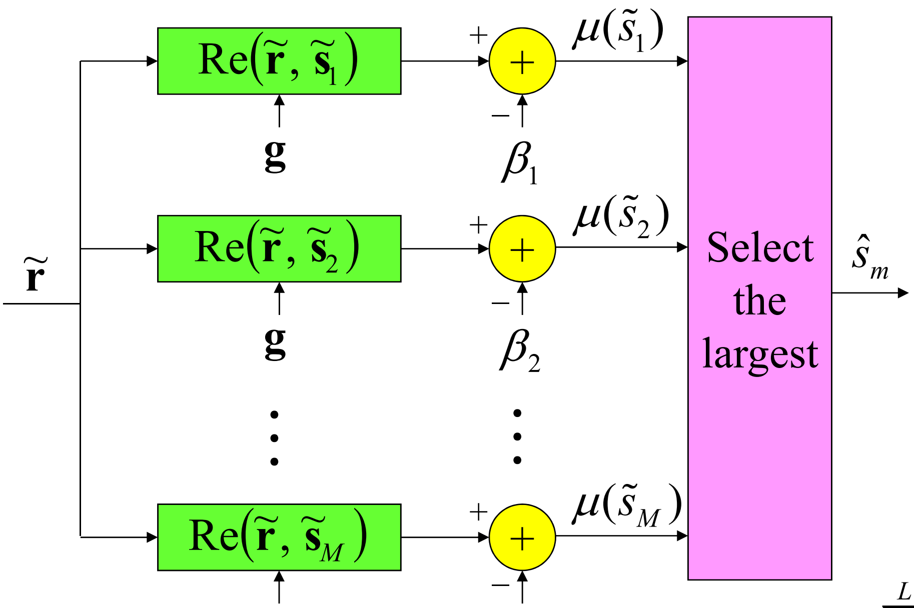
$$\mu_2(\tilde{s}_m) = \sum_{k=1}^{L} \mathrm{Re}\left\{\int_0^T \tilde{r}_k(t) \times g_k^* \tilde{s}_m^*(t)\, dt\right\} - \sum_{k=1}^{L} \left|g_k\right|^2 E_m$$

$$= \sum_{k=1}^{L} \mathrm{Re}\left\{g_k^* \int_0^T \tilde{r}_k(t) \tilde{s}_m^*(t)\, dt\right\} - \underline{E_m \sum_{k=1}^{L} \left|g_k\right|^2}$$

May be different for different S_m

The inner product: $\left(u, v\right) = \int u(t) \times v^*(t)\, dt \qquad \left\|\tilde{s}_m\right\|^2 = 2E_m$

Metric Computer for MRC



$$\beta_m = E_m \sum_{k=1}^{L} \left|g_k\right|^2$$

Maximal Ratio Combining (MRC)

- If all symbols have equal energy, the metric becomes

$$\mu_3(\tilde{s}_m) = \sum_{k=1}^{L} \mathrm{Re}\left\{ g_k^* \int_0^T \tilde{r}_k(t)\tilde{s}_m^*(t)\, dt \right\}$$

- An alternative form of the ML receiver can also be obtained by rewriting the metric

$$\mu_4(\tilde{s}_m) = \int_0^T \mathrm{Re}\left\{ \sum_{k=1}^{L} g_k^* \tilde{r}_k(t) \times \tilde{s}_m^*(t) \right\} dt - E_m \sum_{k=1}^{L} |g_k|^2$$

- Equivalently, the signal used for demodulation is

Weighting factor

$$\tilde{r}(t) = \sum_{k=1}^{L} g_k^* \tilde{r}_k(t) = \sum_{k=1}^{L} g_k^* g_k \tilde{s}(t) + \sum_{k=1}^{L} g_k^* \tilde{n}_k(t)$$

$$\boxed{g_k = \alpha_k e^{j\phi_k}; \quad g_k^* g_k = \alpha_k^2}$$

Channel gain

Maximal Ratio Combining (MRC)

- The channel-gain envelope of the composite signal is

$$\alpha_C = \sum_{k=1}^{L} \alpha_k^2$$

- Assuming that all branches have the same noise power, the sum of the branch **noise powers** is

$$\sigma_{\tilde{n},tot}^2 = N_0 \sum_{k=1}^{L} \alpha_k^2$$

Weighted by the channel gains

- The symbol energy-to-noise ratio with MRC is

$$\gamma_s^{MR} = \frac{\alpha_C^2 E_{av}}{\sigma_{\tilde{n},tot}^2} = \sum_{k=1}^{L} \left(\alpha_k^2 \sum_{j=1}^{L} \alpha_j^2 E_{av} \right) \Bigg/ \left(N_0 \sum_{j=1}^{L} \alpha_j^2 \right) = \sum_{k=1}^{L} \frac{\alpha_k^2 E_{av}}{N_0} = \sum_{k=1}^{L} \gamma_k$$

  - where E_{av} is the average received symbol energy and

$$\gamma_k = \alpha_k^2 E_{av} / N_0$$

Average received symbol energy

Maximal Ratio Combining (MRC)

- γ_s^{MR} is the sum of the bit energy-to-noise ratios of all diversity branches
- If all diversity branches provide the same average power and the branches are uncorrelated, then γ_s^{MR} has a chi-square distribution with $2L$ degrees of freedom

$$p_{\gamma_s^{MR}}(x) = \frac{1}{(L-1)!(\bar{\gamma}_c)^L} x^{L-1} e^{-x/\bar{\gamma}_c}$$

  – where

$$\bar{\gamma}_c = E[\gamma_k], \quad k = 1, \cdots, L$$

- The cdf of γ_s^{MR}:

$$F_{\gamma_s^{MR}}(x) = 1 - e^{-x/\bar{\gamma}_c} \sum_{k=0}^{L-1} \frac{1}{k!} \left(\frac{x}{\bar{\gamma}_c} \right)^k$$

Maximal Ratio Combining (MRC)

- The average symbol energy-to-noise ratio with MRC is

$$\bar{\gamma}_s^{MR} = E\left[\gamma_s^{MR}\right] = E\left[\sum_{k=1}^{L} \gamma_k\right] = \sum_{k=1}^{L} E\left[\gamma_k\right] = \sum_{k=1}^{L} \bar{\gamma}_k = \sum_{k=1}^{L} \bar{\gamma}_c = L\bar{\gamma}_c$$

- For SC, $L = 2$: $F_{\gamma_s^S}(x) = 10^{-4}$ at $\gamma_s^S - \bar{\gamma}_c = -20$ dB
- For MRC, $L = 2$: $F_{\gamma_s^{MR}}(x) = 10^{-4}$ at $\gamma_s^{MR} - \bar{\gamma}_c = -18$ dB
  – MRC is 2 dB more effective than SC
- The bit error probability of BPSK is

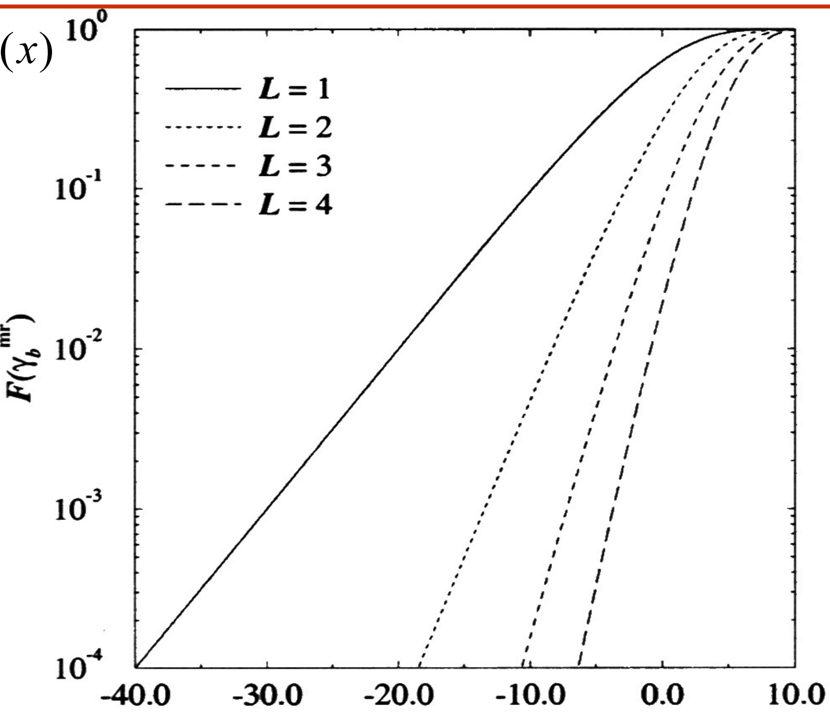
$$P_b = \int_0^\infty P_b(x) p_{\gamma_s^{MR}}(x)\, dx = \int_0^\infty Q\left(\sqrt{2x}\right) \frac{1}{(L-1)!(\bar{\gamma}_c)^L} x^{L-1} e^{-x/\bar{\gamma}_c}\, dx$$

$$= \left(\frac{1-\mu}{2}\right)^L \sum_{k=0}^{L-1} C_k^{L-1+k} \left(\frac{1+\mu}{2}\right)^k$$

$$\boxed{\mu = \sqrt{\frac{\bar{\gamma}_c}{1+\bar{\gamma}_c}}}$$

Maximal Ratio Combining (MRC)

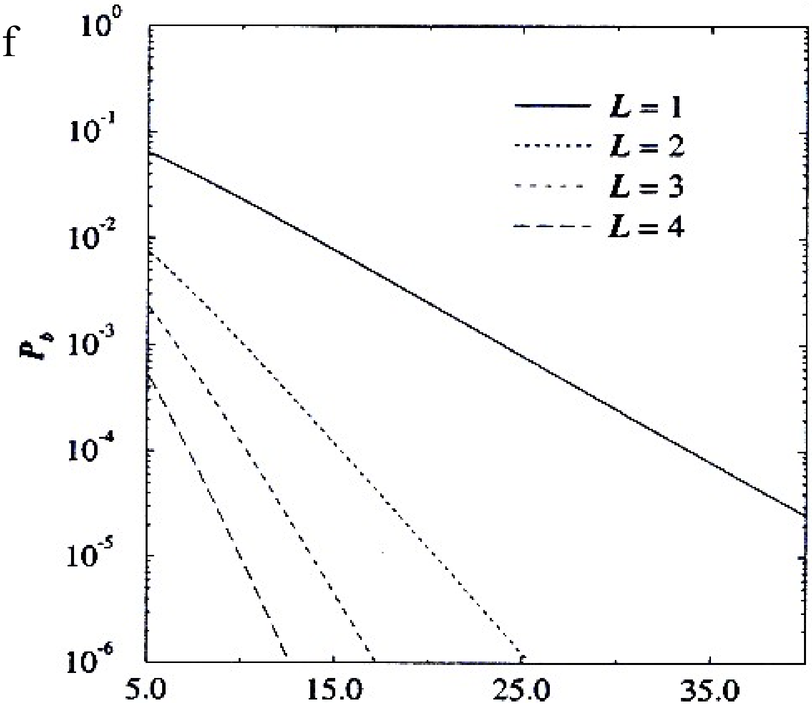
$F_{\gamma_s^{MR}}(x)$



$\gamma_s^{MR} - \overline{\gamma}_c$ (dB)

Maximal Ratio Combining (MRC)

- BER of BPSK



$\overline{\gamma}_c$ (dB)

Coherent Equal Gain Combining (EGC)

- The **coherent** Equal Gain Combining (EGC) is similar to MRC, but the diversity branches are **not weighted**
- EGC is useful for the modulation techniques having **equal energy symbols** (no channel information ***g*** is required), e.g. M-PSK
- For signals with unequal energy, the channel vector ***g*** is required and MRC (**best possible performance**) should be used
- The ML receiver chooses the symbol $\tilde{s}_m$ that maximizes the metric

Phase information is still required

$$\mu(\tilde{s}_m) = \sum_{k=1}^{L} \mathrm{Re}\left\{\tilde{r}_k e^{-j\phi_k}, \tilde{s}_m\right\} = \sum_{k=1}^{L} \mathrm{Re}\left\{ \int_0^T e^{-j\phi_k} \tilde{r}_k(t)\tilde{s}_m^*(t)\, dt\right\}$$

Coherent Equal Gain Combining (EGC)

$$\mu(\tilde{s}_m) = \mathrm{Re}\left\{\sum_{k=1}^{L} e^{-j\phi_k}\tilde{r}_k, \tilde{s}_m\right\} = \int_0^T \mathrm{Re}\left\{\left(\sum_{k=1}^{L} e^{-j\phi_k}\tilde{r}_k(t)\right)\tilde{s}_m^*(t)\right\} dt$$

- For pre-detection combining, the signal after the diversity combiner is

$$\tilde{\mathbf{r}}(t) = \sum_{k=1}^{L} e^{-j\phi_k}\tilde{r}_k(t) = \sum_{k=1}^{L} e^{-j\phi_k} g_k \tilde{s}(t) + \sum_{k=1}^{L} e^{-j\phi_k}\tilde{n}_k(t)$$

- The vector $\tilde{\mathbf{r}}$ is then applied to the metric computer for MRC with $\beta_m = 0, m = 1,\cdots,L$
- The channel-gain envelope of the composite signal is

$$\alpha_E = \sum_{k=1}^{L} \alpha_k$$

The symbol energy E_m is excluded

- Assuming that all branches have the same noise power, the sum of the branch **noise powers** is LN_0

Coherent Equal Gain Combining (EGC)

- The symbol energy-to-noise ratio with EGC is

$$\gamma_s^{EG} = \frac{\alpha_E^2 E_{av}}{LN_0}$$

  – where E_{av} is the average symbol energy

- The cdf and pdf of γ_s^{EG} do not exist in closed form for $L > 2$
- For $L = 2$ and $\bar{\gamma}_1 = \bar{\gamma}_2 = \bar{\gamma}_c$, the cdf and pdf are

$$F_{\gamma_s^{EG}}(x) = 1 - e^{-2x/\bar{\gamma}_c} - \sqrt{\pi \frac{x}{\bar{\gamma}_c}} e^{-x/\bar{\gamma}_c} \left(1 - 2Q\left(\sqrt{2\frac{x}{\bar{\gamma}_c}}\right)\right)$$

$$p_{\gamma_s^{EG}}(x) = \frac{1}{\bar{\gamma}_c} e^{-2x/\bar{\gamma}_c} + \sqrt{\pi} e^{-x/\bar{\gamma}_c} \left(\frac{1}{2\sqrt{x\bar{\gamma}_c}} - \frac{1}{\bar{\gamma}_c}\sqrt{\frac{x}{\bar{\gamma}_c}}\right) \times \left(1 + 2Q\left(\sqrt{2\frac{x}{\bar{\gamma}_c}}\right)\right)$$

Coherent Equal Gain Combining (EGC)

- The average symbol energy-to-noise ratio with EGC is

$$\bar{\gamma}_s^{EG} = \frac{E_{av}}{LN_0} E\left[\alpha_E^2\right] = \frac{E_{av}}{LN_0} E\left[\left(\sum_{k=1}^{L} \alpha_k\right)^2\right] = \frac{E_{av}}{LN_0} \sum_{k=1}^{L} \sum_{\ell=1}^{L} E\left[\alpha_k \alpha_\ell\right]$$

- For an **uncorrelated** Rayleigh fading channel:

$$E\left[\alpha_k^2\right] = 2\sigma^2 \text{ and } E\left[\alpha_k \alpha_\ell\right] = E\left[\alpha_k\right] E\left[\alpha_\ell\right] = \left(\sqrt{\pi/2}\sigma\right)^2 = \pi\sigma^2/2$$

$$\bar{\gamma}_s^{EG} = \frac{E_{av}}{LN_0}\left(2L\sigma^2 + L(L-1)\frac{\pi\sigma^2}{2}\right)$$

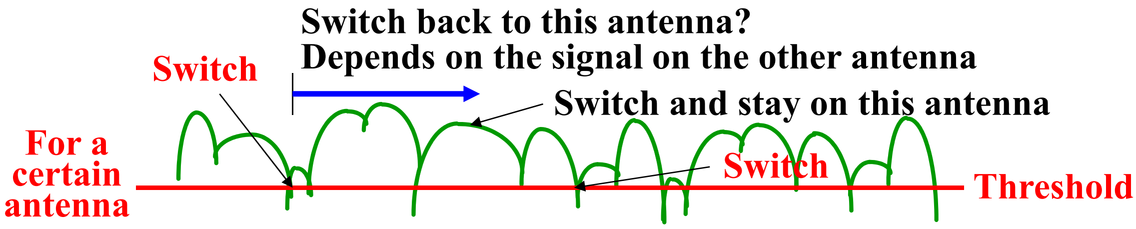
$$= \frac{2\sigma^2 E_{av}}{N_0}\left(1 + (L-1)\frac{\pi}{4}\right) = \bar{\gamma}_c\left(1 + (L-1)\frac{\pi}{4}\right) \le L\bar{\gamma}_c$$

- The bit error probability of coherent BPSK (2-branch):

$$P_b = \int_0^\infty P_b(x) p_{\gamma_s^{EG}}(x)\, dx = \frac{1}{2}\left(1 - \sqrt{1-\mu^2}\right), \quad \mu = \frac{1}{1+\bar{\gamma}_c}$$

Predetection Switched Combining (SW)

- A switched combiner **scans** through the diversity branches until it finds one that has a signal-to-noise ratio **exceeding a specified threshold**
- This diversity branch is used until the signal-to-noise ratio again **drops below the threshold**
- Two-branch **Switch and Stay Combining (SSC)**:
  - The receiver switches to, and stays with, the alternate branch when the bit energy-to-noise ratio drops below the threshold
  - It does this regardless of whether or not the bit energy-to-noise ratio of the alternate branch is **above or below** the threshold



Predetection Switched Combining (SW)

- Assume that the symbol energy-to-noise ratios associated with the two branches are γ_1 and γ_2 and the switching threshold is T
- The pdf of the received bit energy-to-noise ratio is

$$p_{\gamma_k}(x) = \frac{1}{\bar{\gamma}_c} e^{-x/\bar{\gamma}_c}$$

- Define

$$q = \Pr\left[\gamma_i < T\right] = 1 - e^{-T/\bar{\gamma}_c}$$

$$p = \Pr\left[\gamma_i \le S\right] = 1 - e^{-S/\bar{\gamma}_c}$$

- The symbol energy-to-noise ratio at the output of the switched combiner is γ_s^{SW}

$$\Pr\left[\gamma_s^{SW} \le S\right] = \Pr\left[\gamma_s^{SW} \le S \middle| \gamma_s^{SW} = \gamma_1\right] \cup \Pr\left[\gamma_s^{SW} \le S \middle| \gamma_s^{SW} = \gamma_2\right]$$

Predetection Switched Combining (SW)

- γ_1 is statistically identical to $\gamma_2 \Rightarrow$ Assuming that branch 1 is currently in use, we have the cdf of γ_s^{SW} expressed as

$$\Pr\left[\gamma_s^{SW} \le S\right] = \begin{cases} \Pr\left[\{\gamma_1 < T\} \cap \{\gamma_2 \le S\}\right], & S < T \\ \Pr\left[\{T \le \gamma_1 \le S\} \cup \{\gamma_1 < T \cap \gamma_2 \le S\}\right], & S \ge T \end{cases}$$

- Since γ_1 and γ_2 are independent, we have

$$\Pr\left[\{\gamma_1 < T\} \cap \{\gamma_2 \le S\}\right] = qp$$

$$\Pr\left[\{T \le \gamma_1 \le S\} \cup \{\gamma_1 < T \cap \gamma_2 \le S\}\right] = (p - q) + qp$$

$$\Pr\left[\gamma_s^{SW} \le S\right] = \begin{cases} qp, & S < T \\ p - q + qp, & S \ge T \end{cases}$$

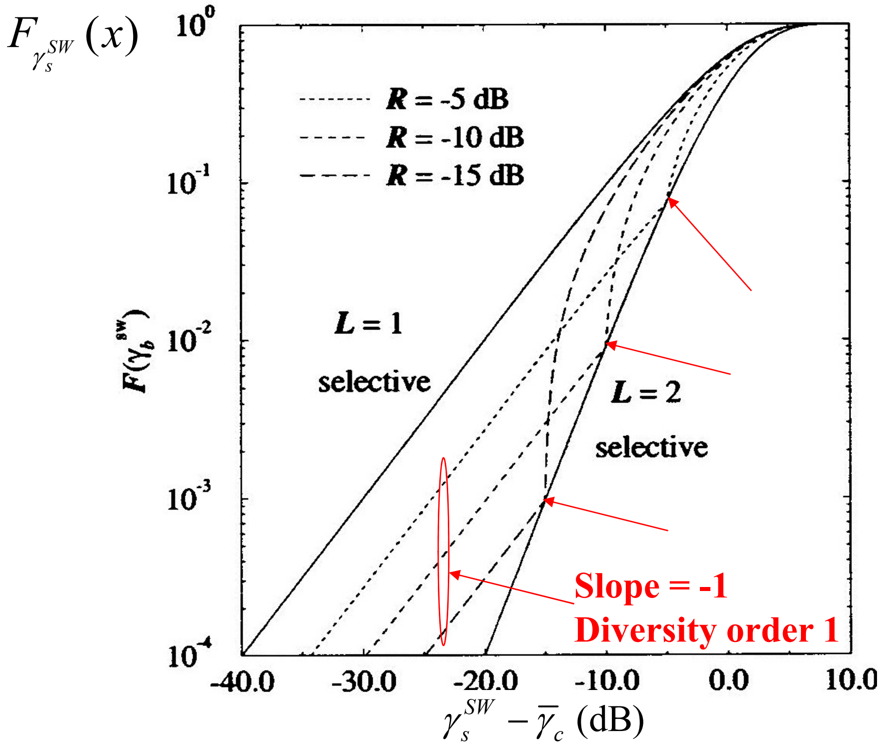
Predetection Switched Combining (SW)

- The normalized threshold is defined as

$$R = 10\log_{10}\left(T/\bar{\gamma}_c\right)(\text{dB})$$

- SSC always performs **worse than SC**
  - Except at the **switching threshold**, where the performance is the same as SC
- The threshold level T should be chosen as γ_{th}
  - The minimum acceptable instantaneous symbol energy-to-noise ratio
- The optimum threshold, $T = R\bar{\gamma}_c$, depends on $\bar{\gamma}_c$
  - Since $\bar{\gamma}_c$ varies due to path loss and shadowing, the normalized threshold R must be adaptive

Predetection Switched Combining (SW)



Predetection Switched Combining (SW)

- The pdf of γ_s^{SW} is

$$p_{\gamma_s^{SW}}(x) = \begin{cases} q\dfrac{1}{\bar{\gamma}_c}e^{-x/\bar{\gamma}_c}, & x < T \\ (1+q)\dfrac{1}{\bar{\gamma}_c}e^{-x/\bar{\gamma}_c}, & x \geq T \end{cases}$$

- If binary DPSK is used, the probability of error is

$$P_b = \int_0^\infty P_b(x) p_{\gamma_s^{SW}}(x)\, dx$$

Diversity order = 1

$$= \frac{1}{2(1+\bar{\gamma}_c)}\underline{\left(q+(1-q)e^{-T}\right)}$$

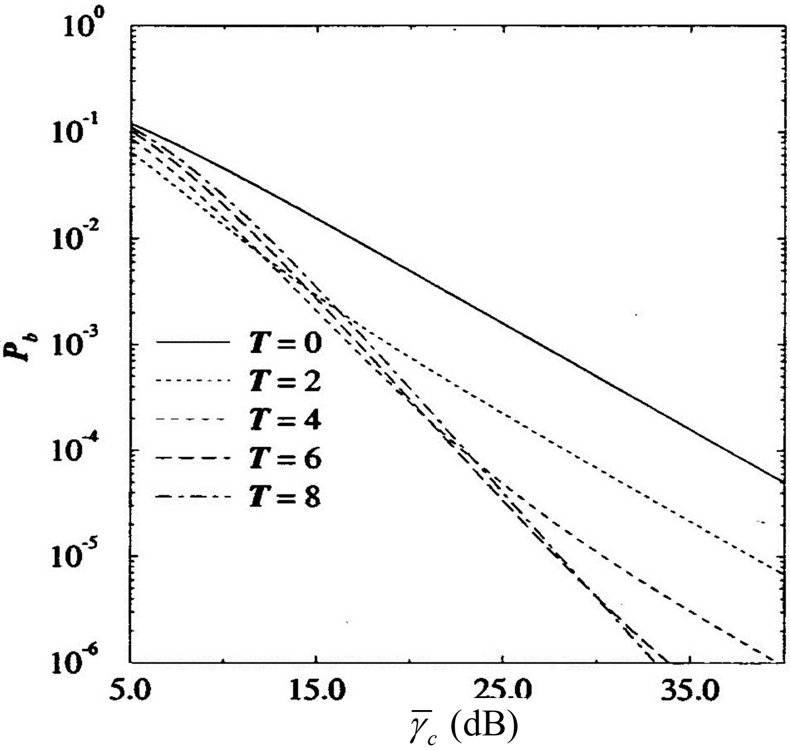
Almost a constant for $\bar{\gamma}_c \gg 0$

- The performance with $T = 0$ (no switching): no diversity at all
- The performance changes little for $T > 6$

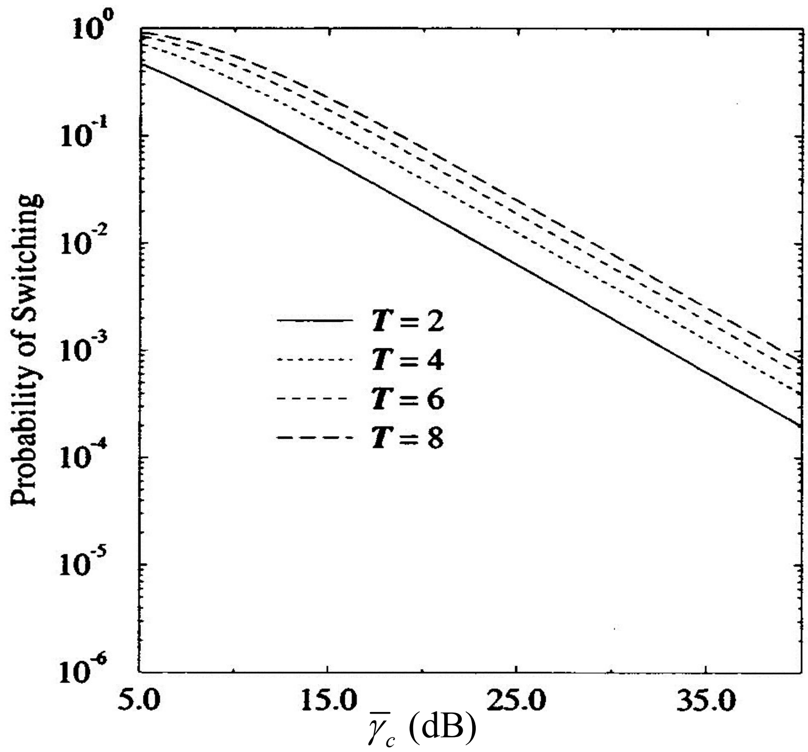
Predetection Switched Combining (SW)

- BER of DPSK
- When $\overline{\gamma}_c \gg 0$, the receiver almost always stays at the same antenna

$\Rightarrow$ **Diversity order = 1**



Predetection Switched Combining (SW)



Question

Transmit Diversity

Transmit Diversity (TD)

- Transmit diversity uses **multiple transmit antennas** to provide the receiver with multiple uncorrelated replicas of the same signal
  - The complexity of having multiple antennas may be shared among many receivers
  - Generally used in **forward link** (base-to-mobile)
- The receiver with only **one** antenna can still benefit from a diversity gain

Space-Time Transmit Diversity

- Assume that two transmit antennas and one receiver antenna are used
- In any given baud period, two data symbols are transmitted simultaneously from the two transmit antennas
  - The symbols transmitted from Antenna 1 and 2 are $\tilde{\mathbf{s}}_{(1)}$ and $\tilde{\mathbf{s}}_{(2)}$
  - During the next baud period, the symbols transmitted from Antenna 1 and 2 are $-\tilde{\mathbf{s}}^*_{(2)}$ and $\tilde{\mathbf{s}}^*_{(1)}$
- The channel gains for the two antennas are $g_1(t)$ and $g_2(t)$
- Assume that the channel stays **constant** over **two** baud intervals

$$g_k(t) = g_k(t+T) = g_k = \alpha_k e^{j\phi_k}$$

  - where T is the baud period

Space-Time Transmit Diversity

- The received complex signal vectors are:

$$\tilde{\mathbf{r}}_{(1)} = g_1\tilde{\mathbf{s}}_{(1)} + g_2\tilde{\mathbf{s}}_{(2)} + \tilde{\mathbf{n}}_{(1)}$$
$$\tilde{\mathbf{r}}_{(2)} = -g_1\tilde{\mathbf{s}}^*_{(2)} + g_2\tilde{\mathbf{s}}^*_{(1)} + \tilde{\mathbf{n}}_{(2)}$$

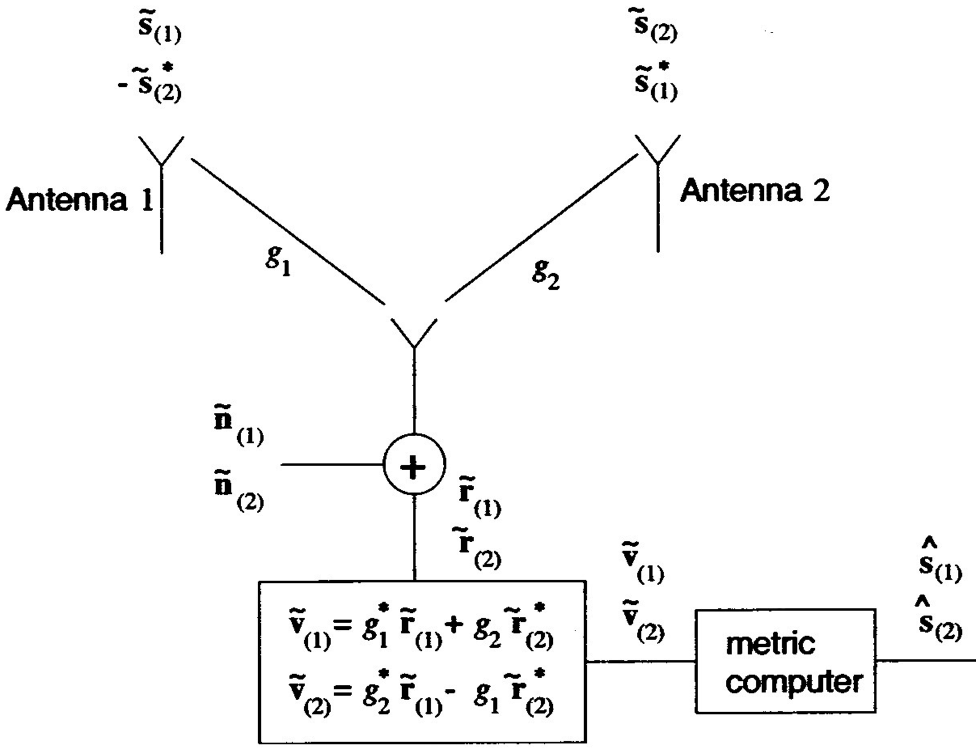
  - where $\tilde{\mathbf{r}}_{(1)}$ and $\tilde{\mathbf{r}}_{(2)}$ represent the received vectors at t and $t+T$
- The diversity combiner constructs the following two signal vectors:

$$\tilde{\mathbf{v}}_{(1)} = g_1^*\tilde{\mathbf{r}}_{(1)} + g_2\tilde{\mathbf{r}}^*_{(2)}$$
$$\tilde{\mathbf{v}}_{(2)} = g_2^*\tilde{\mathbf{r}}_{(1)} - g_1\tilde{\mathbf{r}}^*_{(2)}$$

$$\tilde{\mathbf{v}}_{(1)} = \left(\alpha_1^2 + \alpha_2^2\right)\tilde{\mathbf{s}}_{(1)} + g_1^*\tilde{\mathbf{n}}_{(1)} + g_2\tilde{\mathbf{n}}^*_{(2)}$$
$$\tilde{\mathbf{v}}_{(2)} = \left(\alpha_1^2 + \alpha_2^2\right)\tilde{\mathbf{s}}_{(2)} - g_1\tilde{\mathbf{n}}^*_{(2)} + g_2^*\tilde{\mathbf{n}}_{(1)}$$

$$\boxed{g_1^* g_1 = (\alpha_1 e^{-j\phi_1})(\alpha_1 e^{j\phi_1}) = \alpha_1^2}$$

Space-Time Transmit Diversity (2×1 Diversity)



Space-Time Transmit Diversity

- Applying the metric computer for MRC, we can make the decisions by maximizing the metric

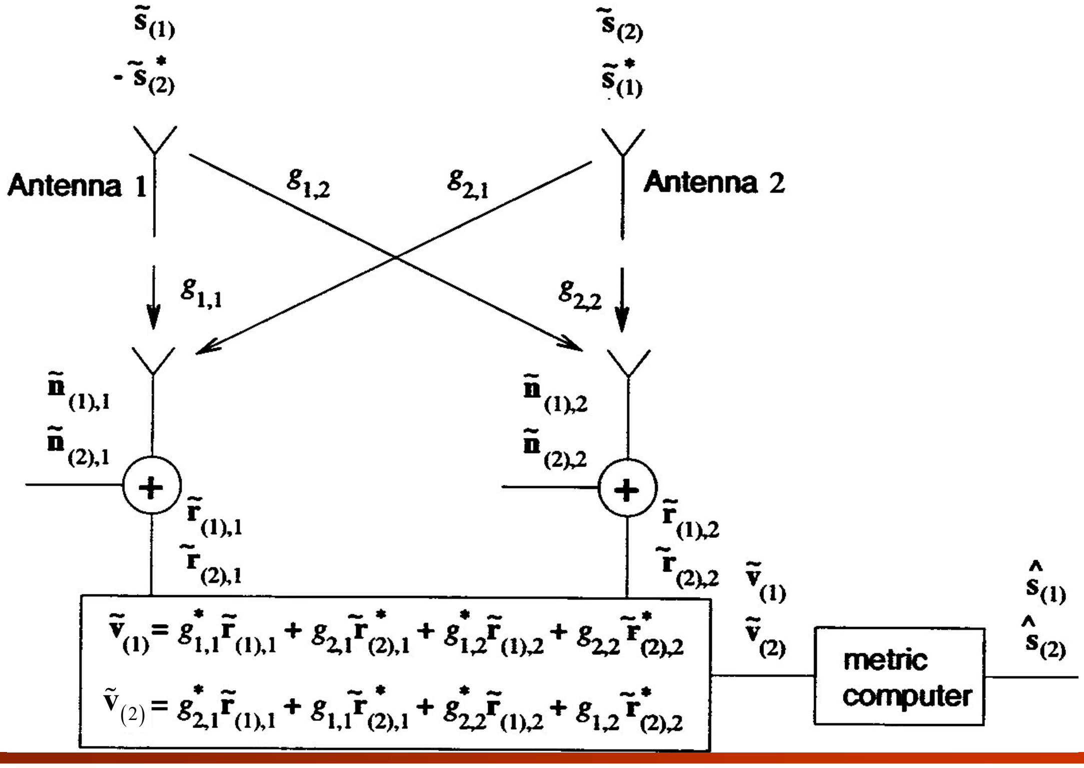
$$\mu\left(\tilde{\mathbf{s}}_{(1),m}\right) = \mathrm{Re}\left(\tilde{\mathbf{v}}_{(1)}, \tilde{\mathbf{s}}_{(1),m}\right) - E_m\left(\left|g_1\right|^2 + \left|g_2\right|^2\right)$$
$$\mu\left(\tilde{\mathbf{s}}_{(2),m}\right) = \mathrm{Re}\left(\tilde{\mathbf{v}}_{(2)}, \tilde{\mathbf{s}}_{(2),m}\right) - E_m\left(\left|g_1\right|^2 + \left|g_2\right|^2\right)$$

- Compared with the output of the MRC with $L = 2$

$$\begin{aligned}\tilde{\mathbf{r}} &= g_1^*\tilde{\mathbf{r}}_1 + g_2^*\tilde{\mathbf{r}}_2 \\ &= \left(\alpha_1^2 + \alpha_2^2\right)\tilde{\mathbf{s}}_m + g_1^*\tilde{\mathbf{n}}_1 + g_2^*\tilde{\mathbf{n}}_2\end{aligned}$$

- The combined signals are the same, except for the **phase rotations of the noise vectors**

Space-Time Transmit Diversity (2×2 Diversity)



Signal Space Diversity

Signal Space Diversity (SSD)

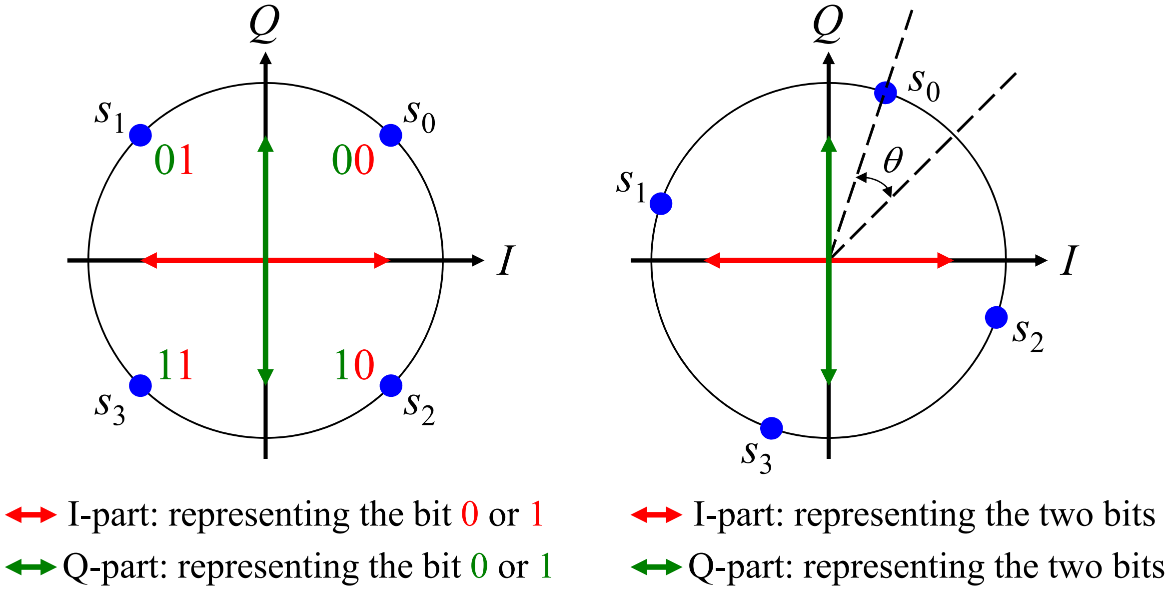
- The SSD technology was proposed to exploit **diversity gain** for the transmission through fading channels.
- At the **transmitter**, the input information bits are modulated to form a set of complex symbols, such as M-PSK or QAM.
- Subsequently, the modulated signals are **phase rotated** by a predetermined angle and followed by an ideal **component interleaving** process.
  - Interleave the **I-part** and **Q-part** components of multiple symbols **to form a set of new symbols**, in each of which **the two components come from two original symbols**.
- Then the new (**interleaved** and **combined**) symbols are transmitted through the propagation channels.

Signal Space Diversity (SSD)

- At the **receiver**, the signals received on all channels are **component de-interleaved** to reconstruct the symbols corresponding to the original symbols.
- Multiple original symbols are involved in using the common channels
  - All related symbols are jointly detected based on the **maximum likelihood (ML)** criterion.
- If L original symbols are involved in SSD, an diversity order L can be achieved.
- No additional power or bandwidth resources are required.

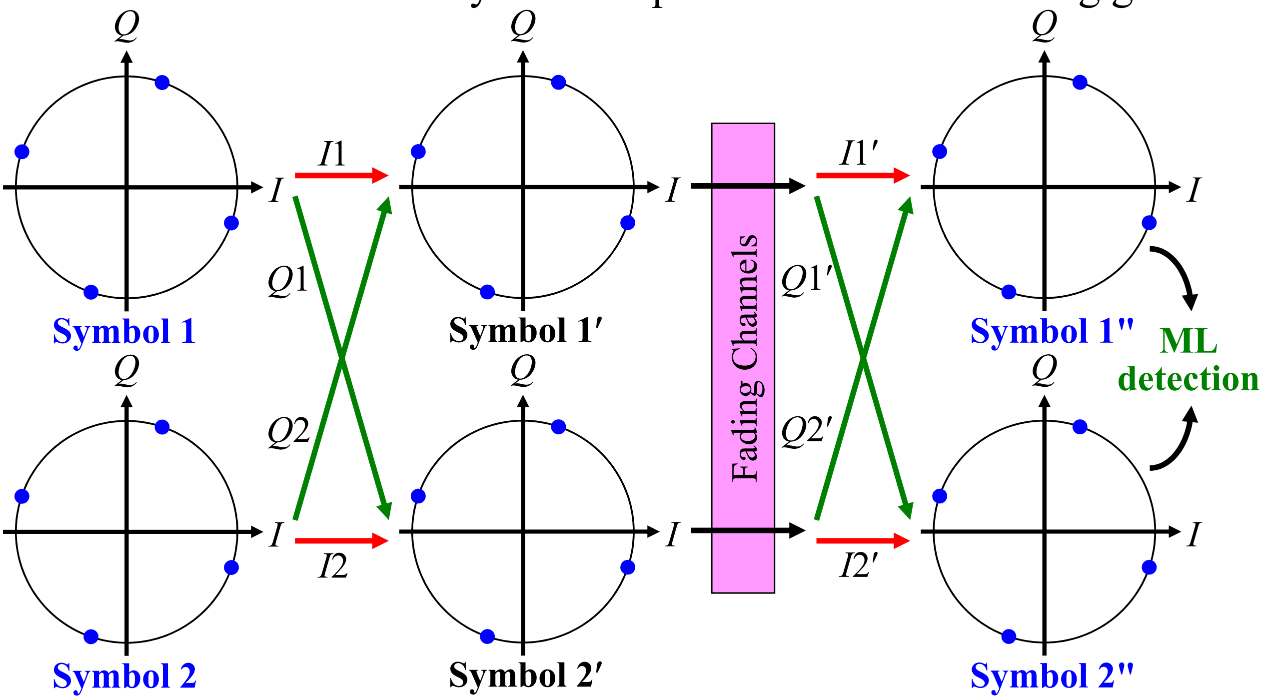
Signal Space Diversity (SSD)

- Without phase rotation, **I-part/Q-part** depends only on one bit
- With phase rotation, **I-part/Q-part** depends on the two bits



Signal Space Diversity (SSD)

- The two transmitted symbols experience **different** fading gains

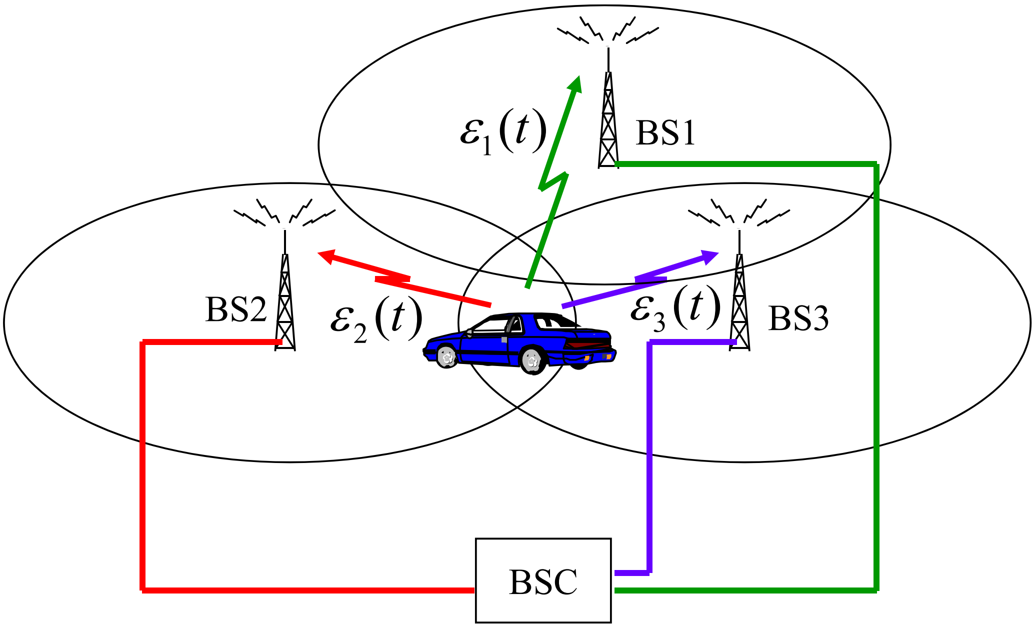


Macroscopic Diversity

Macroscopic Diversity

- **Microscopic diversity** applies some diversity techniques to mitigate the effects of envelope fading (fast fading)
- **Macroscopic diversity** is a diversity technique that is used to combat the effects of **shadow fading**
  - **Up-link** signals from a mobile station are received by **two or more** geographically separated base stations
  - The received signals are diversity combined
  - The diversity advantage is obtained if the signals experience some degree of **uncorrelated shadowing**
- Assume that there are M diversity branches (BSs)
- The average received bit energy-to-noise ratio per diversity branch is $\bar{\gamma}_c$ (in dB)

Macroscopic Diversity



Macroscopic Diversity

- We have $p(\bar{\gamma}_{c(\mathrm{dB})}) = \frac{1}{\sqrt{2\pi}\sigma_{\bar{\gamma}_c}} \exp\left\{ -\left(\bar{\gamma}_{c(\mathrm{dB})} - \mu_{\bar{\gamma}_c} \right)^2 \big/ 2\sigma^2_{\bar{\gamma}_c} \right\}$
  - where $\mu_{\bar{\gamma}_c} = E[\bar{\gamma}_{c(\mathrm{dB})}]$
- The most effective and simplest method for macroscopic diversity is **selective combining**
  - The BS that provides the largest average received bit energy-to-noise ratio is selected as the serving BS
- Let $\bar{\gamma}_{th(\mathrm{dB})}$ be a specified threshold level, and $\bar{\gamma}_{ck(\mathrm{dB})}$ be the average received bit energy-to-noise ratio from BS k

$$\begin{aligned}\Pr(\bar{\gamma}_{ck(\mathrm{dB})} \le \bar{\gamma}_{th(\mathrm{dB})}) &= \int_{-\infty}^{\bar{\gamma}_{th(\mathrm{dB})}} \frac{1}{\sqrt{2\pi}\sigma_{\bar{\gamma}_c}} \exp\left\{ -\left(\bar{\gamma}_{ck(\mathrm{dB})} - \mu_{\bar{\gamma}_{ck}} \right)^2 \big/ 2\sigma^2_{\bar{\gamma}_c} \right\} d\bar{\gamma}_{ck} \\ &= Q\left((\mu_{\bar{\gamma}_{ck}} - \bar{\gamma}_{th(\mathrm{dB})}) / \sigma_{\bar{\gamma}_c} \right)\end{aligned}$$

Macroscopic Diversity

- Let $\mathbf{d} = (d_0, d_1, \cdots, d_{M-1})$ be the set of distances between the MS and the M different BSs
- The outage probability of thermal noise:

$$O(\mathbf{d}) = \prod_{k=0}^{M-1} \Pr(\bar{\gamma}_{ck(\mathrm{dB})} \le \bar{\gamma}_{th(\mathrm{dB})}) = \prod_{k=0}^{M-1} Q\left(\frac{\mu_{\bar{\gamma}_{ck}} - \bar{\gamma}_{th(\mathrm{dB})}}{\sigma_{\bar{\gamma}_c}} \right)$$

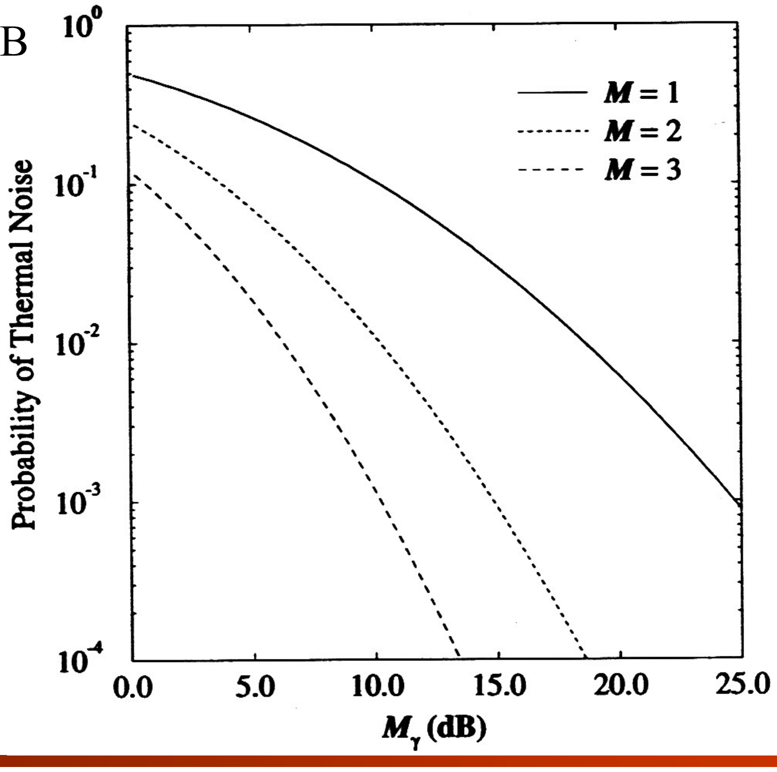
- Assume that $\mu_{\bar{\gamma}_{ck}}$ are all equal to $\mu_{\bar{\gamma}_c}$, i.e., $d_0 = d_1 = \ldots = d_{M-1}$
  - The outage probability becomes

$$O = \left[Q\left(\frac{\mu_{\bar{\gamma}_c} - \bar{\gamma}_{th(\mathrm{dB})}}{\sigma_{\bar{\gamma}_c}} \right) \right]^M$$

- The thermal noise margin: $M_\gamma = \mu_{\bar{\gamma}_c} - \bar{\gamma}_{th(\mathrm{dB})}$

Macroscopic Diversity

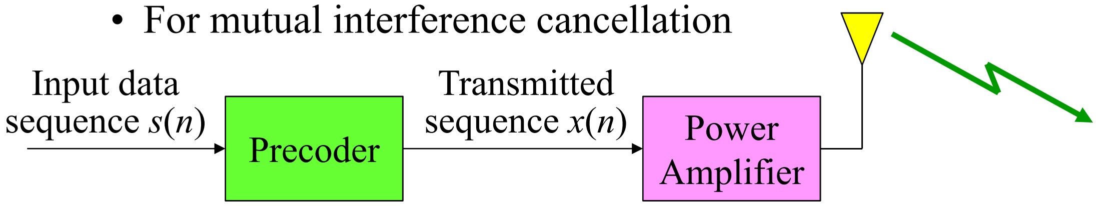
$\sigma_{\bar{\gamma}_c} = 8\text{ dB}$



Precoding

Precoding

- **Precoding** is a modification of a sequence of signals (symbols) before transmission to meet a specific goal/requirement.
- There are various purposes for using precoding in wireless transmissions
  - To decrease the envelope variations of the transmitted signals
  - To mitigate the inter-symbol interference induced by channels
  - To achieve radio resource mapping for different signals
    - For diversity reception
    - For multiple access
    - For mutual interference cancellation



Precoding – Transmit Diversity

- Space-time transmit diversity is an example of using precoding
- The input data signals are transformed to new modified signals for transmission
  - To achieve diversity reception at the receiver
- For 2×1 transmit diversity, the transmitted signals are

Original data signals **Transmitted signals**

$$\begin{bmatrix}\tilde{s}_{(1)}\\ \tilde{s}_{(2)}\end{bmatrix} \xrightarrow{\text{Precoding}} \begin{bmatrix}\tilde{x}_{(1),1} & \tilde{x}_{(1),2}\\ \tilde{x}_{(2),1} & \tilde{x}_{(2),2}\end{bmatrix} = \begin{bmatrix}\tilde{s}_{(1)} & \tilde{s}_{(2)}\\ -\tilde{s}^{*}_{(2)} & \tilde{s}^{*}_{(1)}\end{bmatrix}$$

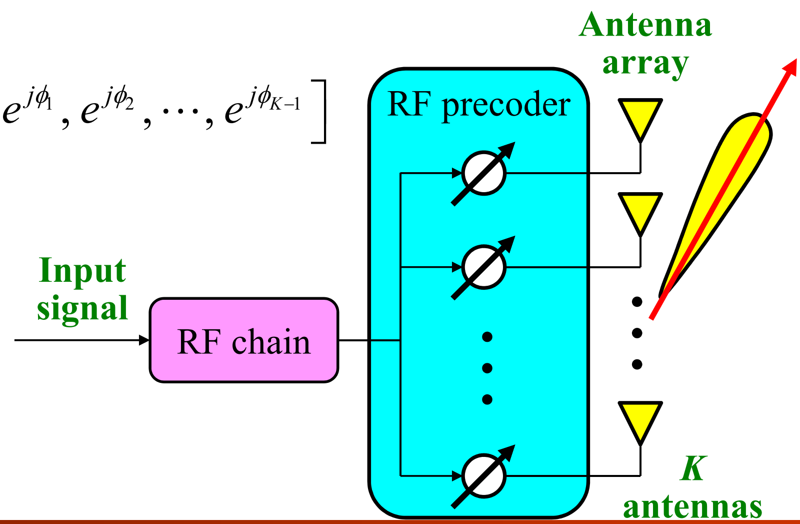
Antenna index

Alamouti's code

Precoding – Beamforming

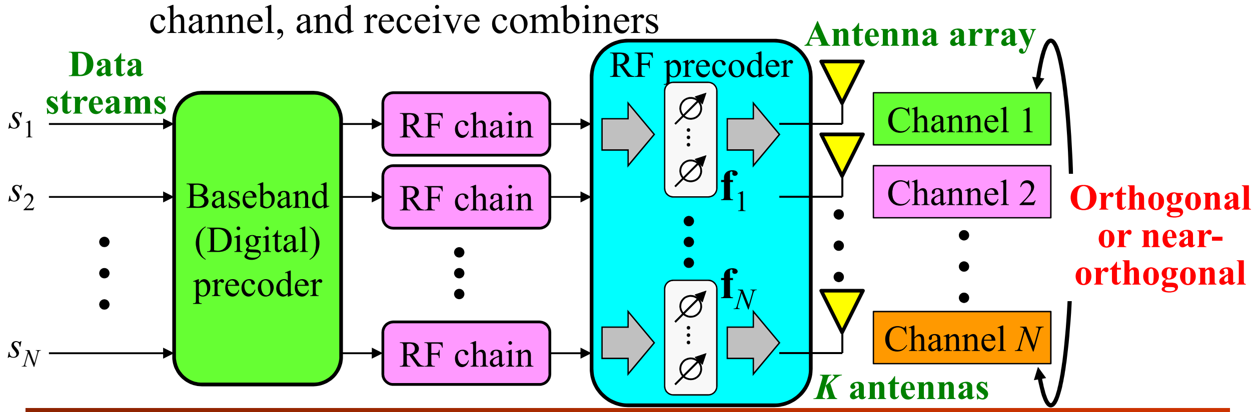
- The same signal is emitted from each of the transmit antennas with appropriately chosen **weighting** (**phase** and/or **gain**)
  - Such that the signal power is **maximized** in a desired direction
  - For RF precoding, a **constant amplitude** constraint is required

$$\mathbf{f} = \left[1, e^{j\phi_1}, e^{j\phi_2}, \cdots, e^{j\phi_{K-1}}\right]$$



Precoding – Spatial Multiplexing

- **Multiple data streams** are emitted from the transmit antennas with **different** and appropriately chosen weightings
  - Such that the transmitter can **simultaneously** transmit multiple data streams to the receiver/receivers using **the same resources**
  - The goal is to **maximize** the total system throughput
  - **Multiple spatial channels** are formed through the precoders, channel, and receive combiners

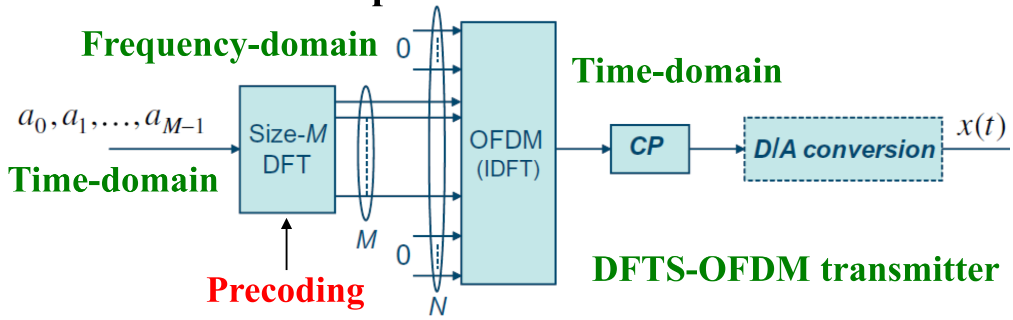


Precoding – DFT-Spread OFDM

- **Flexible bandwidth assignment**: For OFDM-based **uplink** transmission, **dynamically allocating** different number of subcarriers to different terminals depending on their instantaneous channel conditions
- To achieve the goal of **low-PAPR** (peak-to-average power ratio) at a user equipment, "**single-carrier**" transmission is used
  - The **DFT-spread OFDM** (DFTS-OFDM) transmission scheme
  - Small variations in the instantaneous signal power
  - Possibility for low-complexity high-quality **equalization** in the **frequency domain**
  - Possibility for FDMA with **flexible bandwidth** assignment

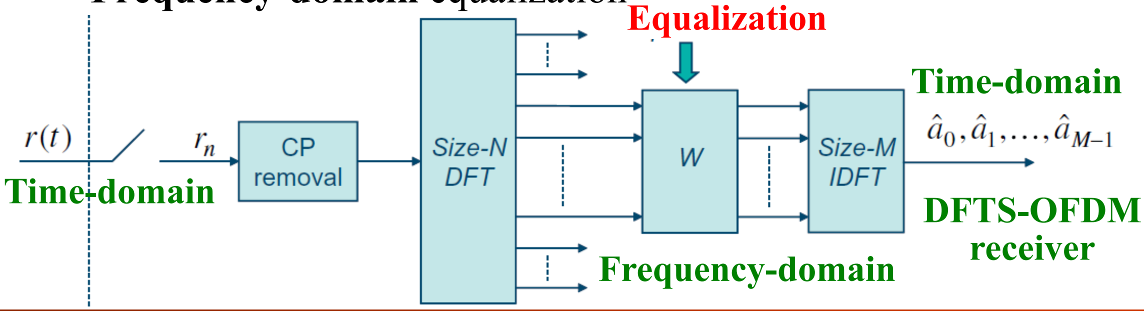
Precoding – DFT-Spread OFDM

- The DFTS-OFDM transmission can be regarded as normal OFDM transmission with a **DFT-based precoding**
  - A block of M modulation symbols (BPSK, QPSK, 16QAM, …) is first applied to a size-M DFT
  - The output is then applied to consecutive inputs (**subcarriers**) of an OFDM modulator, depending on the **bandwidth assignment**
  - The OFDM modulator is a size-N inverse DFT (IDFT) with $N > M$ and the **unused inputs** of the IDFT are set to zero



Precoding – DFT-Spread OFDM

- The demodulation of a DFTS-OFDM signal is similar to the demodulation of an OFDM signal
  - Using size-N DFT processing with removal of the frequency samples **not corresponding to** the desired signal
  - Using size-M inverse DFT processing to obtain data estimate
- An equalizer is needed to compensate for the radio-channel **frequency selectivity**
  - **Frequency-domain** equalization



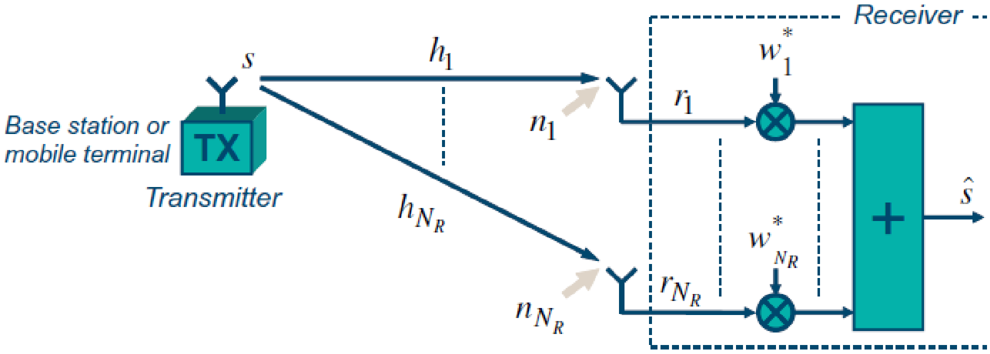
Multi-Antenna Techniques

Antenna Distance

- An important characteristic of multi-antenna configuration is the **distance** between the different antenna elements
  - The **mutual correlation** between the **radio-channel fading** experienced by the signals at the different antennas
  - A **large** inter-antenna distance ⇒ a **low** mutual fading correlation
    - Almost independent fast fading
  - A **small** inter-antenna distance ⇒ a **high** mutual fading correlation
    - Very similar fast fading
- Whether **high or low correlation** is desirable depends on what is to be achieved with the multi-antenna configuration
  - Diversity, beamforming, or spatial multiplexing

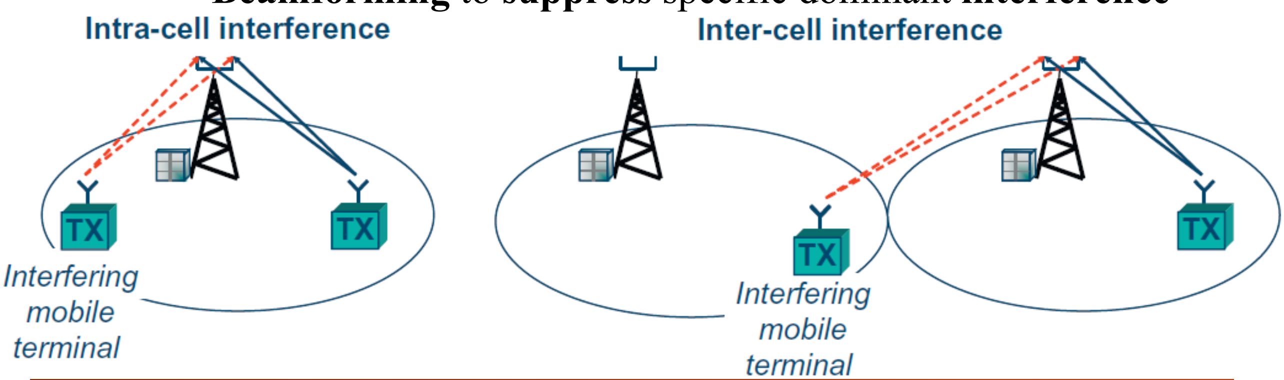
Benefits of Multi-Antenna: Diversity

- Using multiple antennas at the transmitter **and/or** the receiver to provide additional **diversity** against channel fading
  - **Low mutual correlation** between different antennas, implying
    - The need for a sufficiently **large inter-antenna distance** (**spatial diversity**) or
    - The use of **different antenna polarization** directions (**polarization diversity**)



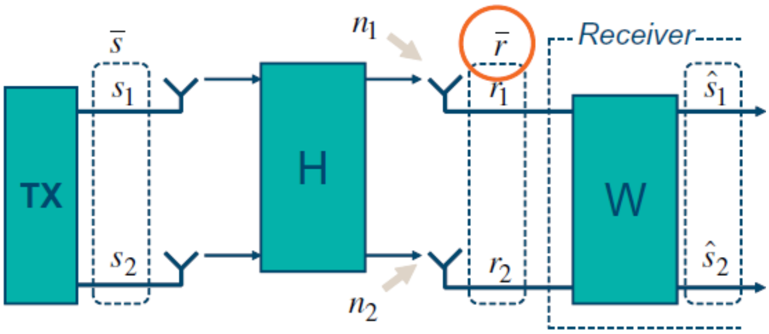
Benefits of Multi-Antenna: Beamforming

- Using multiple antennas at the transmitter **and/or** the receiver to "**shape**" the overall **antenna beam pattern**
  - **Very high mutual correlation** or **low mutual correlation** between different antennas
    - **Beamforming** to **maximize** the overall antenna gain in the direction of the target receiver/transmitter
    - **Beamforming** to **suppress** specific dominant **interference**



Benefits of Multi-Antenna: Spatial Multiplexing

- **Simultaneously** using multiple antennas at the transmitter and the receiver to support **multiple parallel channels** over the radio interface with the same radio resources
  - **Spatial multiplexing** to achieve very high bandwidth efficiency
  - Also referred to as **MIMO (Multi-Input Multi-Output)** antenna processing
- The case of 2×2 antenna configuration



Benefits of Multi-Antenna: Spatial Multiplexing

- The received signals can be expressed as:

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \mathbf{Hs} + \mathbf{n}$$

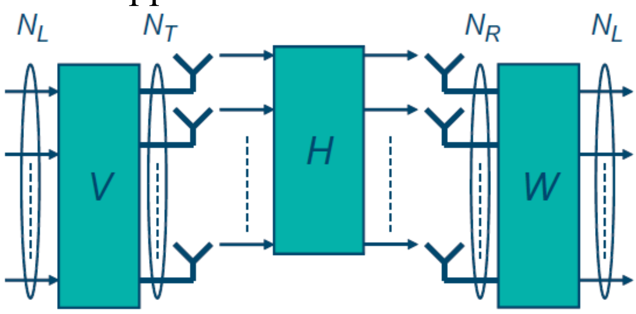
- Assuming no noise and that the channel matrix **H** is **invertible**, both signals s_1 and s_2 can be **perfectly** recovered at the receiver

$$\hat{\mathbf{s}} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \mathbf{Wr} = \mathbf{H}^{-1}\mathbf{r} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \underline{\mathbf{H}^{-1}\mathbf{n}} \leftarrow \text{Noise enhancement}$$

- The properties of **H** will determine the resultant noise level
  - The closer the channel matrix is to being a **singular matrix**, the larger the **increase** in the noise level

Benefits of Multi-Antenna: Spatial Multiplexing

- Consider a multiple-antenna configuration consisting of N_T transmit antennas and N_R receive antennas.
- The number of parallel signals that can be spatially multiplexed is **upper limited** by $N_L = \min\{N_T, N_R\}$
  - No more than N_T different signals can be transmitted from N_T transmit antennas
  - With N_R receive antennas, a maximum of $N_R - 1$ interfering signals can be suppressed

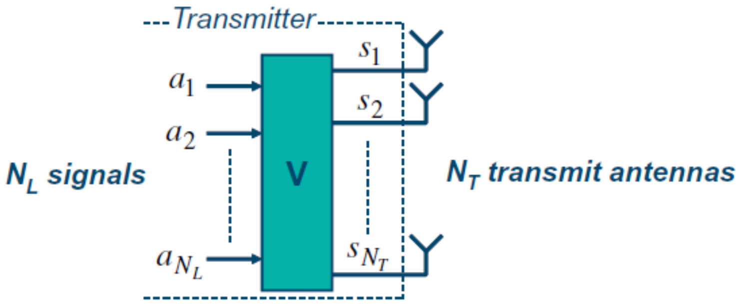


Benefits of Multi-Antenna: Spatial Multiplexing

- However, in many cases, the number of spatially multiplexed signals will be **less than** N_L
- In the case of **very bad** channel conditions (low signal-to-noise ratio, SNR) there is **no gain** of spatial multiplexing
  - The channel capacity is a linear function of the SNR
  - The multiple antennas should be used for **beamforming** to improve the SNR, rather than for spatial multiplexing
- The spatial-multiplexing order should be determined based on the properties of the size $N_R \times N_T$ channel matrix
  - Any **excess antennas** should be used to provide **beamforming**
  - Such combined beamforming and spatial multiplexing can be achieved by means of **precoder-based spatial multiplexing**

Precoder-Based Spatial Multiplexing

- Linear precoding for spatial multiplexing is achieved by using a size $N_T \times N_L$ **precoding matrix** at the transmitter side
- In the general case with $N_L \leq N_T$, N_L signals are spatially multiplexed and transmitted using N_T transmit antennas
  - If $N_L = N_T$, the precoding can be used to "**orthogonalize**" the parallel transmissions, allowing for improved signal isolation
  - If $N_L < N_T$, the precoding provides the mapping including the combination of **spatially multiplexing** and **beamforming**



Precoder-Based Spatial Multiplexing

- In the LTE system, the codebook-based precoding allows for a maximum of **four antenna ports** (output signals) and, as a consequence, a maximum of **four layers** (input signals)
- Precoder matrices for two antenna ports and one and two layers
  - One layer

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{W}x$$

$$\frac{1}{\sqrt{2}}\begin{bmatrix} +1 \\ +1 \end{bmatrix} \qquad \frac{1}{\sqrt{2}}\begin{bmatrix} +1 \\ -1 \end{bmatrix} \qquad \frac{1}{\sqrt{2}}\begin{bmatrix} +1 \\ +j \end{bmatrix} \qquad \frac{1}{\sqrt{2}}\begin{bmatrix} +1 \\ -j \end{bmatrix}$$

  - Two layers

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{Wx} = \mathbf{W}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\frac{1}{\sqrt{2}}\begin{bmatrix} +1 & 0 \\ 0 & +1 \end{bmatrix} \qquad \frac{1}{2}\begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix} \qquad \frac{1}{2}\begin{bmatrix} +1 & +1 \\ +j & -j \end{bmatrix}$$

Equalization

Adaptive Equalizer

- An **adaptive equalizer** is an adjustable filters at the receiver
  - Used to mitigate the combined effect of ISI and noise
- Two broad categories of equalizers:
  - **Symbol-by-symbol** equalizers
    - Include a decision device to make symbol-by-symbol decisions on the received symbol sequence
  - **Sequence estimators**
    - Make decisions on **a sequence of received symbols**
- Sequence estimators are generally more complex than symbol-by-symbol equalizers, but can potentially offer better performance

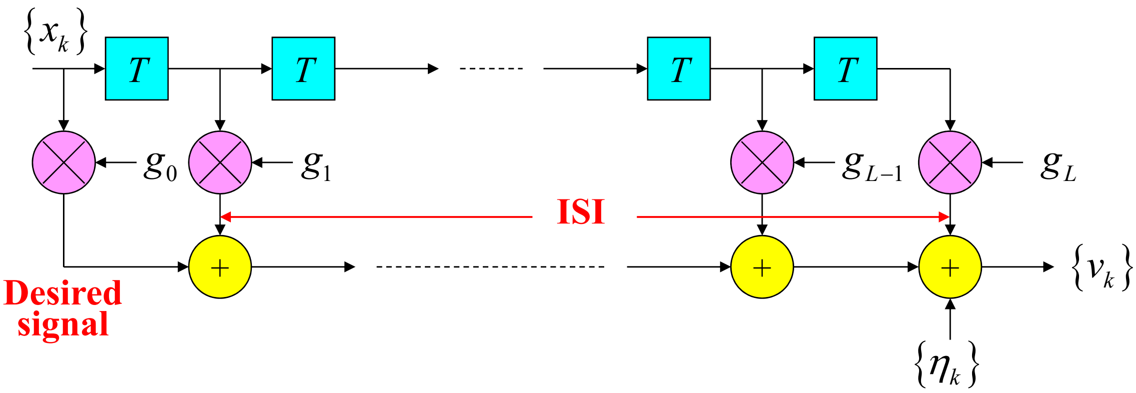
Modeling of ISI Channels

- The multipath channel can be represented by a **discrete-time transversal filter** with coefficients:

$$\mathbf{g} = \left[g_0, g_1, \cdots, g_{L-1}, g_L\right]^T$$

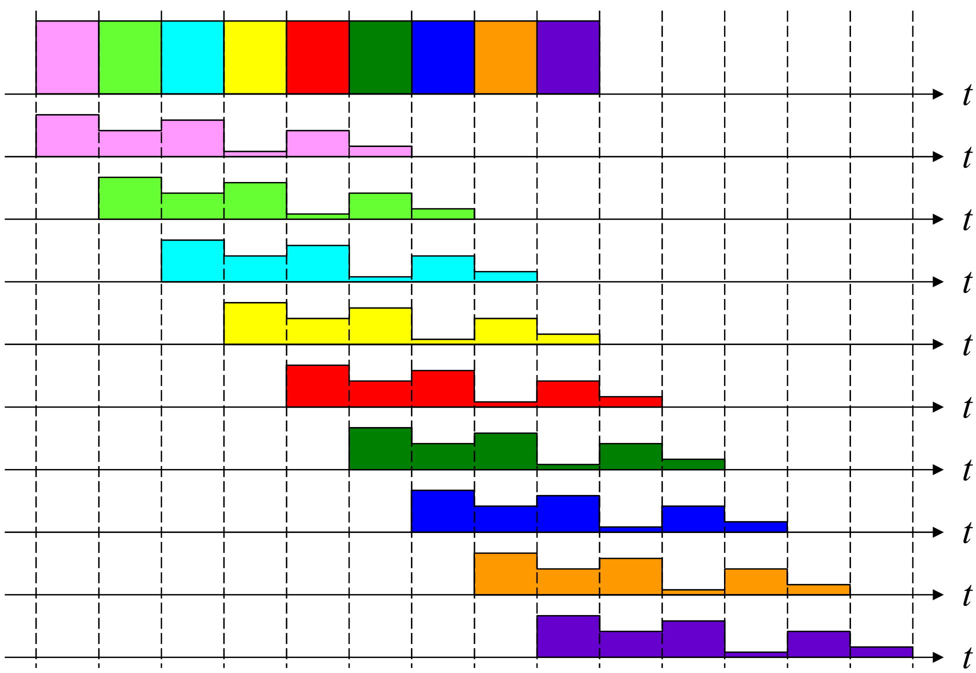
- If the channel gain is time-variant, the coefficients become

$$\mathbf{g}(k) = \left[g_0(k), g_1(k), \cdots, g_{L-1}(k), g_L(k)\right]^T$$



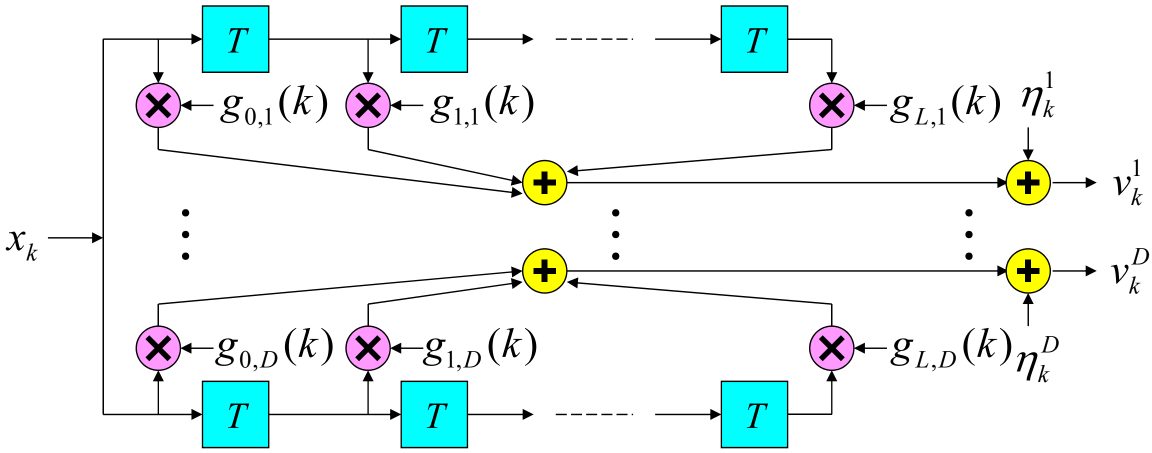
Modeling of ISI Channels

- Each input symbol triggers a channel impulse response in time

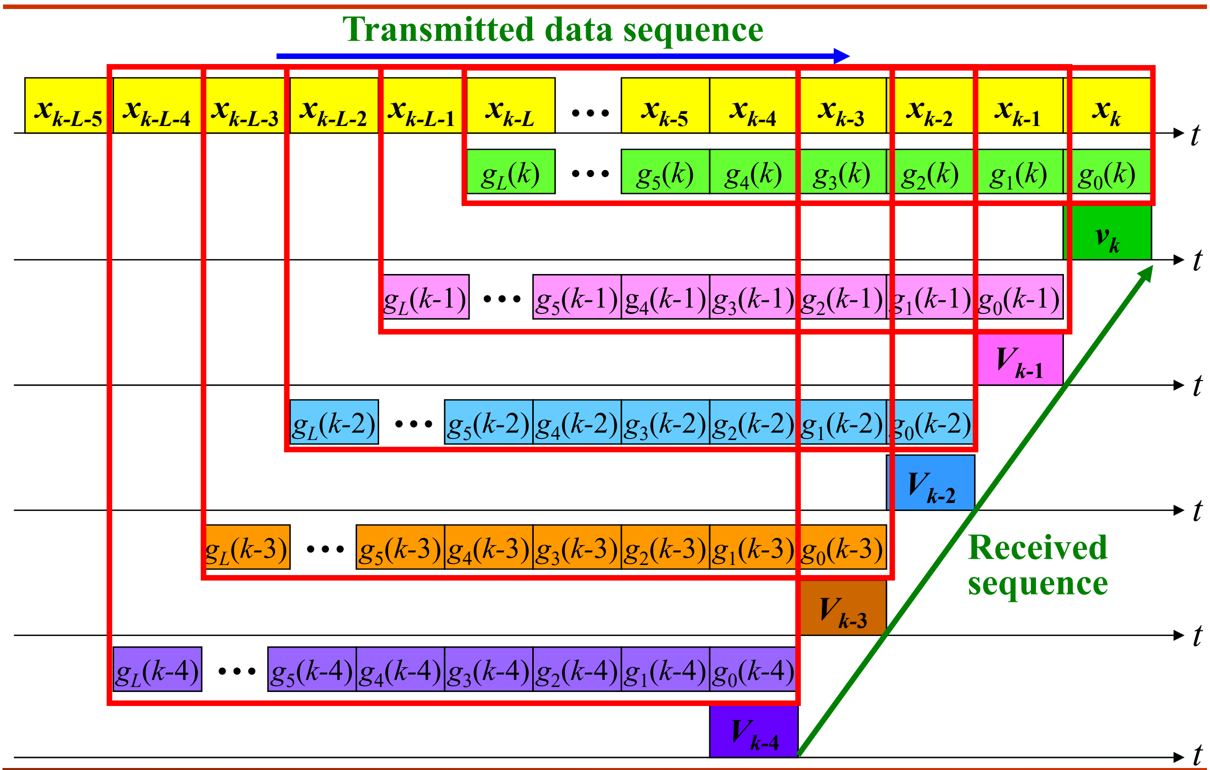


MLSE and the Viterbi Algorithm

- **MLSE: maximum likelihood sequence estimation**
- The overall channel can be modeled by a collection of D transversal filters that are T-spaced and have L+1 taps
  - D is the number of **diversity branches**
  - L is the **delay spread**



MLSE and the Viterbi Algorithm



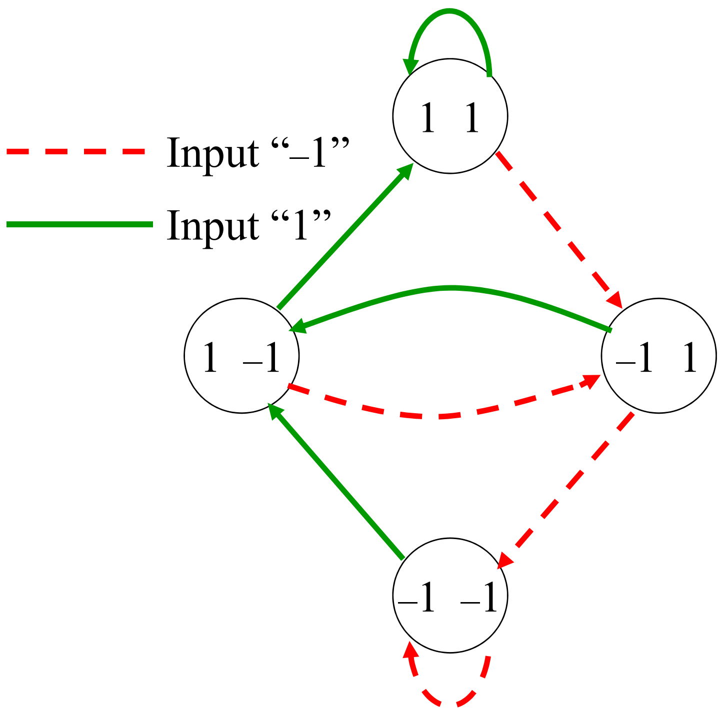
MLSE and the Viterbi Algorithm

- The channel has a finite number of states
  - If the size of the signal constellation is 2^n (M-ary modulation)
  - There are a total of $N_s = 2^{nL}$ states (L shift registers)
- The state at epoch k is:

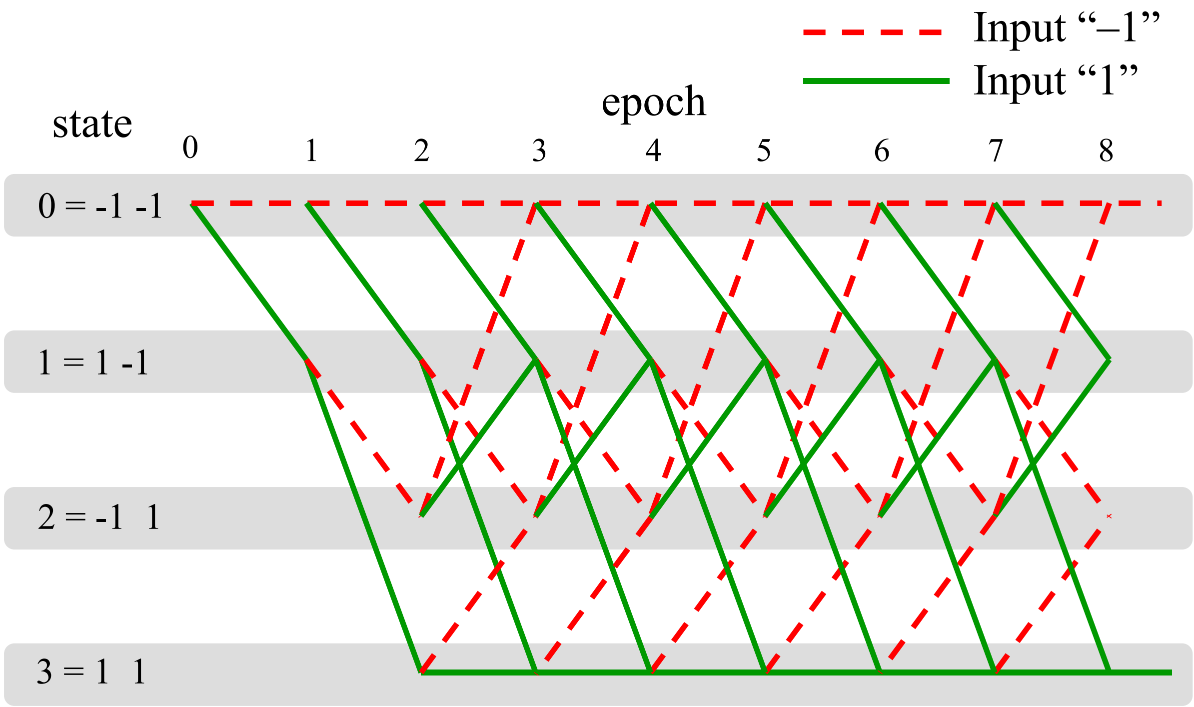
$$S_k = (x_{k-1}, x_{k-2}, \cdots, x_{k-L})$$

- For example: the binary sequence **x**, $x_n \in \{-1, +1\}$, is transmitted over a three-tap static ISI channel with channel vector **g** =(1,1,1)
- There are four states (3-tap ⇒ 2 shift registers ⇒ $2^2 = 4$ states)
- The system can be described by the **state diagram**
- The system state diagram can be used to construct the **trellis diagram**
  - The initial zero state is assumed to be $S_0^{(0)} = (-1, -1)$

State Diagram for Three-Tap ISI Channel



Trellis Diagram for Three-Tap ISI Channel



MLSE and the Viterbi Algorithm

- Let the vector of signals received on all diversity branches at epoch n be

$$\mathbf{V}_n = (v_{n,1}, v_{n,2}, \ldots, v_{n,D})$$

- After receiving the sequence $\{\mathbf{V}_n\}_{n=1}^{k}$, the ML receiver decides in favor of the sequence $\{x_n\}_{n=1}^{k}$ that maximizes the **likelihood function** or **log-likelihood function**

$$p(\mathbf{V}_k, \ldots, \mathbf{V}_1 \mid x_k, \ldots, x_1), \quad \log p(\mathbf{V}_k, \ldots, \mathbf{V}_1 \mid x_k, \ldots, x_1)$$

- Possible input data sequences:

$$\mathbf{x}_1 = \left(x_{k,1}, x_{k-1,1}, \ldots, x_{2,1}, x_{1,1}\right) \Rightarrow p_1 = p(\mathbf{V}_k, \ldots, \mathbf{V}_1 \mid \mathbf{x}_1)$$
$$\mathbf{x}_2 = \left(x_{k,2}, x_{k-1,2}, \ldots, x_{2,2}, x_{1,2}\right) \Rightarrow p_2 = p(\mathbf{V}_k, \ldots, \mathbf{V}_1 \mid \mathbf{x}_2)$$
$$\vdots$$
$$\mathbf{x}_{M^k} = \left(x_{k,M^k}, x_{k-1,M^k}, \ldots, x_{2,M^k}, x_{1,M^k}\right) \Rightarrow p_{M^k} = p(\mathbf{V}_k, \ldots, \mathbf{V}_1 \mid \mathbf{x}_{M^k})$$

MLSE and the Viterbi Algorithm

- Find the maximum value of p_j (or log p_j)

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}_j, 1 \le j \le M^k} \left\{ p(\mathbf{V}_k, \cdots, \mathbf{V}_1 \mid \mathbf{x}_j) \right\}$$

- Since $\mathbf{V}_n$ depends only on the L most recent transmitted symbols, the log-likelihood function at epoch k is denoted as

$$\log p(\mathbf{V}_k, \cdots, \mathbf{V}_1 \mid x_k, \cdots, x_1)$$
$$= \log p(\mathbf{V}_k \mid x_k, \cdots, x_{k-L}) + \underline{\log p(\mathbf{V}_{k-1}, \cdots, \mathbf{V}_1 \mid x_{k-1}, \cdots, x_1)}$$

  – where $x_{k-L} = 0$ for $k - L \le 0$

- Assume that the second term of the right-hand side has been calculated previously at epoch $k - 1$
  – Only the first term **(branch metric)** has to be computed for each incoming signal vector $\mathbf{V}_k$ at epoch k

MLSE and the Viterbi Algorithm

$$\oplus \max\left[\log p(\mathbf{V}_1 \mid x_1)\right]$$
$$\oplus \max\left[\log p(\mathbf{V}_2, \mathbf{V}_1 \mid x_2, x_1)\right]$$
$$\vdots$$
$$\oplus \max\left[\log p(\mathbf{V}_{L+1}, \cdots, \mathbf{V}_1 \mid x_{L+1}, \cdots, x_1)\right]$$
$$\oplus \max\left[\log p(\mathbf{V}_{L+2}, \cdots, \mathbf{V}_1 \mid x_{L+2}, \cdots, x_1)\right]$$
$$= \max\left[\underbrace{\log p(\mathbf{V}_{L+2} \mid x_{L+2}, \cdots, x_2)}_{\text{New item to be calculated}} + \underbrace{\log p(\mathbf{V}_{L+1}, \cdots, \mathbf{V}_1 \mid x_{L+1}, \cdots, x_1)}_{\text{Has been calculated}} \right]$$
$$\vdots$$
$$\oplus \max\left[\log p(\mathbf{V}_k, \cdots, \mathbf{V}_1 \mid x_k, \cdots, x_1)\right]$$
$$= \max\left[\underbrace{\log p(\mathbf{V}_k \mid x_k, \cdots, x_{k-L})}_{\text{New item to be calculated}} + \underbrace{\log p(\mathbf{V}_{k-1}, \cdots, \mathbf{V}_1 \mid x_{k-1}, \cdots, x_1)}_{\text{Has been calculated}} \right]$$

MLSE and the Viterbi Algorithm

- The conditional pdf of the received signal vector:

$$p(\mathbf{V}_k \mid x_k, \cdots, x_{k-L}) = \frac{1}{(\pi N_0)^D} \exp\left\{-\frac{1}{N_0}\sum_{d=1}^{D}\left|v_{k,d} - \sum_{i=0}^{L} g_{i,d}x_{k-i}\right|^2\right\}$$

- For log $p(\mathbf{V}_k \mid x_k, \ldots, x_{k-L})$, the **branch metric** is

$$\mu_k = -\sum_{d=1}^{D}\left|v_{k,d} - \hat{v}_{k,d}(x_k, \cdots, x_{k-i})\right|^2 = -\sum_{d=1}^{D}\left|v_{k,d} - \sum_{i=0}^{L} g_{i,d}x_{k-i}\right|^2$$

- The **Viterbi algorithm** can be used to implement the ML receiver by searching through the N_s–state trellis for most likely transmitted sequence **x**
- This search process is called maximum likelihood sequence estimation (**MLSE**)

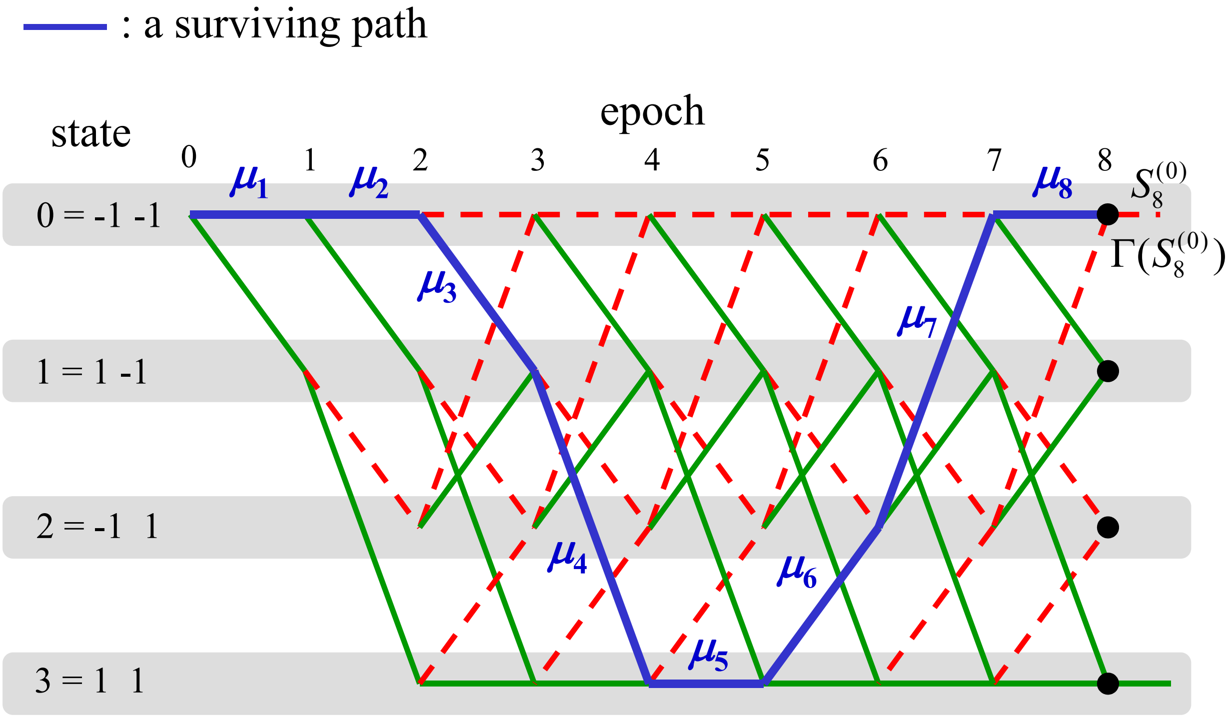
MLSE and the Viterbi Algorithm

- At epoch k, assume that the algorithm has stored N_s **surviving sequences** $\breve{\mathbf{x}}(S_k^{(i)})$ (paths through the trellis)
  - with associated **path metrics** $\Gamma(S_k^{(i)})$ (distances from the received sequence)

$$\Gamma(S_k^{(i)}) = \sum_{\{k\}} \mu_k$$

  - $\{\mu_k\}$ is the set of **branch metrics** along the surviving path $\breve{\mathbf{x}}(S_k^{(i)})$

Trellis Diagram for Three-Tap ISI Channel



MLSE and the Viterbi Algorithm

- After the vector $\mathbf{V}_k$ has been received, the Viterbi algorithm executes the following steps for each state $S_{k+1}^{(j)}, j = 0, \ldots, N_s - 1$
  - Compute the set of path metrics for all possible paths through the trellis that terminate in state $S_{k+1}^{(j)}$

$$\Gamma(S_k^{(i)} \to S_{k+1}^{(j)}) = \Gamma(S_k^{(i)}) + \underline{\mu(S_k^{(i)} \to S_{k+1}^{(j)})}$$

  - Find $\Gamma(S_{k+1}^{(j)}) = \max_i \Gamma(S_k^{(i)} \to S_{k+1}^{(j)})$, where the maximization is over all possible paths through the trellis that terminate in $S_{k+1}^{(j)}$
  - Store $\Gamma(S_{k+1}^{(j)})$ and its associated **surviving sequence** $\breve{\mathbf{x}}(S_{k+1}^{(j)})$, and **drop all other paths**
- $\mu(S_k^{(i)} \to S_{k+1}^{(j)})$ is the branch metric associated with the transition $S_k^{(i)} \to S_{k+1}^{(j)}$ and is computed as

Symbols in the surviving sequence

$$\mu(S_k^{(i)} \to S_{k+1}^{(j)}) = -\sum_{d=1}^{D} \left| v_{k,d} - g_{0,d} x_k(S_k^{(i)} \to S_{k+1}^{(j)}) - \sum_{m=1}^{L} g_{m,d} \underline{x_{k-m}(S_k^{(i)})} \right|^2$$

MLSE and the Viterbi Algorithm

- $x_k(S_k^{(i)} \to S_{k+1}^{(j)})$ is a symbol that is uniquely determined by the transition $S_k^{(i)} \to S_{k+1}^{(j)}$, and the L most recent symbols $\{x_{k-m}(S_k^{(i)})\}_{m=1}^{L}$ are uniquely specified by the previous states $S_k^{(i)}$
- When two paths merged into a state, **why only the path with a larger path metric is stored** (as the **surviving path**)?
- Consider the two sequences s1="1001100" and s2="0101000" merged at State 0, with the path metric M1 > M2
  - Assume that the subsequent sequence with the maximum path metric is "…101101"
  - Then, the metrics "1001100101101" > "0101000101101" > "0101000XXXXXX" $\Rightarrow$ So, the sequence s2 can be discarded
- After all states have been processed, the time index k is **incremented** and the whole algorithm repeats

MLSE and the Viterbi Algorithm

- **Theoretically**, the ML receiver waits until the **entire sequence** $\{\mathbf{V}_n\}_{n=1}^{\infty}$ has been received before making a decision
- **In practice**, a decision about x_{k-Q} is usually made when $\mathbf{V}_k$ is received and processed
  - If $\boldsymbol{Q > 5L}$, the performance degradation caused by the resulting **path metric truncation** is negligible
- The **decision delay** shall be $\boldsymbol{Q > 5L}$
  - Consider a sequence currently with the maximum path metric

…10011000100110111110101

Bits that the receiver can make a decision ← Decision delay $Q > 5L$ → Current symbol

Cumulative Path Metrics

- Suppose that the data sequence **x** = (–1, 1, 1, –1, 1, 1, –1, –1,…) is transmitted
- The noiseless received sequence is **v** = (v_0, v_1, v_2, v_3, v_4, …)

$$\begin{aligned} v_n &= g_0 x_n + g_1 x_{n-1} + g_2 x_{n-2} \\ &= x_n + x_{n-1} + x_{n-2} \end{aligned}$$

$$\boxed{\begin{gathered} \mathbf{g} = (1, 1, 1) \\ S_0^{(0)} = (-1, -1) \end{gathered}}$$

- Then the noiseless received sequence is

$$\mathbf{v} = \left(-3, -1, 1, 1, 1, 1, 1, -1, \cdots\right)$$

  – where the previous two symbols are assumed to be (–1, –1)
- Assume that the noisy received sequence is

$$\mathbf{v} = \left(-3.2, -1.1, 0.9, 0.1, 1.2, 1.5, 0.7, -1.3, \cdots\right)$$

- The path metrics are equal to the square Euclidean distance between surviving sequence $\breve{\mathbf{x}}(S_k^{(i)})$ and received sequence **v**

Cumulative Path Metrics

$$\Gamma(S_0^{(0)}) = 0; \Gamma(S_0^{(1)}) = 0; \Gamma(S_0^{(2)}) = 0; \Gamma(S_0^{(3)}) = 0;$$
$$S_k^{(0)} \to S_{k+1}^{(0)} \Rightarrow v_{k+1} = -3; S_k^{(0)} \to S_{k+1}^{(1)} \Rightarrow v_{k+1} = -1;$$
$$S_k^{(1)} \to S_{k+1}^{(2)} \Rightarrow v_{k+1} = -1; S_k^{(1)} \to S_{k+1}^{(3)} \Rightarrow v_{k+1} = +1;$$
$$S_k^{(2)} \to S_{k+1}^{(0)} \Rightarrow v_{k+1} = -1; S_k^{(2)} \to S_{k+1}^{(1)} \Rightarrow v_{k+1} = +1;$$
$$S_k^{(3)} \to S_{k+1}^{(2)} \Rightarrow v_{k+1} = +1; S_k^{(3)} \to S_{k+1}^{(3)} \Rightarrow v_{k+1} = +3;$$

- Epoch 1

$$\mu(S_0^{(0)} \to S_1^{(0)}) = -\left|-3.2 - (-3)\right|^2 = -0.04; \quad \Rightarrow \Gamma(S_0^{(0)} \to S_1^{(0)}) = -0.04$$
$$\mu(S_0^{(0)} \to S_1^{(1)}) = -\left|-3.2 - (-1)\right|^2 = -4.84; \quad \Rightarrow \Gamma(S_0^{(0)} \to S_1^{(1)}) = -4.84$$

- Epoch 2

$$\mu(S_1^{(0)} \to S_2^{(0)}) = -\left|-1.1 - (-3)\right|^2 = -3.61; \Gamma(S_1^{(0)} \to S_2^{(0)}) = -0.04 - 3.61 = -3.65$$
$$\mu(S_1^{(0)} \to S_2^{(1)}) = -\left|-1.1 - (-1)\right|^2 = -0.01; \Gamma(S_1^{(0)} \to S_2^{(1)}) = -0.04 - 0.01 = -0.05$$
$$\mu(S_1^{(1)} \to S_2^{(2)}) = -\left|-1.1 - (-1)\right|^2 = -0.01; \Gamma(S_1^{(1)} \to S_2^{(2)}) = -4.84 - 0.01 = -4.85$$
$$\mu(S_1^{(1)} \to S_2^{(3)}) = -\left|-1.1 - (+1)\right|^2 = -4.41; \Gamma(S_1^{(1)} \to S_2^{(3)}) = -4.84 - 4.41 = -9.25$$

Cumulative Path Metrics

- Epoch 3

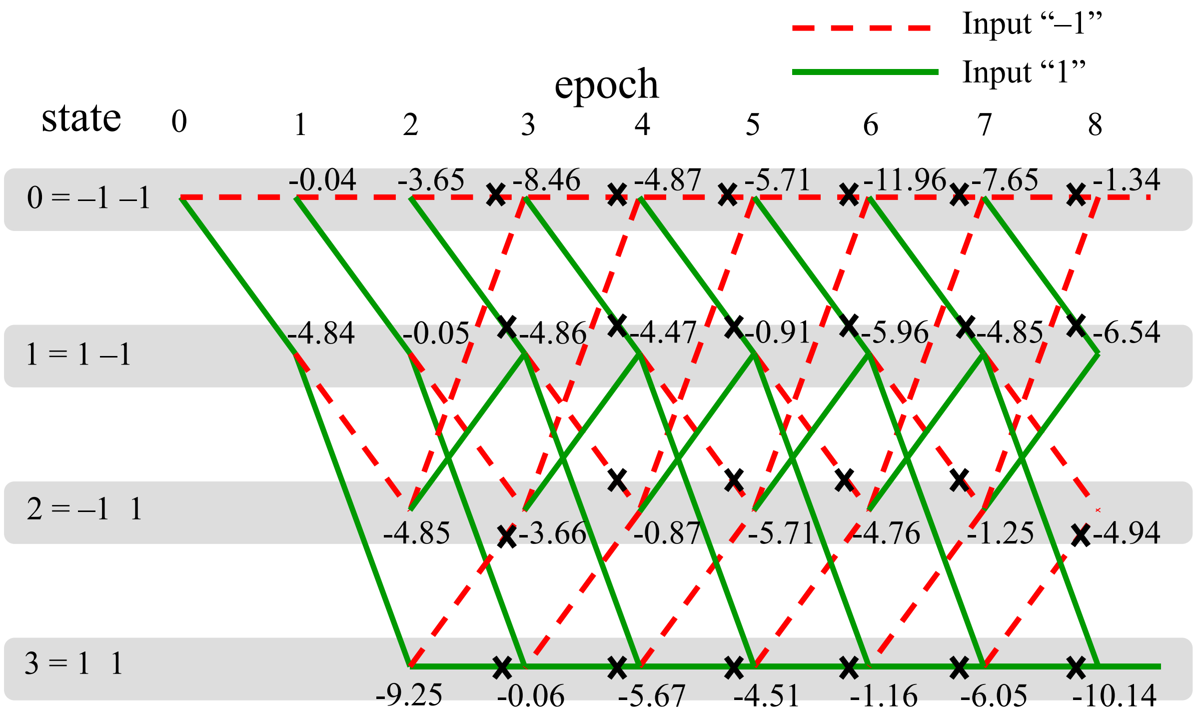
$$\mu(S_2^{(0)} \to S_3^{(0)}) = -\left|0.9-(-3)\right|^2 = -15.21; \Gamma(S_2^{(0)} \to S_3^{(0)}) = -3.65-15.21 = -18.86 \otimes$$
$$\mu(S_2^{(2)} \to S_3^{(0)}) = -\left|0.9-(-1)\right|^2 = -3.61; \Gamma(S_2^{(2)} \to S_3^{(0)}) = -4.85-3.61 = -8.46$$
$$\mu(S_2^{(0)} \to S_3^{(1)}) = -\left|0.9-(-1)\right|^2 = -3.61; \Gamma(S_2^{(0)} \to S_3^{(1)}) = -3.65-3.61 = -7.26 \otimes$$
$$\mu(S_2^{(2)} \to S_3^{(1)}) = -\left|0.9-(+1)\right|^2 = -0.01; \Gamma(S_2^{(2)} \to S_3^{(1)}) = -4.85-0.01 = -4.86$$
$$\mu(S_2^{(1)} \to S_3^{(2)}) = -\left|0.9-(-1)\right|^2 = -3.61; \Gamma(S_2^{(1)} \to S_3^{(2)}) = -0.05-3.61 = -3.66$$
$$\mu(S_2^{(3)} \to S_3^{(2)}) = -\left|0.9-(+1)\right|^2 = -0.01; \Gamma(S_2^{(3)} \to S_3^{(2)}) = -9.25-0.01 = -9.26 \otimes$$
$$\mu(S_2^{(1)} \to S_3^{(3)}) = -\left|0.9-(+1)\right|^2 = -0.01; \Gamma(S_2^{(1)} \to S_3^{(3)}) = -0.05-0.01 = -0.06$$
$$\mu(S_2^{(3)} \to S_3^{(3)}) = -\left|0.9-(+3)\right|^2 = -4.41; \Gamma(S_2^{(3)} \to S_3^{(3)}) = -9.25-4.41 = -13.66 \otimes$$

Cumulative Path Metrics

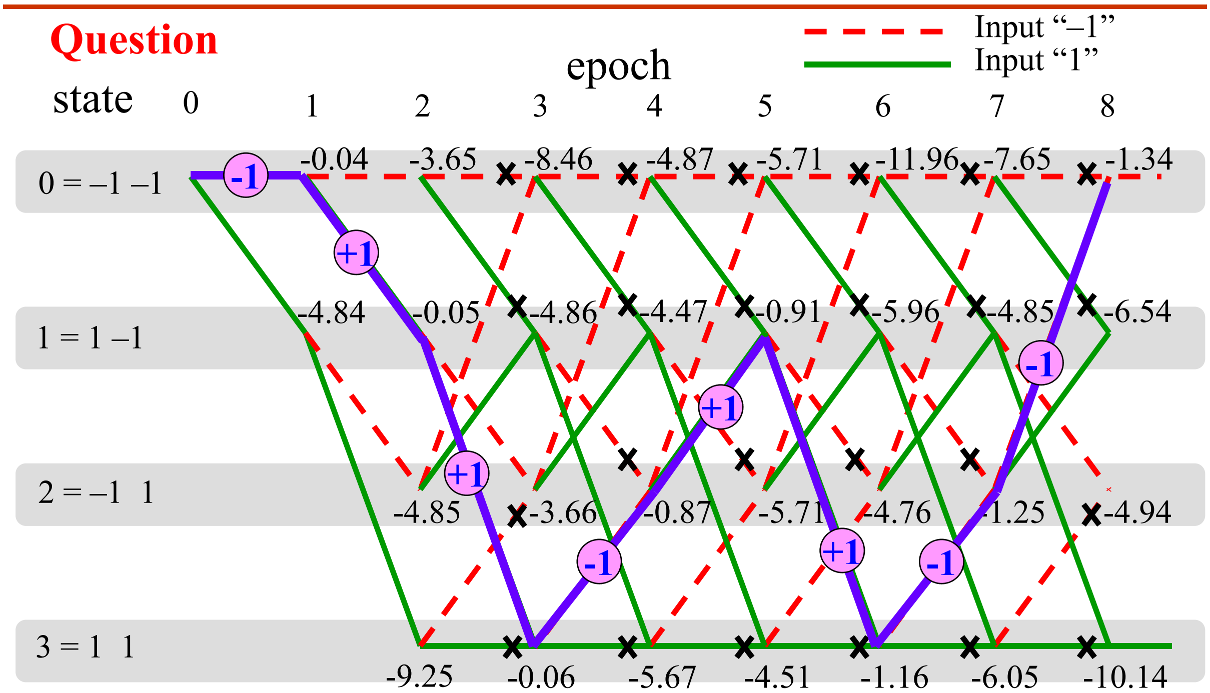
- Epoch 4

$$\mu(S_3^{(0)} \to S_4^{(0)}) = -\left|0.1-(-3)\right|^2 = -9.61; \Gamma(S_3^{(0)} \to S_4^{(0)}) = -8.46-9.61 = -18.07 \otimes$$
$$\mu(S_3^{(2)} \to S_4^{(0)}) = -\left|0.1-(-1)\right|^2 = -1.21; \Gamma(S_3^{(2)} \to S_4^{(0)}) = -3.66-1.21 = -4.87$$
$$\mu(S_3^{(0)} \to S_4^{(1)}) = -\left|0.1-(-1)\right|^2 = -1.21; \Gamma(S_3^{(0)} \to S_4^{(1)}) = -8.46-1.21 = -9.67 \otimes$$
$$\mu(S_3^{(2)} \to S_4^{(1)}) = -\left|0.1-(+1)\right|^2 = -0.81; \Gamma(S_3^{(2)} \to S_4^{(1)}) = -3.66-0.81 = -4.47$$
$$\mu(S_3^{(1)} \to S_4^{(2)}) = -\left|0.1-(-1)\right|^2 = -1.21; \Gamma(S_3^{(1)} \to S_4^{(2)}) = -4.86-1.21 = -6.07 \otimes$$
$$\mu(S_3^{(3)} \to S_4^{(2)}) = -\left|0.1-(+1)\right|^2 = -0.81; \Gamma(S_3^{(3)} \to S_4^{(2)}) = -0.06-0.81 = -0.87$$
$$\mu(S_3^{(1)} \to S_4^{(3)}) = -\left|0.1-(+1)\right|^2 = -0.81; \Gamma(S_3^{(1)} \to S_4^{(3)}) = -4.86-0.81 = -5.67$$
$$\mu(S_3^{(3)} \to S_4^{(3)}) = -\left|0.1-(+3)\right|^2 = -8.41; \Gamma(S_3^{(3)} \to S_4^{(3)}) = -0.06-8.41 = -8.47 \otimes$$

Cumulative Path Metrics



Trace Back



$$\hat{\mathbf{x}} = (-1, 1, 1, -1, 1, 1, -1, -1, \cdots) = \mathbf{x}$$

Adaptive MLSE Receiver

- The Viterbi algorithm requires knowledge of the channel vectors $\mathbf{g}_d$ (channel gain of each diversity branches) to compute the branch metrics
  - An **adaptive channel estimator** is needed
  - A transversal digital filter with LMS algorithm is used
- With the **LMS (Least-Mean-Square)** algorithm, the tap coefficients are updated by

Only $\cdots, \hat{x}^*_{k-Q-1}, \hat{x}^*_{k-Q}$, **are available**

$$\hat{g}_{i,d}(k+1) = \hat{g}_{i,d}(k) + \alpha \underline{\varepsilon_{k-Q,d}}\, \underline{\hat{x}^*_{k-i-Q}}, \quad i = 0, \cdots, L;\ d = 1, \cdots, D$$

  - α is the adaptation step size
  - $\varepsilon_{k\text{-}Q,d}$ is the error associated with branch d, defined as

$$\varepsilon_{k-Q,d} = v_{k-Q,d} - \sum_{i=0}^{L} \hat{g}_{i,d}(k)\hat{x}_{k-i-Q}$$

Adaptive MLSE Receiver

- A major problem with this channel estimator is:
  - It lags behind the true channel vector by the **decision delay *Q*** (in the Viterbi algorithm)
- Actually, the received signal at epoch $k - Q$ is

$$v_{k-Q,d} = \sum_{i=0}^{L} g_{i,d}(k-Q)x_{k-i-Q} + \eta_{k-Q,d}$$

  - The signal $v_{k\text{-}Q,d}$ is influenced by $g_{i,d}(k - Q)$, not $g_{i,d}(k)$
- The error:

$$\varepsilon_{k-Q,d} = \sum_{i=0}^{L} \underline{\left(g_{i,d}(k-Q) - \hat{g}_{i,d}(k)\right)} x_{k-i-Q} + \eta_{k-Q,d}$$

  - which depends on the difference of $g_{i,d}(k - Q)$ (at **epoch** $\boldsymbol{k - Q}$) and the estimated $\hat{g}_{i,d}(k)$ (at **epoch** $\boldsymbol{k}$)

Adaptive MLSE Receiver

- The channel time variations over the decision delay Q will cause the term $\{g_{i,d}(k-Q)-\hat{g}_{i,d}(k)\}\big|_{k=1}^{L}$ to be non-zero, and degrade the tracking performance
- The decision delay Q could be reduced but this will also **reduce the reliability of the decisions** $\hat{x}_{k-i-Q}$ that are used to update the channel estimates
- The overall performance improvement obtained by reducing Q is often **minimal**

Adaptive MLSE Receiver

- The **Per-survivor processing** approach:
  - **Each state** has its own channel estimator that tracks the channel
  - Multiple channel estimators are required
  - The tap coefficients are updated according to

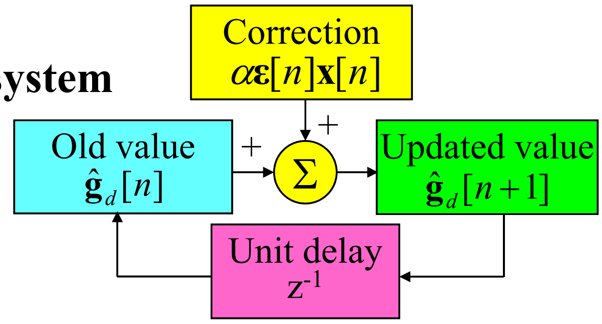
$$\hat{g}_{i,d}(k+1)=\hat{g}_{i,d}(k)+\alpha\varepsilon_{k,d}\,\breve{x}_{k-i}^{*},\quad i=0,\cdots,L;\, d=1,\cdots,D$$

  - where $\breve{\mathbf{x}}$ is the surviving sequence associated with **each state**
  - Each state uses **zero-delay symbols** $\Rightarrow$ good channel tracking performance for the **correct** data sequence

Adaptive MLSE Receiver

- For the **LMS algorithm**:
  - Initialize the algorithm by setting
    $$\hat{\mathbf{g}}_d[1] = [\hat{g}_{0,d}, \hat{g}_{1,d}, \cdots, \hat{g}_{L,d}] = \mathbf{0}$$
  - For n = 1, 2, …, it computes
    $$\hat{\mathbf{v}}_d[n] = \mathbf{x}^T[n]\,\hat{\mathbf{g}}_d[n], \quad \boldsymbol{\varepsilon}_d[n] = \mathbf{v}_d[n] - \hat{\mathbf{v}}_d[n]$$
    $$\hat{\mathbf{g}}_d[n+1] = \hat{\mathbf{g}}_d[n] + \alpha\boldsymbol{\varepsilon}_d[n]\,\mathbf{x}[n]$$
  - Continue the iterative computation **until the equalizer reaches a "steady state"**
- LMS algorithm is a **feedback system**

⇒ It is possible to **diverge**



Training Sequence

- The role of training sequence is to train the adaptive equalizer in **training mode**
- The training sequence is a **fixed-pattern (known)** sequence, for the adaptive equalizer to estimate the channel vector **g**
- For **TDMA** systems, the training sequence should be transmitted **in each slot**
- **Long** training sequences:
  - Good estimation of the channel
  - Wasting the spectrum resource
- **Short** training sequences:
  - Bad estimation of the channel ⇒ degrade the performance

Training Sequence

- The position of the training sequence could be in:
  - The **head** of a frame
  - The **middle** of a frame
  - The **tail** of a frame

Question

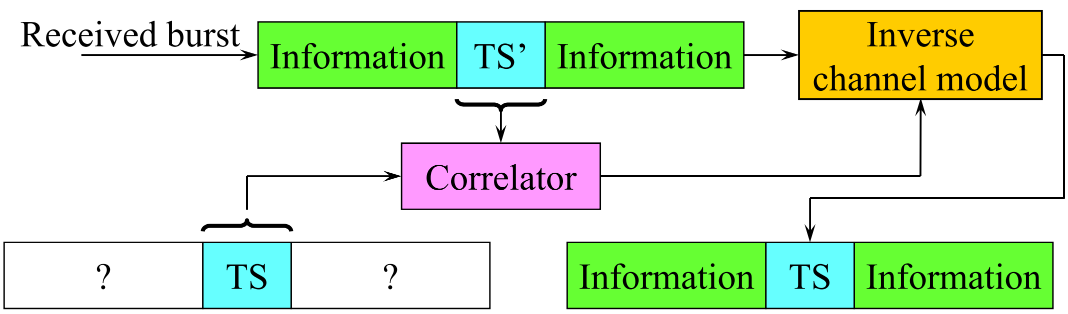
- For GSM systems, the training sequence is located in the **middle** of a burst
  - The length of the training sequence is 26 bits
  - There are a total of 8 training sequence codes (TSC)
- GSM systems use the maximum likelihood sequence estimation (MLSE) with Viterbi algorithm
- It reflects the maximum delay of **15 μs** (4 bits)

Normal Burst (148 Bits)

TB (3)	Information Bits Encrypted (57)	1	Training Sequence (26)	1	Information Bits Encrypted (57)	TB (3)

Training Sequence

TSC 0 (0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1)
TSC 1 (0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1)
TSC 2 (0, 1, 0, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0)
TSC 3 (0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0)
TSC 4 (0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1)
TSC 5 (0, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 1, 0)
TSC 6 (1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1)
TSC 7 (1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0)



Link Adaptation

Link Adaptation

- **Link adaptation** is a technique used to improve the transmission efficiency in wireless communications
  - Adapt the **modulation**, **coding** and other parameters to the transmission conditions (e.g., the propagation loss, the interference, the sensitivity of a receiver, the available transmit power margin, etc.).
- Link adaptation systems generally require some **channel state information (CSI)** at the transmitter.
  - For **TDD** systems, the transmitter can approximate the CSI of the **Tx-to-Rx link** as the CSI estimate of the **Rx-to-Tx link**
    - Under the reciprocal theorem and slow varying assumption
  - For **FDD** systems, the receiver needs to feed back the CSI estimate to the transmitter
    - It may only feed back the decoding results (Success/Failure) or the received signal strength (RSS)

Link Adaptation

- The most popular link adaptation scheme adapts only **modulation** and **coding**
- According to the **channel quality**, the transmitter jointly choose the modulation type and code rate for transmission
  - If the transmission condition is **favorable**, the transmitter uses high-order modulation and high-rate coding
    - To improve the throughput
  - If the transmission condition is **unfavorable**, the transmitter may downgrade the modulation order and/or the code rate
    - To ensure successful reception
- If high-order modulation and high-rate coding are used in an **unfavorable** transmission condition
  - The packet error rate will be very high and the throughput is still degraded

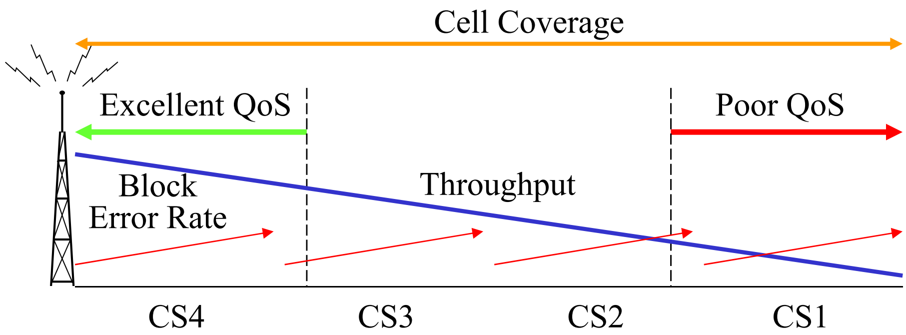
Link Adaptation (GPRS)

- For example, **GPRS (General Packet Radio Service)** in GSM systems uses a rate adaptation algorithm that adapts the **coding scheme (CS)** according to the quality of the radio channel
  - Modulation: GMSK (1 bit/symbol);
  - Code rate: 1.0, 3/4, 2/3, 1/2;
- The code rate adaptation is achieved by **bit puncturing**
  - The **same encoder** (code rate 1/2) is used for different code rates

Scheme	Code Rate	Punctured bits	Data rate (kbps)
CS-1	1/2	0	9.05
CS-2	2/3	132	13.4
CS-3	3/4	220	15.6
CS-4	1	-	21.4

Link Adaptation (GPRS)

- When a receiver is close to the BS (with an **excellent channel quality**), high-rate transmission (e.g., CS4) is used
  - If low-rate transmission is used, throughput will be degraded
- When a receiver is at the cell border (with a **poor channel quality**), low-rate transmission (e.g., CS1) is used
  - If high-rate transmission (e.g., CS4) is used, throughput is also degraded because of **a very high packet error probability**



Link Adaptation (EDGE)

- For example, **EDGE (Enhanced Data rates for GSM Evolution)** uses a rate adaptation algorithm that adapts the **modulation and coding scheme (MCS)** according to the quality of the radio channel
  - Modulation: GMSK (1 bit/symbol); 8PSK (3 bits/symbol);
  - Code rate: 1.0, 0.8, 0.66, 0.53 (GMSK); 1.0, 0.92, 0.76, 0.49, 0.37 (8PSK);

Scheme	**Code rate**	**Modulation**	**Data rate (kb/s)**	**Scheme**	**Code rate**	**Modulation**	**Data rate (kb/s)**
MCS-4	1.0	GMSK	17.6	MCS-9	1.0	8PSK	59.2
MCS-3	0.80		14.8/13.6	MCS-8	0.92		54.4
MCS-2	0.66		11.2	MCS-7	0.76		44.8
MCS-1	0.53		8.8	MCS-6	0.49		29.6/27.2
				MCS-5	0.37		22.4

Link Adaptation (HSDPA)

- **HSDPA (High-Speed Downlink Packet Access)** in 3G CDMA systems uses a rate adaptation algorithm that adapts the modulation and coding scheme, known as the **AMC (Adaptive Modulation and Coding)** scheme, according to the quality of the radio channel
  - Modulation: QPSK (2 bits/symbol); 8-PSK (3 bits/symbol); 16-QAM (4 bits/symbol); 64-QAM (6 bits/symbol);
  - Code rate: 3/4, 1/2, 1/4 (QPSK); 3/4 (8-PSK); 3/4, 1/2 (16-QAM); 3/4 (64-QAM);
- Traditionally, CDMA systems have used **fast power control** as the preferred method for link adaptation.
  - Control the transmission power, not control the transmission rate
- The AMC scheme has offered an alternative link adaptation method that promises to raise the **overall system capacity**.

Link Adaptation (HSDPA)

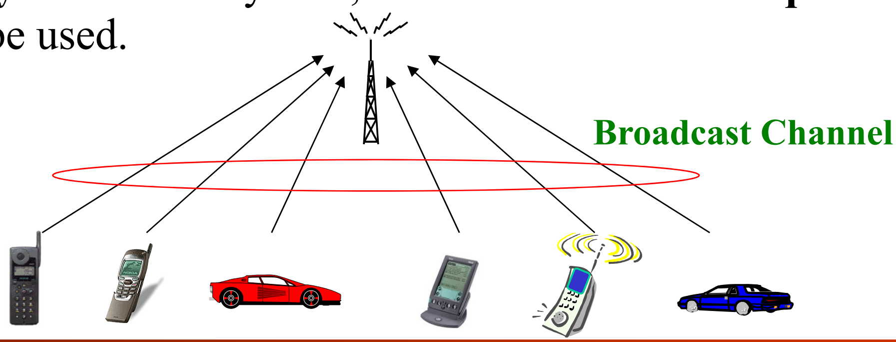
- With AMC, the power of the transmitted signal is held **constant** over a **frame interval**
  - The modulation and coding format is changed to match the current received signal quality or channel conditions.

Scheme	**Modulation**	**Code Rate**	**Data rate (Mbps)**
MCS1	QPSK	1/4	1.2
MCS2	QPSK	1/2	2.4
MCS3	QPSK	3/4	3.6
MCS4	8-PSK	3/4	5.4
MCS5	16-QAM	1/2	4.8
MCS6	16-QAM	3/4	7.2
MCS7	64-QAM	3/4	10.8

Random Access Techniques

Random Access

- In a **broadcast channel** (such as a wireless channel), one of the key issue is to determine who gets the right of using the channel when there is **competition** for it.
- In a **synchronous** system (with a **controller**), the **controlled-access** (multiple access) **techniques**, such as TDMA or CDMA, can be applied to prevent/reduce **signal collision** or **mutual interference**.
- In an **asynchronous** system, **random access techniques** should be used.

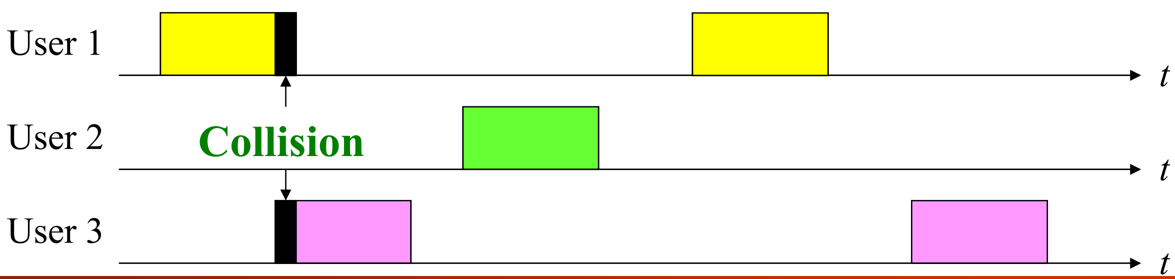


Random Access Schemes

- Random access techniques are completely **decentralized**
- Each user determines its transmission depends solely on a **local random mechanism** and the **local observation**
- It is possible that two or more users decide to transmit their signals in an **overlapping** time interval
  - It is called as a **collision**
- Several random access schemes have been proposed for wireless communications
  - **Pure ALOHA**
  - **Slotted ALOHA**
  - **CSMA/CD**
  - **MACA**

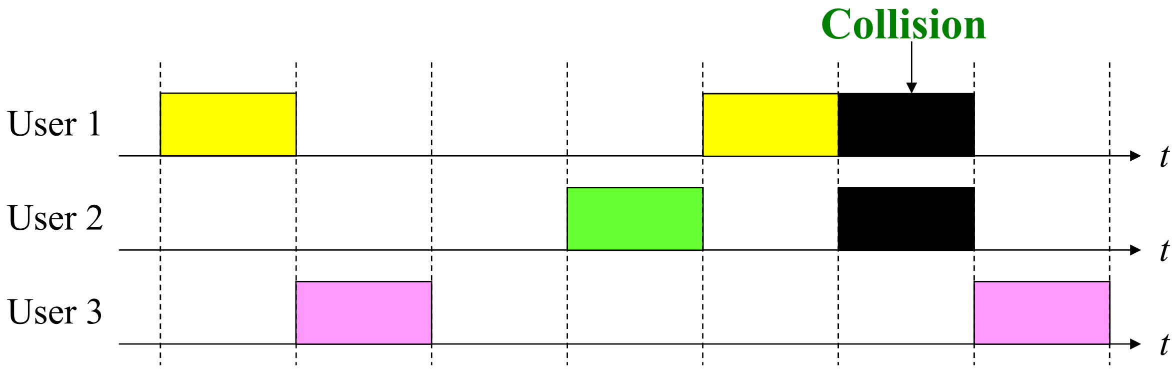
Pure ALOHA

- Pure ALOHA is the **simplest** random access scheme
- Each user transmits the data packet **immediately** whenever it has a packet for transmission
- There is a feedback mechanism to acknowledge the **success/failure** of the transmitted packet
  - Such as the **positive acknowledgement with timeout**
- When a collision occurs, the user will **re-transmit** the packet after a **random backoff time**



Slotted ALOHA

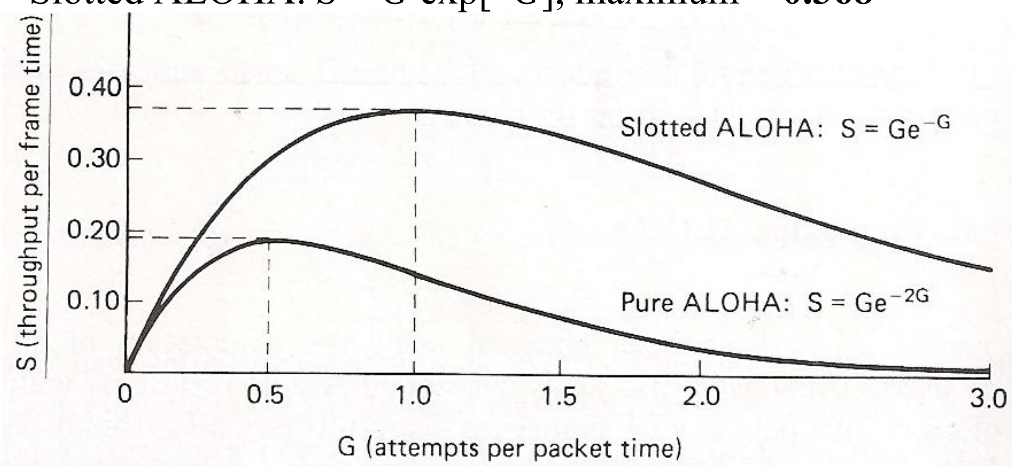
- In slotted ALOHA, the time scale is slotted into **multiple time slots** with the duration equal to the packet length
- Users can transmit data packets at the **beginning** of each time slot only
- The other processes are the same as the pure ALOHA scheme



ALOHA – Average Throughput

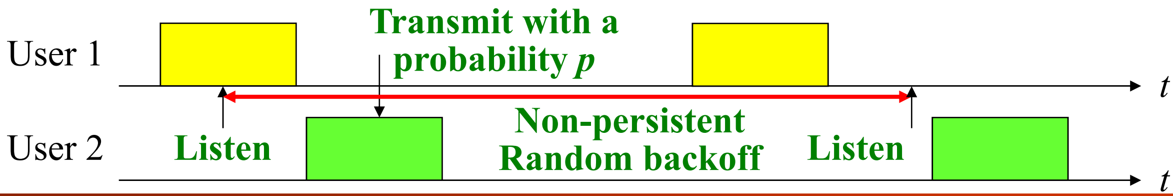
- Denote G as the average **traffic load** (in attempts per packet time) of the channel
- The average throughput (in packets per packet time) is
  - Pure ALOHA: $S = G\exp[-2G]$, maximum = **0.184**
  - Slotted ALOHA: $S = G\exp[-G]$, maximum = **0.368**

inefficient



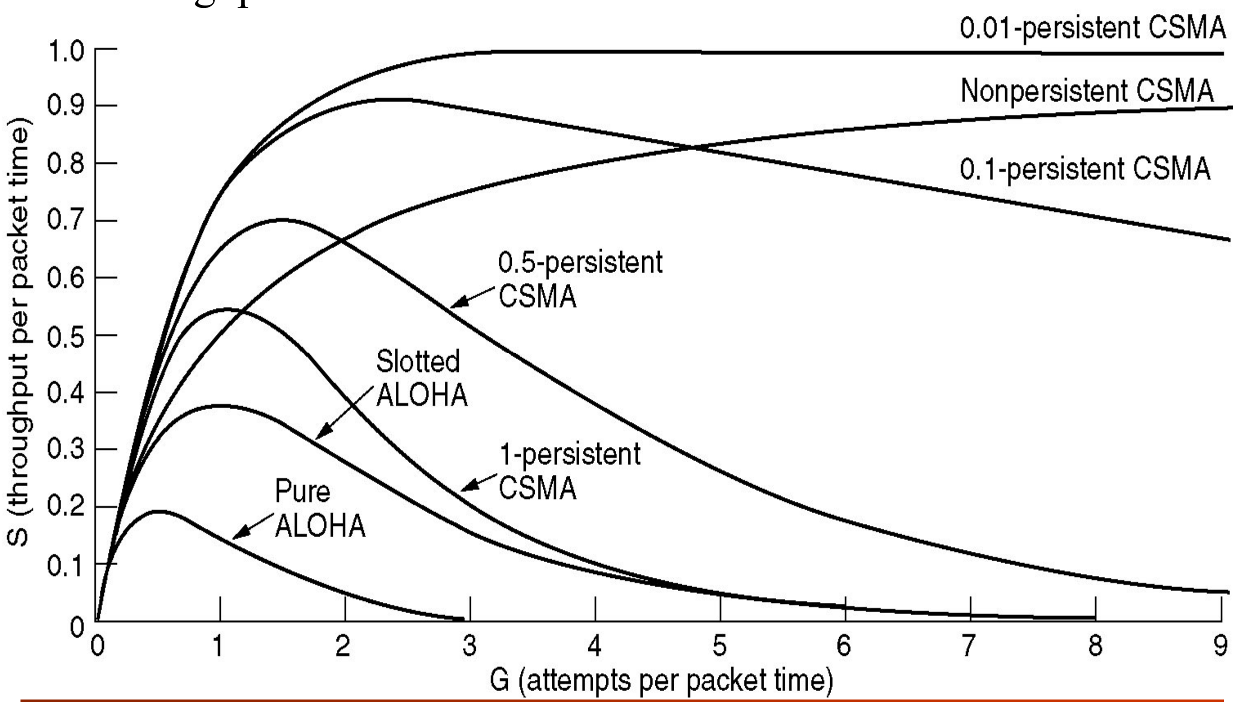
CSMA

- In **Carrier Sense Multiple Access (CSMA)** scheme, each node listens to a carrier and act accordingly.
- When a node has a data packet to send, it listens to the channel to see if there is anyone else under transmission
  - ***p*-persistent**: If there is any, it continually senses and waits for the channel to go idle and then transmits **with a probability $0 \le p \le 1$**, and defers (a random backoff) with probability $q = 1 - p$
  - **Non-persistent**: If there is any, it defers a random backoff
- When a collision occurs, the user **re-transmits** the packet after a **random backoff time**



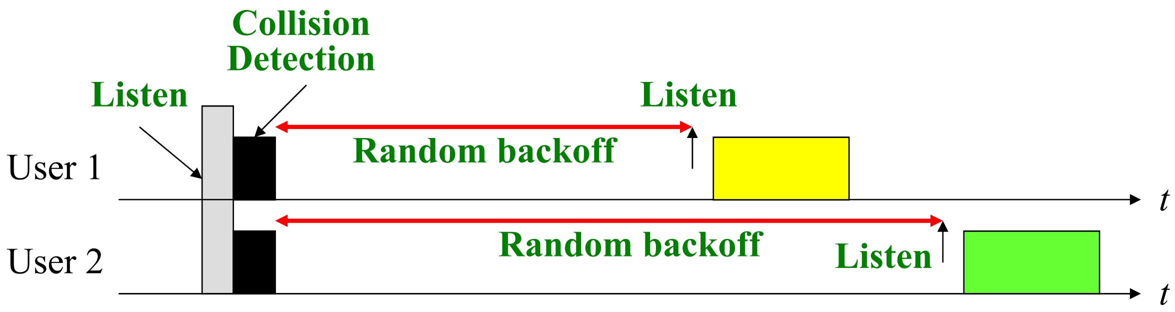
CSMA – Average Throughput

- Throughput versus Traffic load



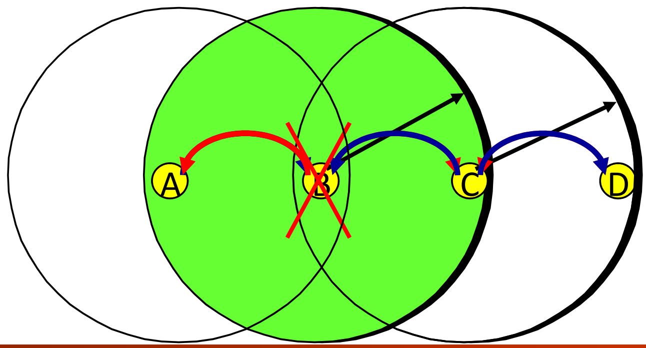
CSMA/CD

- **Carrier Sense Multiple Access with Collision Detection (CSMA/CD)**: based on **CSMA**, collision detection is added for performance improvement
- The stations **abort** their transmission as soon as they **detect a collision**



MACA

- In **mobile ad hoc networks**, random access is more complicated
- **Hidden nodes problem:** A → B are C → B: collide at B
  - Neither A nor C is aware of this collision
- **Exposed node problem:** Suppose that B is sending to A
  - Since C can hear B's transmission, it would be a mistake for C to conclude that it cannot transmit to D
  - C's transmission will not interfere with A's receiving



MACA

- 802.11 addresses these problems with an algorithm called **Multiple Access with Collision Avoidance (MACA)**
  - Before the sender actually transmits any data, the sender and receiver exchange **control frames** with each other
- The sender transmits a **Request to Send (RTS)** frame to the receiver
  - Includes a field that indicates **how long** the sender wants to hold the medium
- The receiver replies with a **Clear to Send (CTS)** frame
  - Echoes the **length field** back to the sender
- Any node that sees the CTS frame knows that it is close to the receiver and **cannot transmit** for the period of time

MACA

- There are two more details for MACA
- The receiver sends an **ACK** to the sender after successfully receiving a frame
  - All nodes must **wait for this ACK** before trying to transmit
- Should two or more nodes detect an idle link and try to transmit an RTS frame at the same time
  - The RTS frames will collide with each other
  - The senders realize the collision has happened when they do not receive the CTS frame after a period of time
  - They wait a random amount of time before trying again
    - **Exponential backoff algorithm** is generally used

MACA

- RTS/CTS/data/ACK

NAV: network allocation vector