通訊系統(II)

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

Prof. Tsai

Chapter 8 Multichannel Modulation

Capacity of AWGN Channel

Prof. Tsai

Capacity of AWGN Channel

 According to Shannon's information capacity law, the capacity of an AWGN channel is defined by

$$C = B \log_2 \left[1 + P/(N_0 B) \right] = B \log_2 \left[1 + SNR \right]$$
 bits/sec

- where B is the channel bandwidth in hertz and SNR is measured at the **channel output**
- Equivalently, the capacity C in bits per channel use is

$$C' = \frac{1}{2}\log_2(1 + P/\sigma^2) = \frac{1}{2}\log_2(1 + SNR)$$
 bits/transmission

• In practice, we usually find that a **physically realizable encoding system** must transmit data at a rate *R* **less than** the maximum possible rate *C* for **reliable** reception.

Signal-to-Noise Ratio Gap

- For an **implementable** system operating at a certain low enough probability of symbol error Actual SNR \Rightarrow capacity C
 - We introduce an **signal-to-noise ratio gap** (or just **gap**), denoted by Γ, which is defined by $\Gamma = \frac{2^{2C} 1}{2^{2R} 1} = \frac{\text{SNR}}{2^{2R} 1}$ Attainable capacity R \Rightarrow equivalent SNR
 - -C: the capacity of the **ideal encoding system**

Depends on the encoding system

- R: the capacity of the corresponding implementable encoding system
- It is a function of the permissible probability of symbol error $P_{\rm e}$ and the encoding system of interest
- It provides a measure of the "efficiency" of an encoding system
 - with respect to the ideal transmission system

Prof. Tsai 5

Signal-to-Noise Ratio Gap (Cont.)

- A small (large) gap corresponds to an efficient (inefficient) encoding system
- Then, we have the attainable transmit data rate

$$R = \frac{1}{2}\log_2(1 + \text{SNR}/\Gamma)$$
 bits/transmission

- For example: the desired **probability of symbol error** $P_{\rm e} = 10^{-6}$
 - For an uncoded PAM or QAM system, the gap is **8.8 dB**
 - Through the use of channel coding (e.g., trellis codes), the gap may be reduced to as low as 1 dB
- Because SNR = P/N_0B , the attainable **data rate** is defined as

$$R = \frac{1}{2}\log_2\left(1 + \frac{P}{\Gamma N_0 B}\right)$$
 bits/transmission

Continuous-Time Channel Partitioning

Prof. Tsai

Continuous-Time Channel Partitioning

- Consider a linear **wideband channel** with an arbitrary frequency response H(f).
 - The magnitude response |H(f)| is approximated by a **staircase** function
 - $-\Delta f$: the width of each **subchannel**
- A subchannel with/ almost no distortion
- In each step, the channel may be assumed to operate as an AWGN channel **free from inter-symbol interference**.
 - Transmitting a wideband signal is transformed into the transmission of a set of narrowband orthogonal signals
 - Each orthogonal narrowband signal, with its own carrier,
 is generated using a modulation technique, e.g., M-ary QAM
 - AWGN is the only transmission impairment (with a **constant response** for each subchannel)

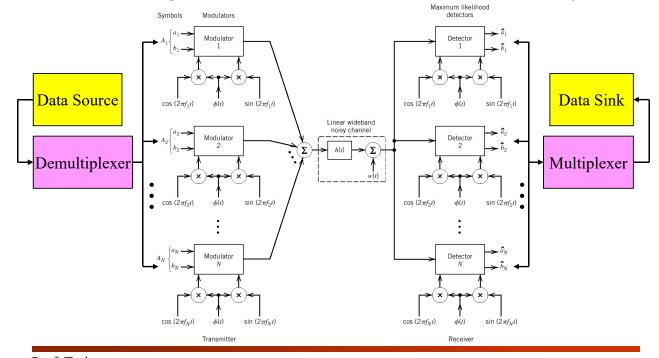
Continuous-Time Channel Partitioning (Cont.)

- Data transmission over each subchannel can be optimized by invoking Shannon's information capacity law
 - The optimization of each subchannel is performed independently of all the others
- The need for **complicated equalization** of a **wideband channel** (due to the **non-constant** response) is replaced by
 - The need of **demultiplexing** and **multiplexing**
 - **Demultiplexing**: Demultiplex the incoming data stream into multiple subchannels
 - **Multiplexing**: Multiplex the demodulated data from multiple subchannels to a single data stream

Prof. Tsai

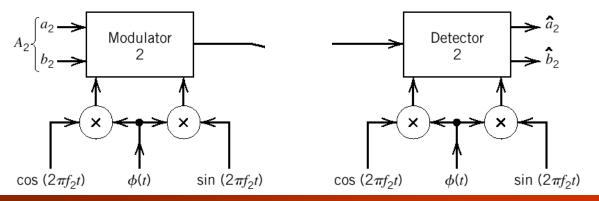
Continuous-Time Channel Partitioning (Cont.)

• A block diagram of the multichannel data transmission system



Continuous-Time Channel Partitioning (Cont.)

- The incoming data stream is first applied to a demultiplexer
 - Produce a set of N substreams
 - Each substream represents a sequence of **two-element** subsymbols, (a_n, b_n) , $n = 1, 2, \dots, N$, for **QAM modulation**
- The detected data of the N substreams are finally applied to a **multiplexer** to restore an output data stream



Prof. Tsai

Geometric Signal-to-Noise Ratio

- In the **multichannel** transmission system, **each subchannel** is characterized by an SNR of its own.
 - However, it is highly desirable to derive a single performance measure of the entire system
- We assume that all of the subchannels are represented by onedimensional constellations
 - The average channel capacity of the entire system is

$$R = \frac{1}{N} \sum_{n=1}^{N} R_n = \frac{1}{2N} \sum_{n=1}^{N} \log_2 \left(1 + \frac{P_n}{\Gamma \sigma_n^2} \right) = \frac{1}{2N} \log_2 \left[\prod_{n=1}^{N} \left(1 + \frac{P_n}{\Gamma \sigma_n^2} \right) \right]$$
$$= \frac{1}{2} \log_2 \left[\prod_{n=1}^{N} \left(1 + \frac{P_n}{\Gamma \sigma_n^2} \right) \right]^{\frac{1/N}{N}}$$
bits/transmission

Geometric Signal-to-Noise Ratio (Cont.)

- Let (SNR)_{overall} denote the **overall SNR** of the entire system.
 - Then, we may express the rate R as

$$R = \frac{1}{2}\log_2\left[1 + \frac{(SNR)_{overall}}{\Gamma}\right]$$
 bits/transmission

• Accordingly, the overall SNR is

$$(SNR)_{overall} = \Gamma \left| \prod_{n=1}^{N} \left(1 + \frac{P_n}{\Gamma \sigma_n^2} \right)^{1/N} - 1 \right|$$

• If the SNR is large enough, we have the approximation

$$(SNR)_{overall} \approx \prod_{n=1}^{N} \left(\frac{P_n}{\sigma_n^2}\right)^{1/N}$$

– It is the **geometric mean** of the SNRs of the individual subchannels and is **independent** of the gap Γ .

Prof. Tsai

Loading of the Multichannel Transmission System

Power Loading

- Define the magnitude response $g_n = |H(f_n)|, n = 1, 2, \dots, N$
- Assuming that the number of subchannels N is large enough
 - $-g_n$ is a **constant** over the entire bandwidth Δf
 - The average channel capacity is

$$R = \frac{1}{2N} \sum_{n=1}^{N} \log_2 \left(1 + \frac{g_n^2 P_n}{\Gamma \sigma_n^2} \right)$$
 bits/transmission

- where g_n and Γ are usually fixed; noise variance is $\Delta f N_0$, $\forall n$
- Goal: Optimize the overall bit rate *R* through a proper allocation of the **total transmit power** among the various subchannels
 - Subject to the total transmit power constraint

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

Prof. Tsai

Power Loading (Cont.)

- **Maximize** the bit rate *R* through an **optimal sharing** of the total transmit power *P* between the *N* subchannels
 - Subject to the total transmit power constraint P
- Through the **method of Lagrange multipliers**, the solution to the **constrained optimization problem** is

$$P_n + \frac{\Gamma \sigma_n^2}{g_n^2} = K, \quad n = 1, 2, \dots, N$$

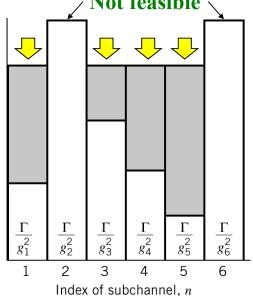
- where K is a prescribed constant to meet the total transmit power constraint P
- The process of allocating the transmit power *P* to the individual subchannels is called **loading**.

Water-Filling Interpretation

• The optimal power allocation must satisfy the condition

 $P_n + \frac{\Gamma \sigma_n^2}{g_n^2} = K, \quad n = 1, 2, \dots, N$

- Consider the case with N = 6
 - The gap Γ is assumed to be constant over all subchannels
 - The average noise power is set to $\sigma_n^2 = N_0 \Delta f = 1$
- We make the following observations
 - With $\sigma_n^2 = 1$, the sum of allocate **power** P_n and the scaled noise **power** Γ/g_n^2 is equal to a constant K for **four subchannels**



Prof. Tsai

Water-Filling Interpretation (Cont.)

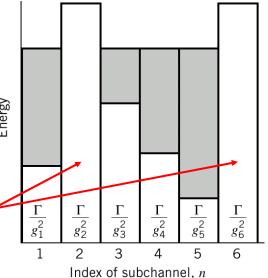
- The sum of power allocations to these four subchannels consumes all the available transmit power P.
- The remaining two subchannels have been eliminated from

consideration

• Because they would each require **negative power** to satisfy the condition (i.e., $P_n < 0$)

The channel response g_n is **too small** \Rightarrow The **scaled noise power** Γ/g_n^2 is **very large**

⇒ Allocating power to these two channels is **inefficient**



Water-Filling Interpretation (Cont.)

- The optimum solution for **loading** is referred to as **water-filling** solution
- This terminology follows from analogy of our optimization problem with
 - A fixed amount of water—standing for transmit power
 - Being poured into a container with a number of connected regions
 - Each having a different depth—standing for noise power
- In such a scenario, the water distributes itself in such a way that
 - A constant water level is attained across the whole container, hence the term "water filling"

Prof. Tsai

Process of Loading

- The allocation of the fixed transmit power *P* among the various subchannels can be formularized as follows:
 - There are a total of (N + 1) unknowns and (N + 1) equations

$$P_{n} + \frac{\Gamma \sigma_{n}^{2}}{g_{n}^{2}} = K, \quad n = 1, 2, \dots, N$$

$$\sum_{i=1}^{N} P_{n} = P$$
Unknowns:
$$P_{n} \text{ and } K$$

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 0 \\ 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ K \end{bmatrix} = \begin{bmatrix} P \\ -\Gamma \sigma^2 / g_1^2 \\ -\Gamma \sigma^2 / g_2^2 \\ \vdots \\ -\Gamma \sigma^2 / g_N^2 \end{bmatrix} \Rightarrow \mathbf{M} \mathbf{u} = \mathbf{c}$$

Process of Loading (Cont.)

- Multiplying the inverse of M on both sides of the equation
 - The unknowns P_1, P_2, \dots, P_N , and K can be obtained

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ K \end{bmatrix} = \mathbf{u} = \mathbf{M}^{-1} \mathbf{c}$$

- K is always positive
- It is possible for some of the P_n values to be negative
 - In such a situation, the negative P_n values are discarded

Prof. Tsai 21

Example

- Consider a linear channel whose squared magnitude response $|H(f)|^2$ has the **piecewise linear form**
- To simplify the example, we have set the gap $\Gamma = 1$ and the noise variance $\sigma_n^2 = 1$

Under this set of values, we have

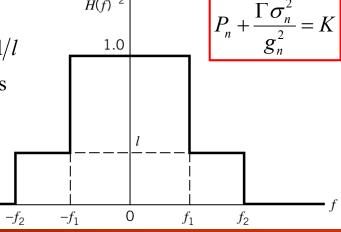
$$P_1 + P_2 = P$$

 $P_1 - K = -1$; $P_2 - K = -1/l$

Solving the three equations for P_1 , P_2 , and K

$$P_1 = (P-1+1/l)/2$$

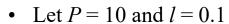
 $P_2 = (P+1-1/l)/2$
 $K = (P+1+1/l)/2$



 $H(f)^{-2}$

Example (Cont.)

- Since 0 < l < 1, it follows that $P_1 > 0$
- But it is possible for P_2 to be **negative**
 - It happens if l < 1/(P+1)
 - Correspondingly, P_1 exceeds the transmit power $P(P_1 > P)$
- Therefore, it follows that, in this example, the only acceptable solution is to have 1/(P+1) < l < 1.

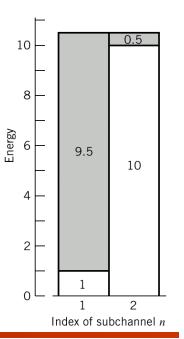


- The desired solution is

$$P_1 = 9.5$$

$$P_2 = 0.5$$

$$K = 10.5$$

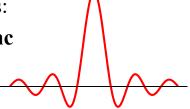


Prof. Tsai

Discrete Multitone

Discrete Multitone

- In the multichannel modulation system, orthogonality of the basis functions and the channel partitioning, is preserved.
- However, the system has two shortcomings:
 - The passband basis functions use a sinc function (rectangular function in the frequency domain)



- Which is nonzero for an **infinite time interval**, resulting in a **non-causal** system
- Practical considerations favor a **finite** time interval
- For a **finite** number of subchannels N, the system is suboptimal
 - Optimality is assured only when N approaches **infinity** (to fit in with the frequency response H(f))

Prof. Tsai 25

Discrete Multitone (Cont.)

- We may overcome these shortcomings by using **discrete multitone** (**DMT**), which transforms a **wideband channel** into a set of N subchannels **operating in parallel**.
- The transformation of DMT is performed in discrete time as well as discrete frequency.
- The transmitter **input-output behavior** of the entire system admits a **linear matrix representation**
 - The discrete Fourier transform (DFT) can be used for implementation of DMT
 - The DFT is the result of discretizing the Fourier transform both in time and frequency.

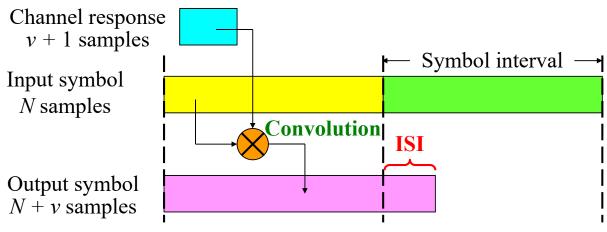
Signal Generation

- Let the **channel impulse response** be h(t), confined to a finite interval $[0, T_h]$
 - The sequence h_0 , h_1 , ..., h_v denote the **baseband equivalent** impulse response of the channel sampled at the rate $1/T_s$ $T_h = (1+v)T_s$
- The sampling rate $1/T_s$ is chosen to be greater than **twice the highest frequency component** of interest (Nyquist rate)
 - In accordance with the sampling theorem
- Let $s[n] = s(nT_s)$ denote a sample of the transmitted symbol s(t), $w[n] = w(nT_s)$ denote a sample of the channel noise w(t), and $x[n] = x(nT_s)$ denote the sample of the channel output.

Prof. Tsai

Signal Generation (Cont.)

- By **linear convolution**, the input symbol sequence $\{s[n]\}$ of length N produces an output sequence $\{x[n]\}$ of length N + v.
 - The extension of v samples is due to the inter-symbol interference produced by the channel (the channel impulse response h(t))



Signal Generation (Cont.)

- To overcome the effect of ISI, we create a cyclically extended guard interval
 - Each symbol sequence is **preceded** by a **periodic extension** of the sequence itself
 - The periodic extension is called as a cyclic prefix (CP)
- Specifically, the **last** *v* samples of the symbol sequence are repeated at the beginning of the sequence being transmitted

$$s[-k] = s[N-k], \quad k = 1, 2, \dots, v$$

• The excess bandwidth factor due to the inclusion of the cyclic prefix is v/N

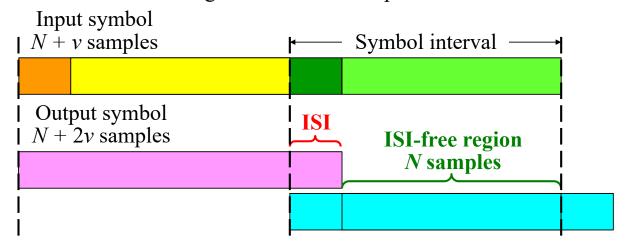
Input symbol N + v samples

Prof. Tsai 29

v samples

Signal Generation (Cont.)

- With insertion of the *v* samples of CP, we have
 - The output symbol contains N + 2v samples
 - The length of ISI is still *v* samples
 - The ISI-free region contains N samples



Signal Generation (Cont.)

• With insertion of the CP, the matrix description of the channel output takes the new form:

$$\begin{bmatrix} x[N-1] \\ x[N-2] \\ \vdots \\ x[v] \\ x[v-1] \\ \vdots \\ x[0] \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 & \cdots & h_{v-1} & h_v & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_{v-1} & h_v & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & h_0 & h_1 & \cdots & h_v \\ h_v & 0 & 0 & \cdots & 0 & 0 & h_0 & \cdots & h_{v-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_1 & h_2 & h_3 & \cdots & h_v & 0 & 0 & \cdots & h_0 \end{bmatrix} \begin{bmatrix} s[N-1] \\ s[V] \\ s[V] \\ s[V] \\ s[V] \end{bmatrix} + \begin{bmatrix} w[N-1] \\ w[V] \\ w[V] \\ w[V] \\ w[V] \end{bmatrix}$$

$$x[N-1] \begin{bmatrix} x[N-1] \\ \vdots \\ x[N-v-2] \end{bmatrix} \begin{bmatrix} x[N-v-1] \\ s[N-v-1] \end{bmatrix} \begin{bmatrix} x[N-v-1] \\ s$$

Prof. Tsai

Signal Generation (Cont.)

• The discrete-time representation of the channel output is

$$x = Hs + w$$

- s: the transmitted symbol vector; w: the channel noise vector;
 x: the received signal vector; all are N-by-1 vectors
- The channel matrix **H** is a **circulant matrix**:
 - Constructed by the **channel response** $[h_0, h_1, \dots, h_v, 0, \dots, 0]^T$

Discrete Fourier Transform

- Consider an *N*-by-1 vector $\mathbf{x} = [x[N-1], x[N-2], \dots, x[0]]^T$
- Let the DFT of **x** be $X = [X[N-1], X[N-2], \dots, X[0]]^T$
 - where the element is defined by

$$X[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \exp(-j2\pi kn/N), \quad k = 0, 1, \dots, N-1$$

- The term $\exp(-j2\pi kn/N)$ is the **kernel** of the DFT
- Correspondingly, the IDFT (i.e., inverse DFT) of X is

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] \exp(j2\pi kn/N), \quad n = 0, 1, \dots, N-1$$

Prof. Tsai

Channel Representation

For a circulant matrix, it permits the spectral decomposition:

$$\mathbf{H} = \mathbf{Q}^{\dagger} \mathbf{\Lambda} \mathbf{Q}$$

- where the superscript † denotes Hermitian transposition
- The matrix **Q** is a **square matrix** defined in terms of the kernel of the *N*-point DFT as

$$\mathbf{Q} = \frac{1}{\sqrt{N}} \begin{bmatrix} \exp\left[-j\frac{2\pi}{N}(N-1)(N-1)\right] & \cdots & \exp\left[-j\frac{2\pi}{N}2(N-1)\right] & \exp\left[-j\frac{2\pi}{N}(N-1)\right] & 1 \\ \exp\left[-j\frac{2\pi}{N}(N-1)(N-2)\right] & \cdots & \exp\left[-j\frac{2\pi}{N}2(N-2)\right] & \exp\left[-j\frac{2\pi}{N}(N-2)\right] & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \exp\left[-j2\pi(N-1)/N\right] & \cdots & \exp\left[-j2\pi \times 2/N\right] & \exp\left[-j2\pi/N\right] & 1 \\ 1 & \cdots & 1 & 1 \end{bmatrix}$$

$$q_{kl} = \frac{1}{\sqrt{N}} \exp\left[-j\frac{2\pi}{N}kl\right], \quad k,l \in \{0,1,\dots,N-1\}$$

Channel Representation (Cont.)

- The matrix \mathbf{Q} is an orthonormal matrix or unitary matrix $\mathbf{Q}^{\dagger}\mathbf{Q} = \mathbf{I}$
- The matrix Λ is a **diagonal matrix** that contains the N **DFT** values of the sequence h_0, h_1, \dots, h_v
- Denoting these transform values by $\lambda_{N-1}, \dots, \lambda_1, \lambda_0$

$$- [\lambda_{N-1}, \cdots, \lambda_1, \lambda_0]^{T} = DFT([h_0, h_1, \cdots, h_v, 0, \cdots, 0]^{T})$$

$$\mathbf{\Lambda} = \operatorname{diag}(\lambda_{N-1}, \dots, \lambda_1, \lambda_0) = \begin{bmatrix} \lambda_{N-1} & 0 & \cdots & 0 \\ 0 & \lambda_{N-2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_0 \end{bmatrix}$$

Prof. Tsai

Frequency-Domain Representation

- Define the time-domain transmit signal vector as Time domain transmitted samples $\mathbf{s} = \mathbf{Q}^{\dagger}\mathbf{S}$ Frequency domain parallel signals
 - where **S** is the **frequency-domain** transmit signal vector
 - Each element of the *N*-by-1 vector **S** is a **complex-valued** point in a two-dimensional QAM signal constellation.
- Given the channel output vector **x**, we define its corresponding **frequency-domain** representation as

$$X = Qx$$

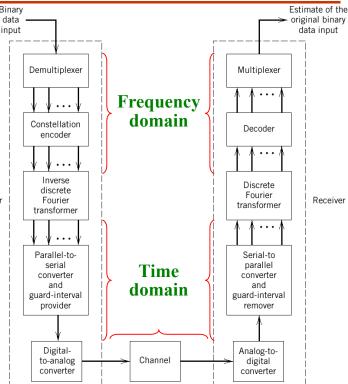
• Then, we may rewrite **X** in the equivalent form

$$\mathbf{X} = \mathbf{Q}\mathbf{x} = \mathbf{Q}(\mathbf{H}\mathbf{s} + \mathbf{w}) = \mathbf{Q}(\mathbf{Q}^{\dagger}\mathbf{\Lambda}\mathbf{Q}\mathbf{Q}^{\dagger}\mathbf{S} + \mathbf{w}) = \mathbf{\Lambda}\mathbf{S} + \mathbf{W}$$

- where W = Qw is the **frequency-domain** noise vector
- That is, $X_k = \lambda_k S_k + W_k$, $k = 0, 1, \dots, N-1$ Parallel subchannels

DFT-Based DMT System

- Block diagram of the DFT-based DMT system
- The **transmitter** consists of the functional blocks:
 - Demultiplexer,
 Constellation encoder,
 IDFT, Parallel-to- Transmitter
 serial converter,
 Digital-to-analog
 converter (DAC)
- The receiver performs the inverse operations of the transmitter



Prof. Tsai

DFT-Based DMT System (Cont.)

- **Demultiplexer**: converts the incoming **serial data stream** into **parallel form** (N constellations in the **frequency-domain**)
- Constellation encoder: maps the parallel data into all multibit subchannels with each subchannel being represented by a QAM signal constellation
 - Bit allocation among the subchannels is performed in accordance with a loading algorithm.
- **IDFT**: transforms the **frequency-domain** parallel data at the constellation encoder output into parallel **time-domain** data
 - For efficient implementation of the IDFT using the **fast Fourier transform** (**FFT**) algorithm, we need to choose $N = 2^k$, where k is a positive integer

DFT-Based DMT System (Cont.)

- Parallel-to-serial converter: converts the parallel timedomain samples into serial form (discrete time samples)
 - Guard intervals stuffed with cyclic prefixes are inserted into the serial data on a periodic basis before conversion into analog form.
- **Digital-to-analog converter (DAC)**: converts the digital data into analog form ready for transmission over the channel

Prof. Tsai

Orthogonal Frequency Division Multiplexing (OFDM)

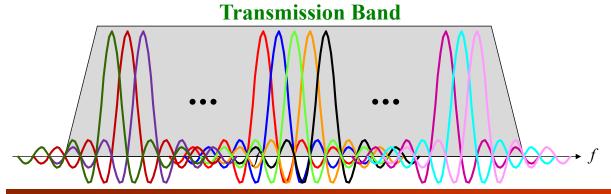
OFDM Concept

- Orthogonal frequency division multiplexing (OFDM) is a form of multi-carrier modulation.
 - OFDM is particularly well suited for high data-rate transmission over delay-dispersive channels.
- Specifically, a large number of closely spaced **orthogonal subcarriers (tones)** is used to support the transmission.
 - The incoming data stream is divided into a number of low data-rate sub-streams, one for each subcarrier
- In addition, two other changes have to be made for OFDM:
 - In the transmitter, an upconverter is included after the DAC to translate the signal to the transmission band
 - In the receiver, a downconverter is included before the ADC to translate the signal to the baseband

Prof. Tsai

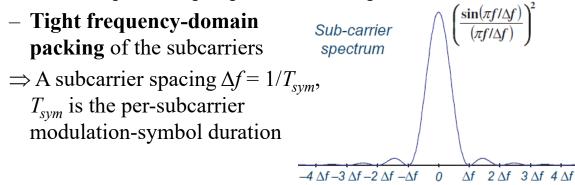
OFDM Concept (Cont.)

- Orthogonal frequency division multiplexing (OFDM) is a promising technique because of its
 - High bandwidth efficiency and
 - Resistance to multipath fading
- Orthogonality is maintained among the subcarriers
- Narrowband transmission for each digitally modulated signal



OFDM Concept (Cont.)

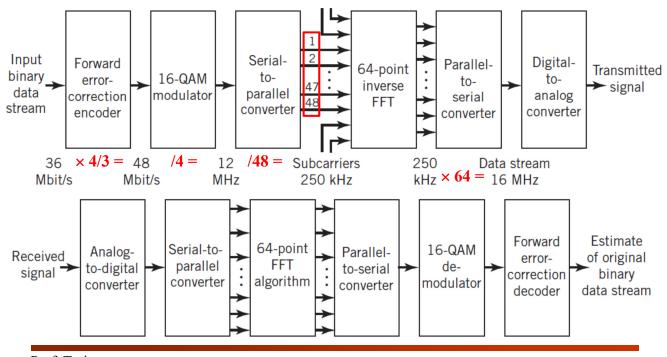
- However, the following characteristics distinguish OFDM from a straightforward multi-carrier extension:
 - The use of a typically very large number of relatively narrowband subcarriers (e.g., several hundred subcarriers)
 - Simple rectangular pulse shaping is used
 - ⇒ A sinc-square-shaped per-subcarrier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



Prof. Tsai

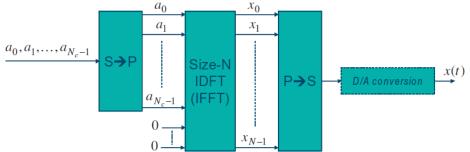
OFDM Transmitter/Receiver

Block diagrams of transmitter/receiver for a 36 Mbits/s system

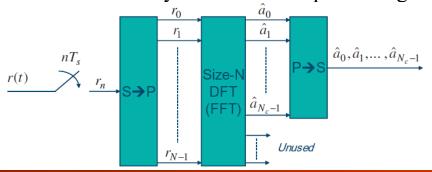


OFDM Implementation

• OFDM modulation by means of IFFT processing

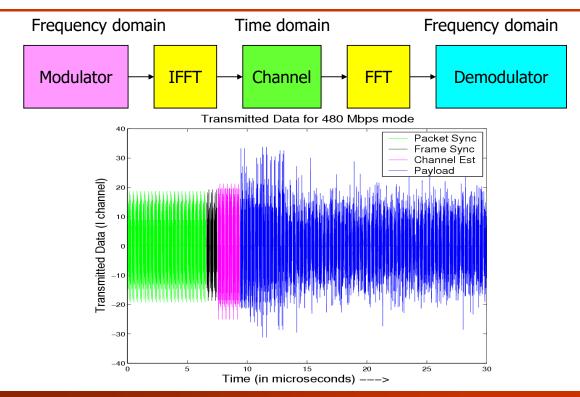


• OFDM demodulation by means of FFT processing



Prof. Tsai 45

OFDM Transmission



Homework

- You must give detailed derivations or explanations, otherwise you get no points.
- Communication Systems, Simon Haykin (4th Ed.)

• 6.43;