

# I225E Statistical Signal Processing

## 13. Signal Processing I

---

MAWALIM and UNOKI

[candylin@jaist.ac.jp](mailto:candylin@jaist.ac.jp) and [unoki@jaist.ac.jp](mailto:unoki@jaist.ac.jp)

School of Information Science

# Signal processing

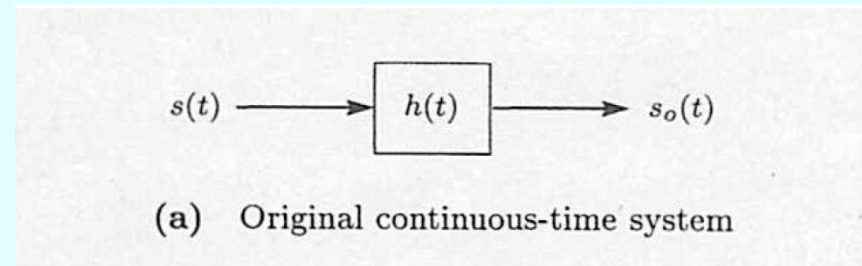
---

- Continuous and discrete-time systems
- Nyquist-Shannon sampling theorem
- Wiener filter
- Kalman filter

# 1. Continuous-Time System

- **Continuous-time signal**  $s(t)$  is the input to a linear time invariant system with **impulse response**  $h(t)$ .
- The output  $s_o(t)$  is

$$s_o(t) = \int_{-\infty}^{\infty} h(t - \tau)s(\tau)d\tau$$



## ■ Assumptions

- $s(t)$ : bandlimited to  $B$  Hz
- $s_o(t)$ : bandlimited to  $B$  Hz
- Frequency response  $\mathcal{F}\{h(t)\}$ : bandlimited to  $B$  Hz

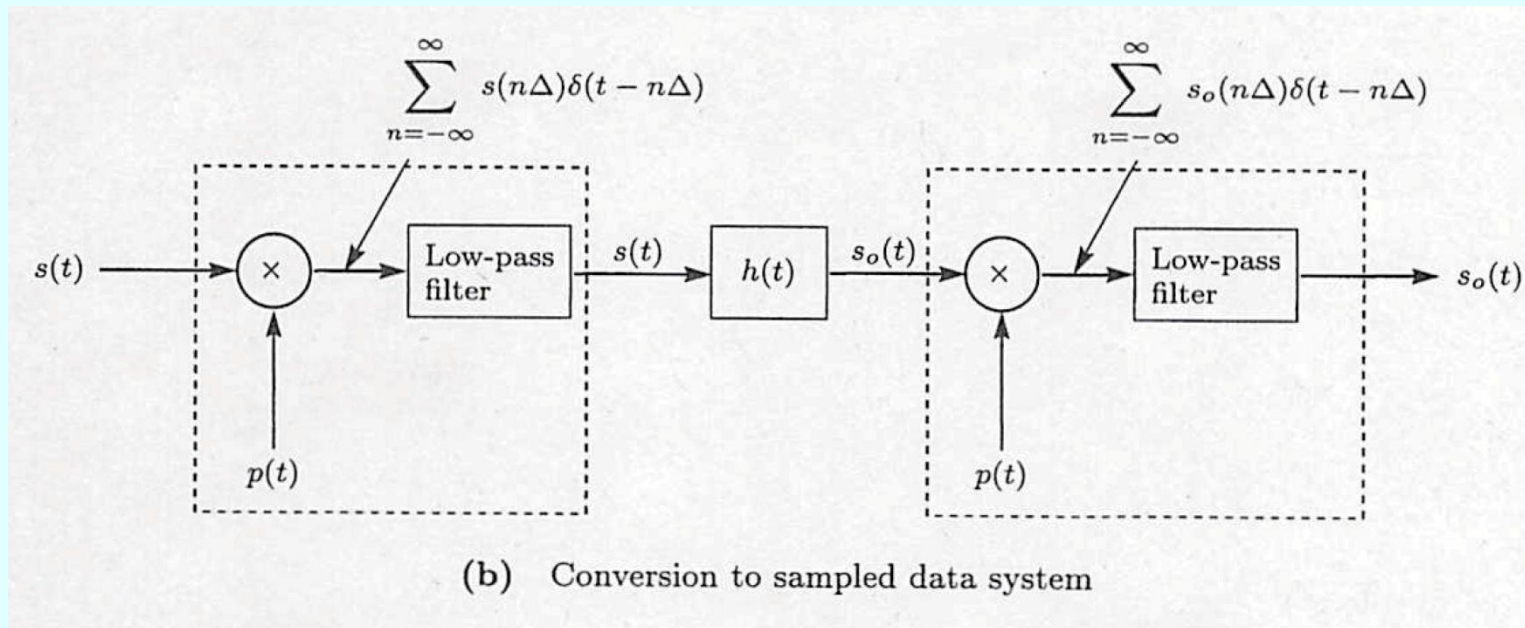
# Sampling

- Sampling function is

$$p(t) = \sum_{n=-\infty}^{\infty} \delta(t - n\Delta).$$

- Frequency response of the low-pass filter is

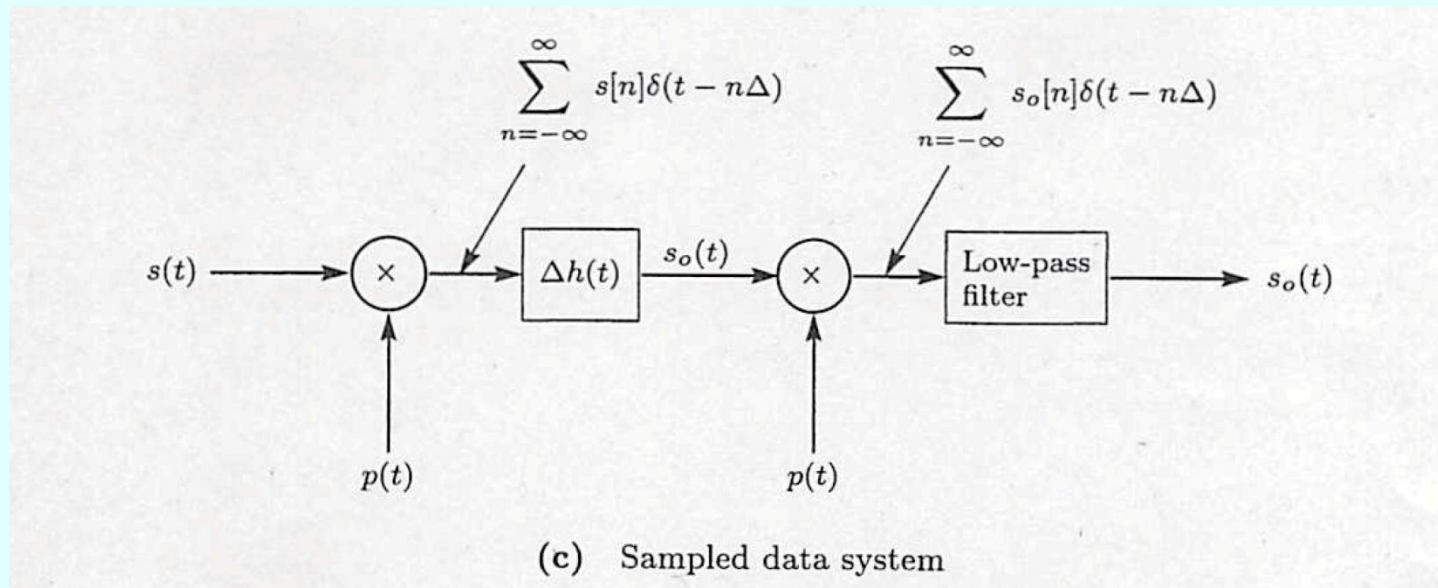
$$H_{\text{lpf}}(F) = \begin{cases} \Delta & |F| < B \\ 0 & |F| > B \end{cases}.$$



## 2. Discrete-Time System

### ■ Reconstruction of $s_o(t)$

$$\begin{aligned} s_o(t) &= \int_{-\infty}^{\infty} \Delta h(t - \tau) \sum_{m=-\infty}^{\infty} s(m\Delta) \delta(\tau - m\Delta) d\tau \\ &= \Delta \sum_{m=-\infty}^{\infty} s(m\Delta) \int_{-\infty}^{\infty} h(t - \tau) \delta(\tau - m\Delta) d\tau \\ &= \Delta \sum_{m=-\infty}^{\infty} s(m\Delta) h(t - m\Delta) \end{aligned}$$



## ■ Output samples

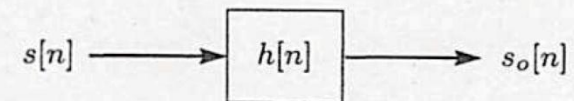
$$s_o(n\Delta) = \sum_{m=-\infty}^{\infty} \Delta h(n\Delta - m\Delta) s(m\Delta)$$

$$s_o[n] = s_o(n\Delta), h[n] = \Delta h(n\Delta), s[n] = s(n\Delta)$$

## ■ Discrete-time output signal

$$s_o[n] = \sum_{m=-\infty}^{\infty} h[n - m] s[m]$$

$$\sum_{n=-\infty}^{\infty} s_o[n] \delta(t - n\Delta)$$



$$h[n] = \Delta h(n\Delta)$$

(d) Discrete-time system

# Nyquist-Shannon sampling theorem (1/2)

## ■ Sampling (Definition)

The process of converting a signal (a function of continuous time or space) into a numeric sequence (a function of discrete time or space).

- A function is called bandlimited to  $B$  Hz if its Fourier transform

$$X(\omega) = F[x(t)] = \int \frac{d\omega}{2\pi} x(t) e^{-i\omega t}$$

satisfies

$$|X(\omega)| = 0 \text{ for } |\omega| > \frac{B}{2\pi}$$

## ■ Shannon's sampling theorem

If a function  $x(\omega)$  contains no frequency higher than  $B$  Hz (bandlimited to  $B$  Hz), it is completely determined by giving its ordinates at a series of points spaced  $1/2B$  seconds apart.

# Nyquist-Shannon sampling theorem (2/2)

## ■ Whittaker-Shannon interpolation formula

If a continuous-time signal  $x(t)$  is bandlimited to  $B$  Hz and  $x[n]$ 's are samples from  $x(t)$  at interval  $1/2B$ , then  $x(t)$  is reconstructed by interpolating  $x[n]$ :

$$x(t) = \sum_{n=-\infty}^{\infty} x[n] \frac{\sin \pi(2Bt - n)}{\pi(2Bt - n)} = \sum_{n=-\infty}^{\infty} x[n] \operatorname{sinc} \pi(2Bt - n)$$



# Note on the definition of frequency

## ■ Angular frequency

$$\begin{cases} x(t) = \int \frac{d\omega}{2\pi} e^{i\omega t} X(\omega) \\ X(\omega) = \int dt e^{-i\omega t} x(t) \end{cases}$$

$$\delta(t) = \int \frac{d\omega}{2\pi} e^{i\omega t}$$

## ■ Frequency

$$\begin{cases} x(t) = \int df e^{2\pi i f t} X(f) \\ X(f) = \int dt e^{-2\pi i f t} x(t) \end{cases}$$

$$\delta(t) = \int df e^{2\pi i f t}$$

$$\omega = 2\pi f$$

# Exercise 1

Consider a continuous-time signal  $x(t) = \cos(2\pi \cdot 10t)$ .

- a) What is the regular frequency ( $f$ ) of this signal in Hertz (Hz)?
- b) What is the angular frequency ( $\omega$ ) of this signal in radians per second (rad/s)?
- c) Find the Fourier Transform  $X(f)$  of  $x(t)$

(Hint: You might find Euler's formula helpful:  $\cos(\theta) = \frac{e^{j\theta} + e^{-j\theta}}{2}$ )

- d) Now, find the Fourier Transform  $X(\omega)$  of  $x(t)$ .

# 3. Signal Processing Example

- There are many signal processing problems that fit the Bayesian linear model.
- **Problem:** transmit a signal  $s(t)$  through a channel with impulse response  $h(t)$ .
- To estimate  $s(t)$  over the interval  $0 \leq t \leq T_s$ .  
Channel will distort and lengthen the signal and  $x(t)$  is observed over the longer interval  $0 \leq t \leq T$ .
- ***Deconvolution problem (blind deconvolution)***  
To deconvolve  $s(t)$  from a noise corrupted version of  $s_o(t) = s(t) * h(t)$ .

■  $h(t)$ : known

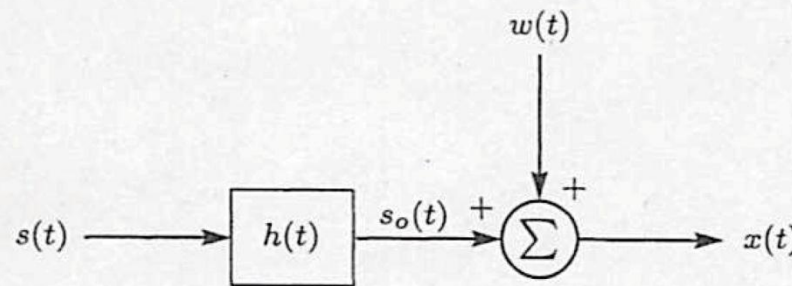
■  $s(t)$ : realization of a random

$$x(t) = \int_0^{T_s} h(t - \tau)s(\tau)d\tau + w(t), \quad 0 \leq t \leq T$$

$s(t)$  has nonzero over the interval  $[0, T_s]$

$h(t)$  has nonzero over the interval  $[0, T_h]$

$$T = T_s + T_h$$



(a) System model

- Discrete time signals for  $s(t)$ ,  $s_o(t)$ : bandlimited to  $B$  Hz
- $w(t)$ : WSS (wide-sense stationary) Gaussian random

process with

$$P_{ww}(F) = \begin{cases} \frac{N_0}{2} & |F| < B \\ 0 & |F| > B \end{cases}$$

- Sampling

$$t = n\Delta, \Delta = 1/(2B), n = 0, 1, \dots, N - 1$$

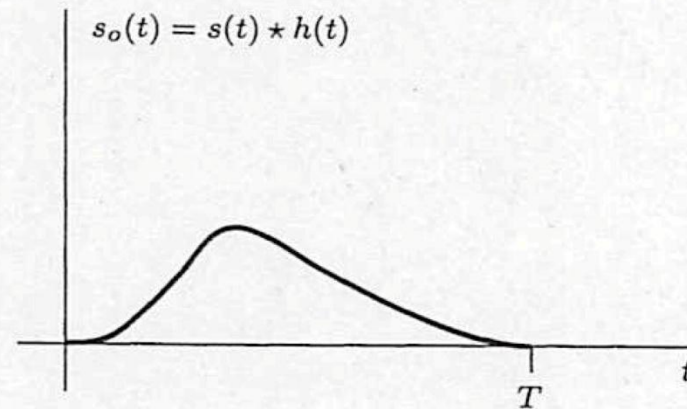
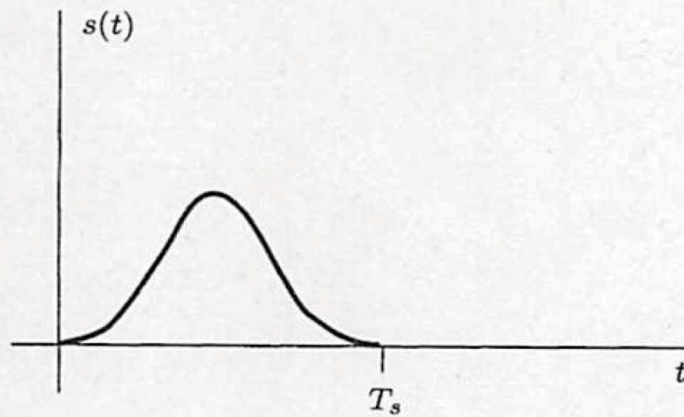
$$x[n] = \sum_{m=0}^{n_s-1} h[n-m]s[m] + w[n]$$

$s[n]$  has nonzero over the interval  $[0, n_s - 1]$

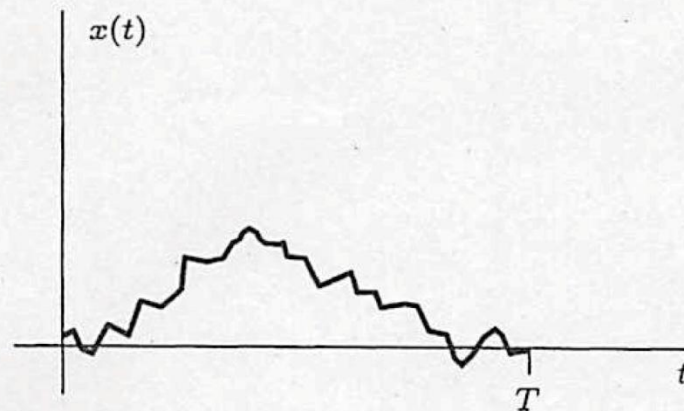
$h[n]$  has nonzero over the interval  $[0, n_h - 1]$

$$N = n_s + n_h - 1$$


$w[n]$  is White Gaussian Noise (WGN) with  $\sigma^2 = N_0B$



(b) Typical signals



(c) Typical data waveform



$$x[n] = \sum_{m=0}^{n_s-1} h[n-m]s[m] + w[n]$$

Component form:

$$\begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{bmatrix} = \begin{bmatrix} h[0] & 0 & \cdots & 0 \\ h[1] & h[0] & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ h[N-1] & h[N-2] & \cdots & h[N-n_s] \end{bmatrix} \begin{bmatrix} s[0] \\ s[1] \\ \vdots \\ s[n_s-1] \end{bmatrix} + \begin{bmatrix} w[0] \\ w[1] \\ \vdots \\ w[N-1] \end{bmatrix}$$

Matrix form:

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{w}$$

## Note

- $h[n] = 0$  for  $n > n_h - 1$
- $\mathbf{s} \sim N(\mathbf{0}, \mathbf{C}_s)$
- $[\mathbf{C}_s]_{ij} = r_{ss}[i - j]$ ,  $r_{ss}[k]$ : Autocorrelation
- MMSE estimator

$$\hat{\mathbf{s}} = \mathbf{C}_s \mathbf{H}^T (\mathbf{H} \mathbf{C}_s \mathbf{H}^T + \sigma^2 \mathbf{I})^{-1} \mathbf{x}$$



# MAP derivation of MMSE estimate

- Find  $\mathbf{s}$  that maximizes the posterior distribution:

$$\begin{aligned}\hat{\mathbf{s}} &= \arg \max_{\mathbf{s}} p(\mathbf{s} | \mathbf{x}) \\ &= \arg \max_{\mathbf{s}} p(\mathbf{x} | \mathbf{s}) p(\mathbf{s}) \\ &= \arg \max_{\mathbf{s}} \log(p(\mathbf{x} | \mathbf{s}) p(\mathbf{s})) \\ &= \arg \max_{\mathbf{s}} \left[ -\frac{1}{2}(\mathbf{x} - \mathbf{H}\mathbf{s})^T \sigma^{-2} (\mathbf{x} - \mathbf{H}\mathbf{s}) - \frac{1}{2} \mathbf{s}^T \mathbf{C}_s^{-1} \mathbf{s} \right]\end{aligned}$$

by taking the derivative and setting it to zero:

$$0 = \frac{\partial}{\partial \mathbf{s}} \left[ \frac{1}{2}(\mathbf{x} - \mathbf{H}\mathbf{s})^T \sigma^{-2} (\mathbf{x} - \mathbf{H}\mathbf{s}) + \frac{1}{2} \mathbf{s}^T \mathbf{C}_s^{-1} \mathbf{s} \right]_{\hat{\mathbf{s}}}$$

$$\begin{aligned}\hat{\mathbf{s}} &= (\mathbf{H}^T \sigma^{-2} \mathbf{H} + \mathbf{C}_s^{-1})^{-1} \mathbf{H}^T \sigma^{-2} \mathbf{x} \\ &= \mathbf{C}_s \mathbf{H}^T (\mathbf{H} \mathbf{C}_s \mathbf{H}^T + \sigma^2 \mathbf{I})^{-1} \mathbf{x}\end{aligned}$$

(another) Woodbury formula

$$(\mathbf{A} + \mathbf{BCD})^{-1} \mathbf{BC} = \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{DA}^{-1} \mathbf{B})^{-1}$$

# Exercise 2

- Consider a communication system where a 2-dimensional signal vector  $\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$  is transmitted through a channel represented by the matrix  $\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  (an identity channel for simplicity).
- The received signal vector is  $\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{w} = \mathbf{s} + \mathbf{w}$ , where  $\mathbf{w} = \begin{bmatrix} w_1 \\ w \end{bmatrix}$  is additive white Gaussian noise with variance  $\sigma^2 = 1$  for each component, and the noise components are independent.
- The prior distribution of the signal  $\mathbf{s}$  is zero-mean Gaussian with a covariance matrix  $\mathbf{C}_s = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ .

## Exercise 2 (Cont.)

- a) Write down the likelihood function  $p(x | s)$ .
- b) Write down the prior distribution  $p(s)$ .
- c) Find the MAP estimate  $\hat{s}_{MAP}$  of  $s$  given  $x$  using the formula:  
$$\hat{s}_{MAP} = (H^T H + \sigma^2 C_s^{-1})^{-1} H^T x$$

Calculate the  $C_s^{-1}$  and then compute  $\hat{s}_{MAP}$ .
- d) Now, compute the MMSE estimate  $\hat{s}_{MMSE}$  using the formula:  
$$\hat{s}_{MMSE} = C_s H^T (H C_s H^T + \sigma^2 I)^{-1} x$$

Verify if  $\hat{s}_{MAP} = \hat{s}_{MMSE}$  in this case.

## ■ Wiener filter

$$\hat{\mathbf{s}} = \mathbf{A}\mathbf{x} = \mathbf{C}_s \mathbf{H}^T (\mathbf{H} \mathbf{C}_s \mathbf{H}^T + \sigma^2 \mathbf{I})^{-1} \mathbf{x}$$

Scalar case

$$\hat{s}[0] = \frac{C_s}{C_s + \sigma^2} x[0] = \frac{r_{ss}[0]}{r_{ss}[0] + \sigma^2} x[0] = \frac{\eta}{\eta + 1} x[0]$$
$$\eta = r_{ss}[0] / \sigma^2$$

$$s[n] = -a[1]s[n-1] + u[n]$$

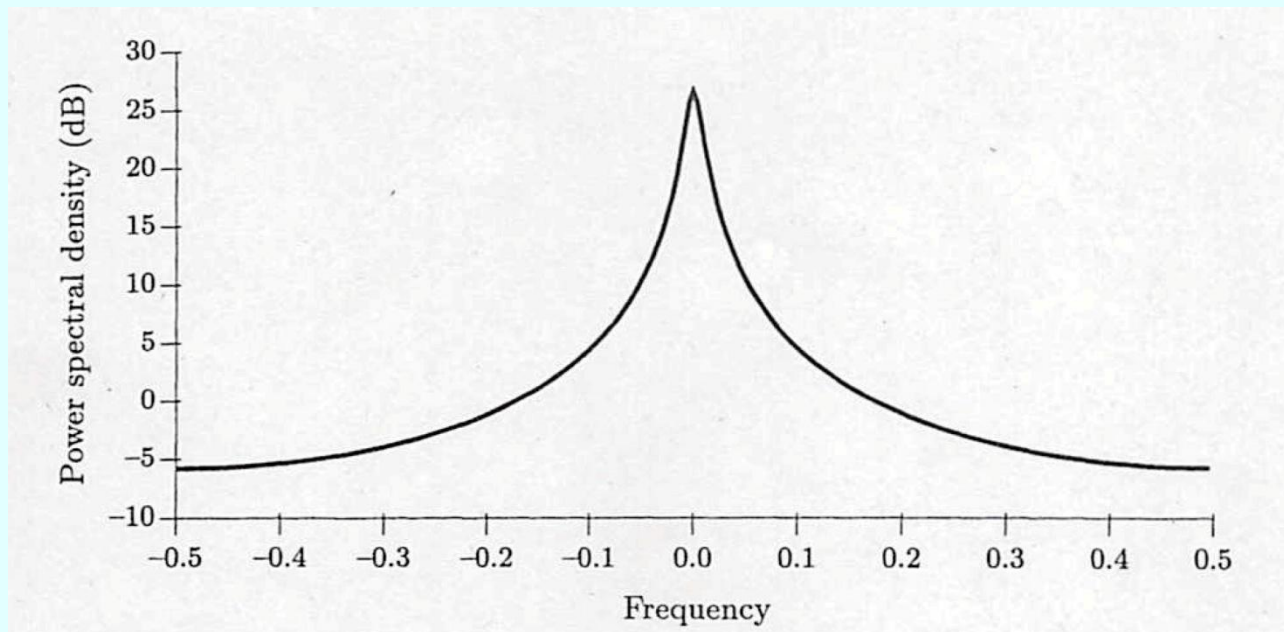
$u[n]$ : WGN with variance  $\sigma^2$

$$r_{ss}[k] = \frac{\sigma_u^2}{1 - a^2[1]} (-a[1])^{|k|}$$

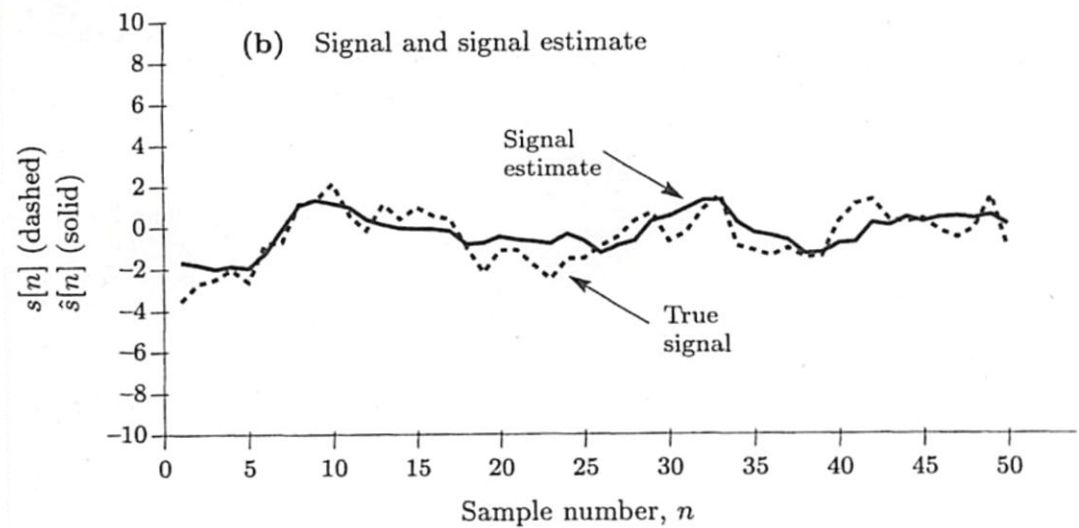
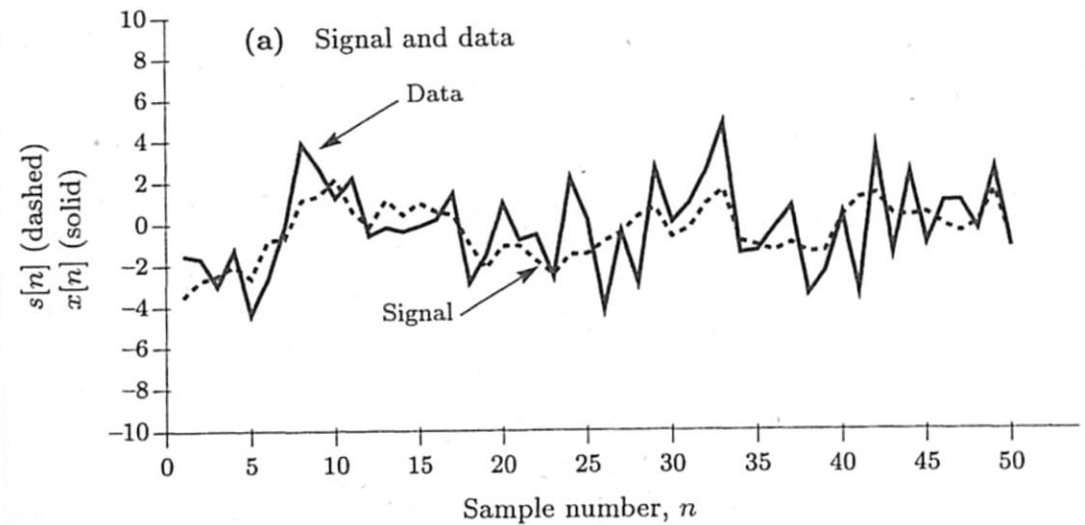
$$P_{ss}(f) = \frac{\sigma_u^2}{|1 + a[1]\exp(-2\pi i f)|^2}$$

# Example

- $a[1] < 0$ , the power density function is that of a low-pass process.
- In case of  $a[1] = -0.95$  and  $\sigma_u^2 = 1$



## Realization of $s[n]$



# 4. Wiener Filtering

- LMMSE estimator

- Data  $\{x[0], x[1], \dots, x[n-1]\}$  is WSS with zero mean.

- Covariance matrix:

$$\mathbf{C}_{xx} = \begin{bmatrix} r_{xx}[0] & r_{xx}[1] & \cdots & r_{xx}[N-1] \\ r_{xx}[1] & r_{xx}[0] & \cdots & r_{xx}[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[N-1] & r_{xx}[N-2] & \cdots & r_{xx}[0] \end{bmatrix} = \mathbf{R}_{xx}$$

(=Autocorrelation matrix)

## ■ **Wiener filters:**

### ■ **Filtering**

$\theta = s[n]$  is to be estimated based on

$$x[m] = s[m] + w[m] \text{ for } m = 0, 1, \dots, n.$$

### ■ **Smoothing**

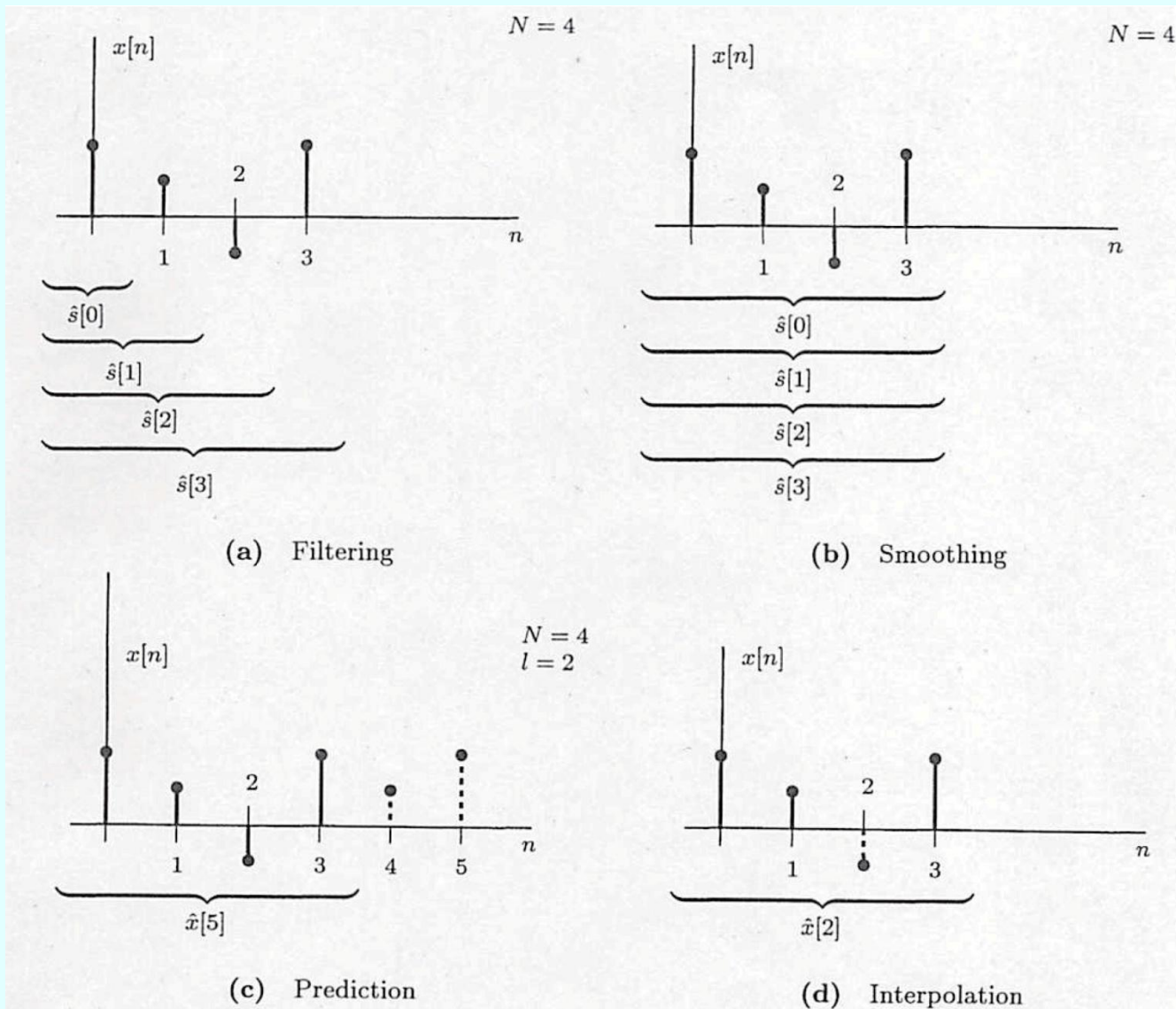
$\theta = s[n]$  is to be estimated for  $n = 0, 1, \dots, N - 1$  based on the data set  $\{x[0], x[1], \dots, x[N - 1]\}$ , where  $x[n] = s[n] + w[n]$ .

### ■ **Prediction**

$\theta = x[N - 1 + l]$  for  $l$  a positive integer is to be estimated base on  $\{x[0], x[1], \dots, x[N - 1]\}$ .

→ ***l-step prediction***





# Exercise 3

- Consider a discrete-time signal  $s[n]$  that is corrupted by additive white noise  $w[n]$ . We observe  $x[n] = s[n] + w[n]$ .
  - Assume the signal  $s[n]$  has an autocorrelation function  $r_{ss}[k] = (0.8)^{|k|}$  and the noise  $w[n]$  is white with variance  $\sigma_w^2 = 0.36$ . The signal and noise are uncorrelated. We want to design a non-causal Wiener filter to estimate  $s[n]$  from  $x[n]$ .
- a) Find the PSD of the signal  $s[n]$  ( $S_{ss}(\omega)$ ). (Hint: The DTFT of  $a^{|k|}$  is  $\frac{1-a^2}{1-2a \cos(\omega)+a^2}$ ).
- b) Find the PSD of the noise  $w[n]$  ( $S_{ww}(\omega)$ ).
- c) Determine the frequency response  $W(\omega)$  of the non-causal Wiener filter that estimates  $s[n]$  from  $x[n]$ . The formula is:

$$W(\omega) = \frac{S_{ss}(\omega)}{S_{ss}(\omega) + S_{ww}(\omega)}$$

# Wiener filter and Kalman filter

Kalman filter is an important generalization of Wiener filter.

## ■ *Wiener Filter*

- WSS (Wide-sense-stationary) Process
- Data from infinite past
- Scalar signals
- Non-adaptive

## ■ *Kalman Filter*

- Gauss-Markov Process
- Data from a specific point in time
- Vector signals
- Adaptive (model may evolve over time)