Statistical signal processing (5CTA0)

Lecture 3

Lecturer: Simona Turco

Electrical Engingeering, Signal Processing Systems group

Part 1: Random variables and Random Signals

Part 1

Random Variables and Random Signals

Lecture 1: Probability and Random Variables

Lecture 2: Random vectors, Random processes and random signals

Lecture 3: Linear random signal models

Outline

- Introduction: random signal models
- Recap LTI systems
- LTI with random inputs
- Innovation representation of a random signal
- Spectral factorization
- Autoregressive moving-average models

Introduction

- A stochastic process can be described by a stochastic model governed by a set of parameters
- The stochastic model describes the statistical properties of the process
- Linear random signal models are a special class of stationary random sequences modeled by driving a linear, time-invariant system with white noise

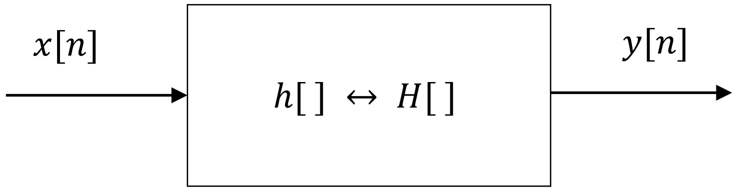
$$x[n] \longrightarrow \boxed{h[n] \leftrightarrow H(z)} \longrightarrow y[n]$$

$H(z)$ rational polynomial $\Rightarrow$ Autoregressive moving average (ARMA)

Recap: Linear-time invariant systems

Lecture 3

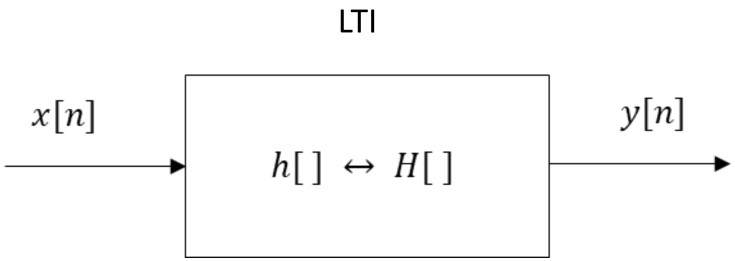
Linear-time invariant (LTI) systems



- *System:* any physical device or algorithm that transforms a signal (input) into another signal (output)

- *System model:* mathematical relationship between input and output

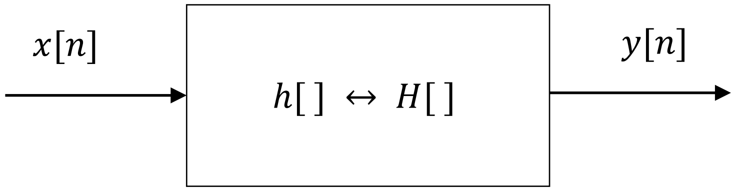
Properties LTI



Linearity $x[n] = Ax_1[n] + Bx_2[n] \rightarrow y[n] = Ay_1[n] + By_2[n]$ *A,B constants*

Time-invariance $x[n] \rightarrow y[n] \quad \Rightarrow \quad x[n-n_0] \rightarrow y[n-n_0]$

LTI systems



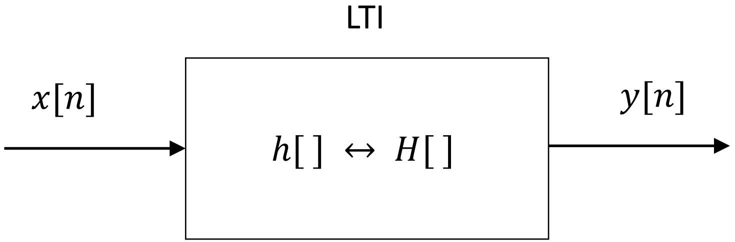
Time-domain analysis

$$y[n] = x[n] * h[n] \triangleq \sum_{k=-\infty}^{\infty} x[k]h[n-k]$$

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k] = h[n] * x[n]$$

$h[n]$, **impulse response** of the system

Properties LTI



Causality: the output signal depends only on the present and/or past values of the input

- Sufficient condition: $$h[n] = 0, \quad for \;\; n < 0$$

Stability: each bounded input produces a bounded output (BIBO-stability)

- Sufficient condition: $$\sum_{n=-\infty}^{\infty} |h[n]| < \infty$$

Transform-domain analysis

Z-transform (discrete-time)

$$Y(z) = \underbrace{H(z)}X(z)$$

$$H(z) = Z\{h[n]\}$$ *system transfer function*

If unit circle inside ROC of $H(z)$, then the system is **stable** and $H(z = e^{j\theta})$ represents the **frequency response** of the system

LTI systems: pole-zero description

Describe $H(z)$ as a rational polynomial

$\boxed{a_0 = 1}$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{Q} b_k z^{-k}}{1 + \sum_{k=1}^{P} a_k z^{-k}} = G \frac{\prod_{k=1}^{Q} (1 - z_k z^{-1})}{\prod_{k=1}^{P} (1 - p_k z^{-1})} = \frac{B(z)}{A(z)}$$

Rational approximation of functions: *<u>any continuous function</u> can be approximated by a rational polynomial as closely as we want by increasing the degree of the numerator and denominator*

LTI systems: pole-zero description

Description system function by **poles** and **zeros**:

$\boxed{a_0 = 1}$

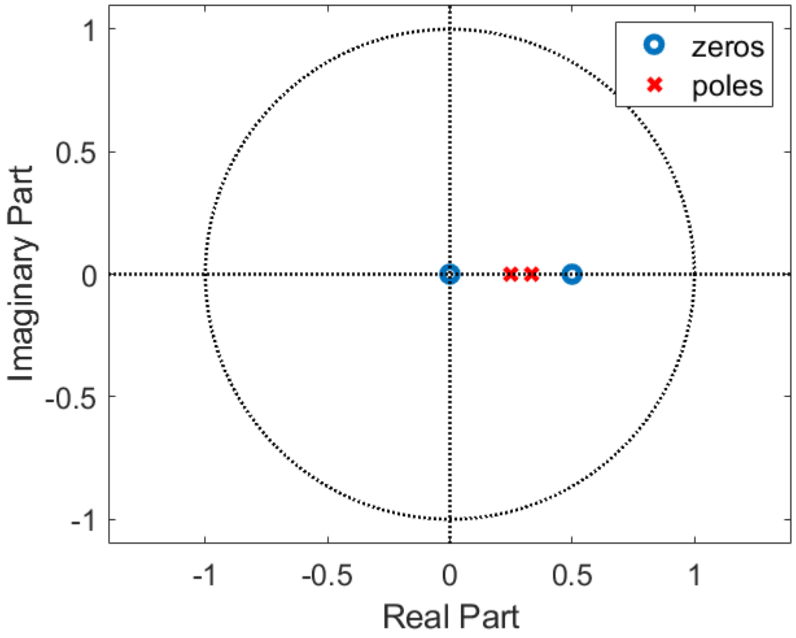
$$H(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_Q z^{-Q}}{1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_P z^{-P}} = \frac{(1 - z_1 z^{-1})(1 - z_2 z^{-1})\ldots(1 - z_Q z^{-1})}{(1 - p_1 z^{-1})(1 - p_2 z^{-1})\ldots(1 - p_P z^{-1})}$$

Zeros: $z_1, z_2, .., z_Q$ zeros of numerator $B(z)$

Poles: $p_1, p_2, .., p_P$ zeros of denominator $A(z)$

Example: pole-zero description

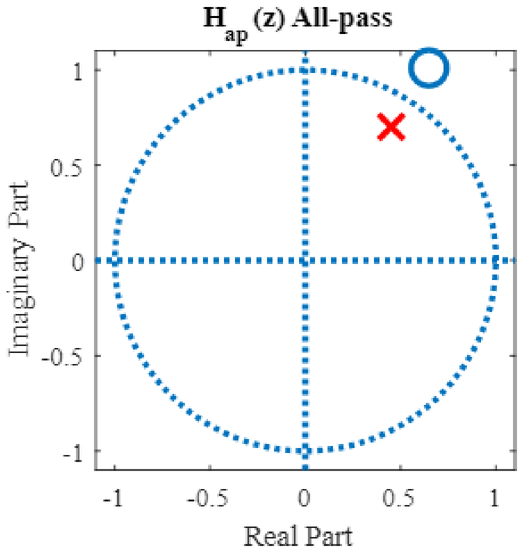
$$H(z)=\frac{1-\frac{1}{2}z^{-1}}{\left(1-\frac{1}{3}z^{-1}\right)\left(1-\frac{1}{4}z^{-1}\right)}=\frac{1-\frac{1}{2}z^{-1}}{1-\frac{7}{12}z^{-1}+\frac{1}{12}z^{-2}}$$



LTI systems: pole-zero description

Description system function by poles and zeros: $H(z)=\frac{Y(z)}{X(z)}=\frac{\sum_{k=0}^{Q} b_k z^{-k}}{1+\sum_{k=1}^{P} a_k z^{-k}}=G\frac{\prod_{k=1}^{Q}(1-z_k z^{-1})}{\prod_{k=1}^{P}(1-p_k z^{-1})}=\frac{B(z)}{A(z)}$

- **Stable**: all poles in $|z| = 1$
- **Causal**: #poles $\geq$ #zeros (P $\geq$ Q)
- **Minimum phase**: all poles and zeros inside $|z| = 1$
- **Maximum phase**: all poles and zeros outside $|z| = 1$
- **Stable all-pass**:
  - all poles in $|z| = 1$
  - Poles and zeros “mirrored pairs”



LTI systems: difference equation

$$H(z)=\frac{Y(z)}{X(z)}=\frac{b_0+b_1z^{-1}+b_2z^{-2}+\ldots+b_Qz^{-Q}}{1+a_1z^{-1}+a_2z^{-2}+\ldots+a_Pz^{-P}}=\frac{B(z)}{A(z)}$$

$$x[n] \longrightarrow \boxed{H(z)=\frac{B(z)}{A(z)}} \longrightarrow y[n]$$

$$Y(z)(1+a_1z^{-1}+a_2z^{-2}+\ldots+a_Pz^{-P})=X(z)(b_0+b_1z^{-1}+b_2z^{-2}+\ldots+b_Qz^{-Q})$$

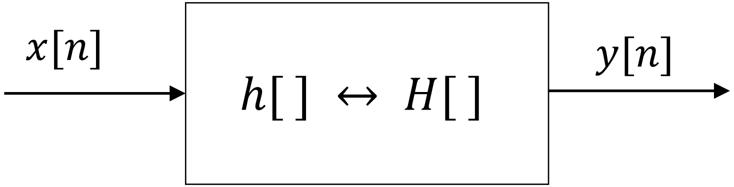
IZT

Time-shifting property: $z^{-n_0}X(z) \Leftrightarrow x[n-n_0]$

$$y[n]+a_1y[n-1]+a_2y[n-2]+\ldots+a_Py[n-P]=x[n]+b_1x[n-1]+b_2x[n-2]+\ldots+b_Qx[n-Q]$$

DE: $$y[n]=\underbrace{x[n]}_{\textit{Current input}}+\underbrace{b_1x[n-1]+b_2x[n-2]+\ldots+b_Qx[n-Q]}_{\textit{Linear combination of past \underline{input} samples}}\underbrace{-a_1y[n-1]-a_2y[n-2]-\ldots-a_Py[n-P]}_{\textit{Linear combination of past \underline{output} samples}}$$

LTI systems: deterministic signals



Input-output relationships

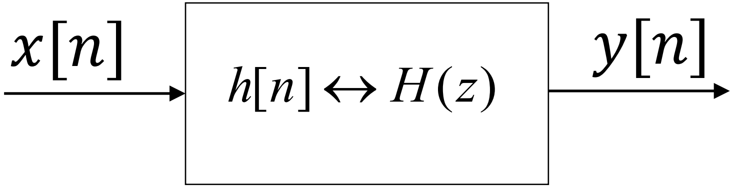
Time-domain $$y[n] = x[n] * h[n] = h[n] * x[n]$$

Z-domain $$Y(z) = H(z)X(z) = X(z)H(z)$$

LTI with random input

What happens when an LTI system is fed with a random input?

- **Assumption**: $x[n]$ is WSS with zero-mean



LTI with random input

Can we calculate the **second-order statistics** of the output if we know the input second-order statistics and the system impulse response/transfer function?

$$P_x(e^{j\theta}), r_x[l] \longrightarrow \boxed{h[n] \leftrightarrow H(z)} \longrightarrow P_y(e^{j\theta}), r_y[l]?$$

TU/e

LTI with random input: autocorrelation

$$r_x[l] \longrightarrow \boxed{h[l]} \xrightarrow{r_{yx}[l]}$$

$$r_{yx}[l] = h[l] * r_x[l]$$

$$r_{xy}[l] = h[-l] * r_x[l]$$

Proof...

$$r_{xy}[l] = E\{x[n]y[n-l]\} = E\{x[n+l]y[n]\}$$

$$E\{x[n+l]y[n]\} = E\left\{x[n+l]\sum_{k=-\infty}^{\infty} h[k]x[n-k]\right\}$$

Sum does not depend on n nor l, filter coefficients are deterministic

$$E\left\{x[n+l]\sum_{k=-\infty}^{\infty} h[k]x[n-k]\right\} = E\left\{\sum_{k=-\infty}^{\infty} h[k]x[n+l]x[n-k]\right\} = \sum_{k=-\infty}^{\infty} h[k]E\{x[n+l]x[n-k]\}$$

$$r_{xy}[l] = \sum_{k=-\infty}^{\infty} h[k]\, r_x[n+l-n+k] = \sum_{k=-\infty}^{\infty} h[k]\, r_x[l+k] = h[-l] * r_x[l]$$

$$\boxed{r_{xy}[l] = h[-l] * r_x[l]}$$

Proof...

$$r_{yx}[l] = E\{y[n]x[n-l]\}$$

$$E\{y[n]x[n-l]\} = E\left\{\sum_{k=-\infty}^{\infty} h[k]x[n-k]x[n+l]\right\}$$

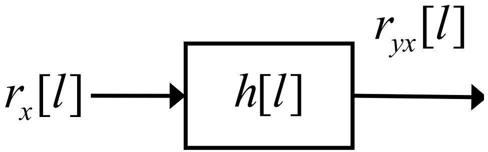
Sum does not depend on n nor l, filter coefficients are deterministic

$$E\left\{\sum_{k=-\infty}^{\infty} h[k]x[n-k]x[n+l]\right\} = \sum_{k=-\infty}^{\infty} h[k]E\{x[n-k]x[n-l]\}$$

$$r_{yx}[l] = \sum_{k=-\infty}^{\infty} h[k]\, r_x[n-k-n+k] = \sum_{k=-\infty}^{\infty} h[k]\, r_x[l-k] = h[l] * r_x[l]$$

$$\boxed{r_{yx}[l] = h[l] * r_x[l]}$$

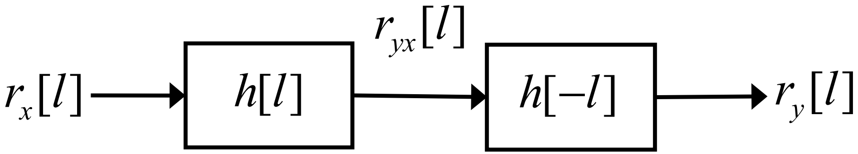
LTI with random input: autocorrelation



$$r_{yx}[l] = h[l] * r_x[l]$$

$$r_{xy}[l] = h[-l] * r_x[l]$$

LTI with random input: autocorrelation



$$r_{yx}[l] = h[l] * r_x[l]$$

$$r_{xy}[l] = h[-l] * r_x[l]$$

$$r_y[l] = h[-l] * r_{yx}[l]$$

Proof…

$$r_y[l] = E\{y[n+l]y[n]\} = E\{y[n+l]x[n]*h[n]\} = E\left\{y[n+l]\sum_{k=-\infty}^{\infty} h[k]x[n-k]\right\}$$

filter coefficients $h[k]$ *are deterministic…*

$$\sum_{k=-\infty}^{\infty} h[k]E\{y[n+l]x[n-k]\} = \sum_{k=-\infty}^{\infty} h[k]\, r_{yx}[(n+l)-n+k] = \sum_{k=-\infty}^{\infty} h[k]\, r_{yx}[l+k]$$

$$\boxed{r_y[l] = h[-l]*r_{yx}[l]}$$

LTI with random input: autocorrelation

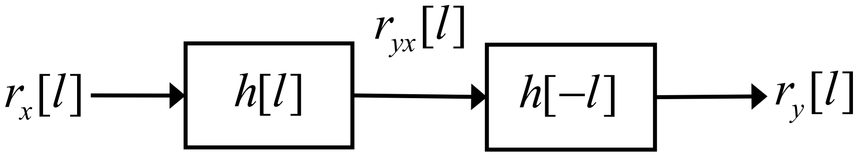
$$r_x[l] \longrightarrow \boxed{h[l]} \xrightarrow{r_{yx}[l]} \boxed{h[-l]} \longrightarrow r_y[l]$$

$$r_{yx}[l] = h[l] * r_x[l]$$

$$r_{xy}[l] = h[-l] * r_x[l]$$

$$r_y[l] = h[-l] * r_{yx}[l] = h[-l] * h[l] * r_x[l] = r_h[l] * r_x[l]$$

LTI with random input: autocorrelation



$$r_{yx}[l] = h[l] * r_x[l]$$

$$r_{xy}[l] = h[-l] * r_x[l]$$

Autocorrelation function of the system

$$\boxed{r_y[l] = r_h[l] * r_x[l]}$$

with $r_h[l] = \sum_{n=-\infty}^{\infty} h[k]h[k-l] = h[l] * h[-l]$

LTI with random input: PSD

Taking the z-transform on the unit circle, and for real $h[n]$:

$$r_{yx}[l] = h[l] * r_x[l] \quad \xleftrightarrow{\mathcal{F}} \quad P_{yx}(e^{j\theta}) = H(e^{j\theta})P_x(e^{j\theta})$$

$$r_{xy}[l] = h[-l] * r_y[l] \quad \xleftrightarrow{\mathcal{F}} \quad P_{xy}(e^{j\theta}) = H^*(e^{j\theta})P_x(e^{j\theta})$$

$$r_y[l] = h[-l] * r_{yx}[l] = \quad \xleftrightarrow{\mathcal{F}} \quad P_y(e^{j\omega}) = H^*(e^{j\omega})H(e^{j\omega})P_x(e^{j\omega}) =$$

$$= h[-l] * h[l] * r_x[l] \qquad\qquad = |H(e^{j\omega})|^2 P_x(e^{j\omega})$$

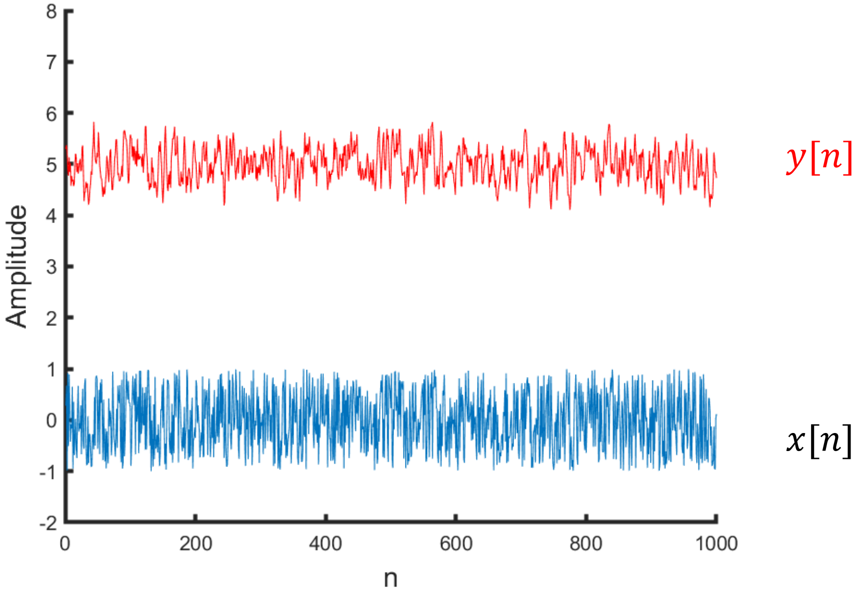
$$\boxed{r_y[l] = r_h[l] * r_x[l]}$$

$$\boxed{P_y(e^{j\theta}) = |H(e^{j\theta})|^2 P_x(e^{j\theta})}$$

Random signals and LTI

From amplitude-time plot

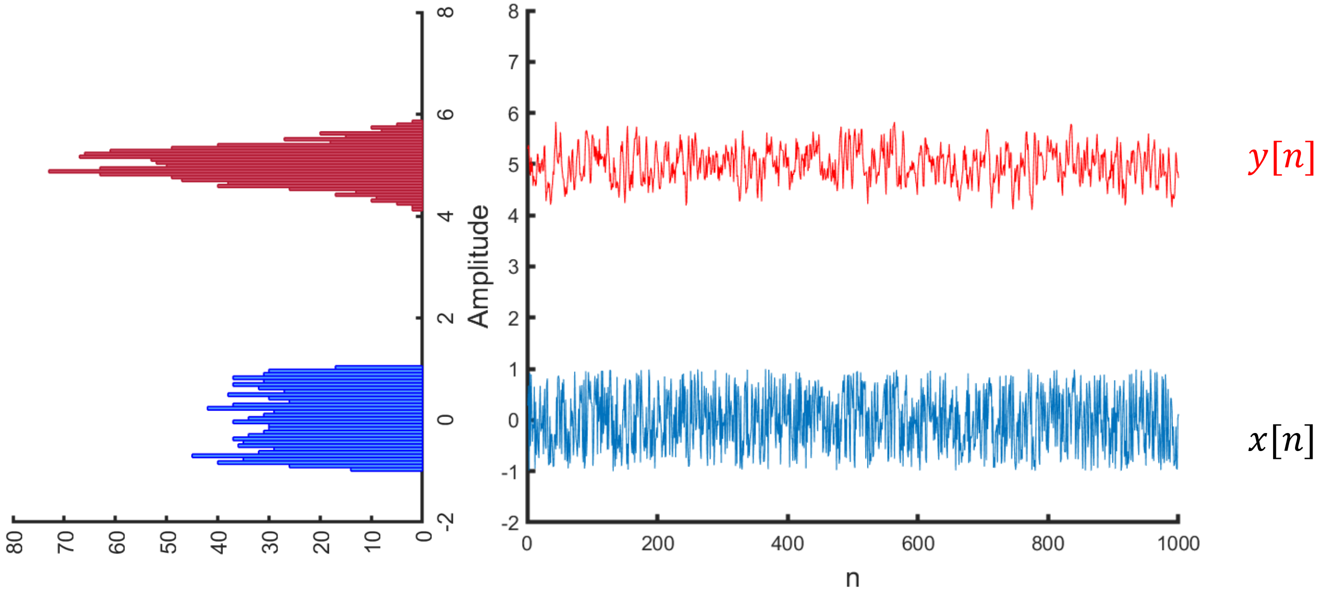
- Frequency of occurrence of various signal amplitude (histogram)



Random signals and LTI

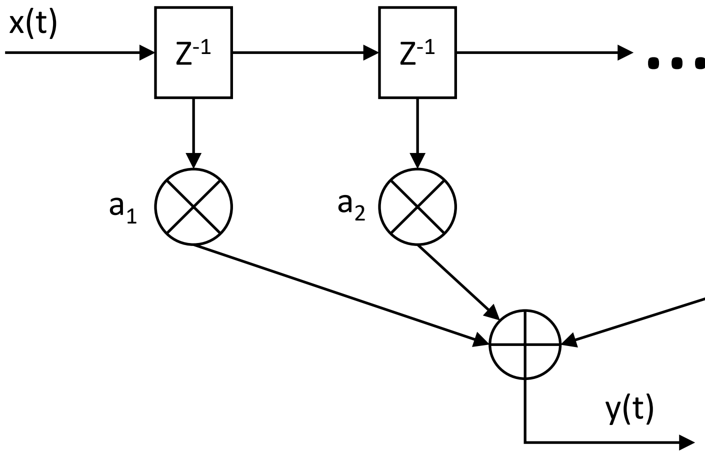
From amplitude-time plot

- Frequency of occurrence of various signal amplitude (histogram)



Central limit theorem

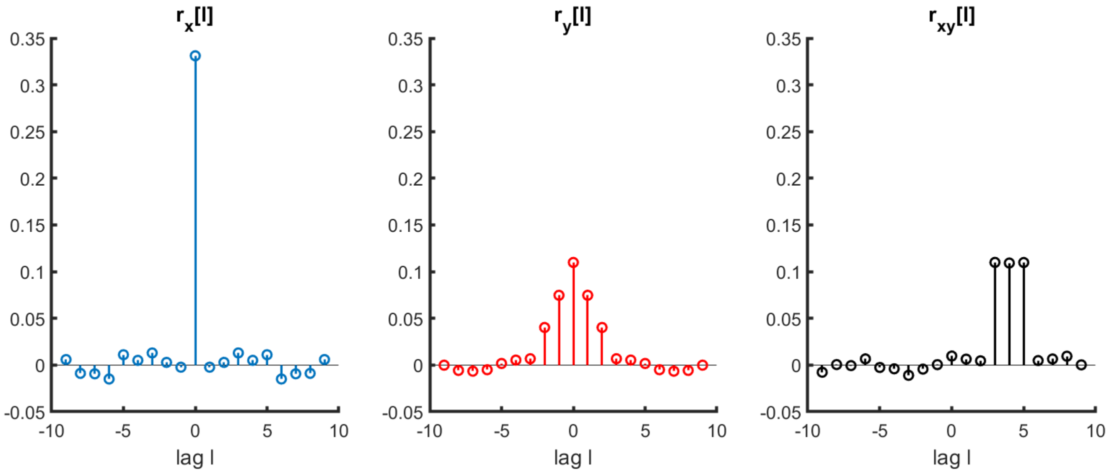
If the filter is long enough and filter's coefficients are of comparable values, then the output is Gaussian distributed, regardless of the distribution of the input



Random signals and LTI

From amplitude-time plot

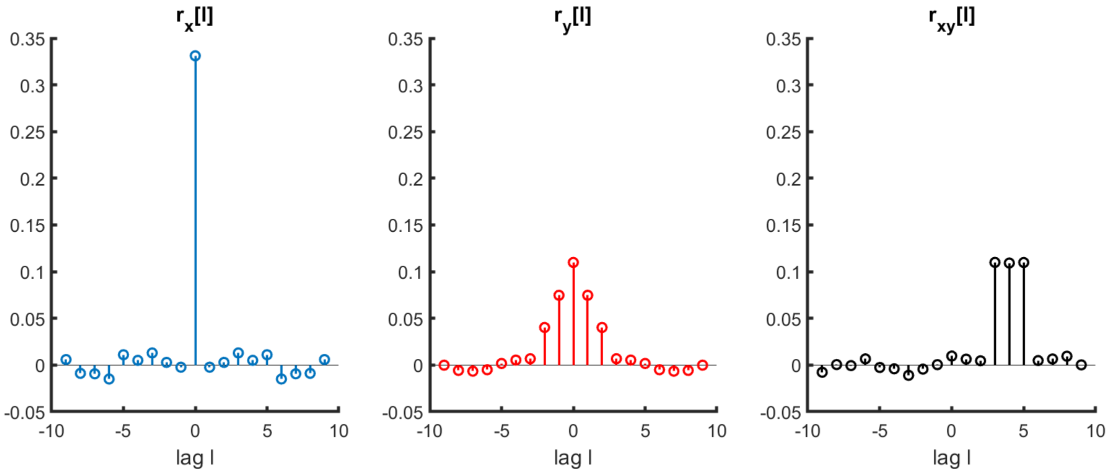
- Degree of dependence between signal samples (auto-correlation/covariance)
- Degree of dependence between two signals (cross-correlation/covariance)



Random signals and LTI

From amplitude-time plot

- Degree of dependence between signal samples (auto-correlation/covariance)
- Degree of dependence between two signals (cross-correlation/covariance)

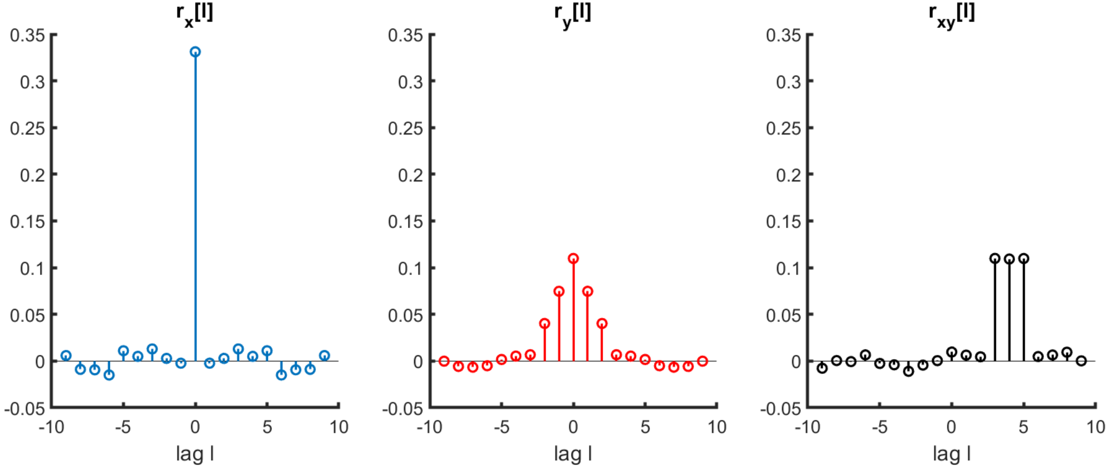


$$x[n] \sim \text{uniform}(-1,1)$$

Random signals and LTI

From amplitude-time plot

- Degree of dependence between signal samples (auto-correlation/covariance)
- Degree of dependence between two signals (cross-correlation/covariance)



$$x[n] \sim \text{uniform(-1,1)}$$

$$y[n] = h[n] * x[n-5] + 5$$

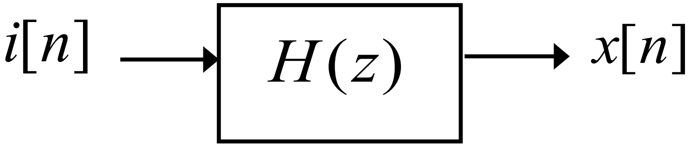
$$h[n] = \frac{1}{3}(\delta[n] + \delta[n-1] + \delta[n-2])$$

Spectral factorization

Lecture 3

Innovation representation

Most random processes with a continuous PSD can be described as the output of a causal filter driven by white noise, the so-called innovation representation of the random process

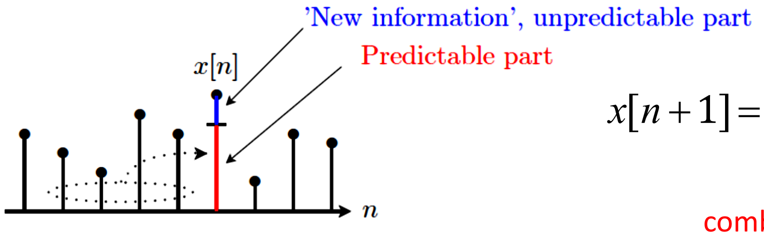


TU/e

Innovation representation

Most random processes with a continuous PSD can be generated as the output of a causal filter driven by white noise, the so-called innovation representation of the random process

Wold's decomposition theorem:
Every wide sense stationary (WSS) signal can be written as the sum of two components, one deterministic and one stochastic



$$x[n+1] = \underbrace{\sum_{k=-\infty}^{n} h_L[n+1-k]x[k]}_{\text{Predictable part: linear combination of past information}} + \underbrace{i[n+1]}_{\text{New information (innovation)}}$$

Innovation representation

Most random processes with a continuous PSD can be generated as the output of a causal filter driven by white noise, the so-called innovation representation of the random process

Synthesis or coloring filter [innovation filter]

$$i[n] \longrightarrow \boxed{L(z)} \longrightarrow x[n]$$

- **Innovation $i[n]$: zero mean, white noise, variance σ_i^2**

Innovation representation

Most random processes with a continuous PSD can be generated as the output of a causal filter driven by white noise, the so-called innovation representation of the random process

Synthesis or coloring filter [innovation filter]

$$i[n] \longrightarrow \boxed{L(z)} \longrightarrow x[n]$$

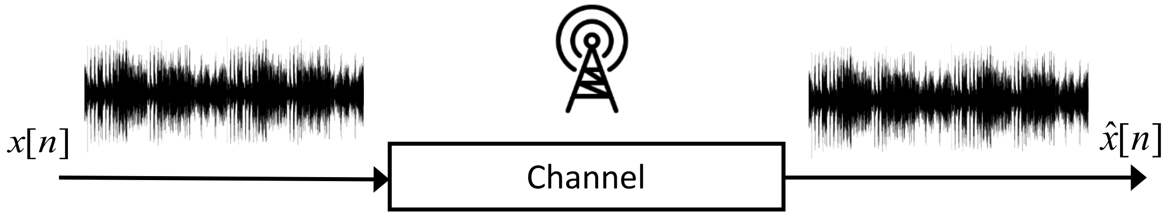
Analysis or whitening filter

$$x[n] \longrightarrow \boxed{\Gamma(z)} \longrightarrow i[n]$$

- **Innovation $i[n]$: zero mean, white noise, variance σ_i^2**
- **Inverse system $\Gamma(z) = L^{-1}(z)$: causal and stable**

Motivation

- Data compression



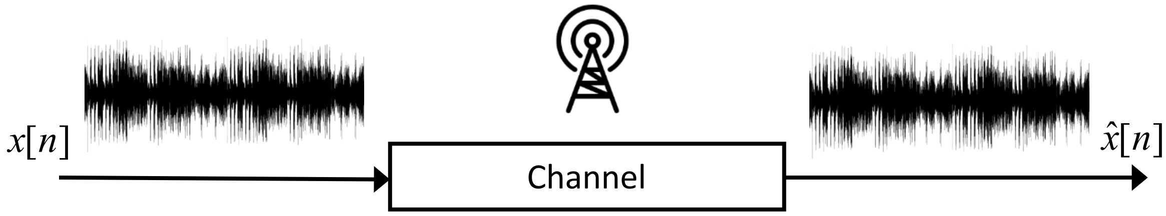
$x[n], 0 \leq n < N$

$N = 10^9$

$10^9 \cdot 64$ bit ~ 8 Gb

Motivation

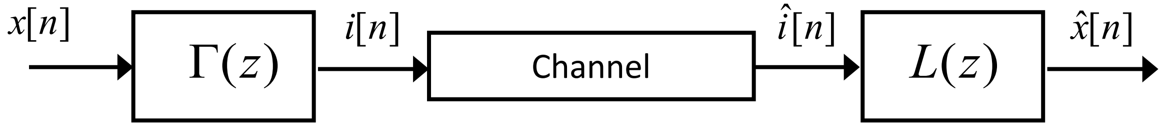
- Data compression



$x[n], 0 \leq n < N$

$N = 10^9$

$10^9 \cdot 64$ bit ~8 Gb



$h_L, 0 \leq n < N$

$N = 100$

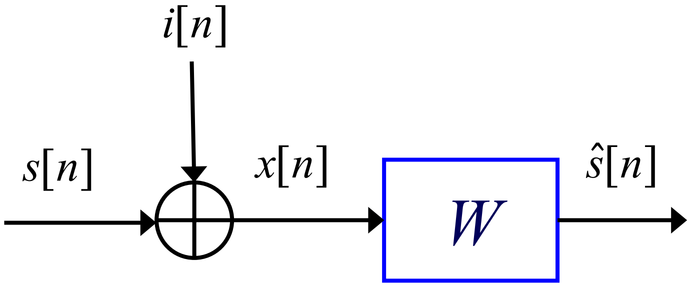
$i[n]$, zero-mean white noise, σ_i^2

$(100 + 1) \cdot 64$ bit $<$ 0.1 Kb

Motivation

- Optimal (Wiener) filtering

(out of scope)



$s[n]$: original signal

$\hat{s}\,[n]$: estimated signal

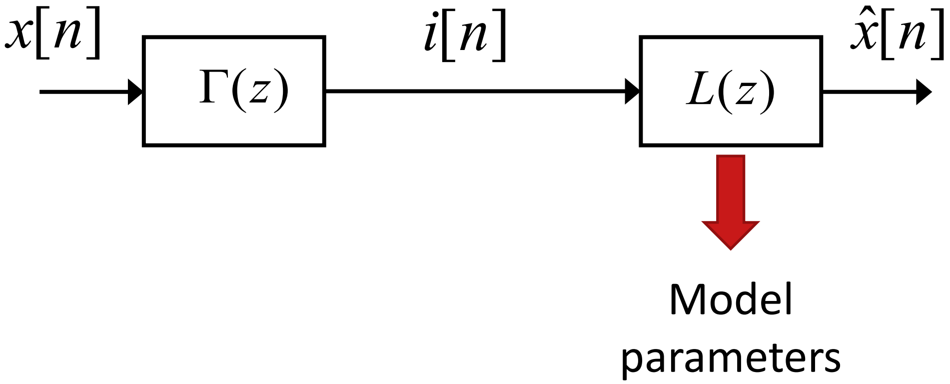
$i[n]$: white noise

$x[n]$: observed signal

W: filter to be designed

Motivation

- Parametric approach to spectral estimation



$$P_x(e^{j\theta}) = \sigma_i^2 \, |L(e^{j\theta})|^2$$

Spectral factorization

Modeling a random sequence $y[n]$ as LTI driven by white noise, can we determine $H(z)$ and σ^2 only knowing the second-order statistics of $y[n]$?

$$P_x(e^{j\theta}) = \sigma^2, \quad r_x[l] = \sigma^2 \delta[n] \longrightarrow \boxed{H(z)?} \longrightarrow P_y(e^{j\theta}), r_y[l]$$

- Problem: we only know the magnitude response $|H(e^{j\theta})|^2$

$$P_y(e^{j\theta}) = \sigma_x^{\ 2} \left| H(e^{j\theta}) \right|^2$$

Spectral factorization: example

$$i_1[n] \longrightarrow \boxed{H(z) = 1 - \frac{1}{2}z^{-1}} \longrightarrow x_1[n]$$

$$\sigma_{i_1}^2 = 1 \quad \Rightarrow \quad P_{x_1}(e^{j\theta}) = \left|1 - \frac{1}{2}e^{-j\theta}\right|^2 = \frac{5}{4} - \cos(\theta)$$

Spectral factorization: example

$$i_1[n] \longrightarrow \boxed{H(z) = 1 - \frac{1}{2} z^{-1}} \longrightarrow x_1[n]$$

$$\sigma_{i_1}^2 = 1 \quad \Rightarrow \quad P_{x_1}(e^{j\theta}) = \left|1 - \frac{1}{2} e^{-j\theta}\right|^2 = \frac{5}{4} - \cos(\theta)$$

$$i_2[n] \longrightarrow \boxed{H(z) = 1 - 2z^{-1}} \longrightarrow x_2[n]$$

$$\sigma_{i_2}^2 = \frac{1}{4} \quad \Rightarrow \quad P_{x_2}(e^{j\theta}) = \frac{1}{4}\left|1 - 2e^{-j\theta}\right|^2 = \frac{5}{4} - \cos(\theta)$$

Spectral factorization: example

$$i_1[n] \longrightarrow \boxed{H(z)=1-\frac{1}{2}z^{-1}} \longrightarrow x_1[n] \qquad \sigma_{i_1}^2=1 \quad \Rightarrow \quad P_{x_1}(e^{j\theta})=\left|1-\frac{1}{2}e^{-j\theta}\right|^2=\frac{5}{4}-\cos(\theta)$$

$$i_2[n] \longrightarrow \boxed{H(z)=1-2z^{-1}} \longrightarrow x_2[n] \qquad \sigma_{i_2}^2=\frac{1}{4} \quad \Rightarrow \quad P_{x_2}(e^{j\theta})=\frac{1}{4}\left|1-2e^{-j\theta}\right|^2=\frac{5}{4}-\cos(\theta)$$

- Additional constraint: minimum-phase system

Spectral factorization: determination of a minimum-phase system from its magnitude response or from its autocorrelation function

Spectral factorization

If $P(z)$ is rational, the it can be factored as

$$P(z) = \sigma_i^2 L(z) L(z^{-1}),$$

with $L(z)$ causal, stable, minimum-phase

- Note:
  - No poles on the unit circle
  - Innovation filter $L(z)$ causal: $L(z) = \frac{B(z)}{A(z)} = \sum_{k=0}^{\infty} l_k z^{-k}$
  - $L(z)$ stable, minimum-phase: all poles and zeros within |z|=1
  - To overcome ambiguity we choose σ_i^2 such that $l[0]$=1

Spectral factorization

What is the spectral factorization of $P_x\left(e^{j\theta}\right)$? $P_x(e^{j\theta}) = \frac{5}{4} - \cos(\theta)$

(1) $\sigma_i^2 = 1, \quad L(z) = 1 - \frac{1}{2} z^{-1}$

(2) $\cancel{\sigma_i^2 = \frac{1}{4}, \quad L(z) = 1 - 2z^{-1}}$ Zero outside unit circle $\rightarrow$ not minimum phase

TU/e

Spectral factorization

What is the spectral factorization of $P_x(e^{j\theta})$? $\quad P_x(e^{j\theta}) = \frac{5}{4} - \cos(\theta)$

(1) $\sigma_i^2 = 1, \quad L(z) = 1 - \frac{1}{2}z^{-1}$

(2) $\cancel{\sigma_i^2 = \frac{1}{4}, \quad L(z) = 1 - 2z^{-1}}$ Zero outside unit circle $\rightarrow$ not minimum phase

$$P_x(e^{j\theta}) = \frac{5}{4} - \cos(\theta) = 1 \cdot \left(1 - \frac{1}{2}e^{-j\theta}\right)\left(1 - \frac{1}{2}e^{j\theta}\right) \Rightarrow P(z) = \underbrace{1}_{\sigma_i^2} \cdot \underbrace{\left(1 - \frac{1}{2}z^{-1}\right)}_{L(z)}\underbrace{\left(1 - \frac{1}{2}z\right)}_{L(z^{-1})}$$

Spectral factorization

What is the spectral factorization of $P\left(e^{j\theta}\right)$? $\quad P_x(e^{j\theta}) = \frac{5}{4} - \cos(\theta)$

$$i[n],\ \sigma_i^2 = 1 \longrightarrow \boxed{L(z) = 1 - \frac{1}{2} z^{-1}} \longrightarrow x[n]$$

$$P_x(e^{j\theta}) = \frac{5}{4} - \cos(\theta) = 1 \cdot \left(1 - \frac{1}{2} e^{-j\theta}\right)\left(1 - \frac{1}{2} e^{j\theta}\right) \Rightarrow P(z) = \underbrace{1}_{\sigma_i^2} \cdot \underbrace{\left(1 - \frac{1}{2} z^{-1}\right)}_{L(z)} \underbrace{\left(1 - \frac{1}{2} z\right)}_{L(z^{-1})}$$

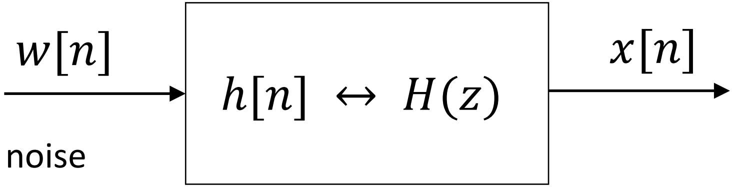
Autoregressive moving-average models

Lecture 3

EINDHOVEN UNIVERSITY OF TECHNOLOGY

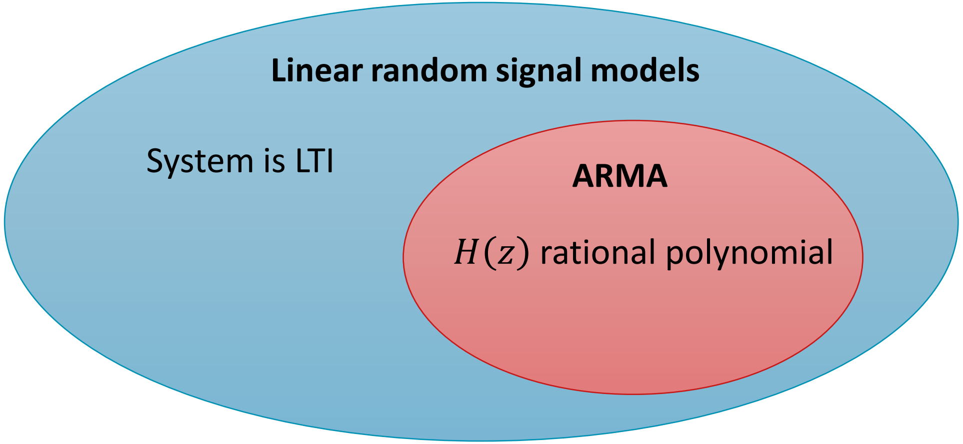
Introduction

Linear random signal models: random signals described as the output of an LTI driven by white noise



$w[n]$, zero-mean white noise

$H(z)$ rational polynomial $\Rightarrow$ Autoregressive moving average (ARMA)

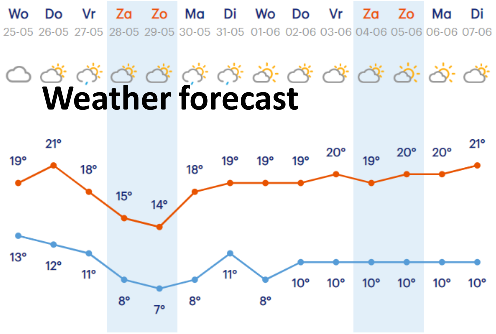


Autoregressive (AR) models

Autoregressive

*Predict future values from past values of the **same** signal*

predicting some values from other values



- Demand of goods
- Weekly/Monthly sales
- Model pendulum oscillation in viscous medium
- …

AR models: Difference equation

$$x[n] = \underbrace{w[n]}_{\text{Unpredictable part (error term)}} \underbrace{- a_1 x[n-1] - a_2 x[n-2] - \ldots - a_p x[n-p]}_{\text{Linear combination of past \textbf{output} samples}}$$

Unpredictable part
(error term)

Linear combination of past
output *samples*

LTI systems: difference equation

$$H(z)=\frac{Y(z)}{X(z)}=\frac{b_0+b_1z^{-1}+b_2z^{-2}+\dots+b_Qz^{-Q}}{1+a_1z^{-1}+a_2z^{-2}+\dots+a_Pz^{-P}}=\frac{B(z)}{A(z)}$$

$$x[n]\longrightarrow\boxed{H(z)=\frac{B(z)}{A(z)}}\longrightarrow y[n]$$

DE: $$y[n]=\underbrace{x[n]}_{\textit{Current input}}\underbrace{\cancel{+b_1x[n-1]+b_2x[n-2]+\dots+b_Qx[n-Q]}}_{\textit{Linear combination of past input samples}}\underbrace{-a_1y[n-1]-a_2y[n-2]-\dots-a_Py[n-P]}_{\textit{Linear combination of past output samples}}$$

AR $$x[n]\longrightarrow\boxed{H(z)=\frac{1}{A(z)}}\longrightarrow y[n]$$

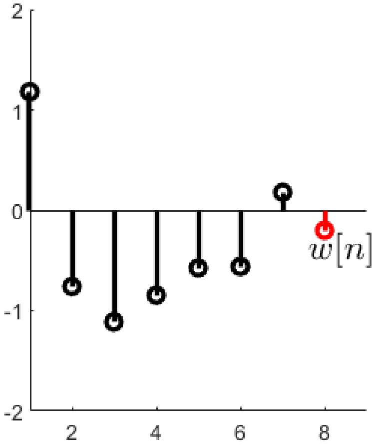
AR models: Difference equation

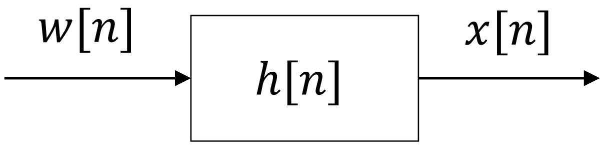
$$x[n] = w[n] - \underbrace{a_1 x[n-1] - a_2 x[n-2] - \ldots - a_p x[n-p]}$$

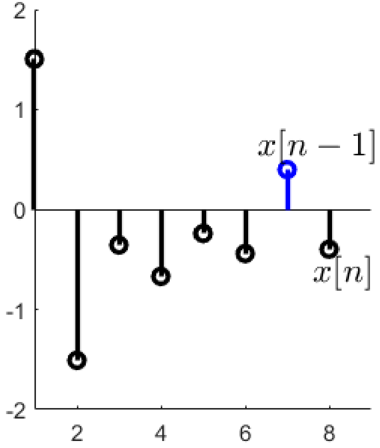
Unpredictable part (error term)

Linear combination of past ***output*** *samples*

AR(1): $x[n] = w[n] - a_1 x[n-1]$







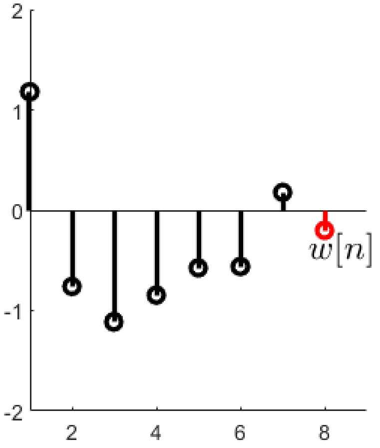
AR models: Difference equation

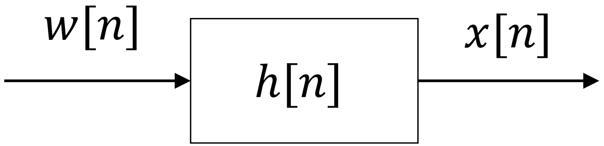
$$x[n] = w[n] \underbrace{- a_1 x[n-1] - a_2 x[n-2] - \ldots - a_p x[n-p]}$$

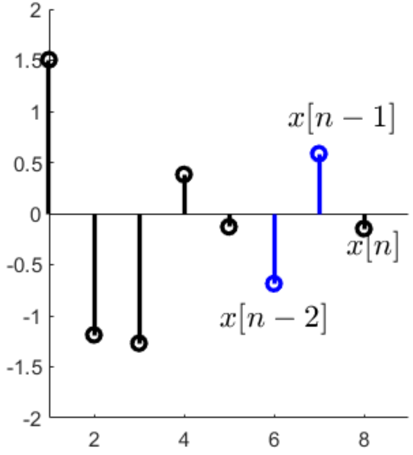
Unpredictable part
(error term)

Linear combination of past
output *samples*

AR(2): $x[n] = w[n] - a_1 x[n-1] - a_2 x[n-2]$







AR models: Difference equation

$$x[n] = w[n] - \underbrace{a_1 x[n-1] - a_2 x[n-2] - \ldots - a_p x[n-p]}$$

Unpredictable part
(error term)

Linear combination of past ***output*** *samples*

$$a_1^{\,2} x[n-2]$$

AR(1): $x[n] = w[n] - a_1 \underline{x[n-1]} = w[n] - a_1(\underline{w[n-1] - a_1 x[n-2]})$

$$= w[n] - a_1(w[n-1] - a_1(\underline{w[n-2] - a_1 x[n-3]}))$$

$$a_1^{\,3} x[n-3]$$

$|p_1| < 1$ for stability

$|a_1| < 1$

"decaying" memory

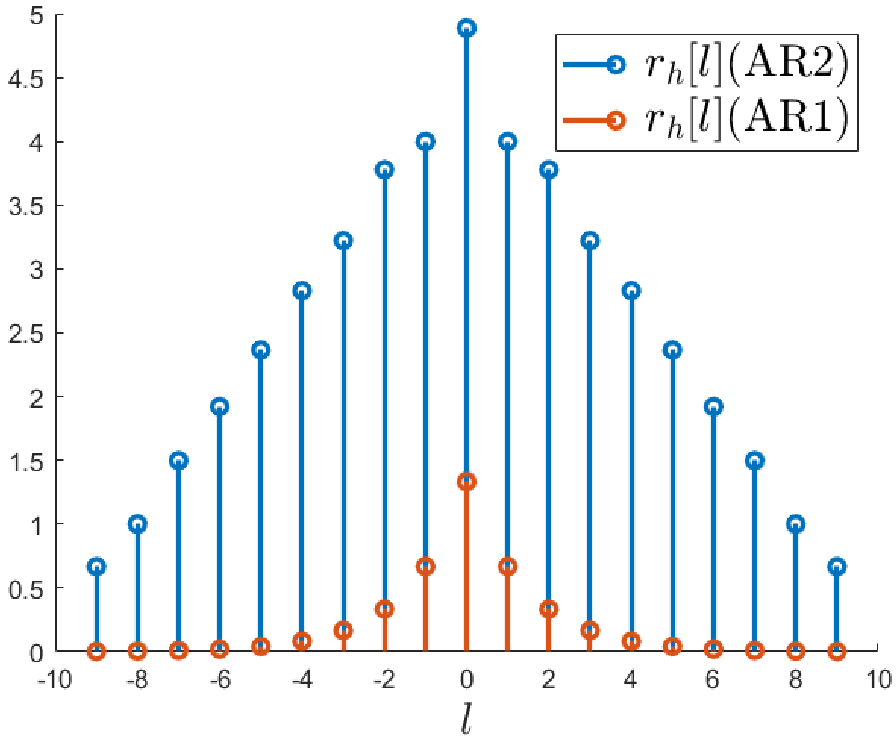
$a_1^3 < a_1^2 < a_1$

AR models: long memory models

AR(1): $a_1 = \frac{1}{2}$

AR(2): $a_1 = a_2 = \frac{1}{2}$

System autocorrelation



AR models: autocorrelation

modified Yule-Walker equations

$$r_x[l] = \begin{cases} \sigma_w^2 - \sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } l = 0 \\ -\sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } l > 0 \end{cases}$$

Model parameters $\begin{cases} \sigma_w^2 \\ a_1, \ldots, a_p \end{cases}$

In practice, the autocorrelation of ARMA processes can be calculated in two different ways:

1. Using modified Yule-Walker equations
2. Using difference equation and definition of autocorrelation

Example AR(1)

$$r_x[l] = \begin{cases} \sigma_w^2 - \sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } l = 0 \\ -\sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } l > 0 \end{cases}$$

Via Yule-Walker

$$x[n] = w[n] - \frac{1}{2} x[n-1]$$

$$\boxed{\sigma_w^{\ 2} = 1,\ a_1 = \frac{1}{2}}$$

$$\left.\begin{aligned} r[0] &= \sigma_w^{\ 2} - a_1 r[-1] = 1 - \frac{1}{2} r[1] \\ r[1] &= -a_1 r[0] = -\frac{1}{2} r[0] \end{aligned}\right\} \Rightarrow r[0] = \frac{4}{3},\ r[1] = -\frac{2}{3}$$

$$r[2] = -a_1 r[1] = -\frac{1}{2} r[1] = -\frac{1}{2}\left(-\frac{1}{2}\right) r[0] = \frac{1}{4} r[0]$$

$$r[3] = -a_1 r[2] = -\frac{1}{2} r[1] = -\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) r[0] = -\frac{1}{8} r[0]$$

$$\Rightarrow r[l] = \frac{4}{3}\left(-\frac{1}{2}\right)^{|l|}$$

AR models: Power spectral density

$$x[n] \longrightarrow \boxed{h[n] \leftrightarrow H(z)} \longrightarrow y[n]$$

$$P_y(e^{j\theta}) = |H(e^{j\theta})|^2 \, P_x(e^{j\theta})$$

$$w[n] \longrightarrow \boxed{H(z)} \longrightarrow x[n]$$

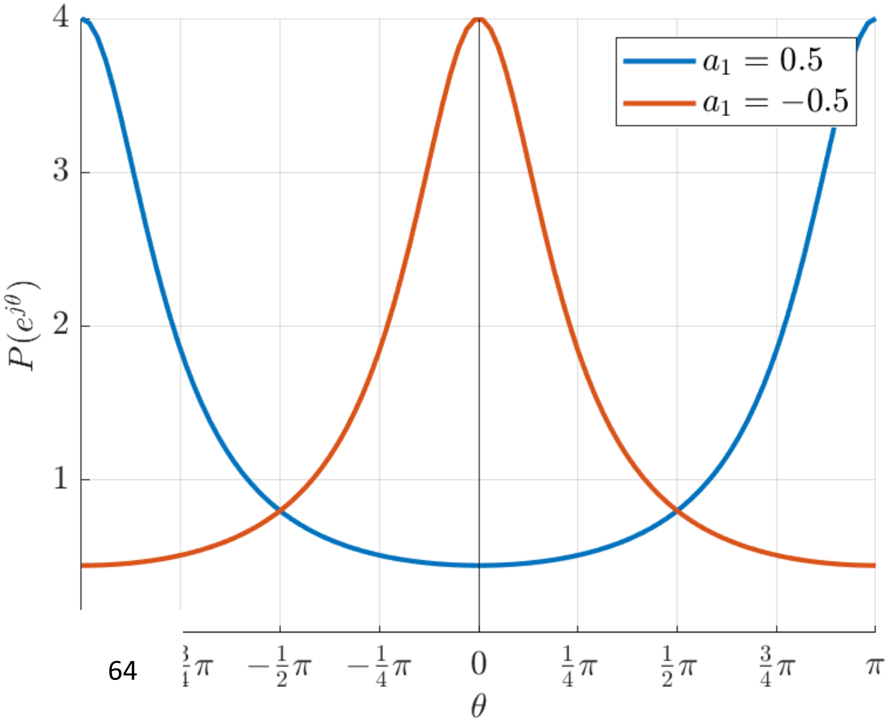
$$P_w(e^{j\theta}) = \sigma_w^{\;2}$$

$$H(z) = \frac{B(z)}{A(z)} = \frac{1}{1 + a_1 z^{-1} + \ldots + a_P z^{-P}}$$

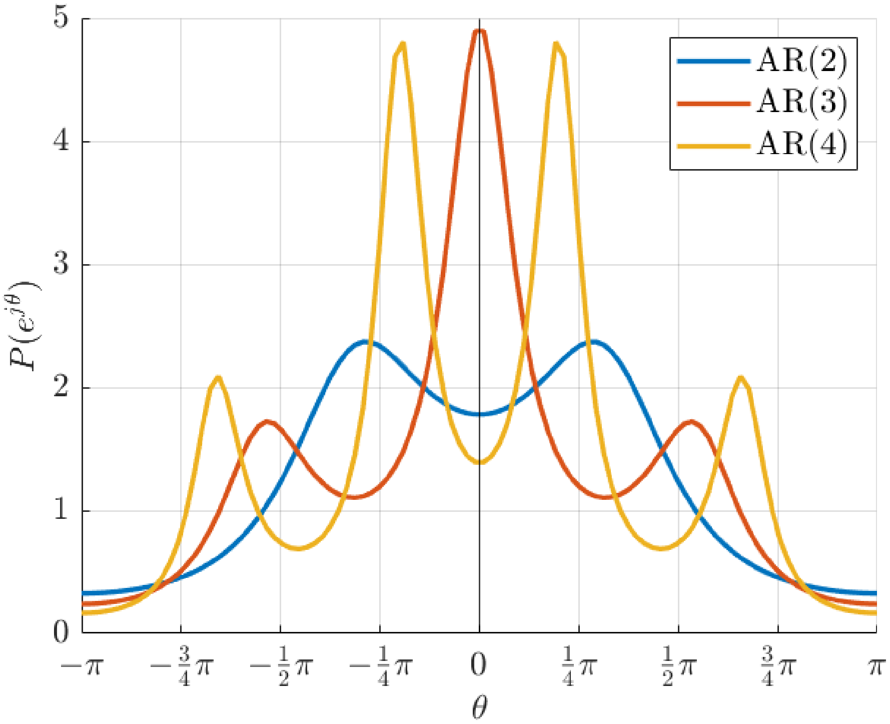
$$\boxed{P_x(e^{j\theta}) = \frac{\sigma_w^2}{\left|1 + \sum_{k=1}^{P} a_k e^{-jk\theta}\right|^2}}$$

AR models: PSD examples

AR(1)



AR(3), AR(3), AR(4)



Moving average (MA) models

Moving average

Each sample is predicted from adjacent previous samples

Weighted sum of some values

- **Smoothing**: filtering out the noise from random short-term fluctuations
- **Prediction**: estimating future values based on previous *input* samples

MA models: Difference equation

$$x[n] = w[n] + \underbrace{b_1 w[n-1] + b_2 w[n-2] + \ldots + b_q w[n-q]}_{\text{Linear combination of past input samples}}$$

unpredictable part

*Linear combination of past **input** samples*

LTI systems: difference equation

$$H(z)=\frac{Y(z)}{X(z)}=\frac{b_0+b_1z^{-1}+b_2z^{-2}+\ldots+b_Qz^{-Q}}{1+a_1z^{-1}+a_2z^{-2}+\ldots+a_Pz^{-P}}=\frac{B(z)}{A(z)}$$

$$x[n] \longrightarrow \boxed{H(z)=\frac{B(z)}{A(z)}} \longrightarrow y[n]$$

DE: $$y[n]=\underbrace{x[n]}_{\text{Current input}}+\underbrace{b_1x[n-1]+b_2x[n-2]+\ldots+b_Qx[n-Q]}_{\text{Linear combination of past input samples}}\cancel{\underbrace{-a_1y[n-1]-a_2y[n-2]-\ldots-a_Py[n-P]}_{\text{Linear combination of past output samples}}}$$

Current input

*Linear combination of past **input** samples*

*Linear combination of past **output** samples*

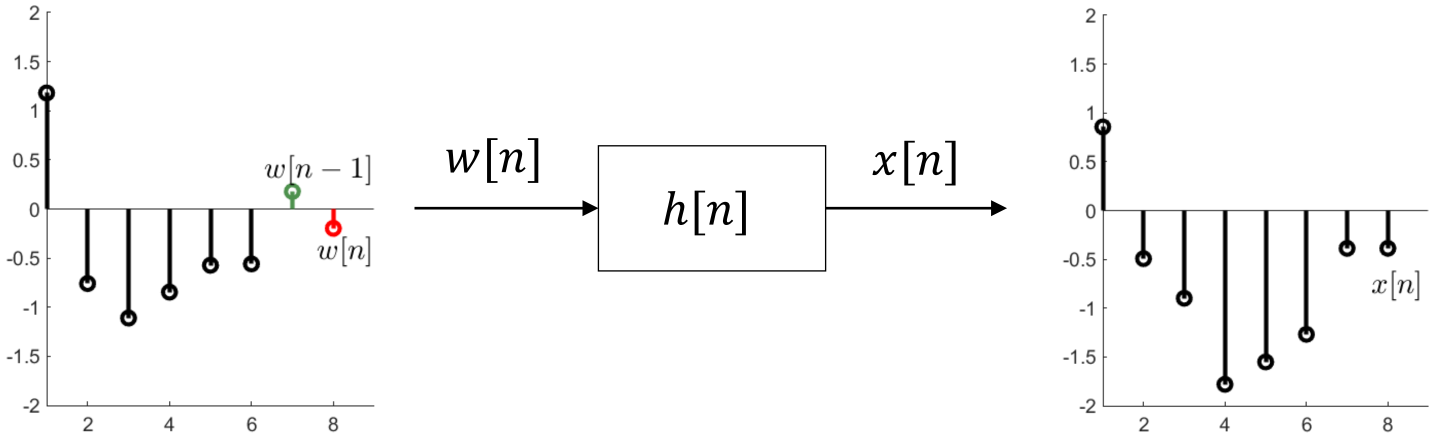
MA $$x[n] \longrightarrow \boxed{H(z)=\frac{B(z)}{1}} \longrightarrow y[n]$$

TU/e

MA models: Difference equation

$$x[n] = \underbrace{w[n]}_{\text{unpredictable part}} + \underbrace{b_1 w[n-1] + b_2 w[n-2] + \ldots + b_q w[n-q]}_{\text{Linear combination of past \textbf{input} samples}}$$

MA(1): $x[n] = w[n] + b_1 w[n-1]$



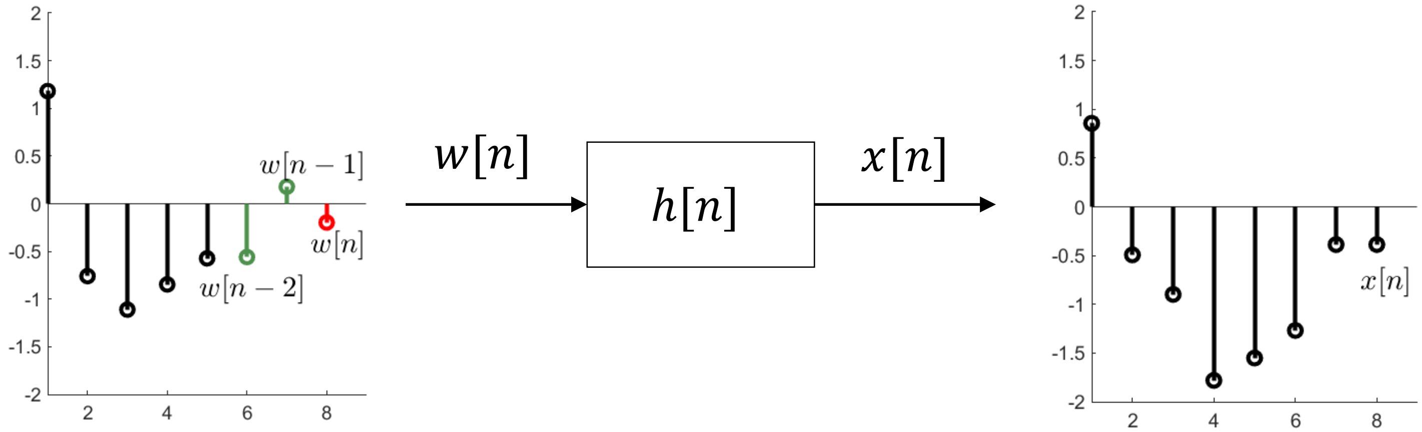
MA models: Difference equation

$$x[n] = w[n] + \underbrace{b_1 w[n-1] + b_2 w[n-2] + \ldots + b_q w[n-q]}_{\text{Linear combination of past input samples}}$$

unpredictable part

*Linear combination of past **input** samples*

MA(2): $x[n] = w[n] + b_1 w[n-1] + b_2 w[n-2]$



MA models: Autocorrelation

$$r_x[l] = \sigma_w^2 \sum_{k=|l|}^{q} b_k b_{k-|l|}$$

Model parameters $\begin{cases} \sigma_w^2 \\ b_1, \ldots, b_q \end{cases}$

Autocorrelation of MA models has finite length determined by the number of zeros

$$r_x[l] = 0 \quad \text{for} \quad |l| > Q$$

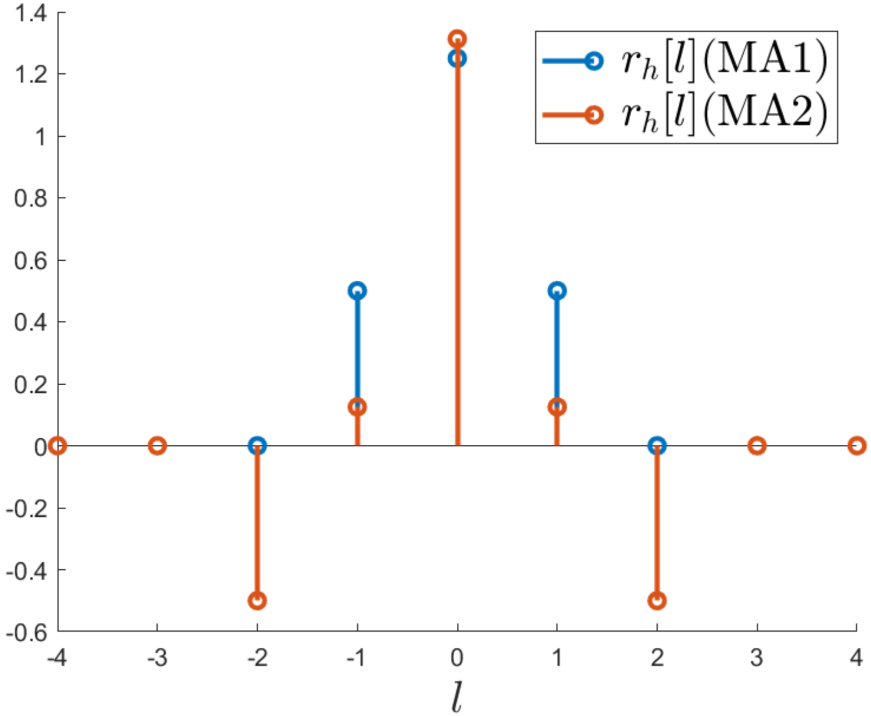
Length: $2Q + 1$

MA models: short-range correlation

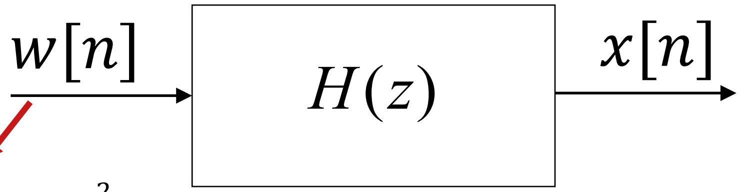
MA(1): $b_1 = \frac{1}{2}$

MA(2): $b_1 = \frac{1}{4}$; $b_2 = -\frac{1}{2}$

System autocorrelation



MA models: Power spectral density



$$P_y(e^{j\theta}) = |H(e^{j\theta})|^2 \, P_x(e^{j\theta})$$

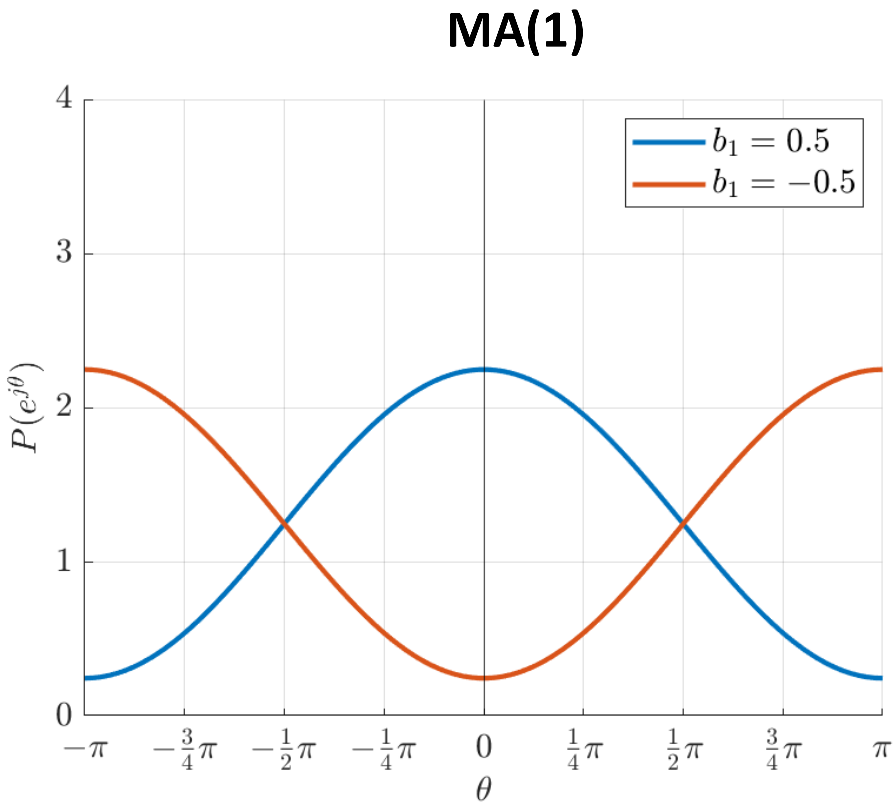
$$P_w(e^{j\theta}) = \sigma_w^{\ 2}$$

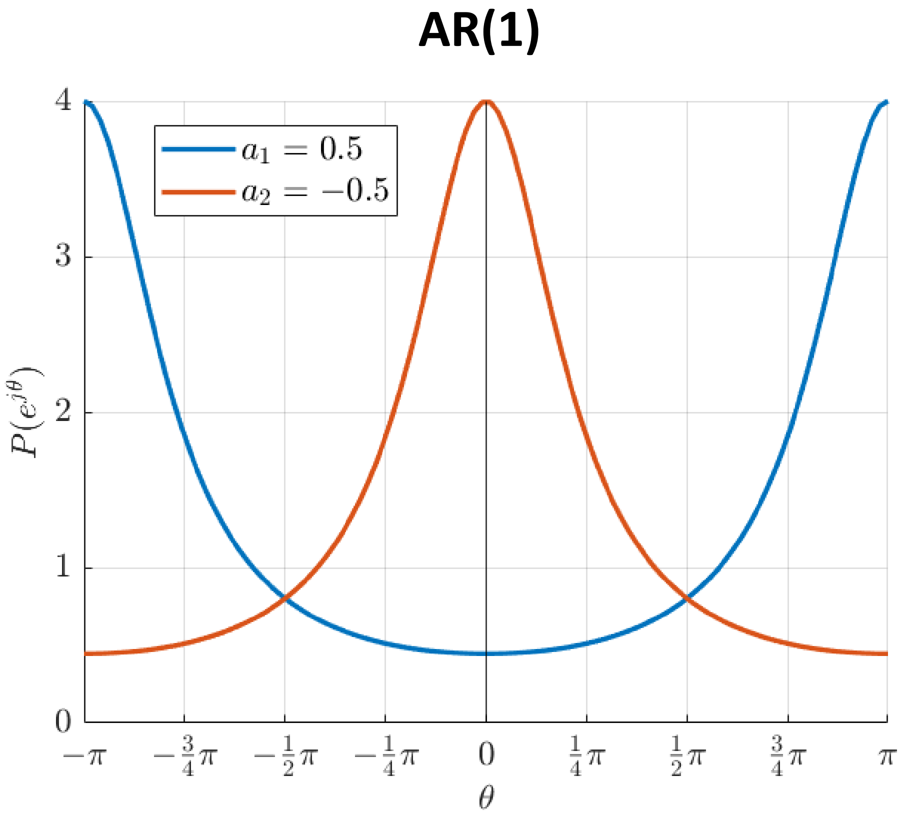
Innovation filter: $b_0 = l_0 = 1$

$$H(z) = \frac{B(z)}{A(z)} = b_0 + b_1 z^{-1} + \ldots + b_Q z^{-Q} = 1 + b_1 z^{-1} + \ldots + b_Q z^{-Q}$$

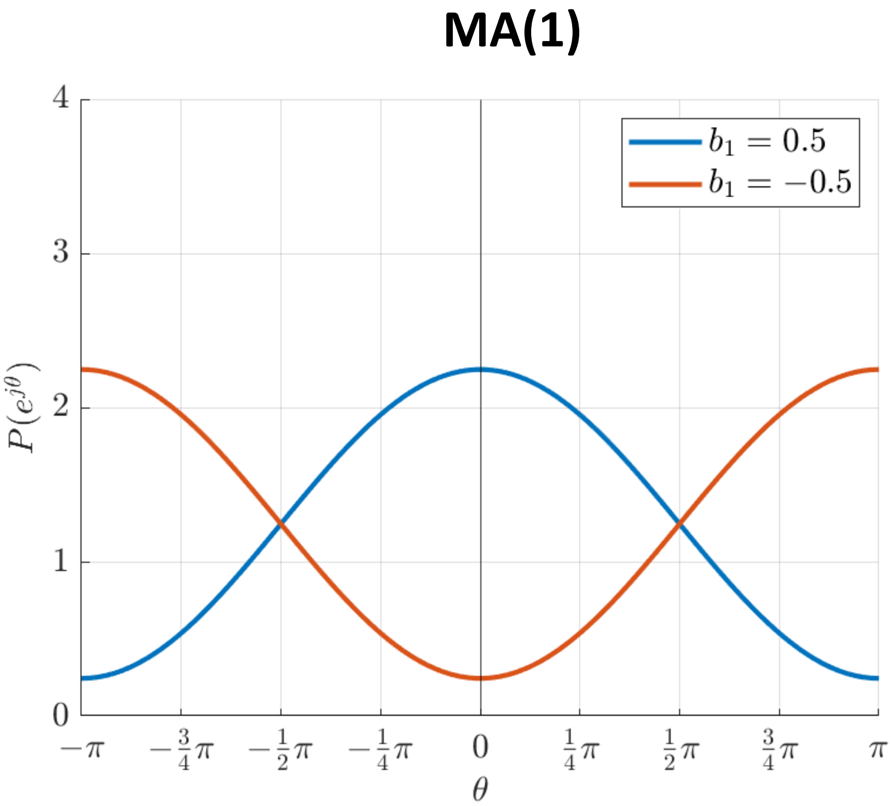
$$\boxed{P_x(e^{j\theta}) = \sigma_w^{\ 2} \left| 1 + \sum_{k=1}^{Q} b_k e^{-jk\theta} \right|^2}$$

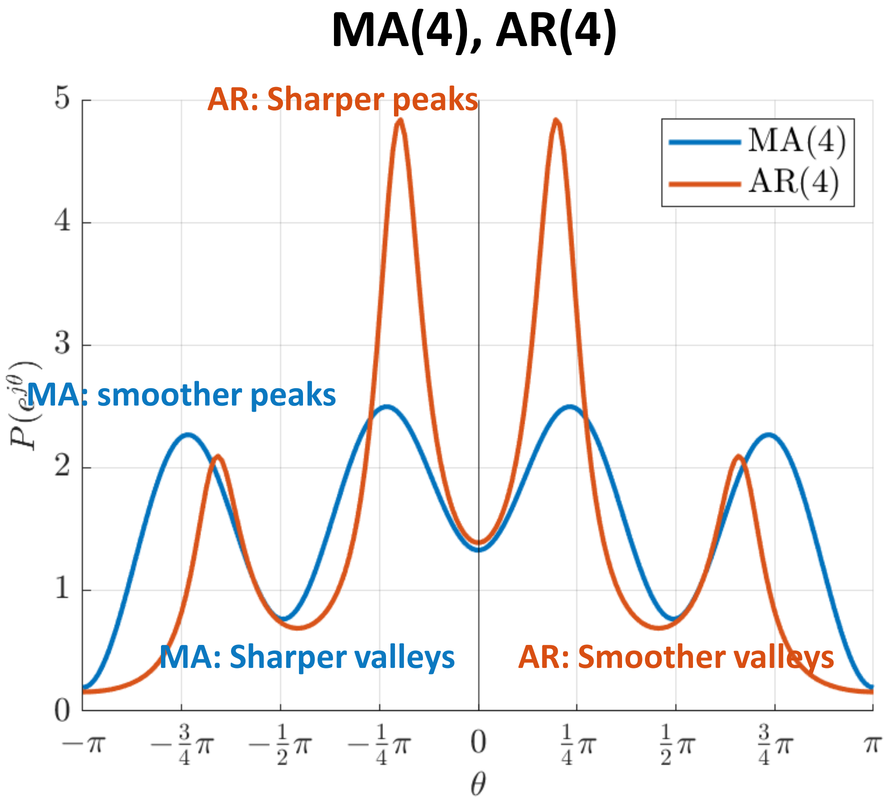
MA models: PSD examples





MA models: PSD examples





Autoregressive moving average (ARMA) models

Autoregressive **Moving Average**

Models a "memory" that decades with time

Smooths the input by linear combination of the present and previous samples

$$x[n] = w[n] + \underbrace{b_1 w[n-1] + b_2 w[n-2] + \ldots + b_q w[n-q]}_{\text{Linear combination of past input samples}} \underbrace{- a_1 x[n-1] - a_2 x[n-2] - \ldots - a_p x[n-p]}_{\text{Linear combination of past output samples}}$$

unpredictable part

*Linear combination of past **input** samples*

*Linear combination of past **output** samples*

TU/e

ARMA: Autocorrelation function

modified Yule-Walker equations

$$r_x[l] = \begin{cases} \sigma_w^2 \sum_{k=|l|}^{q} b_k h[k-|l|] - \sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } 0 \leq |l| \leq q \\ -\sum_{k=1}^{p} a_k r_x[|l|-k]. & \text{for } |l| > q \end{cases}$$

Model parameters

$$\begin{cases} \sigma_w^2 \\ a_{1,} \ldots, a_p \\ b_1, \ldots, b_q \end{cases}$$

ARMA: Autocorrelation function

(Modified) Yule-Walker equations

$$r_x[l] = \begin{cases} \underbrace{\sigma_w^2 \sum_{k=|l|}^{q} b_k h[k-|l|]}_{\text{MA}} - \underbrace{\sum_{k=1}^{p} a_k r_x[|l|-k]}_{\text{AR}} & \text{for } \underline{0 \leq |l| \leq q} \\ -\sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } |l| > q \end{cases}$$

Model parameters $\begin{cases} \sigma_w^2 \\ a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{cases}$

For lags up to the number of zeros, both MA and AR contribute to AC function

ARMA: Autocorrelation function

modified Yule-Walker equations

$$r_x[l] = \begin{cases} \sigma_w^2 \sum_{k=|l|}^{q} b_k h[k-|l|] - \sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } 0 \le |l| \le q \\ \underbrace{-\sum_{k=1}^{p} a_k r_x[|l|-k]}_{\text{AR}} & \text{for } |l| > q \end{cases}$$

Model parameters

$$\begin{cases} \sigma_w^2 \\ a_{1,} \dots, a_p \\ b_1, \dots, b_q \end{cases}$$

For lags larger than the number of zeros, only AR contributes to AC function

$\downarrow$

recursive (decaying) structure

ARMA: Autocorrelation function

modified Yule-Walker equations

$$r_x[l] = \begin{cases} \sigma_w^2 \sum_{k=|l|}^{q} b_k h[k-|l|] - \sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } 0 \leq |l| \leq q \\ -\sum_{k=1}^{p} a_k r_x[|l|-k] & \text{for } |l| > q \end{cases}$$

Model parameters

$$\begin{cases} \sigma_w^2 \\ a_{1,} \dots, a_p \\ b_1, \dots, b_q \end{cases}$$

$h[k]$, impulse response of the system

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_Q z^{-Q}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_P z^{-P}}$$

ARMA: Autocorrelation function

By polynomial division…

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_Q z^{-Q}}{1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_P z^{-P}} = 1 + h_1 z^{-1} + h_2 z^{-2} + \ldots$$

$$\Big\Downarrow \text{ IZT}$$

Impulse response

$$h[k] = \delta[k] + h_1 \delta[k-1] + h_2 \delta[k-2] + \ldots$$

MA: finite
AR, ARMA: infinite

$\longrightarrow$

MA: FIR filter
AR, ARMA: IIR filter

ARMA: Autocorrelation function

$$r_x[l] = \begin{cases} \sigma_w^2 \sum_{k=|l|}^{q} b_k h[k - |l|] - \sum_{k=1}^{p} a_k r_x[|l| - k] & \text{for } 0 \le |l| \le q \\ -\sum_{k=1}^{p} a_k r_x[|l| - k] & \text{for } |l| > q \end{cases}$$

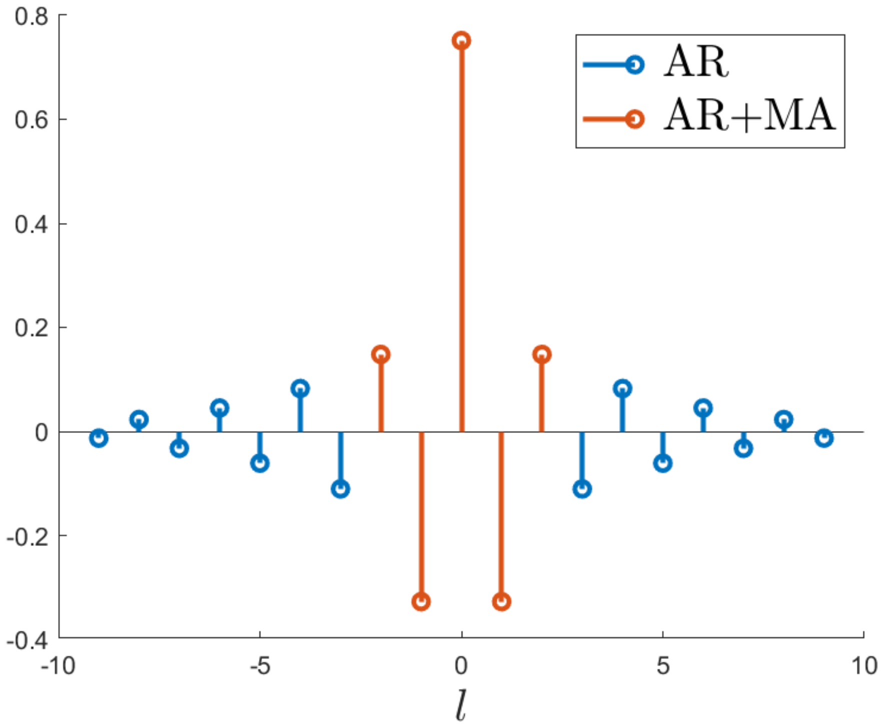
<u>AC: $h[k]$ coefficients needed only for number of lags equal to number of zeros</u>

ARMA models: system autocorrelation

ARMA(1,2):

$$b_1 = \frac{1}{4}, b_2 = -\frac{1}{4}, a_1 = \frac{3}{4}$$

System autocorrelation



ARMA models: Power spectral density

$$w[n] \longrightarrow \boxed{H(z)} \longrightarrow x[n]$$

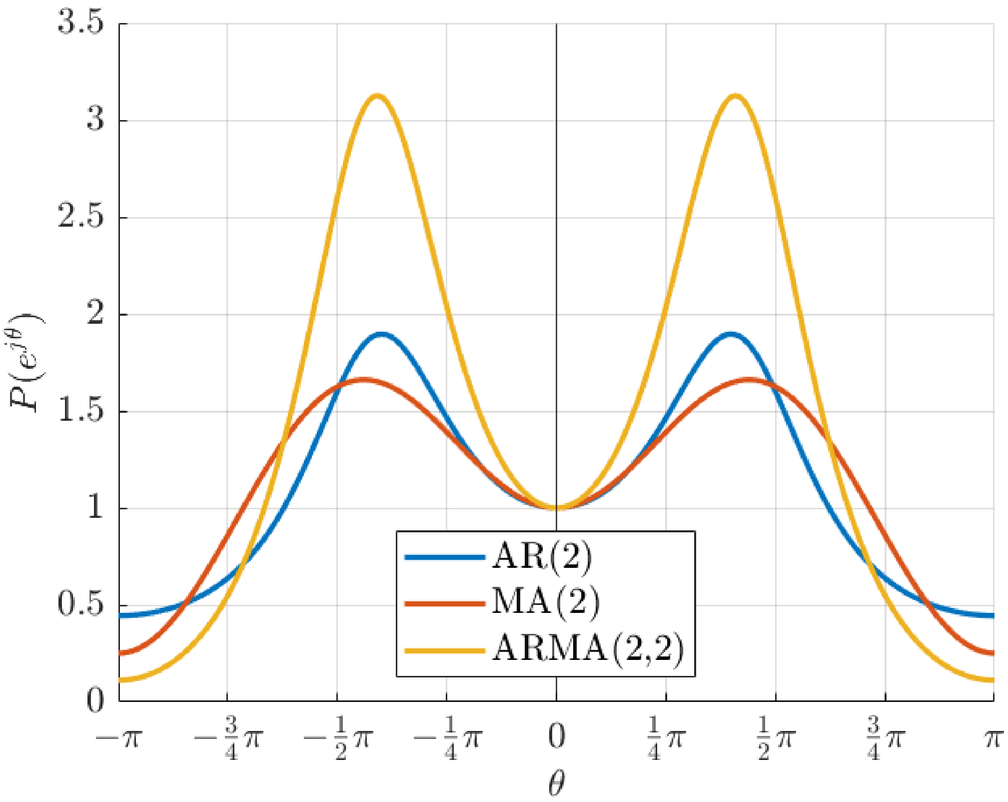
$$P_y(e^{j\theta}) = |H(e^{j\theta})|^2 \, P_x(e^{j\theta})$$

$$P_w(e^{j\theta}) = \sigma_w{}^2$$

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_Q z^{-Q}}{1 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_P z^{-P}}$$

$$\boxed{P_x(e^{j\theta}) = \sigma_w^2 \frac{|1 + b_1 e^{-j\theta} + \ldots + b_q e^{-jq\theta}|^2}{|1 + a_1 e^{-j\theta} + \ldots + a_p e^{-jp\theta}|^2}}$$

ARMA models: PSD examples



Wrap-up (I)

- Linear-time invariant systems are described in the time domain by the impulse response, $h[n]$, and in the z-domain by the transfer function, $H(z)$
- Systems represented by a rational $H(z)$ are described by their poles and zeros, from which properties as causality, stability, and minimum-phase can be easily determined
- For deterministic signals, the output of a LTI system can be fully determined by the input-output relationships in the time- and z-domains
- For stochastic signals, input-output relationships can be found for the first and second order statistics (focus on WSS)

Wrap-up (II)

- Spectral factorization allows to model any WSS process as a LTI driven by white noise, based on its second-order statistics
- When we choose a rational polynomial for the LTI system function, the resulting process is auto-regressive moving average
- Autoregressive models use a linear combination of past output samples: useful to model a variety of time series
- AR autocorrelation function shows a decaying memory structure
- AR power spectral density is useful to model sharp peaks

Wrap-up (III)

- Moving average models use a linear combination of past input samples: they filter out short-term fluctuations
- MA autocorrelation has finite length determined by the order of the model (number of zeros)
- MA power spectral density presents smoother peaks and sharper valleys
- Autoregressive moving-average models present the characteristics of both AR and MA models: more flexible

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Statistical signal processing (5CTA0)

Lecture 3

Lecturer: Simona Turco

Electrical Engingeering, Signal Processing Systems group