EE3731C: Signal Processing Methods

Lecture II-5: Linear Stochastic Models

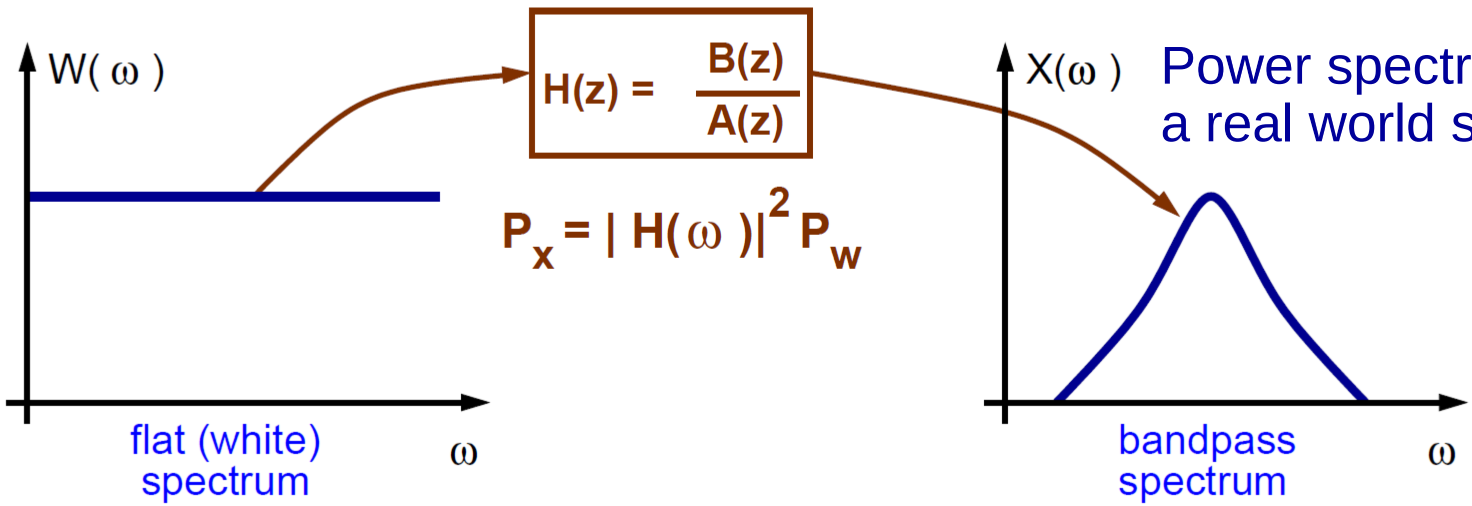
National University of Singapore

Outline

- Motivation: Wold Decomposition Theorem
- Linear Stochastic Processes
  - Autoregressive moving average (ARMA) process
  - Autoregressive (AR) process
  - Moving average (MA) process
- Whitening Filter
- Linear Prediction

How to Model a Real World Signal?

- Model first and second order statistics of a real world signal by shaping the white noise spectrum using some transfer function
  - Can describe a long signal with very few parameters
  - Can use the linear stochastic model for prediction



Wold Decomposition Theorem

- A general random process can be written as a sum of two processes

$$x[n] = x_p[n] + x_r[n]$$

$\Rightarrow x_r[n]$ – regular random process
$\Rightarrow x_p[n]$ – predictable process, with $x_r[n] \perp x_p[n]$,

$E\{x_r[m]x_p[n]\} = 0$ orthogonality

- The predictable process (i.e., a deterministic signal) and the random process (i.e., a random signal) can be treated separately.

Linear Stochastic Processes

- Random processes generated by filtering white noise with a linear shift invariant filter that has a rational transfer function
  - –Autoregressive moving average (ARMA) process
  - –Autoregressive (AR) process (all pole)
  - –Moving average (MA) process (all zero)

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

$$R_w[m] = \sigma_w^2 \delta[m]$$

$$P_w\left(e^{j\omega}\right) = \sigma_w^2$$

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{\sum_{k=0}^{q} b_q[k] z^{-k}}{1 + \sum_{k=1}^{p} a_p[k] z^{-k}}$$

Consider real case

ARMA Process

- The power spectrum of the output $x[n]$ is

$$P_x(z) = \sigma_w^2 \frac{B_q(z)B_q(z^{-1})}{A_p(z)A_p(z^{-1})} \qquad P_x(e^{j\omega}) = \sigma_w^2 \frac{\left|B_q(e^{j\omega})\right|^2}{\left|A_p(e^{j\omega})\right|^2}$$

  – H(z) shapes the spectrum of white noise

  – known as an autoregressive moving average process of order (p, q) and is referred to as an ARMA(p,q) process

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

$$R_w[m] = \sigma_w^2 \delta[m]$$

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

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{\sum_{k=0}^{q} b_q[k] z^{-k}}{1 + \sum_{k=1}^{p} a_p[k] z^{-k}}$$

ARMA: Difference Equation

- In the time domain, $x[n]$ and $w[n]$ are related by

$$x[n]+\sum_{k=1}^{p} a_p[k]x[n-k]=\sum_{k=0}^{q} b_q[k]w[n-k]$$

- Multiplying both sides by $x[n-m]$ and take expectation, we have

$$x[n]x[n-m]+\sum_{k=1}^{p} a_p[k]x[n-k]x[n-m]=\sum_{k=0}^{q} b_q[k]w[n-k]x[n-m]$$

$$R_x[-m]+\sum_{k=1}^{p} a_p[k]R_x[-m+k]=\sum_{k=0}^{q} b_q[k]R_{wx}[-m+k]$$

H(z) is causal $\quad R_{wx}[l]=\begin{cases} 0 & l<0 \\ \sigma_w^2 h[l] & l\geq 0 \end{cases}$

ARMA: Auto-correlation Sequence

$$R_x[-m] + \sum_{k=1}^{p} a_p[k] R_x[-m+k] = \sum_{k=0}^{q} b_q[k] R_{wx}[-m+k]$$

$$R_{wx}[l] = \begin{cases} 0 & l < 0 \\ \sigma_w^2 h[l] & l \geq 0 \end{cases}$$

$$R_x[m] + \sum_{k=1}^{p} a_p[k] R_x[m-k] = \begin{cases} 0 & m > q \\ \sigma_w^2 \sum_{k=m}^{q} b_q[k] h[k-m] & 0 \leq m \leq q \end{cases}$$

For lags $m = 0,\ldots,L$, we can write a set of $L+1$ equations in matrix-vector form.

ARMA Process: Example

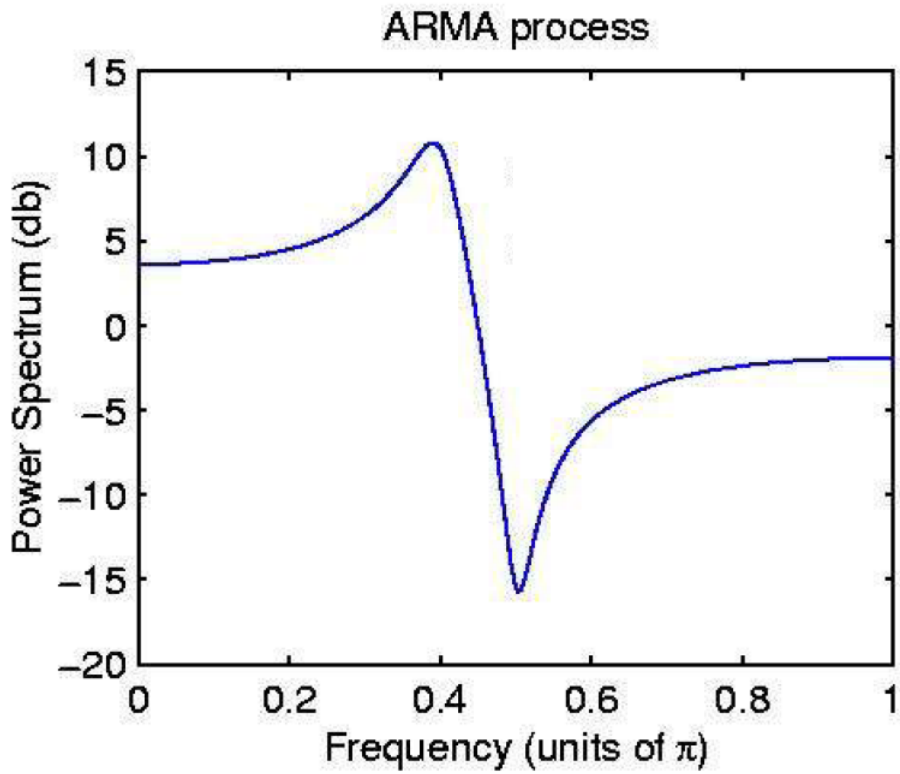
The transfer function of an ARMA(2,2) process is:

$$H(z) = \frac{1 + 0.9025z^{-2}}{1 - 0.5562z^{-1} + 0.81z^{-2}}$$

Zeros: $z = 0.95e^{\pm j\pi/2}$

Poles: $z = 0.9e^{\pm j2\pi/5}$

ARMA is used to model PSD with both peaks and valleys.



AR Process

- An autoregressive (AR) process is a special type of an ARMA process for $b_q[0] = 1$; $b_q[k] = 0$; $k > 0$
- The power spectrum of the output $x[n]$ is

$$P_x(z) = \sigma_w^2 \frac{1}{A_p(z)A_p(z^{-1})} \qquad P_x(e^{j\omega}) = \sigma_w^2 \frac{1}{\left|A_p(e^{j\omega})\right|^2}$$

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

$$R_w[m] = \sigma_w^2 \delta[m]$$

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

$$H(z) = \frac{1}{1 + \sum_{k=1}^{p} a_p[k] z^{-k}}$$

AR: Difference Equation

- In the time domain, $x[n]$ and $w[n]$ are related by

$$x[n]+\sum_{k=1}^{p} a_p[k]x[n-k]=w[n]$$

- Multiplying both sides by $x[n\text{-}m]$ and take expectation, we have

$$R_x[-m]+\sum_{k=1}^{p} a_p[k]R_x[-m+k]=R_{wx}[-m]$$

$$R_x[m]+\sum_{k=1}^{p} a_p[k]R_x[m-k]=R_{wx}[-m]$$

AR: Auto-correlation Sequence

$$x[n] = -\sum_{k=1}^{p} a_p[k]x[n-k] + w[n]$$

$$R_x[m] + \sum_{k=1}^{p} a_p[k]R_x[m-k] = R_{wx}[-m] = \begin{cases} \sigma_w^2 & m = 0 \\ 0 & m > 0 \end{cases}$$

For lags $m = 0,\dots,p$, we can write a set of $p+1$ equations in matrix form.

$$\begin{bmatrix} R_x[0] & R_x[1] & \cdots & R_x[p] \\ R_x[1] & R_x[0] & & R_x[p-1] \\ \vdots & \vdots & \ddots & \vdots \\ R_x[p] & R_x[p-1] & \cdots & R_x[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_p[1] \\ \vdots \\ a_p[p] \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

AR Modelling

$$\begin{bmatrix} R_x[0] & R_x[1] & \cdots & R_x[p] \\ R_x[1] & R_x[0] & & R_x[p-1] \\ \vdots & \vdots & \ddots & \vdots \\ R_x[p] & R_x[p-1] & \cdots & R_x[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_p[1] \\ \vdots \\ a_p[p] \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

If we omit the first equation, we get

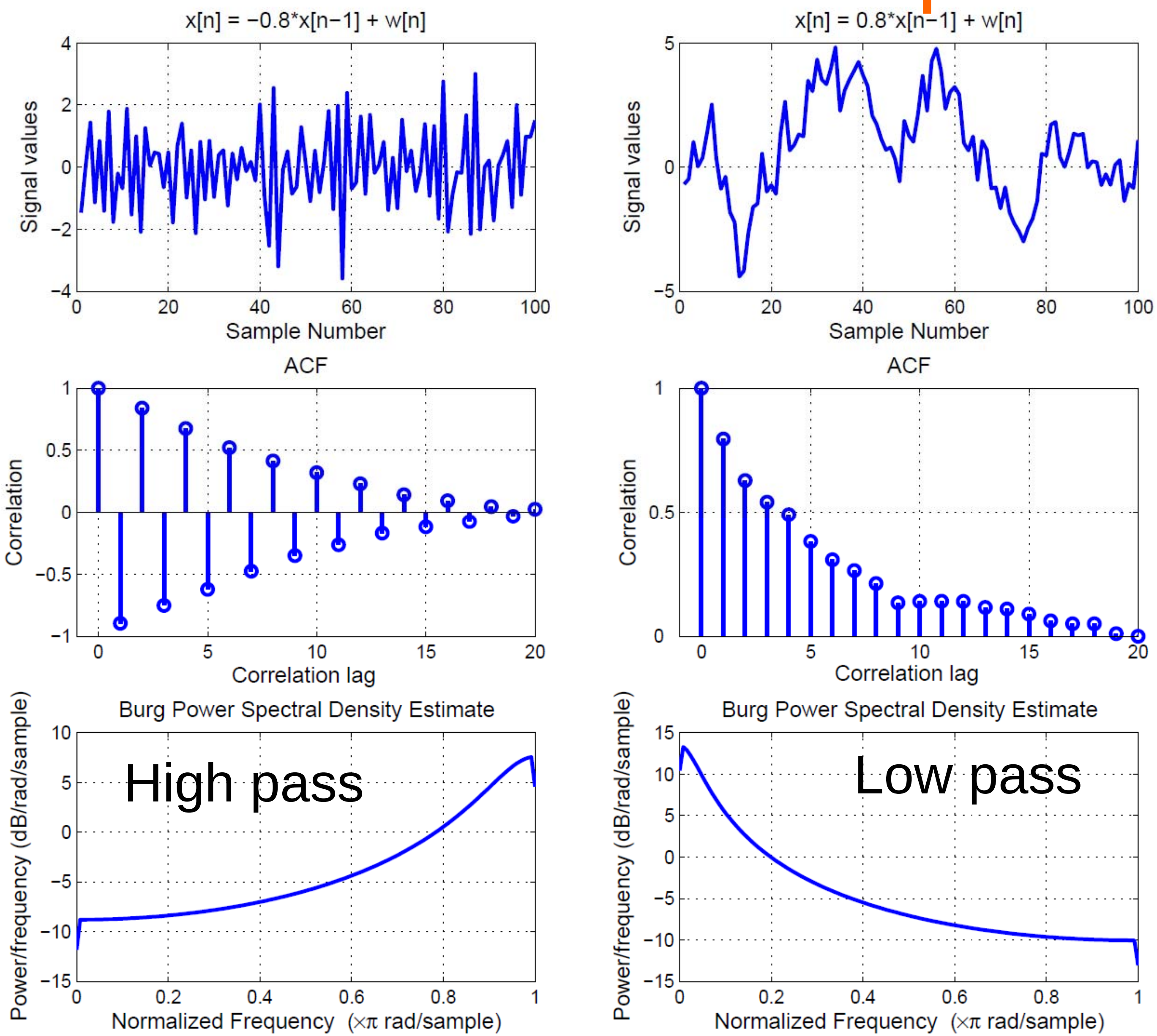
$$\begin{bmatrix} R_x[0] & R_x[1] & \cdots & R_x[p-1] \\ R_x[1] & R_x[0] & & R_x[p-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_x[p-1] & R_x[p-2] & \cdots & R_x[0] \end{bmatrix} \begin{bmatrix} a_p[1] \\ a_p[2] \\ \vdots \\ a_p[p] \end{bmatrix} = -\begin{bmatrix} R_x[1] \\ R_x[2] \\ \vdots \\ R_x[p] \end{bmatrix}$$

Yule-Walker equations in matrix-vector notation:

$$\mathbf{R}_x \mathbf{a} = -\mathbf{r}_x$$

$$\mathbf{a} = -\mathbf{R}_x^{-1} \mathbf{r}_x$$

AR Process: Example

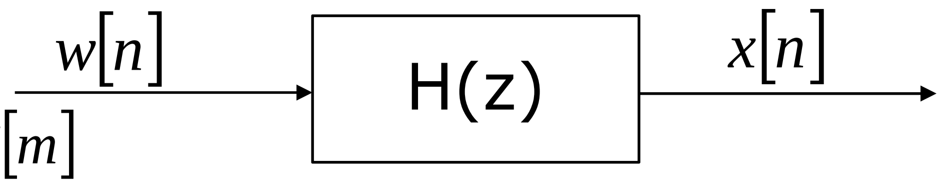


MA Process

- An moving average (MA) process is a special type of an ARMA process for $p=0$.

- The power spectrum of the output $x[n]$ is

$$P_x(z)=\sigma_w^2 B_q(z) B_q\left(z^{-1}\right) \qquad P_x\left(e^{j\omega}\right)=\sigma_w^2\left|B_q\left(e^{j\omega}\right)\right|^2$$



$$R_w[m]=\sigma_w^2\delta[m]$$

$$P_w\left(e^{j\omega}\right)=\sigma_w^2$$

$$H(z)=\sum_{k=0}^{q} b_q[k] z^{-k}$$

MA: Difference Equation

- In the time domain, $x[n]$ and $w[n]$ are related by

$$x[n] = \sum_{k=0}^{q} b_q[k] w[n-k]$$

- The impulse response is

$$h[n] = \sum_{k=0}^{q} b_q[k] \delta[n-k]$$

- Multiplying both sides by $x[n-m]$ and take expectation, we have

$$R_x[-m] = \sum_{k=0}^{q} b_q[k] R_{wx}[-m+k]$$

MA: Auto-correlation Sequence

$$R_x[-m] = \sum_{k=0}^{q} b_q[k] R_{wx}[-m+k]$$

$$R_x[m] = \sum_{k=0}^{q} b_q[k] R_{wx}[k-m]$$

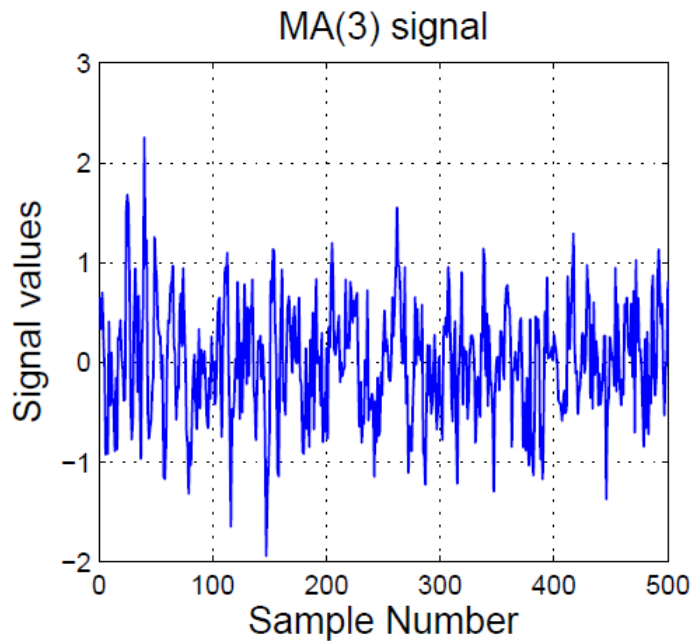
$$R_{wx}[l] = E\left(w[n]x[n+l]\right)$$

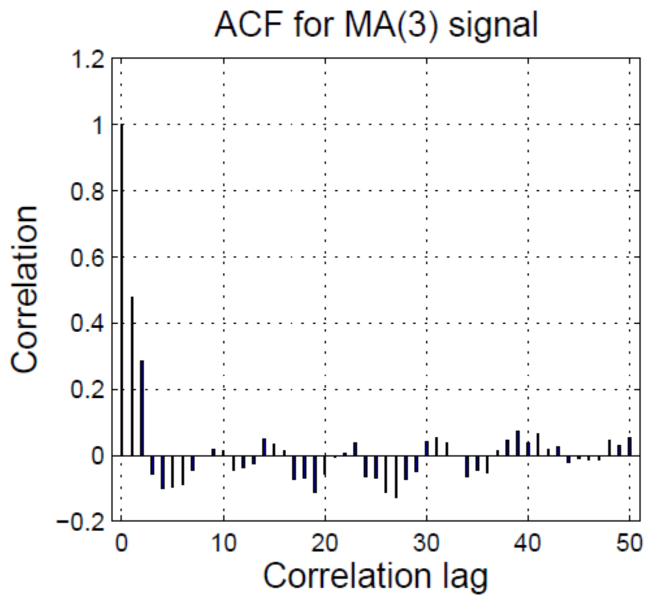
$$= E\left(w[n] \sum_{k=0}^{q} b_q[k] w[n+l-k] \right) = \begin{cases} 0 & l < 0 \\ b_q[l]\sigma_w^2 & 0 \le l \le q \end{cases}$$

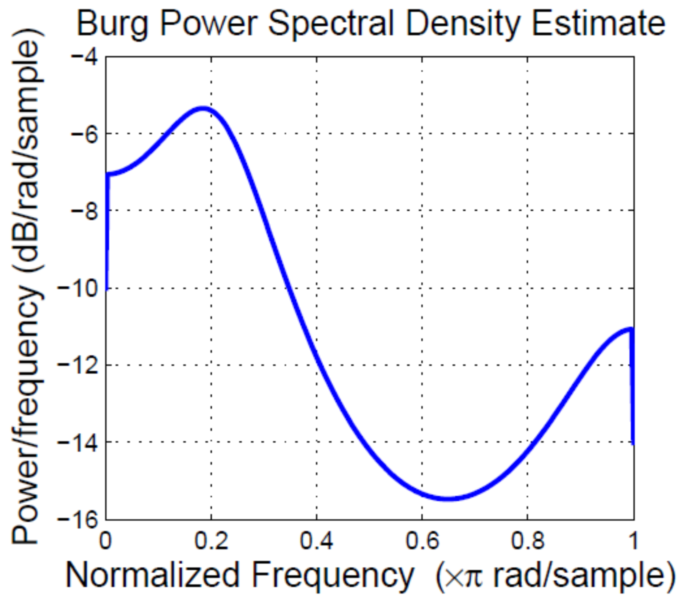
$$R_x[m] = \sigma_w^2 \sum_{k=0}^{q} b_q[k] b_q[k-m] = \sigma_w^2 \sum_{k=0}^{q-|m|} b_q[k] b_q[k+|m|]$$

The variance of the process: $R_x[0] = \sigma_w^2 \sum_{k=0}^{q} \left(b_q[k]\right)^2$

MA Process: Example







- MA models valley PSD better
- All zeros: struggle to model PSD with peaks

Duality between AR and MA Processes

- A stationary finite AR(p) process can be represented as an infinite order MA process.

$$x[n] = a_1 x[n-1] + w[n] \quad \Leftrightarrow \quad \sum_{j=0}^{\infty} b_j w[n-j]$$

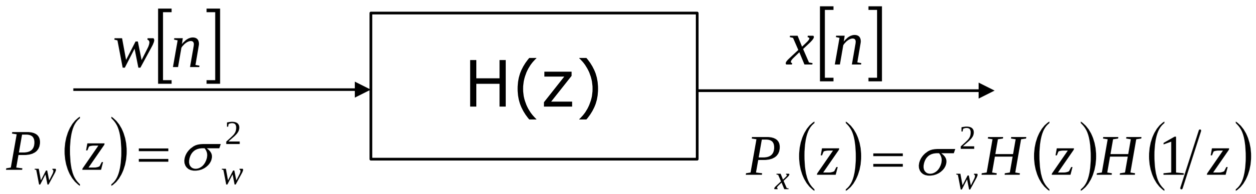
- A finite MA process can be represented as an infinite AR process.

$$\sum_{k=0}^{M} b_k z^{-k} = \frac{1}{A_\infty(z)}$$

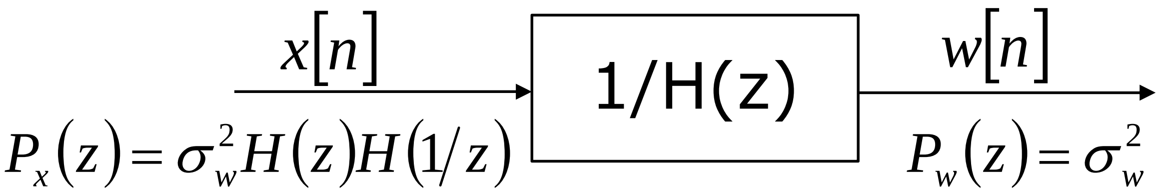
Follows the duality between IIR and FIR filters

Whitening Filter

- A random process $x[n]$ is represented as the output of a filter H(z) with an input of white noise $w[n]$.



- A whitening filter for the random process $x[n]$ is the inverse of H(z), i.e., 1/H(z).

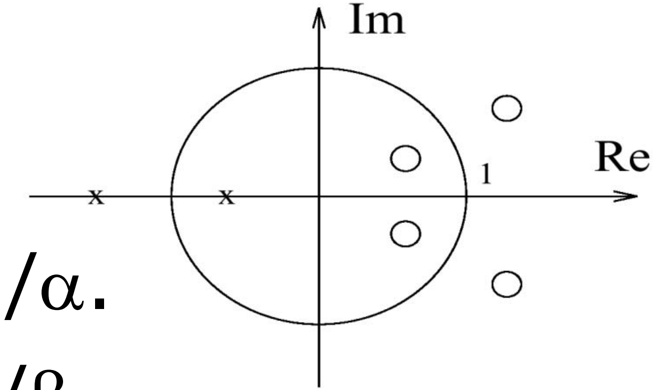


Different random processes have different whitening filters.

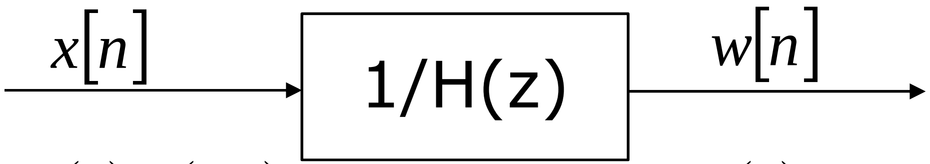
Spectral Factorization and Whitening Filter

- A random process $x[n]$ has a power spectrum

$$P(z) = \sigma^2 \frac{B(z)B(\frac{1}{z})}{A(z)A(\frac{1}{z})}$$



- If α is a zero of P(z), so is $1/\alpha$.
- If β is a pole of P(z), so is $1/\beta$.
- Since all coefficients are real, the poles and zeroes of P(z) occur in complex conjugate pairs.

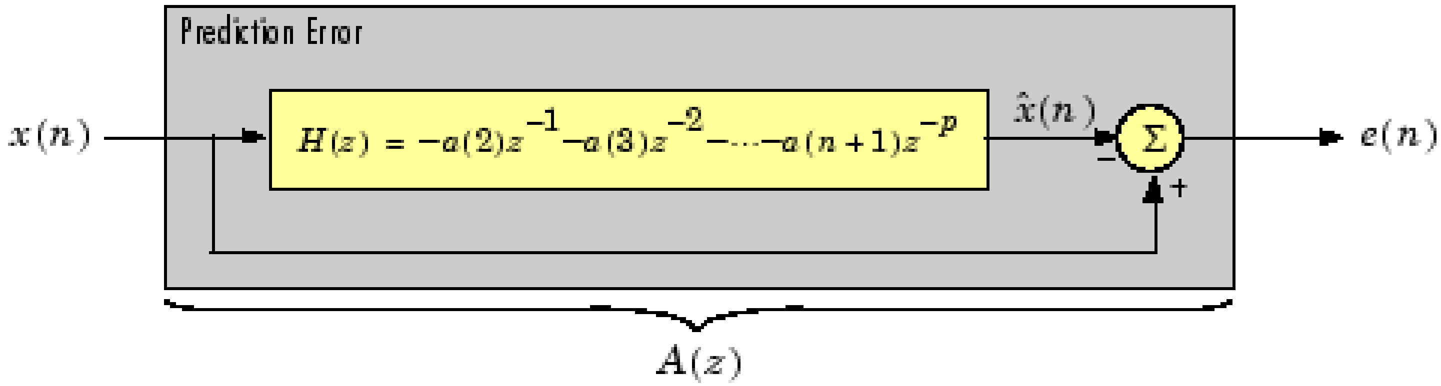


$$P_x(z) = \sigma_w^2 H(z) H(1/z)$$

$$P_w(z) = \sigma_w^2$$

Linear Prediction

- Linear prediction is a mathematical operation where future values of a discrete-time signal are estimated as a linear function of previous samples.
- The goal of linear prediction is to predict sample $x[n]$ from its past p samples.



Linear Prediction with an AR Model

- Assume that a time series can be modelled by a *p*-order model.
  - x[n] is a function of the *p* previous values plus an error term e[n]

$$x[n] = -\sum_{k=1}^{p} a[k]x[n-k] + e[n]$$

  - Prediction: $\hat{x}[n] = -\sum_{k=1}^{p} a[k]x[n-k]$

- The prediction error is: $e[n] = x[n] - \hat{x}[n]$

Minimize the Mean Square Error (MSE), $E\{(e[n])^2\}$ we obtain the same linear equations as the AR model parameters - Yule-Walker equations

Linear Prediction Filter Coefficients

In matlab, lpc determines the coefficients of a forward linear predictor by minimizing the prediction error in the least squares sense.

```
>> a = lpc(x,P);
```

The error process $e[n]$ represents the "new" or "innovative" part over the linear prediction, hence the name innovation process.

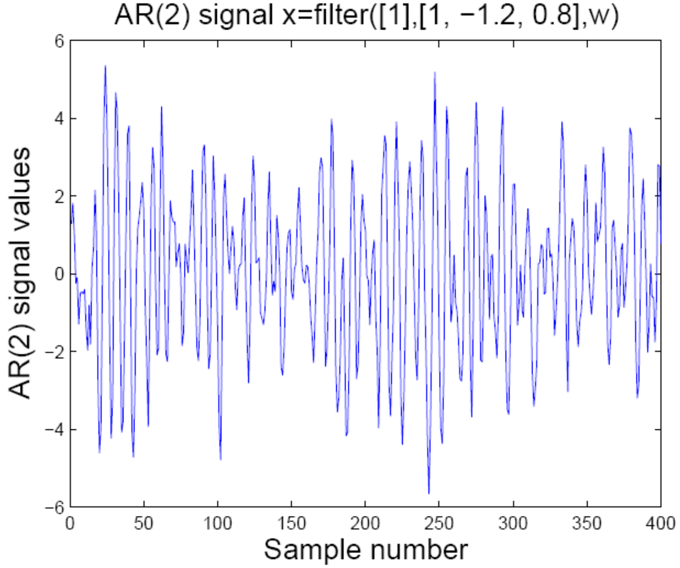
One can show that the innovation $e[n]$ in linear prediction is white and uncorrelated to the prediction $\hat{x}[n]$.

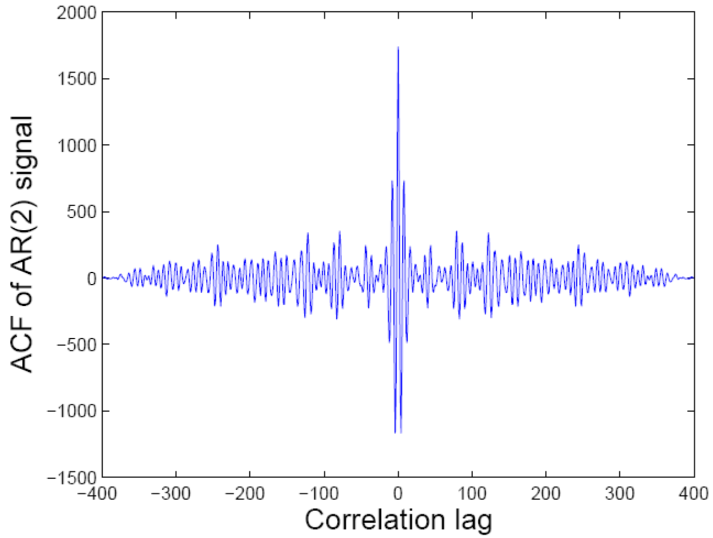
From Data to AR(p) Model

- The auto-correlation are estimated from the data and the Yule-Walker equations are solved to find the AR model parameters.
  - –Record data x(k)
  - –Find the autocorrelation of the data ACF(x)
  - –Write down Yule-Walker equations
  - –Solve for the vector of AR parameters
- Problem: we do not know the model order *p* beforehand.
- Solution: model order selection

Example

AR(2) process: $x[n] = 1.2x[n-1] - 0.8x[n-2] + w[n]$






```
for i=1:6; [a,e]=aryule(x,i); display(a);end
```

$\mathbf{a}^{(1)} = [0.6689]$ $\mathbf{a}^{(2)} = [1.2046, -0.8008]$

$\mathbf{a}^{(3)} = [1.1759, -0.7576, -0.0358]$

$\mathbf{a}^{(4)} = [1.1762, -0.7513, -0.0456, 0.0083]$

$\mathbf{a}^{(5)} = [1.1763, -0.7520, -0.0562, 0.0248, -0.0140]$

$\mathbf{a}^{(6)} = [1.1762, -0.7518, -0.0565, 0.0198, -0.0062, -0.0067]$

Application Example: MA Filters

- Problem:
  - Propose a time-domain technique to remove random noise given only one realization of the signal or the event of interest.
- Solution:
  - Since ensemble of multiple realizations of an event is not available, synchronized averaging is not possible.
  - Temporal averaging for noise removal
  - With the assumption that processes are ergodic, temporal window of samples is moved to obtain output at various points of time: moving-window averaging or moving-average (MA) filter.

MA Filters

- The general form:

$$y[n]=\sum_{k=0}^{N} b_k x[n-k]$$

$y(n)$: Output (filtered signal)

$x(n)$: Input signal

b_k : Filter coefficients

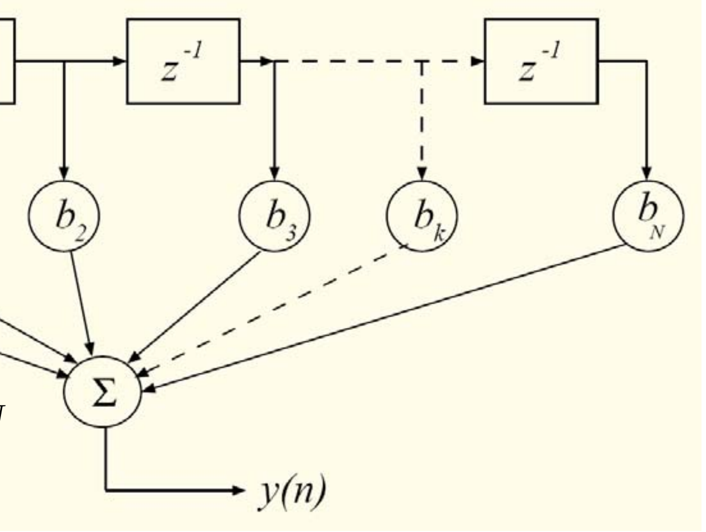
N : order of filter

Applying the z-transform,

$$Y(z)=X(z)\sum_{k=0}^{N} b_k z^{-k}$$

The transfer function is:

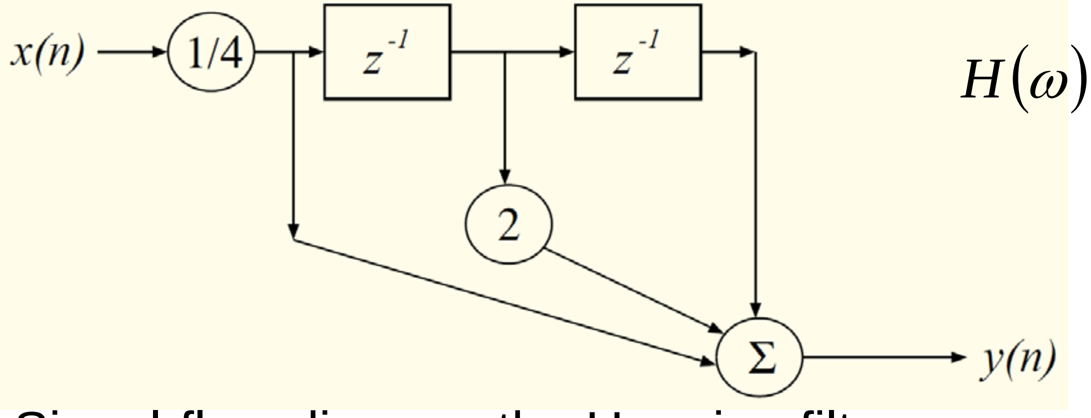
$$H(z)=\frac{Y(z)}{X(z)}=\sum_{k=0}^{N} b_k z^{-k}=b_0+b_1 z^{-1}+\ldots+b_N z^{-N}$$



Signal-flow diagram of a moving-average filter of order N

Hanning Filter

- A simple MA filter: $y(n) = \frac{1}{4}\left(x[n] + 2x[n-1] + x[n-2]\right)$
- Impulse response: $h(n) = \frac{1}{4}\left[\delta(n) + 2\delta(n-1) + \delta(n-2)\right]$
- Transfer function: $H(z) = \frac{1}{4}\left[1 + 2z^{-1} + z^{-2}\right]$
- Frequency response: $H(\omega) = H(z)\big|_{z=e^{j\omega}}$

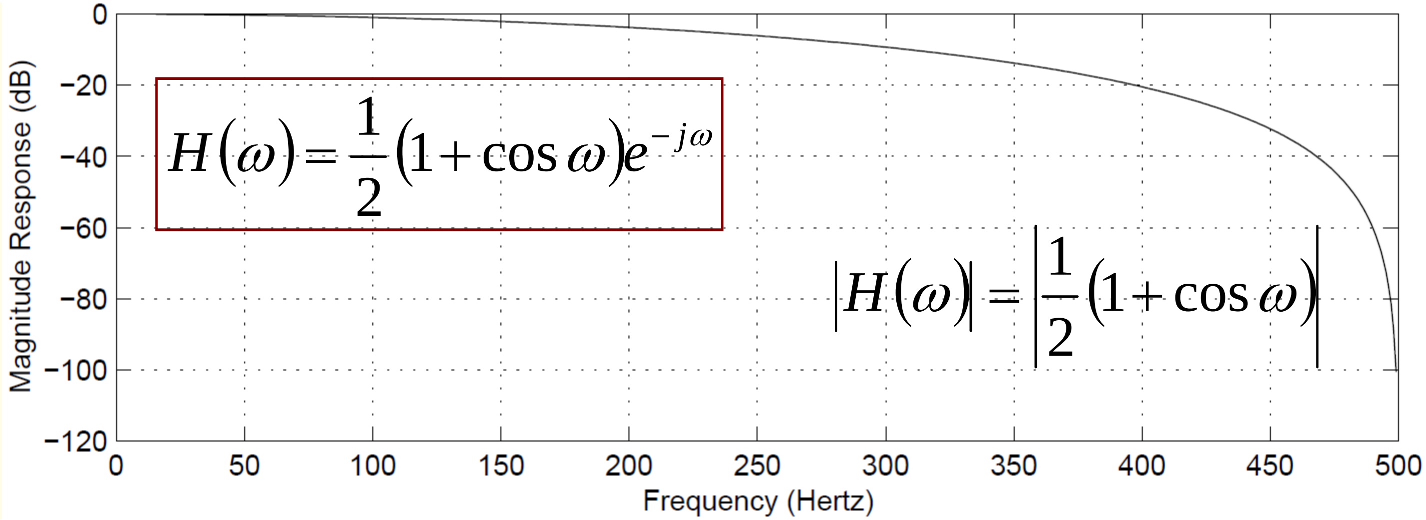


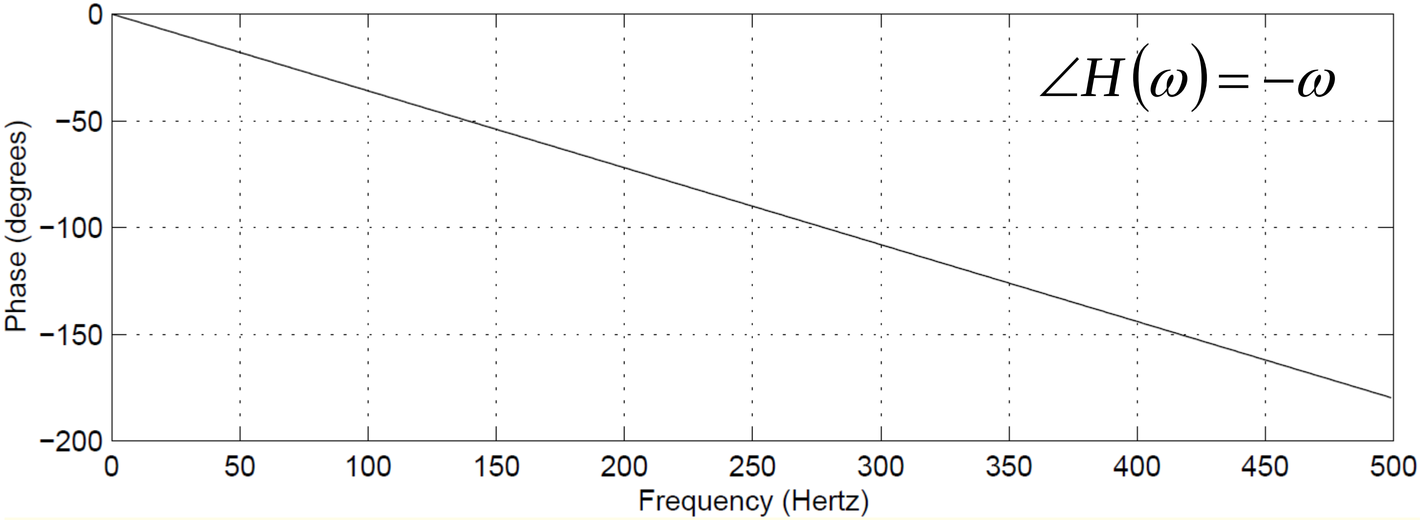
Signal-flow diagram the Hanning filter

$$H(\omega) = \frac{1}{4}\left[1 + 2e^{-j\omega} + e^{-j2\omega}\right]$$

$$H(\omega) = \frac{1}{2}\left(1 + \cos\omega\right)e^{-j\omega}$$

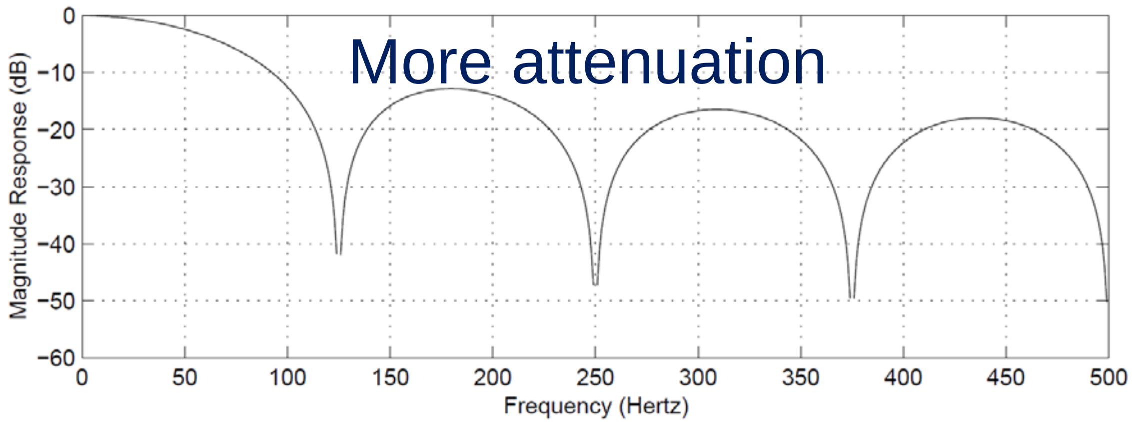
Magnitude and Phase Response of the Hanning Filter





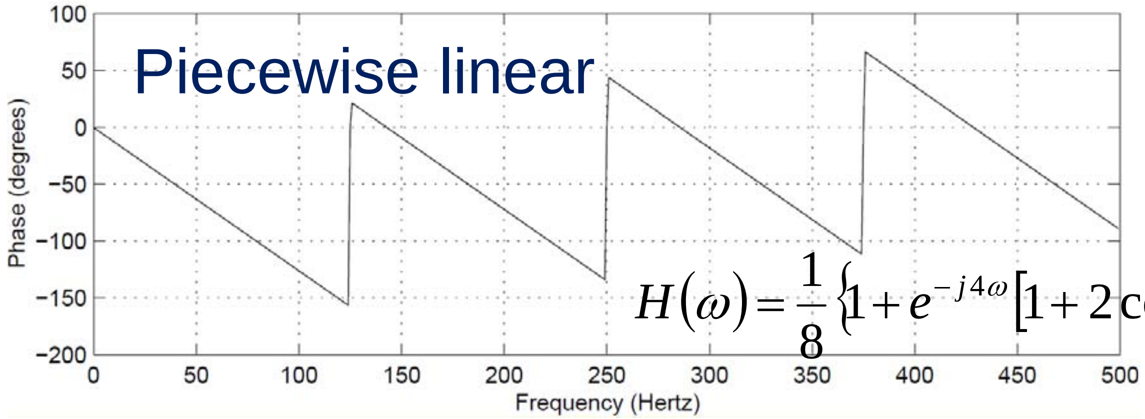
MA Filters

- Increased smoothing is achieved by averaging signal samples over longer time windows, at the expense of increased filter delay.



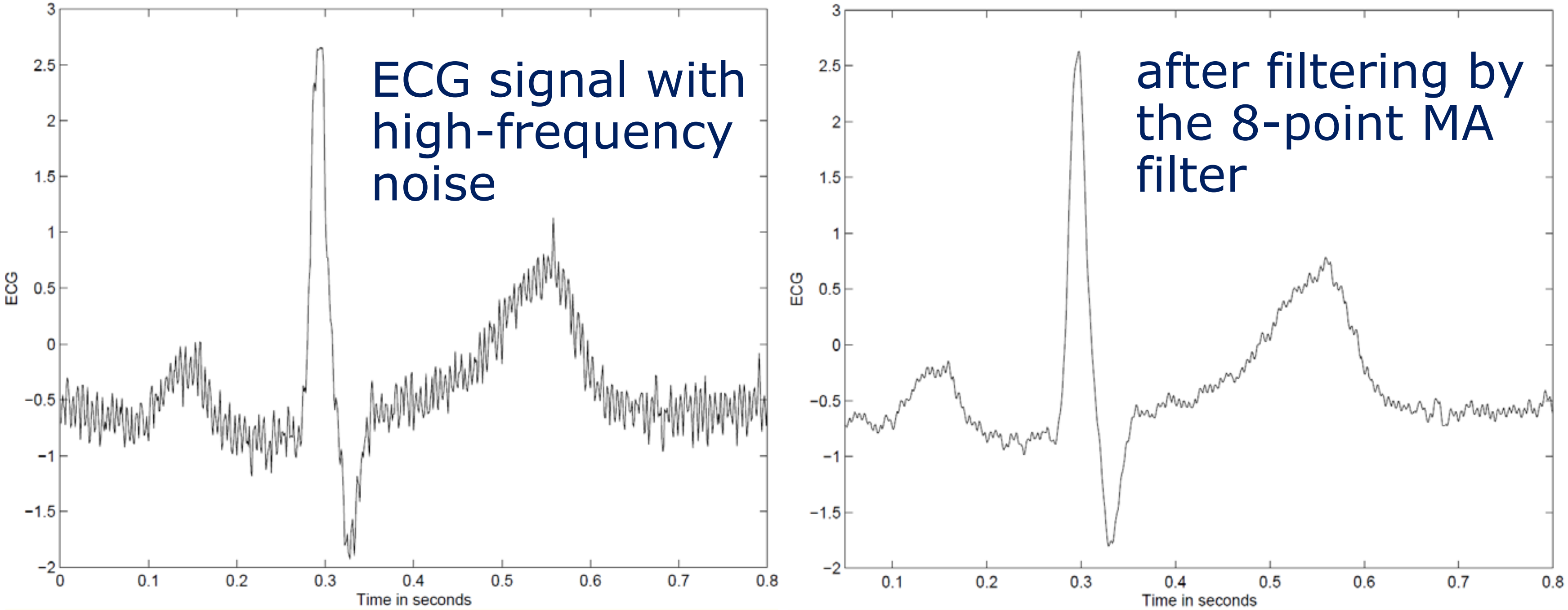
8-point MA (smoothing) filter

$$y(n)=\frac{1}{8}\sum_{k=0}^{7}x(n-k)$$



$$H(\omega)=\frac{1}{8}\left\{1+e^{-j4\omega}\left[1+2\cos\omega+2\cos(2\omega)+2\cos(3\omega)\right]\right\}$$

Example: ECG Signal



- Although the noise level has been reduced, some noise is still present in the result.
- This is because the attenuation of the simple 8-point MA filter is not more than -20dB at most frequencies (except near the zero of the filter)