EE3731C: Signal Processing Methods

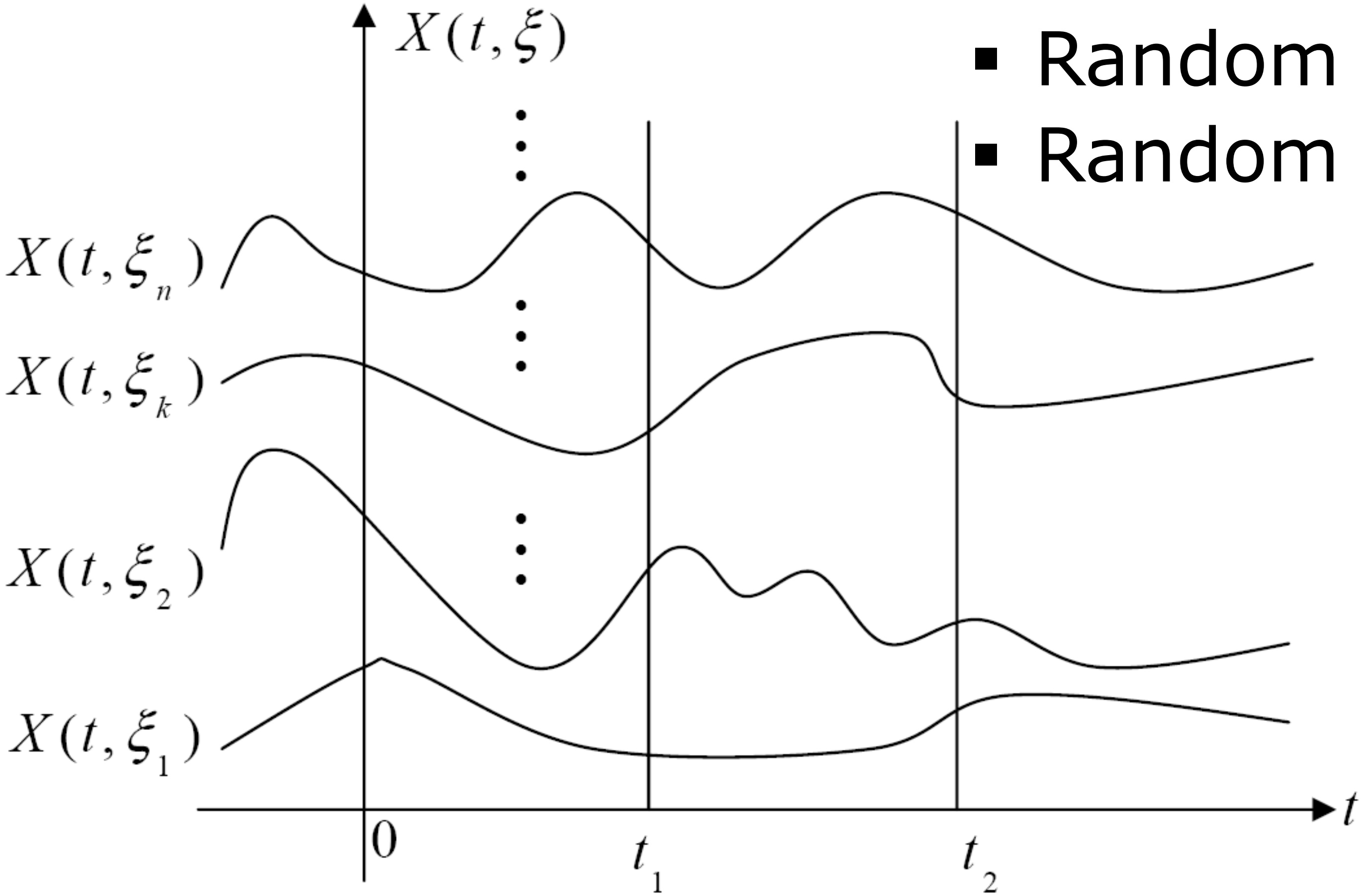
Lecture II-4: Random Signals II

Dr. Sun Ying

Outline

- Stochastic Processes
- Statistical Independence
- Linear Time-Invariant System
- Random Signals and Linear Systems
- Power Spectrum
- Cross Spectrum
- Wiener Filtering

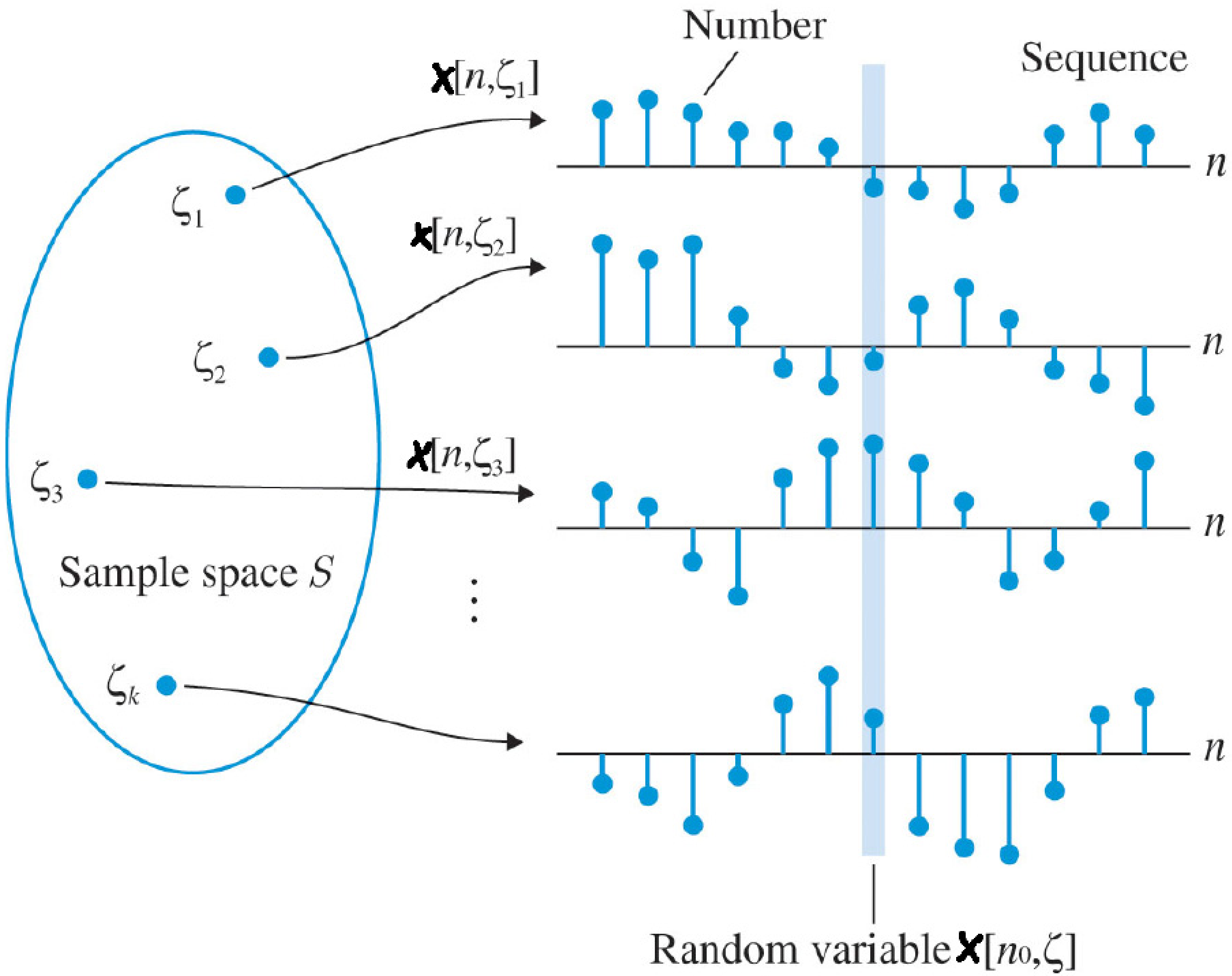
Stochastic Processes (Revisit)



- Random Signals
- Random Processes

Can be completely specified by the collection of joint CDF/PDF among the random variables $\{X(t_1), X(t_2), \ldots, X(t_n)\}$ for any set of $\{t_1, t_2, \ldots, t_n\}$ and any order *n*.

Stochastic Processes



Statistical Independence

- Two processes X(t) and Y(t) are independent if any two vectors of time samples, one from each process, are independent. That is, $[X(t_1), X(t_2), \cdots X(t_n)]$ and $[Y(t_1), Y(t_2), \cdots Y(t_n)]$ are independent random vectors.
- If X(t) and Y(t) are independent then they are uncorrelated (the reverse is not always true).

Linear Time-Invariant System

- Let $x[n]$, $x_1[n]$, and $x_2[n]$ be inputs to a linear system and let $y[n]$, $y_1[n]$, and $y_2[n]$ be their corresponding outputs.
- A linear system satisfies
  - Additivity: $x_1[n] + x_2[n] \Rightarrow y_1[n] + y_2[n]$
  - Homogeneity: $\alpha\, x[n] \Rightarrow \alpha\, y[n]$ for any constant α
- Let $x[n]$ be the input to time-invariant system and $y[n]$ be its corresponding output. Then,
  $x[n - m] \Rightarrow y[n - m]$, for any integer m
- Examples: $y_1[n] = \left(x[n]\right)^2 \qquad y_2[n] = x[Mn]$

Is the system LTI?

Linear Time-Invariant System

- Any linear time-invariant system (LTI) system can be uniquely characterized by its
  - Impulse response: response of system to an impulse
  - Frequency response: response of system to a complex exponential $e^{j2\pi f}$ for all possible frequencies f
  - Transfer function: Laplace transform of impulse response
- Given one of the three, we can find other two provided that they exist.

Discrete-time Convolution

- Output $y[n]$ for input $x[n]$

$$y[n] = T\{x[n]\}$$

- Any signal can be decomposed into sum of discrete impulses

$$y[n] = T\left\{\sum_{m=-\infty}^{\infty} x[m]\,\delta[n-m]\right\}$$

- Apply linear properties

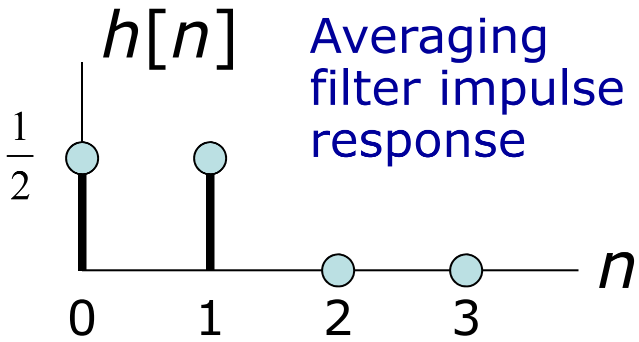
$$y[n] = \sum_{m=-\infty}^{\infty} x[m]\,T\{\delta[n-m]\}$$

- Apply shift-invariance

$$y[n] = \sum_{m=-\infty}^{\infty} x[m]\,h[n-m]$$

- Apply change of variables

$$y[n] = \sum_{m=-\infty}^{\infty} h[m]\,x[n-m]$$

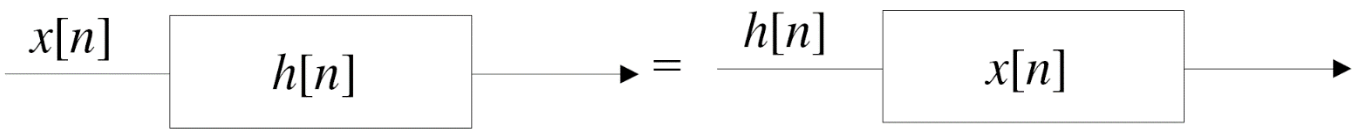


Averaging filter impulse response

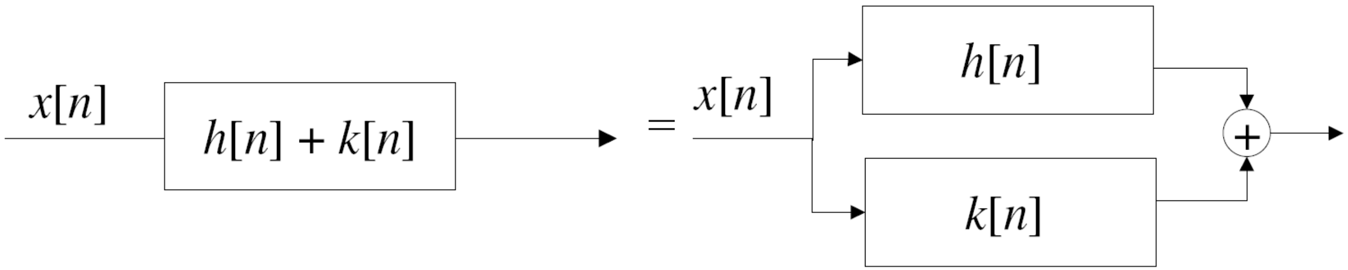
$$y[n] = h[0]\,x[n] + h[1]\,x[n-1]$$
$$= (\,x[n] + x[n-1]\,)\,/\,2$$

Properties

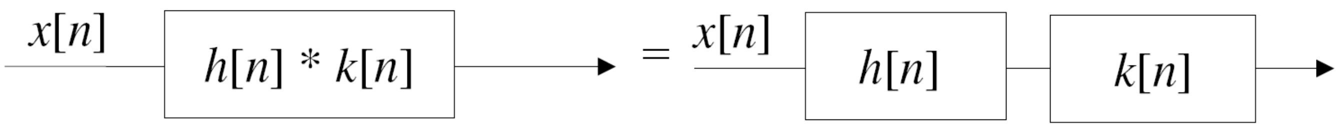
- Commutative: $x[n]*h[n]=h[n]*x[n]$



- Distributive: $x[n]*(h[n]+k[n])=x[n]*h[n]+x[n]*k[n]$



- Associative: $x[n]*(h[n]*k[n])=(x[n]*h[n])*k[n]$



Linear Difference Equations

- Discrete-time LTI systems can be characterized by difference equations.
- For example,

$$y[n] = (1/2)\, y[n-1] + (1/4)\, y[n-2] + x[n]$$

- Taking z-transform of difference equation gives description of system in z-domain.

Z-transform

- A polynomial representation of a sequence {x(n)} in terms of a complex-valued variable z^{-1}:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n}$$

- If X(z) = B(z)/A(z)

  B(z), A(z): polynomials in z, then z_i such that $B(z_i) = 0$ are <u>zeros</u> of X(z), and

  p_j such that $A(p_j) = 0$ are <u>poles</u> of X(z).

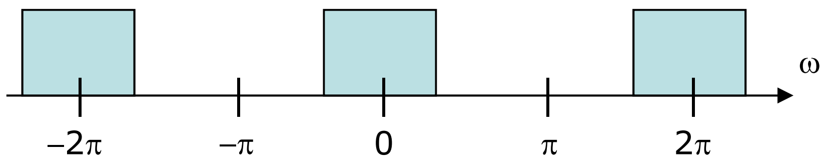
- Frequency domain representation of {x(n)} (discrete time Fourier transform, DTFT):

$$X(e^{j\omega}) = X(z)\Big|_{z=e^{j\omega}} = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}$$

- Convolution formula:

$$y(n) = x(n) * h(n)$$
$$\Leftrightarrow Y(z) = X(z)H(z)$$

- Digital Frequency Axis:
  - Periodic, period = 2π



Transfer Function

- A general *N*th order difference equation

$$y[n]+a_1 y[n-1]+\cdots+a_N y[n-N]=b_0 x[n]+b_1 x[n-1]+\cdots+b_N x[n-N]$$

- Take Z-transform on both sizes

$$\left(1+a_1 z^{-1}+\cdots+a_{N-1} z^{-(N-1)}+a_N z^{-N}\right)Y[z]=\left(b_0+b_1 z^{-1}+\cdots+b_{N-1} z^{-(N-1)}+b_N z^{-N}\right)X[z]$$

- Transfer function

$$H[z]\equiv\frac{Y[z]}{X[z]}=\frac{b_0+b_1 z^{-1}+\cdots+b_{N-1} z^{-(N-1)}+b_N z^{-N}}{1+a_1 z^{-1}+\cdots+a_{N-1} z^{-(N-1)}+a_N z^{-N}}$$

Relation between $h[n]$ and $H[z]$

- Either one can be used to describe an LTI system
  - Having one is equivalent to having the other since they are a z-transform pair
  - By definition, impulse response $h[n]$ is

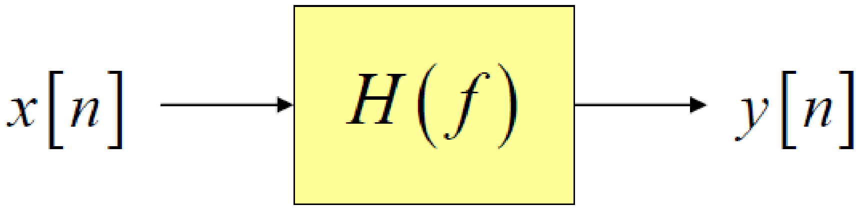
$$y[n] = h[n] \text{ when } x[n] = \delta[n]$$

$$Z\{h[n]\} = H[z]Z\{\delta[n]\} \Rightarrow H[z] = H[z] \cdot 1$$

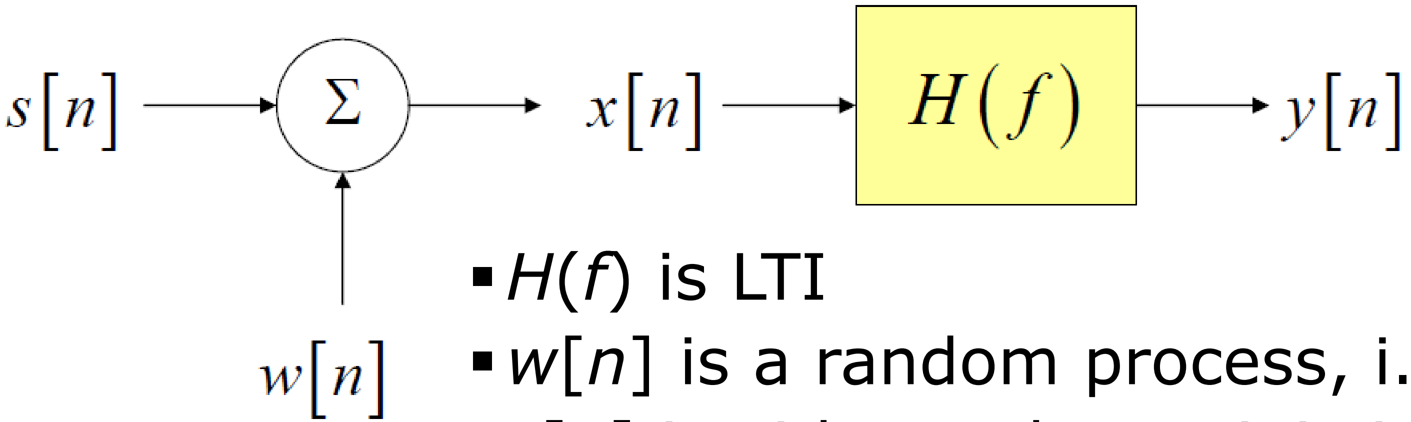
$$h[n] \Leftrightarrow H[z]$$

- Since discrete-time signals can be built up from unit impulses, knowing the impulse response or the transfer function completely characterizes the LTI system.

Random Processes & LTI Systems



- $H(f)$ is LTI
- $x[n]$ is a random process



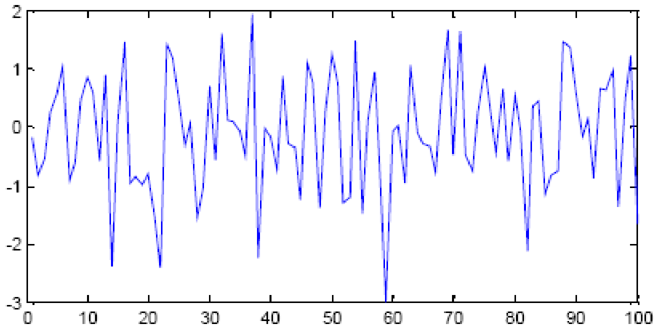
- $H(f)$ is LTI
- $w[n]$ is a random process, i.e., noise
- $s[n]$ is either a deterministic signal or a random process

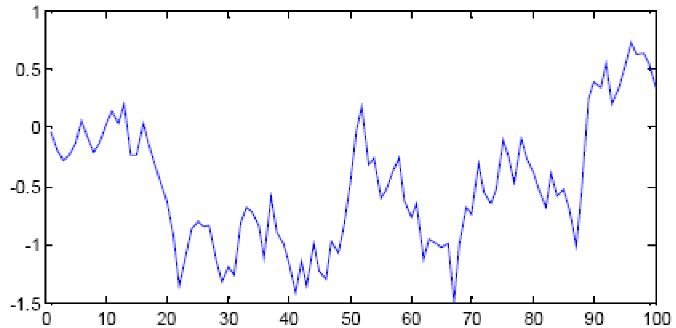
What can we say about y[n]?

Random Sequences & LTI Systems

$$x[n] \longrightarrow \boxed{h[n] \leftrightarrow H[k]} \longrightarrow y[n]$$

LTI



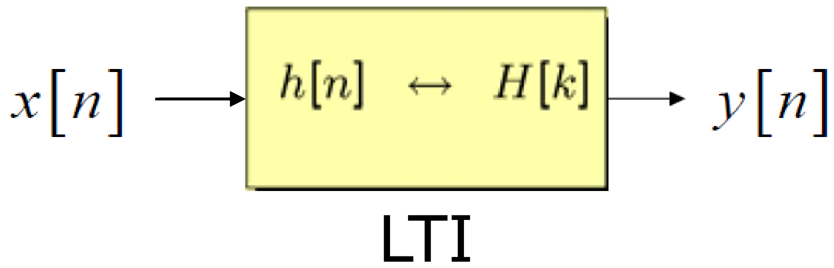


2nd order statistical properties

$$x[n] \Rightarrow R_x[m]$$
$$R_x[m], h[n] \Rightarrow R_y[m]$$

Response of an LTI System

Given a WSS ergodic random process, *x*[*n*], and LTI system with impulse response *h*[*n*] we can compute the time-average of the output *y*[*n*].



$$m_y = E\{y[n]\}$$

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

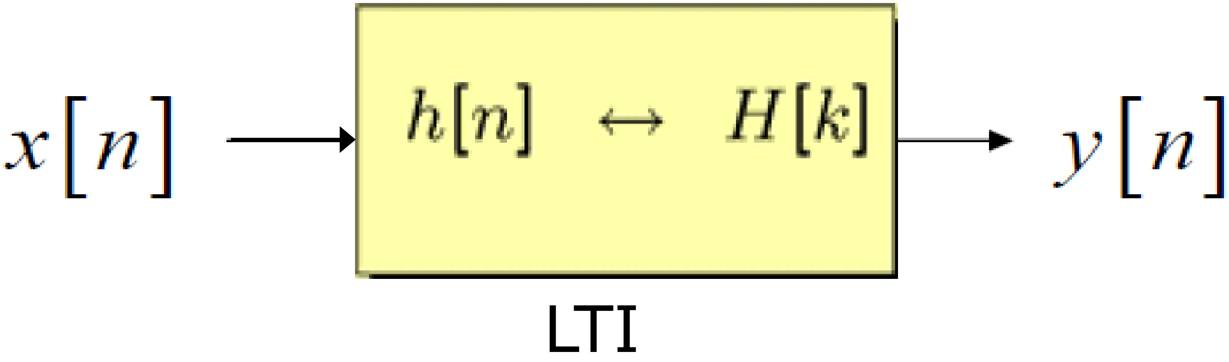
$$= \sum_{k=-\infty}^{\infty} h[k]m_x = m_x \sum_{k=-\infty}^{\infty} h[k] = m_x H[0]$$

Response of an LTI System

Given a WSS ergodic random process, $x[n]$, and LTI system with impulse response $h[n]$ we can compute the autocorrelation of the output $y[n]$.

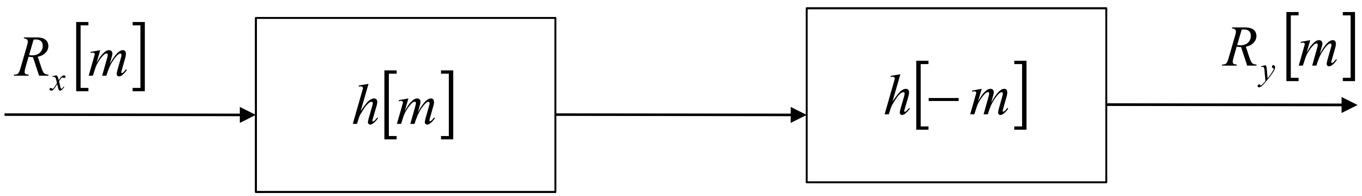
$$
\begin{aligned}
R_y[m] &= E\{y[n]y[n+m]\} \\
&= E\left\{\sum_{k=-\infty}^{\infty} h[k]x[n-k]\sum_{l=-\infty}^{\infty} h[l]x[n+m-l]\right\} \\
&= \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} h[l]h[k]E\{x[n-k]x[n+m-l]\} \\
&= \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} h[l]h[k]R_x[m-l+k] \\
&= R_x[m]*h[m]*h[-m]
\end{aligned}
$$

Stochastic Process - Filtering



Filtering corresponds to convolution of the autocorrelation

$$R_y[m] = R_x[m] * h[m] * h[-m]$$



Power Spectral Density

The autocorrelation of the system response function is computed as the convolution of the $h[m]$ with a time-reversed version $h[-m]$.

$$R_h[m] = h[-m] * h[m]$$

It can be computed in the Fourier domain:

$$\begin{aligned} h[-m] * h[m] &= \mathcal{F}^{-1}\{H^*(f)H(f)\} \\ &= \mathcal{F}^{-1}\left\{|H(f)|^2\right\} \end{aligned}$$

$$R_h[m] \leftrightarrow |H(f)|^2$$

Power Spectrum

The power spectrum of a random process is defined as the Fourier transform of the autocorrelation function:

$$R_x[m] \leftrightarrow S_x(f)$$

$$S_x(f) = \sum_{k=-\infty}^{\infty} R_x[k] e^{-j2\pi fk}$$

$$S_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} R_x[k] e^{-j\omega k}$$

The output power spectrum is given by:

$$S_y(f) = |H(f)|^2 S_x(f)$$

$$S_y(e^{j\omega}) = |H(e^{j\omega})|^2 S_x(e^{j\omega})$$

$$R_y[m] = R_x[m] * h[m] * h[-m]$$

White Noise Sequences

White noise sequences: WSS processes whose autocorrelation function is:

$$\boxed{\begin{aligned} R_w[m] &= \langle w[n]w[n+m]\rangle \\ &= \sigma_w^2 \delta[m] \end{aligned}}$$

where

$$\boxed{\begin{aligned} \langle w[n]\rangle &= 0 \\ \left\langle w[n]^2 \right\rangle &= \sigma_w^2 \end{aligned}}$$

Why is this called "white" noise?

- The white noise sequence is uncorrelated from sample to sample.
- All i.i.d. sequences satisfy this constraint.
- It only requires the samples to be uncorrelated, but not independent.

Power Spectrum: White Noise

Autocorrelation function

$$\begin{aligned} R_w[m] &= \langle w[n]w[n+m] \rangle \\ &= \sigma_w^2 \delta[m] \end{aligned}$$

Its power spectrum is a constant equal to the variance:

$$\boxed{S_w(f) \;=\; \sigma_w^2}$$

- White noise has equal power at all frequencies.
- By analogy with white light which is a uniform mixture of all visible frequencies.

Power Spectrum Estimation

Direct estimate of the power spectrum from the first N observations of a realization $x[n]$:

$$\begin{aligned}\hat{S}_x(f) &= \frac{1}{N}|X_N(f)|^2 \\ &= \frac{1}{N}\left|\sum_{n=0}^{N-1} x[n]e^{-j2\pi fn}\right|^2\end{aligned}$$

- It does not converge to $S_x(f)$ as N grows large. ☹
- The variance of the estimate does not converge to 0 for increasing N.

Power Spectrum Estimation

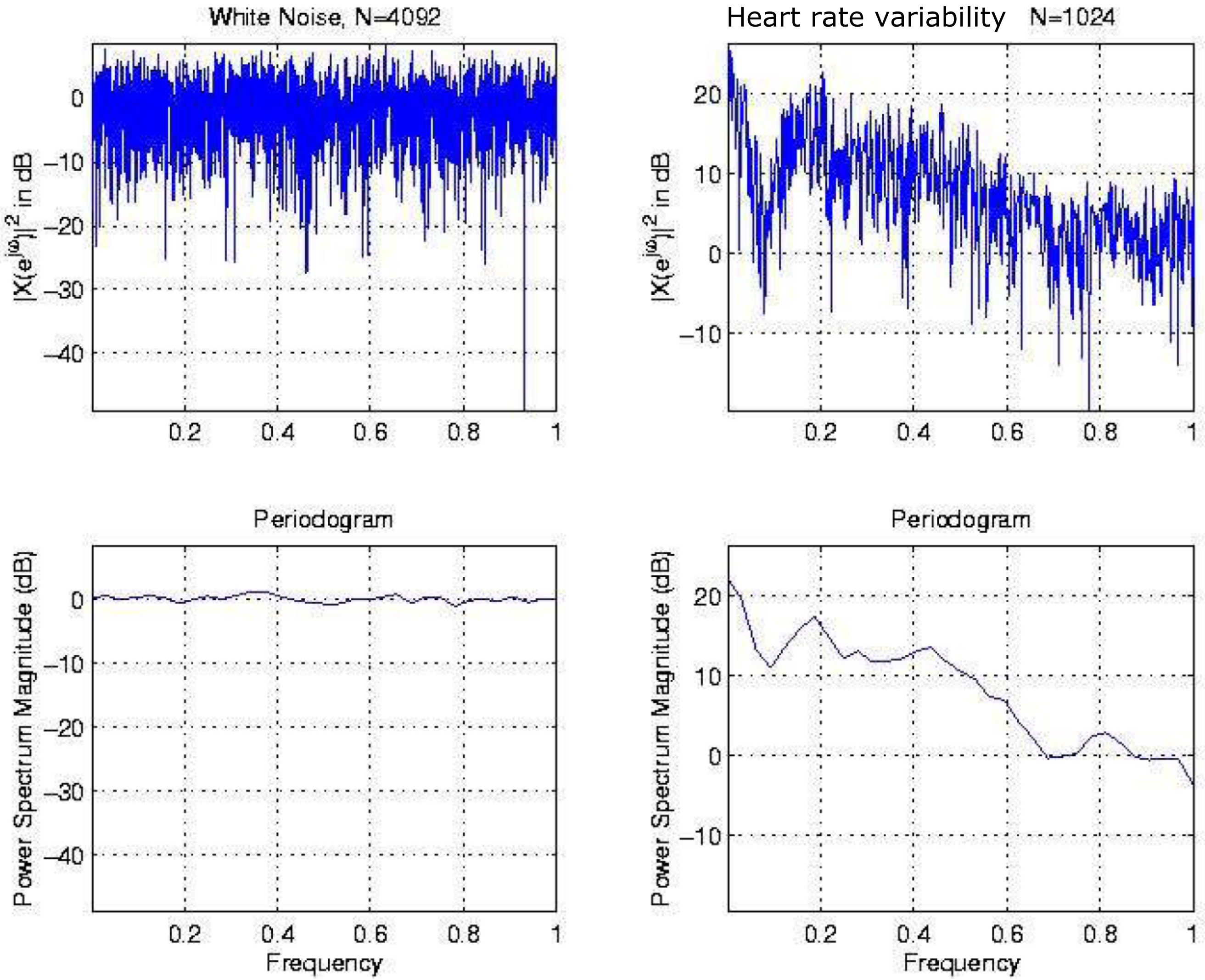
Averaging Periodogram can be used to smooth neighboring frequencies:

Compute Fourier transform for window of size *N* and average over *M* windows.

$$\hat{S}_x(f) = \frac{1}{NM} \sum_{m=0}^{M-1} \left| X_{N,M}(f) \right|^2$$

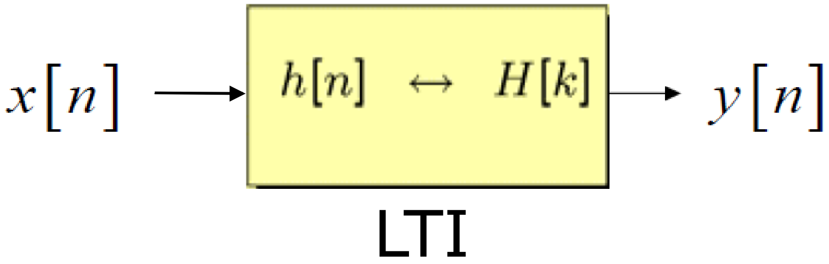
The estimate converges to the true Power Spectrum as *M* and *N* grow large.

Power Spectrum: Examples



Compute FT for window of size N/K and average over K windows.

Cross-correlation of Input and Output



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

$$R_{xy}[m] = E\{x[n]y[n+m]\}$$

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

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

$$= \sum_{l=-\infty}^{\infty} h[l]R_x[m-l] = h[m] * R_x[m]$$

$$S_{xy}(z) = H(z)S_x(z)$$

$$S_{xy}(e^{j\omega}) = H(e^{j\omega})S_x(e^{j\omega})$$

Estimating System Response

If the input to the system is a white-noise sequence:

$$
\begin{aligned}
R_{wy}[m] &= h[m] * R_w[m] \\
&= h[m] * \sigma_w^2 \delta[m] \\
&= \sigma_w^2 h[m]
\end{aligned}
$$

then the 2nd order statistics of the output process can be used to estimate the system response function (i.e., system identification).

Cross-correlation of Input and Output

$$x[n] \longrightarrow \boxed{h[n] \;\leftrightarrow\; H[k]} \longrightarrow y[n]$$

LTI

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

$$S_{yx}(z) = H\left(z^{-1}\right)S_x(z)$$

$$S_{yx}\left(e^{j\omega}\right) = H\left(e^{-j\omega}\right)S_x\left(e^{j\omega}\right)$$

$$\begin{aligned}
R_{yx}[m] &= E\{y[n]x[n+m]\} \\
&= E\left\{\sum_{l=-\infty}^{\infty} h[l]x[n-l]x[n+m]\right\} \\
&= \sum_{l=-\infty}^{\infty} h[l]E\{x[n-l]x[n+m]\} \\
&= \sum_{l=-\infty}^{\infty} h[l]R_x[m+l] \\
&= \sum_{l=-\infty}^{\infty} h[-(-l)]R_x[m-(-l)] \\
&= \sum_{k=-\infty}^{\infty} h[-k]R_x[m-k] = h[-m] * R_x[m]
\end{aligned}$$

$k = -l$

Cross Spectrum

The cross spectrum of the signals *x*[*n*] and *y*[*n*] is the Fourier transform of the cross-correlation function:

$$S_{xy}(f) \triangleq \sum_{k=-\infty}^{\infty} R_{xy}[k]\, e^{-j2\pi f k}$$

Relationships:

$$R_{xy}[m] = h[m] * R_x[m] \rightarrow S_{xy}(f) = H(f)\, S_x(f)$$

$$R_{yx}[m] = R_{xy}[-m] \rightarrow S_{yx}(f) = S_{xy}(-f) = S^*_{xy}(f)$$

When the input is zero-mean, white noise: $S_{wy}(f) = \sigma_w^2\, H(f)$

Power Spectrum: Summary

$$S_x(f) = \sum_{k=-\infty}^{\infty} R_x[k] e^{-j2\pi fk}$$

$$R_{xy}[n] = \sum_{k=-\infty}^{\infty} h[k] R_x[n-k] = h[n] * R_x[n]$$

$$R_{xy}[n] \leftrightarrow S_{xy}(f)$$
$$h[n] \leftrightarrow H(f)$$
$$R_x[n] \leftrightarrow S_x(f)$$

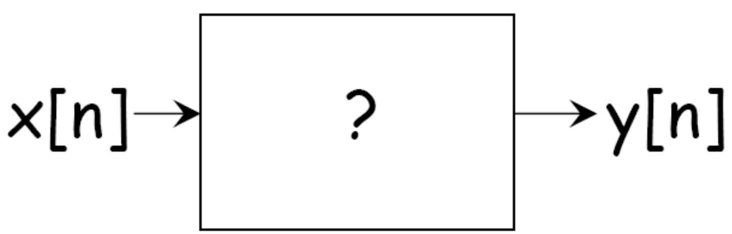
$$H(f) = \frac{S_{xy}(f)}{S_x(f)}$$

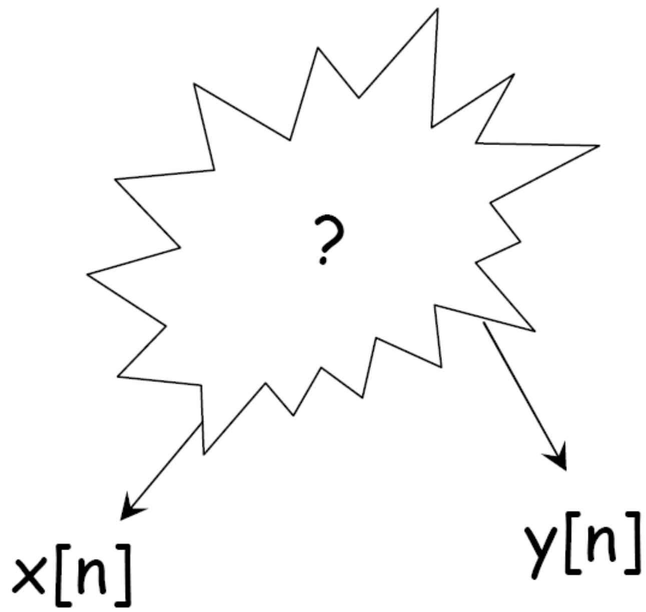
$$S_y(f) = |H(f)|^2 S_x(f)$$

Each frequency is independent of all other frequencies.

Wiener Filtering

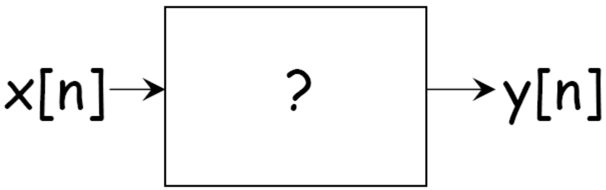
Suppose we have two jointly WSS random processes whose joint second-order statistics are known, and we want to estimate one from an observation of the other.





Wiener Filter: Example

Given two realizations of random processes, *y*[*n*] and *x*[*n*], how to construct a filter which predicts *y*[*n*] from observations of *x*[*n*]?



Minimize the mean square error (MSE) between $y[n]$ and the prediction $\hat{y}[n]$

$$
\begin{aligned}
\left\langle e[n]^2 \right\rangle &= \left\langle \left(y[n] - \hat{y}[n]\right)^2 \right\rangle \\
&= \left\langle \left(y[n] - h[n] * x[n]\right)^2 \right\rangle \\
&= \left\langle \left(y[n] - \sum_k h[k]\, x[n-k]\right)^2 \right\rangle \\
&= \sigma_y^2 - 2\sum_k h[k]\, R_{xy}[k] + \sum_k \sum_l h[k]\, h[l]\, R_x[k-l]
\end{aligned}
$$

Wiener Filter: Example

The error is a quadratic function of the filter coefficients $h[k]$. Taking the first order derivative of the MSE with respect to $h[k]$ and setting it to zero, we obtain a system of linear equations.

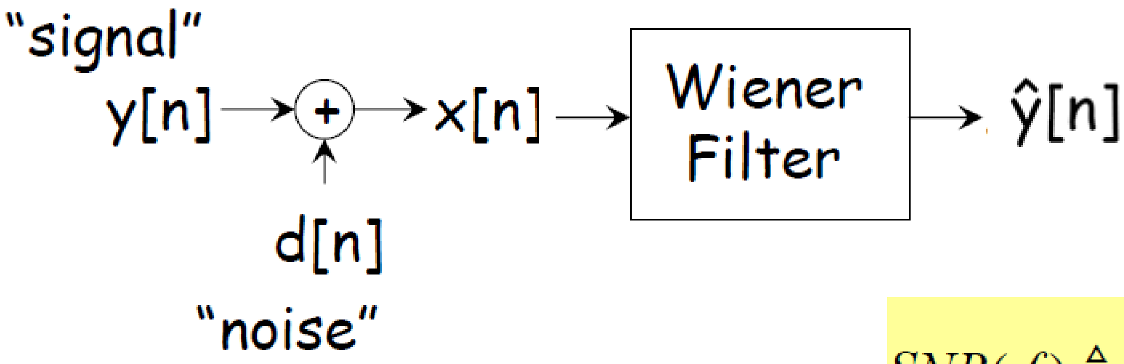
$$R_{xy}[k] \; = \sum_{l} \; h[l] \; R_x[k-l]$$

Known as Wiener-Hopf or Yule-Walker equations.

If we restrict to Nth-order FIR filters, then it becomes

$$\begin{bmatrix} R_x[0] & R_x[1] & \cdots & R_x[N-1] \\ R_x[1] & R_x[0] & & R_x[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_x[N-1] & R_x[N-2] & \cdots & R_x[0] \end{bmatrix} \begin{bmatrix} h[0] \\ \vdots \\ \vdots \\ h[N-1] \end{bmatrix} = \begin{bmatrix} R_{xy}[0] \\ \vdots \\ \vdots \\ R_{xy}[N-1] \end{bmatrix}$$

Wiener Filter for Noise Removal



Uncorrelated noise:

$R_{yd}[n] = 0 \longleftrightarrow S_{yd}(f) = 0$

$$H(f) = \frac{S_{xy}(f)}{S_x(f)} = \frac{S_y(f)}{S_y(f) + S_d(f)}$$

$$SNR(f) \triangleq \frac{S_y(f)}{S_d(f)}$$

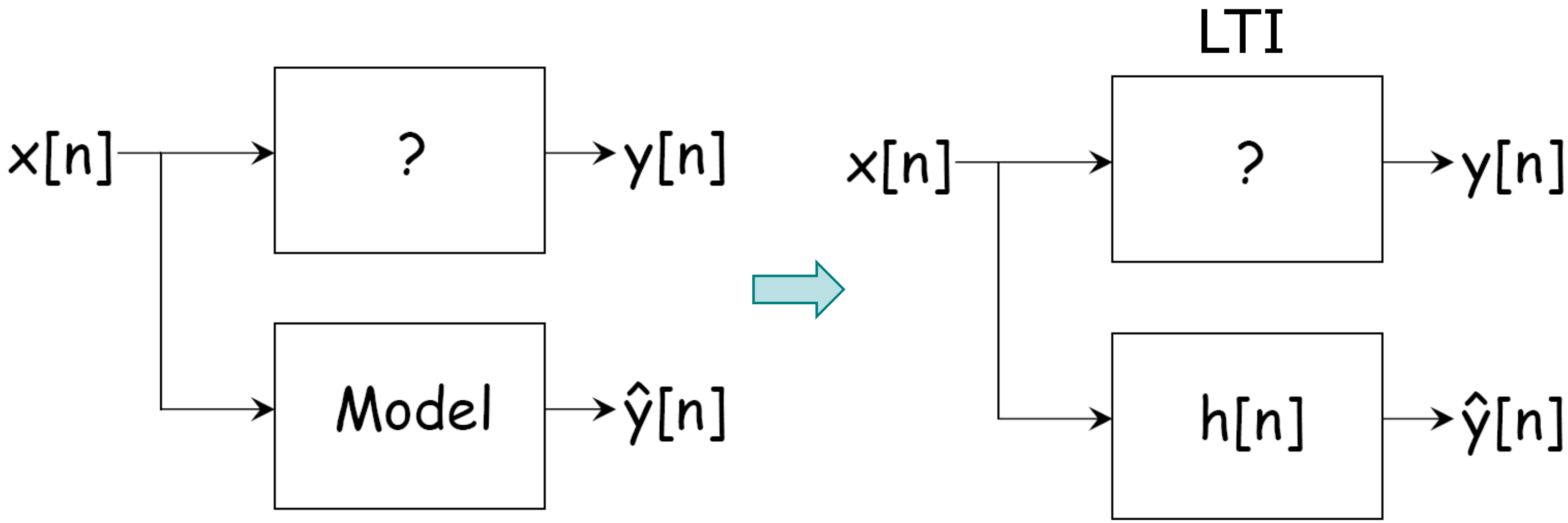
$$H(f) = \frac{S_y(f)}{S_y(f) + S_d(f)} = \frac{SNR(f)}{SNR(f) + 1}$$

SNR(f) >> 1 : H(f)≈1

SNR(f) << 1 : H(f)≈0

Wiener Filter for System Identification

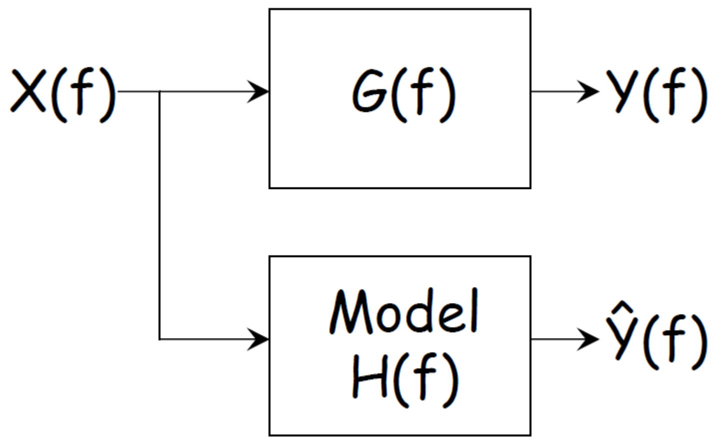
System Identification: Estimate a model of an unknown system based on observations of inputs and outputs.



Find the LTI system that best predicts the output from the input.

Example: LTI System

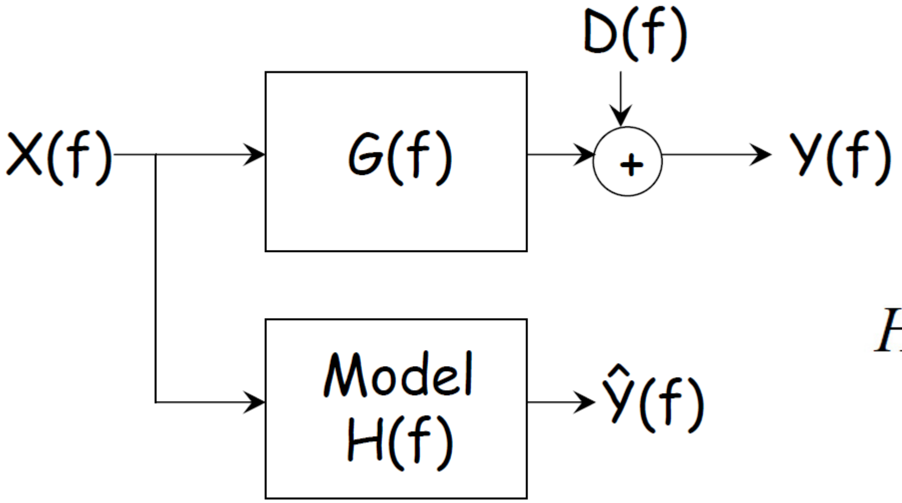
If the unknown system is LTI, we hope that Wiener filtering recovers that LTI system exactly.



$$H(f) = \frac{S_{xy}(f)}{S_x(f)} = \frac{S_x(f)G(f)}{S_x(f)} = G(f)$$

Example: LTI System + Noise

We can only observe y[n] which is the output of the LTI system + noise



d[*n*] is uncorrelated with *x*[*n*].

$$H(f) = \frac{S_{xy}(f)}{S_x(f)}$$

$$= \frac{S_x(f)G(f) + S_{xd}(f)}{S_x(f)}$$

$$= \frac{S_x(f)G(f)}{S_x(f)} = G(f)$$

We can still recover the LTI system exactly!

Example: Spectral Factorization

Suppose that we know the autocorrelation function $R_x[m]$ of the input $x[n]$ but do not have access to $x[n]$ and therefore cannot determine the cross-correlation $R_{xy}[m]$ with the output $y[n]$, but can determine the output autocorrelation $R_y[m]$.

If $R_x[m] = \delta[m]$, then $R_y[m] = h[m] * h[-m]$

$$S_y[z] = H[z]H[z^{-1}]$$

Additional assumptions or constraints, for instance on the stability and causality of the system and its inverse, may allow one to recover $H(z)$ from knowledge of $H(z)H(z^{-1})$.

Example: Causal Wiener Filter

$$\begin{bmatrix} R_x[0] & R_x[1] & \cdots & R_x[N-1] \\ R_x[1] & R_x[0] & & R_x[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_x[N-1] & R_x[N-2] & \cdots & R_x[0] \end{bmatrix} \begin{bmatrix} h[0] \\ \vdots \\ \vdots \\ h[N-1] \end{bmatrix} = \begin{bmatrix} R_{xy}[0] \\ \vdots \\ \vdots \\ R_{xy}[N-1] \end{bmatrix}$$

If $R_x[m] = \delta[m]$, then the system is simplified to

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} h[0] \\ \vdots \\ \vdots \\ h[N-1] \end{bmatrix} = \begin{bmatrix} R_{xy}[0] \\ \vdots \\ \vdots \\ R_{xy}[N-1] \end{bmatrix}$$

Causal Wiener filter: $\boxed{h[n] = \begin{cases} R_{xy}[n] & n \geq 0 \\ 0 & n < 0 \end{cases}}$

Wiener Filter: Summary

- A Wiener filter finds the MMSE (minimum mean squared error) estimate of one random process as a linear function of another random process.
- Applications include noise removal and system identification.
- How do we get R_x and R_{xy}?
  - Sample statistics for auto- and cross-correlations
  - Periodogram estimates of power spectral densities

Acknowledgements

- John Fisher. Course materials for HST.582J / 6.555J / 16.456J, Biomedical Signal and Image Processing, Spring 2007. MIT Open CourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology.
- Bertrand Delgutte. Course materials for HST.582J / 6.555J / 16.456J, Biomedical Signal and Image Processing, Spring2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology.