System Identification & Parameter Estimation

Wb2301: SIPE lecture 7

Overview SI, Intro PE

Erwin de Vlugt, Dept. of Biomechanical Engineering (BMechE), Fac. 3mE
March 23, 2010

System Identification

- System identification involves the **estimation of relationships between signals**
- **Noise is the common enemy**: introduces uncertainty in terms of variance (random errors)
- Main learning goals:
  - How to estimate relationships
    - Correlation Functions, IRF, FRF
  - How to minimize uncertainty
    - Averaging, input design

Cross-correlation functions

- Measure for common structure in two signals:
  - Cross-correlation

    $$\Phi_{xy}(\tau) = E\left[x(t-\tau)y(t)\right]$$

  - Cross-covariance

    $$C_{xy}(\tau) = E\left[\left(x(t-\tau)-\mu_x\right)\left(y(t)-\mu_y\right)\right] = \Phi_{xy}(\tau) - \mu_x\mu_y$$

  - Cross-correlation coefficient

    $$r_{xy}(\tau) = E\left[\left(\frac{x(t-\tau)-\mu_x}{\sigma_x}\right)\left(\frac{y(t)-\mu_y}{\sigma_y}\right)\right] = \frac{C_{xy}(\tau)}{\sqrt{C_{xx}(0)C_{yy}(0)}}$$

Estimates of Correlation Functions

- Unbiased estimator: $$\hat{\Phi}_{xy}(\tau) = \frac{1}{N-\tau}\sum_{i=\tau}^{N} x(i-\tau)y(i)$$

- Variance of the estimator increases with lag!
  To avoid this, divide by N (biased estimator):

$$\boxed{\hat{\Phi}_{xy}(\tau) = \frac{1}{N}\sum_{i=\tau}^{N} x(i-\tau)y(i)}$$

- Or, use large N to minimize variance: $\frac{N}{N-\tau} \rightarrow 1$

- Similar estimators can be derived for the covariance and correlation coefficient functions.

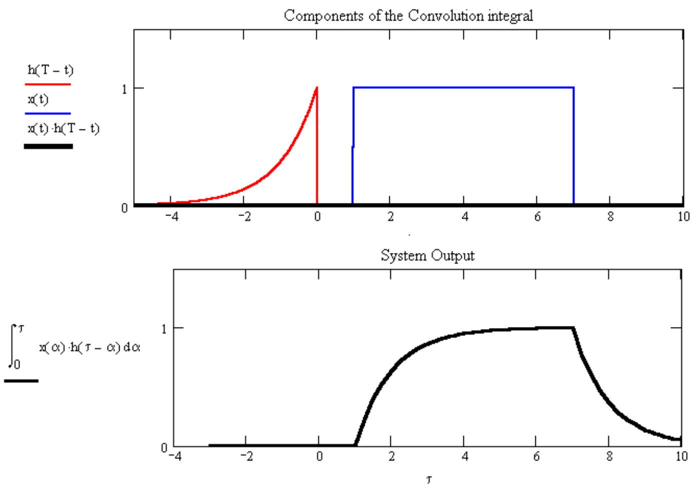
Time domain models

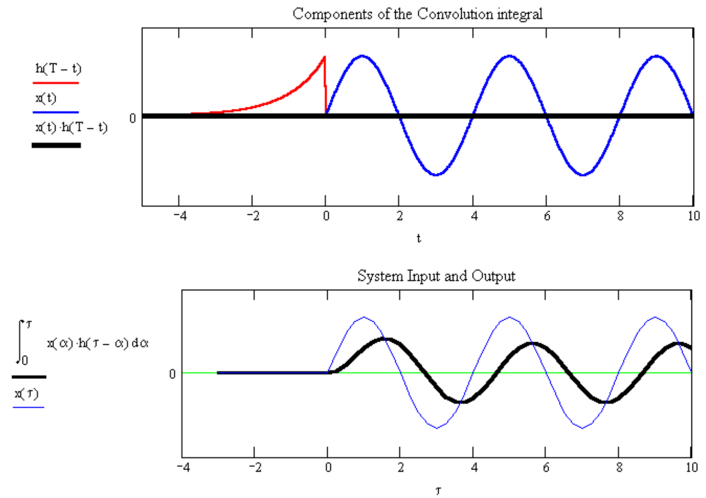
- System:

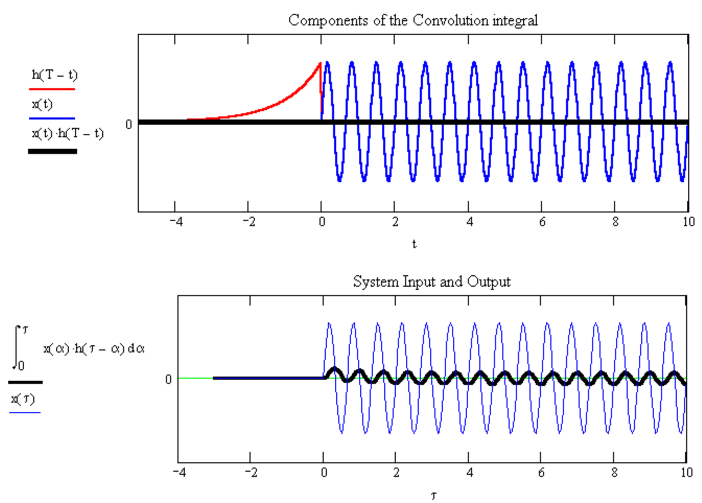
$$y(t) = \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau$$

- h(t) is the impulse response function (IRF) and the output the convolution integral of the input with h(t)
- h(t) contains the system's dynamics or 'memory'.

Output by Convolution



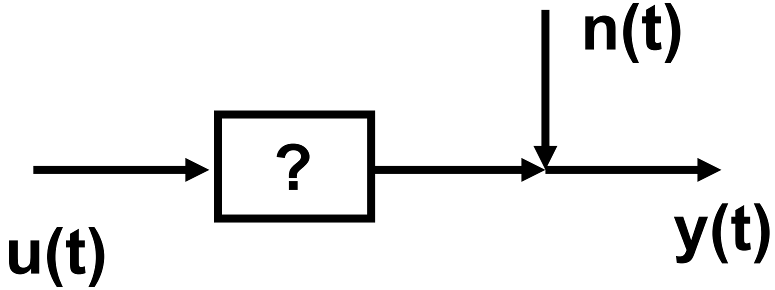




Time domain models

- Causal system: h(t)=0 for t<0
- Finite memory: h(t)=0 for t>T
- Continuous $$y(t) = \int_{o}^{T} h(\tau)u(t-\tau)d\tau$$
- Discrete $$y(t) = \sum_{\tau=0}^{T-1} h(\tau)u(t-\tau)\Delta\tau$$

Basic identification with cross-covariance



$$y(t) = n(t) + \int h(t')u(t-t')dt'$$

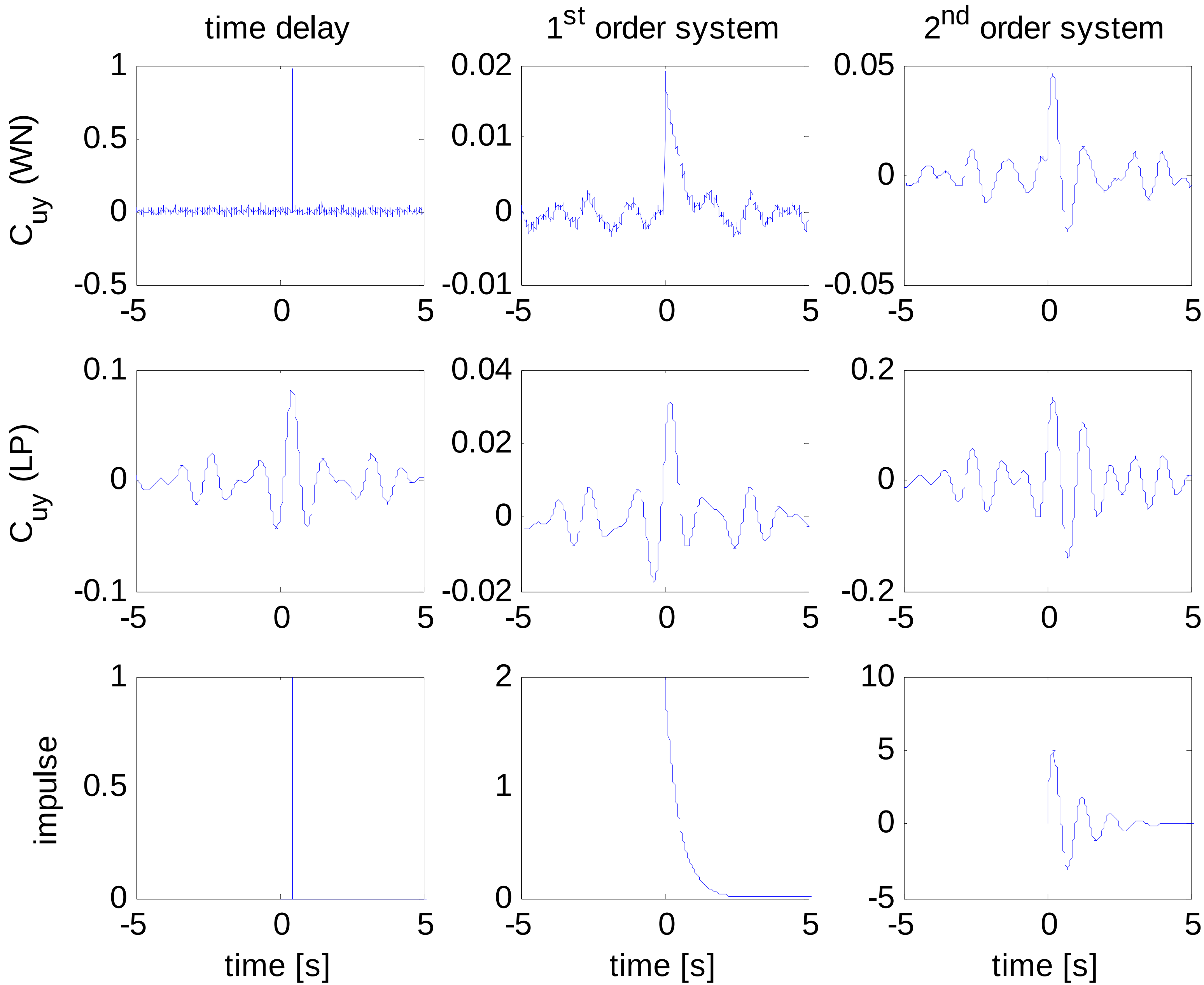
multiply with $u(t-\tau)$: $u(t-\tau)y(t) = u(t-\tau)n(t) + \int h(t')u(t-\tau)u(t-t')dt'$

$$C_{uy}(\tau) = C_{un}(\tau) + \int h(t')C_{uu}(\tau - t')dt'$$

white noise: $C_{uu}(\tau) = 0$ for $\tau \neq 0$; $\quad C_{uu}(0) = 1$

$$C_{uy}(\tau) = C_{un}(\tau) + h(\tau)$$

Other 'tricks' needed when u(t) is not white


time delay
1st order system
2nd order system
C_uy (WN)
C_uy (LP)
impulse
time [s]
time [s]
time [s]

IRF via least squares regression

$$y(t) = \sum_{\tau=0}^{T-1} h(\tau)u(t-\tau)\Delta\tau$$

- Rewrite convolution as matrix

$$y = Uh$$

with

$$U = \begin{bmatrix} u(1) & 0 & 0 & \dots & 0 \\ u(2) & u(1) & 0 & \dots & 0 \\ u(3) & u(2) & u(1) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ u(N) & u(N-1) & u(N-2) & \dots & u(N-T+1) \end{bmatrix}$$

IRF via least squares regression

- z is measurement of y, with noise n(t):

$$z(t) = y(t) + v(t)$$

$$z = Uh + v$$

- Solution via linear least-squares regression (see W&K p.26):

$$\hat{h} = \left(U^T U\right)^{-1} U^T z$$

Estimator for the cross-spectrum

- Fourier transform of the cross-correlation function is called the cross-spectrum:

$$\hat{S}_{uy}(f) = \sum_{\tau=0}^{N-1} \hat{\Phi}_{uy}(\tau) e^{-j2\pi \frac{f\tau}{N}}$$

$$\boxed{\hat{S}_{uy}(f) = \frac{1}{N} U^*(n\Delta f) Y(n\Delta f)}$$

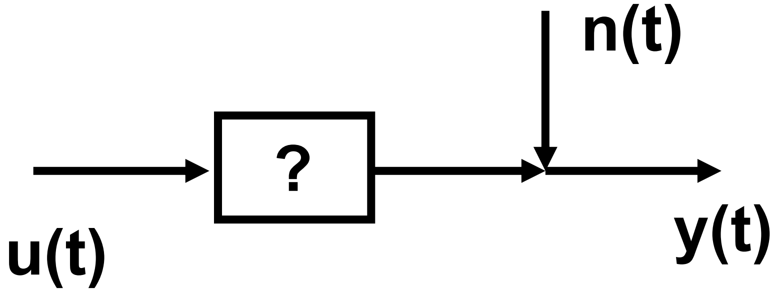
Frequency domain models

- Time-domain: convolution integral $$y(t)=\int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau$$
- Fourier transform: $$Y(f)=\Im\big(y(t)\big)=\Im\left(\int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau\right)$$

$$Y(f)=H(f)U(f)$$

'Convolution in time domain is multiplication in frequency domain' (and vice-versa)

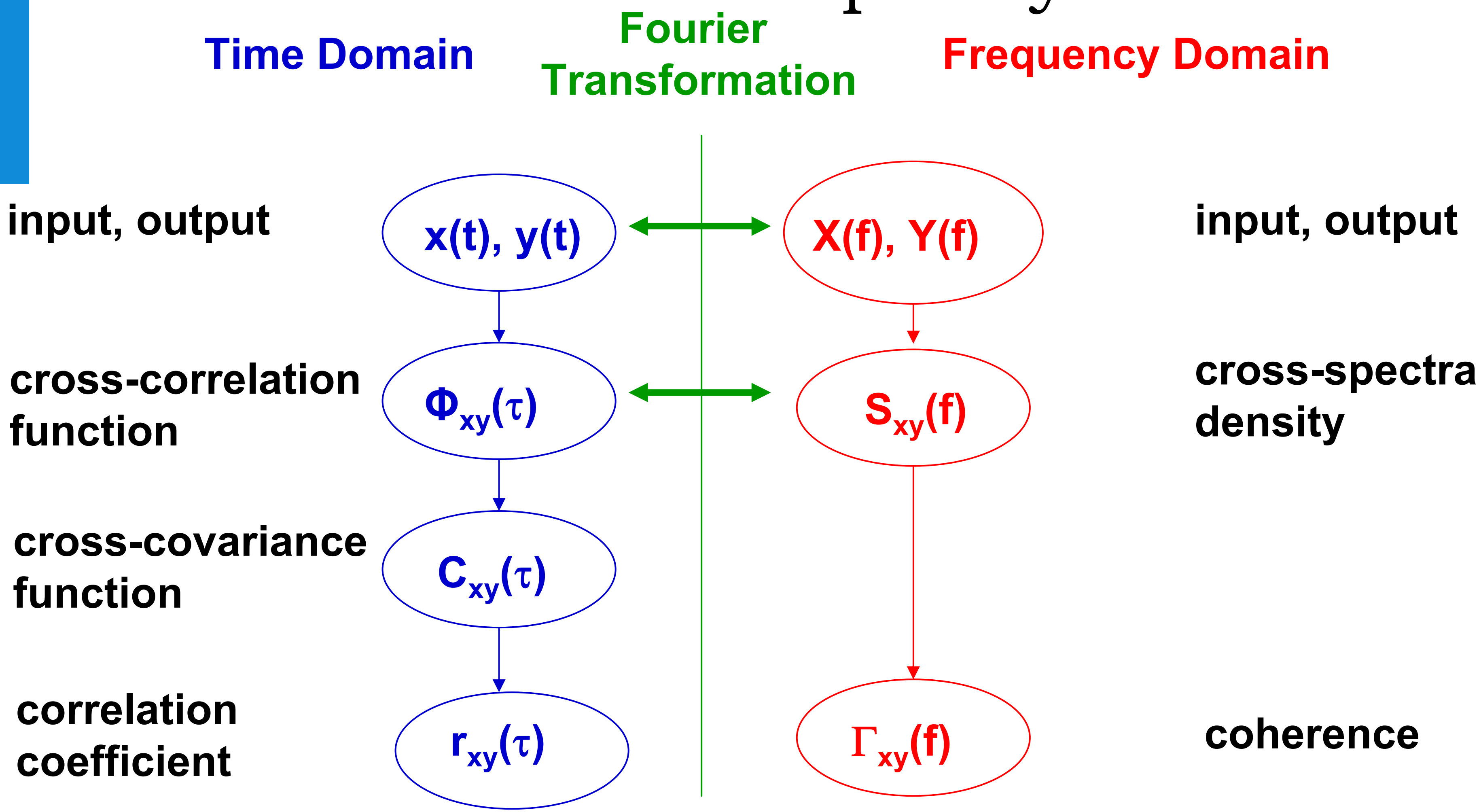
Basic identification with spectral densities



Fourier transform: $S_{uy}(f) = S_{un}(f) + H(f) S_{uu}(f)$

if $S_{un}(f) = 0$: $H(f) = \dfrac{S_{uy}(f)}{S_{uu}(f)}$

Time-domain vs. Frequency-domain



Reduce of Variance

- Longer recordings
  - More ‘information’ but same amount of noise

- Repeat the experiment
  - Noise cancels out by **averaging** from exactly the same inputs

- Improve Signal-to-Noise Ratio
  - Concentrate **power** at specified frequencies, assuming the noise power remains the same

Averaging

- Reduce variance of spectral estimator by averaging
  - either over multiple repetitions,
  - or over adjacent frequencies

- Basic ‘idea’:
  - Multiple repetitions: each realizations has the same frequency content.
  - Frequency averaging: the spectral density is often smooth, i.e. adjacent frequencies contain (approx.) the same information.

- => Averaging will reduce the effect of noise, as the noise has zero mean.

Welch method

- Divide data in multiple segments
- Calculate spectral density for each segment
- Average over the segments
- D-time segments
- Drawback: reduced spectral resolution with factor D

$$\hat{S}_{uu}(f) = \frac{1}{D}\sum_{d=1}^{D} S_{uu}(f) = \frac{1}{DN_D}\sum_{d=1}^{D} U(-f)U(f)$$

Frequency averaging

- Calculate the raw spectral density
- Average over adjacent frequencies over bandwidth D

$$\hat{S}_{uu}(f_c) = \frac{1}{D}\sum_{d=1}^{D}\hat{S}_{uu}(f_d); \quad f_c = \frac{1}{D}\sum_{d=1}^{D} f_d$$

- Drawback: introduces bias at sharp transitions in FRF

Coherence and coherency

- Coherency

$$\gamma_{uy}(f) = \frac{S_{uy}(f)}{\sqrt{S_{yy}(f) S_{uu}(f)}}$$

- Coherence

$$\gamma_{uy}^{2}(f) = \frac{\left|S_{uy}(f)\right|^2}{S_{yy}(f) S_{uu}(f)} = \frac{1}{1 + \frac{S_{nn}(f)}{\left|H(f)\right|^2 S_{uu}(f)}}$$

- Coherence
  - Real valued, between 0 and 1
  - Squared coherency

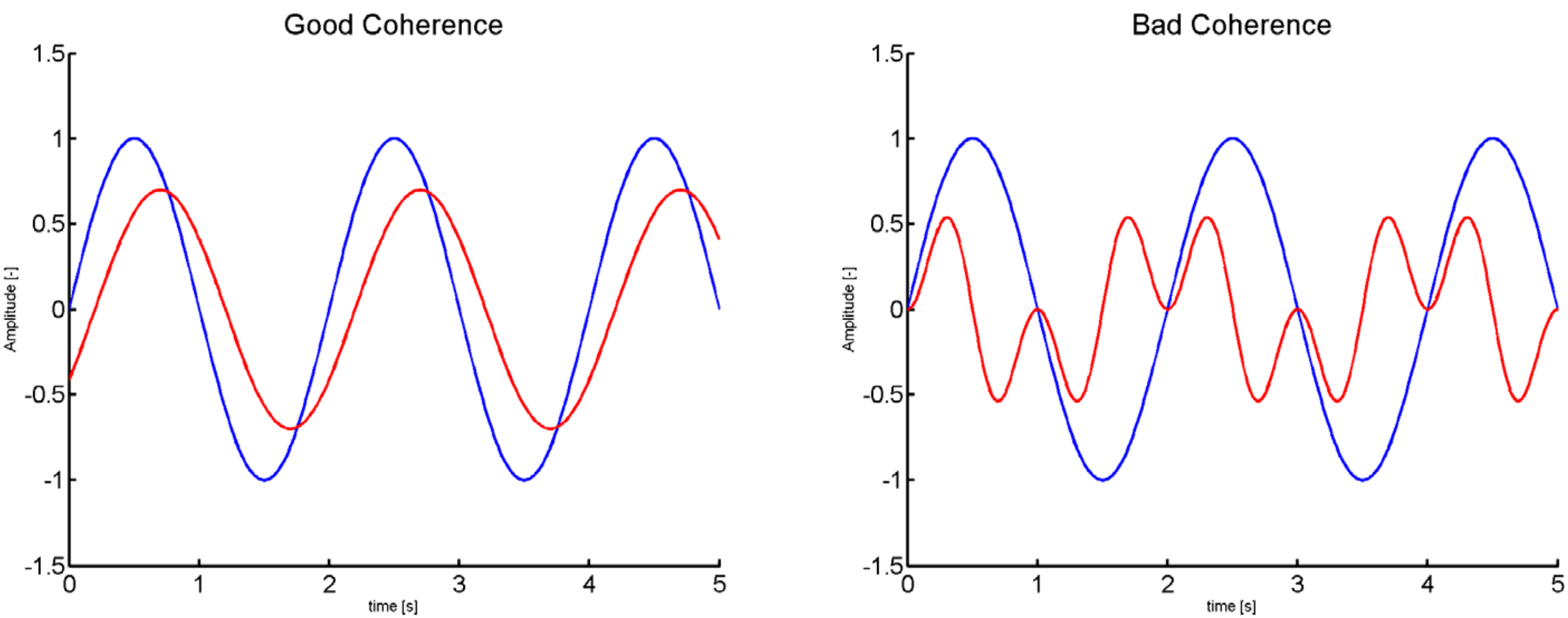
- Coherency has a phase, which represent the relative delay between the signals

Coherence γ_{uy}^2

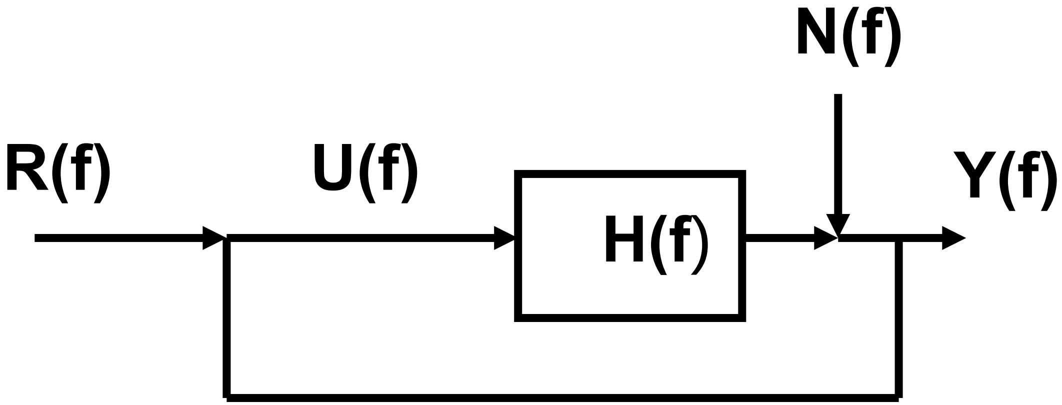
Coherence:

$$\gamma_{uy}^{2}(f) = \frac{\left|S_{uy}(f)\right|^{2}}{S_{yy}(f)S_{uu}(f)} \qquad 0 \le \gamma_{yu}^{2}(f) \le 1$$

Coherence: Linear relationship between input and output, irrespective of the type of system in between.



Identification of a system in closed loop

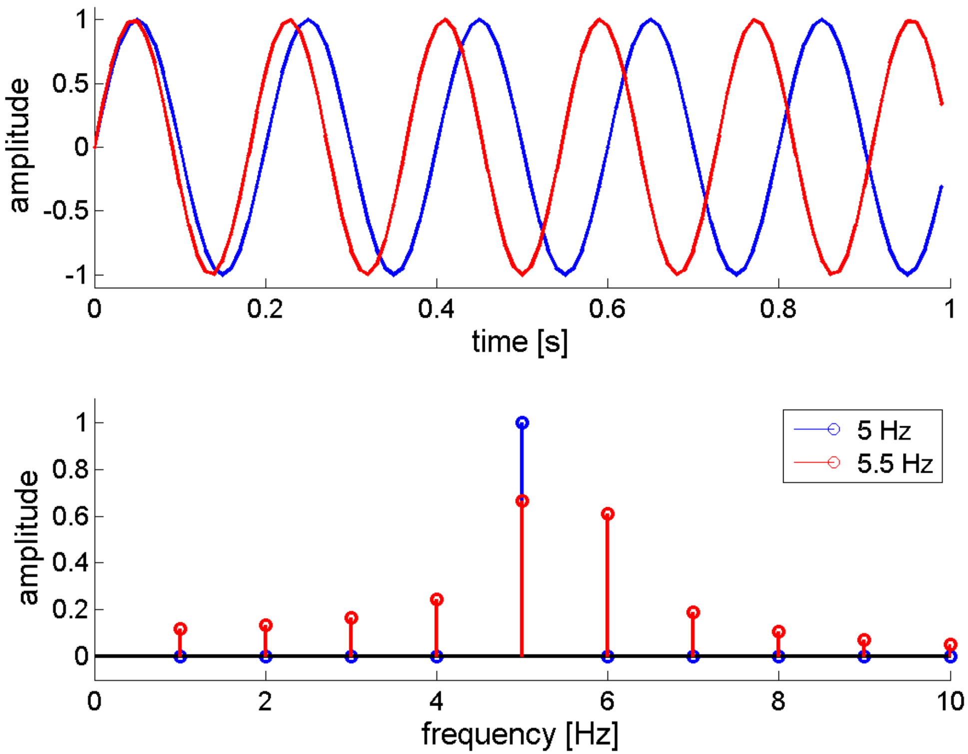


- Closed loop: $S_{un}(f) \neq 0$!
- Consequently: $H(f) \neq S_{uy}(f)/S_{uu}(f)$

Input signal design

- Variance due to noise can be minimized by increasing input power
- Bias due to leakage can be avoided using signals with period equal to (or integer ratios of) the observation time.

Example leakage



- Matlab: Lec4_LeakageDemo.m

- 1 second observation => resolution: 1Hz
- 5 Hz and 5.5 Hz sine

- Only frequencies with integer number of period are correctly observed!
- Other frequencies will ‘leak’ to neighboring frequencies

Improving the estimate 1

Bias (structural errors):

- Main causes
  - finite observation of stochastic input (leakage!)
  - frequency averaging
- Cure
  - application of 'leakage free' deterministic perturbation signals
  - moderate frequency averaging
  - apply method that do not need frequency averaging

Improving the estimate 2

Variance (random errors):

- Main causes
  - noise
- Cure
  - improve SNR
  - averaging (time or frequency domain)

Multisine signals

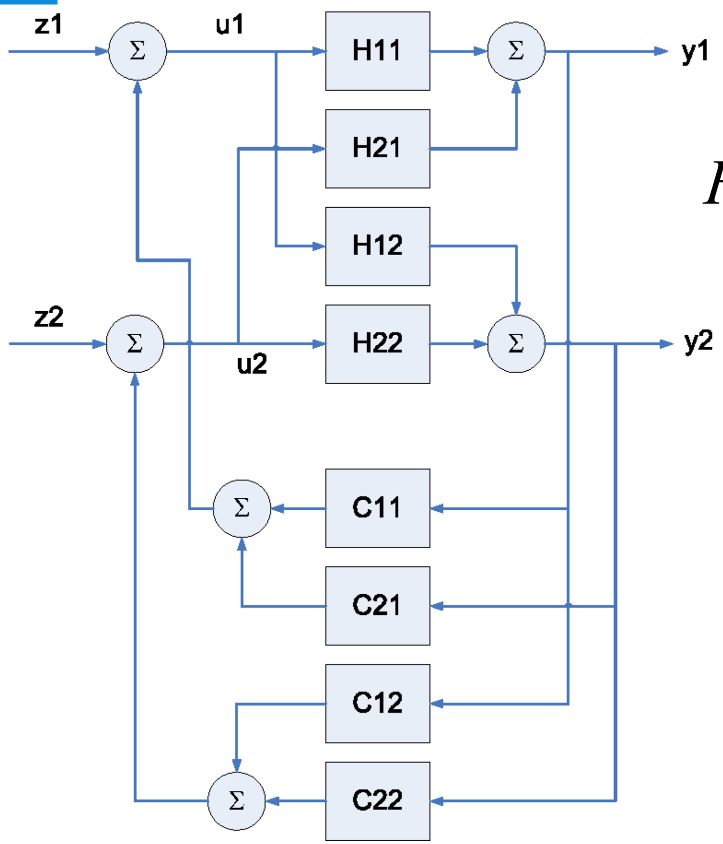
- With a limited observation time only the frequencies with an integer number of period can be ‘seen’ by spectral estimators.

- Idea: construct signal with all frequencies with integer number op period

- => Multisine signals

- Advantages:
  - no leakage, no aliasing
  - better SNR compared to white noise

Closed loop case



$$\hat{H} = \begin{bmatrix} \dfrac{S_{z1y1} - S_{z2y1}\dfrac{S_{z1u2}}{S_{z2u2}}}{S_{z1u1} - S_{z2u1}\dfrac{S_{z1u2}}{S_{z2u2}}} & \dfrac{S_{z2y1} - S_{z1y1}\dfrac{S_{z2u1}}{S_{z1u1}}}{S_{z2u2} - S_{z1u2}\dfrac{S_{z2u1}}{S_{z1u1}}} \\ \dfrac{S_{z1y2} - S_{z2y2}\dfrac{S_{z1u2}}{S_{z2u2}}}{S_{z1u1} - S_{z2u1}\dfrac{S_{z1u2}}{S_{z2u2}}} & \dfrac{S_{z2y2} - S_{z1y2}\dfrac{S_{z2u1}}{S_{z1u1}}}{S_{z2u2} - S_{z1u2}\dfrac{S_{z2u1}}{S_{z1u1}}} \end{bmatrix}$$

Open loop: particular case

$$Z = U: \qquad \hat{H} = \begin{bmatrix} \dfrac{S_{u1y1} - S_{u2y1}\dfrac{S_{u1u2}}{S_{u2u2}}}{S_{u1u1} - S_{u2u1}\dfrac{S_{u1u2}}{S_{u2u2}}} & \dfrac{S_{u2y1} - S_{u1y1}\dfrac{S_{u2u1}}{S_{u1u1}}}{S_{u2u2} - S_{u1u2}\dfrac{S_{u2u1}}{S_{u1u1}}} \\ \dfrac{S_{u1y2} - S_{u2y2}\dfrac{S_{u1u2}}{S_{u2u2}}}{S_{u1u1} - S_{u2u1}\dfrac{S_{u1u2}}{S_{u2u2}}} & \dfrac{S_{u2y2} - S_{u1y2}\dfrac{S_{u2u1}}{S_{u1u1}}}{S_{u2u2} - S_{u1u2}\dfrac{S_{u2u1}}{S_{u1u1}}} \end{bmatrix}$$

- If inputs are decoupled, i.e. $S_{u1u2}=0$, then:

$$\hat{H} = \begin{bmatrix} \dfrac{S_{u1y1}}{S_{u1u1}} & \dfrac{S_{u2y1}}{S_{u2u2}} \\ \dfrac{S_{u1y2}}{S_{u1u1}} & \dfrac{S_{u2y2}}{S_{u2u2}} \end{bmatrix}$$