System Identification
Lecture 3

Identification of nonparametric frequency response models

Paul Van den Hof

Control Systems Group
Department of Electrical Engineering
Eindhoven University of Technology

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Contents

Introduction

F-transform of system

ETFE

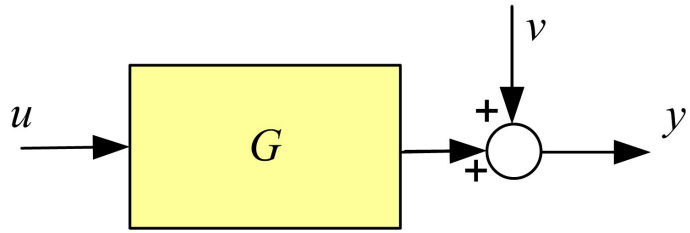
ETFE smoothing

Discussion

Chapter 3: Non-parametric Identification

Contents

- Transforming the systems relation to the f-domain
- Estimation of Frequency responses (FRFs) through ETFE
  - Periodic inputs
  - Non-periodic inputs
- Smoothing the ETFE for non-periodic inputs: windowing
- Alternative interpretation through spectral analysis



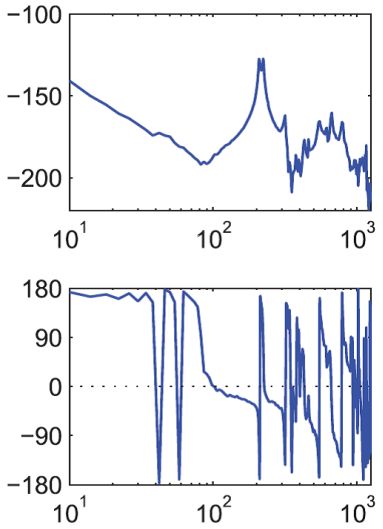
Identify a model in the form of

- Frequency response function (FRF) $G(e^{i\omega})$
  Bode plots: $|G(e^{i\omega})|$ and $\arg G(e^{i\omega})$ as function of ω.

Experiments:

- Sinusoidal signals (long experiments)
- Periodic signals
- General i/o data

Analyze the results in terms of bias and variance.



Transforming the systems relation to the f-domain

Transform the system's relation

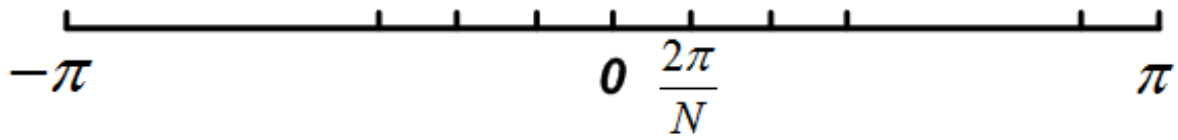
$$y(t) = G(q)u(t)$$

on the basis of N-length signals, to the frequency domain.

$$G(q) \rightarrow G(e^{i\omega})$$

through analysis of the DFT:

$$Y_N(e^{i\omega}),\ \omega = \frac{2\pi k}{N},\ k = 1, 2, ...$$



$$Y_N(e^{-i\omega}) = Y_N(e^{i\omega})^*$$

Result

Fourier-transformed system

It holds that

$$Y_N(\omega) = G(e^{i\omega})U_N(e^{i\omega}) + R_N(\omega)$$

with

- $R_N(\omega) = 0$ for all $\omega = \frac{2\pi k}{N}$, $k \in \mathbb{Z}$ if u is periodic extended outside the measurement interval.
- $|R_N(\omega)| \leq c_1$, for all ω, for general input signals

Proof (line of reasoning) for periodic input:

With $G(q) = \sum_{k=0}^{\infty} g_k q^{-k}$, and $y(t) = G(q)u(t)$, then

$$\begin{aligned} Y_N(\omega) &= \sum_{t=0}^{N-1} \sum_{k=0}^{\infty} g(k)u(t-k)e^{-i\omega t} \qquad (t-k \to \tau) \\ &= \sum_{k=0}^{\infty} g(k)e^{-i\omega k} \sum_{\tau=-k}^{N-1-k} u(\tau)e^{-i\omega\tau}. \end{aligned}$$

If:

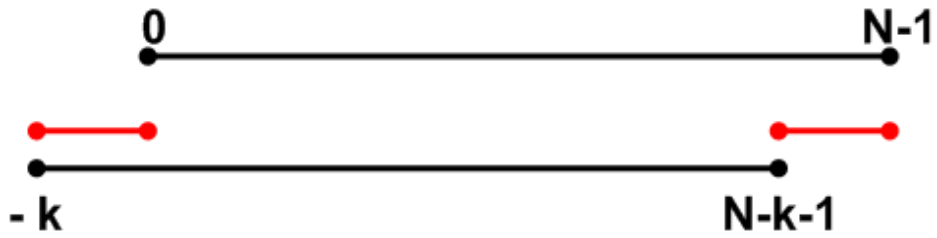
- u is periodic with period N, i.e. $u(\tau) = u(\tau + N)$

then

$$u(\tau)e^{-i\omega\tau} = u(\tau+N)e^{-i\omega\tau} = u(\tau+N)e^{-i\omega(\tau+N)}$$

for frequencies $\omega = 2\pi k/N$.

In other words: for frequencies $\omega = 2\pi k/N$ the function $u(\tau)e^{-i\omega\tau}$ is periodic with period N.



Summation interval can be shifted in time:

$$\sum_{\tau=-k}^{N-1-k} u(\tau)e^{-i\omega\tau} = \sum_{\tau=0}^{N-1} u(\tau)e^{-i\omega\tau} = U_N(e^{i\omega}).$$

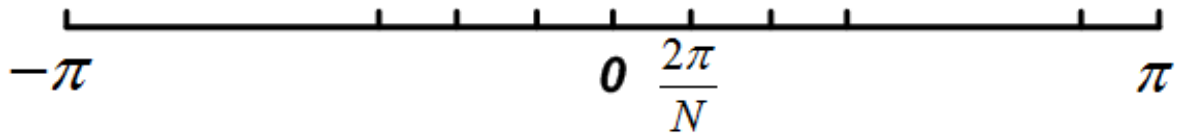
And as a result:

$$Y_N(\omega) = G(e^{i\omega})U_N(\omega)$$

(end proof)

The relevant frequency grid of DFT's

All information in $\{\ u(t) \mid t = 0...(N-1)\ \}$ is mapped to $U_N(\omega)$ at the $\frac{N}{2}$ frequencies $\omega_k = k\frac{2\pi}{N}$, $k = 0, 1, ...$ located in $[0\ \pi]$
For signals of length N this is the relevant frequency grid.



However, if $u(t)$ is periodic with fundamental frequency $\omega_0 = \frac{2\pi}{N_0}$. (i.e. periodic with period N_0), then the relevant DFT frequency grid becomes

$$\omega_k = k\omega_0 = k\frac{2\pi}{N_0},\ k = 1, 2, \cdots$$

located in $[0\ \pi]$.

Frequency Response Identification - ETFE

For

$$y(t) = G_0(q)u(t) + v(t)$$

it holds that

$$Y_N(\omega) = G_0(e^{i\omega})U_N(\omega) + R_N(\omega) + V_N(\omega)$$

(u considered to be fixed and deterministic)

Definition ETFE

The Empirical Transfer Function Estimate (ETFE) of G_0 is defined as:

$$\check{G}_N(e^{i\omega}) := \frac{Y_N(\omega)}{U_N(\omega)}$$

Analysis:

$$Y_N(\omega) = G_0(e^{i\omega})U_N(\omega) + R_N(\omega) + V_N(\omega)$$

$$\breve{G}_N(e^{i\omega}) = G_0(e^{i\omega}) + \frac{R_N(\omega)}{U_N(\omega)} + \frac{V_N(\omega)}{U_N(\omega)}$$

Basic result to relate DFT's (2)

Let v be a stationary stochastic process with $\mathbb{E}v = 0$ and spectral density $\Phi_v(\omega)$, then

$$EV_N(\omega) \quad = \quad 0 \quad \text{for all } \omega = \frac{2\pi k}{N}, k \in \mathbb{Z}.$$

$$\mathbb{E}\breve{G}_N(e^{i\omega}) = G_0(e^{i\omega}) + \underbrace{\frac{R_N(\omega)}{U_N(\omega)}}_{\textit{Leakage}}$$

$$\left|\frac{R_N(\omega)}{U_N(\omega)}\right| \leq \frac{c_1/\sqrt{N}}{\frac{1}{\sqrt{N}}U_N(\omega)}$$

Properties of $\breve{G}_N(e^{i\omega})$

- For periodic inputs $R_N(\omega) = 0$ at $\omega = k\frac{2\pi}{N_0}$, and so $\breve{G}_N(e^{i\omega})$ is unbiased in those frequencies.
- For general signals $\breve{G}_N(e^{i\omega})$ is asymptotically unbiased in those frequencies where $U_N(\omega) \neq 0$.

Illustration

$$y(t) = G_0(q)u(t) + H_0(q)e(t)$$

$$G_0(q) = \frac{B(q^{-1})}{A(q^{-1})}; \qquad H_0(q) = \frac{1}{A(q^{-1})}$$

$$\begin{aligned} B(q^{-1}) &= q^{-3}\left(0.103 + 0.181q^{-1}\right) \\ A(q^{-1}) &= 1 - 1.991q^{-1} + 2.203q^{-2} - 1.841q^{-3} + 0.894q^{-4} \end{aligned}$$

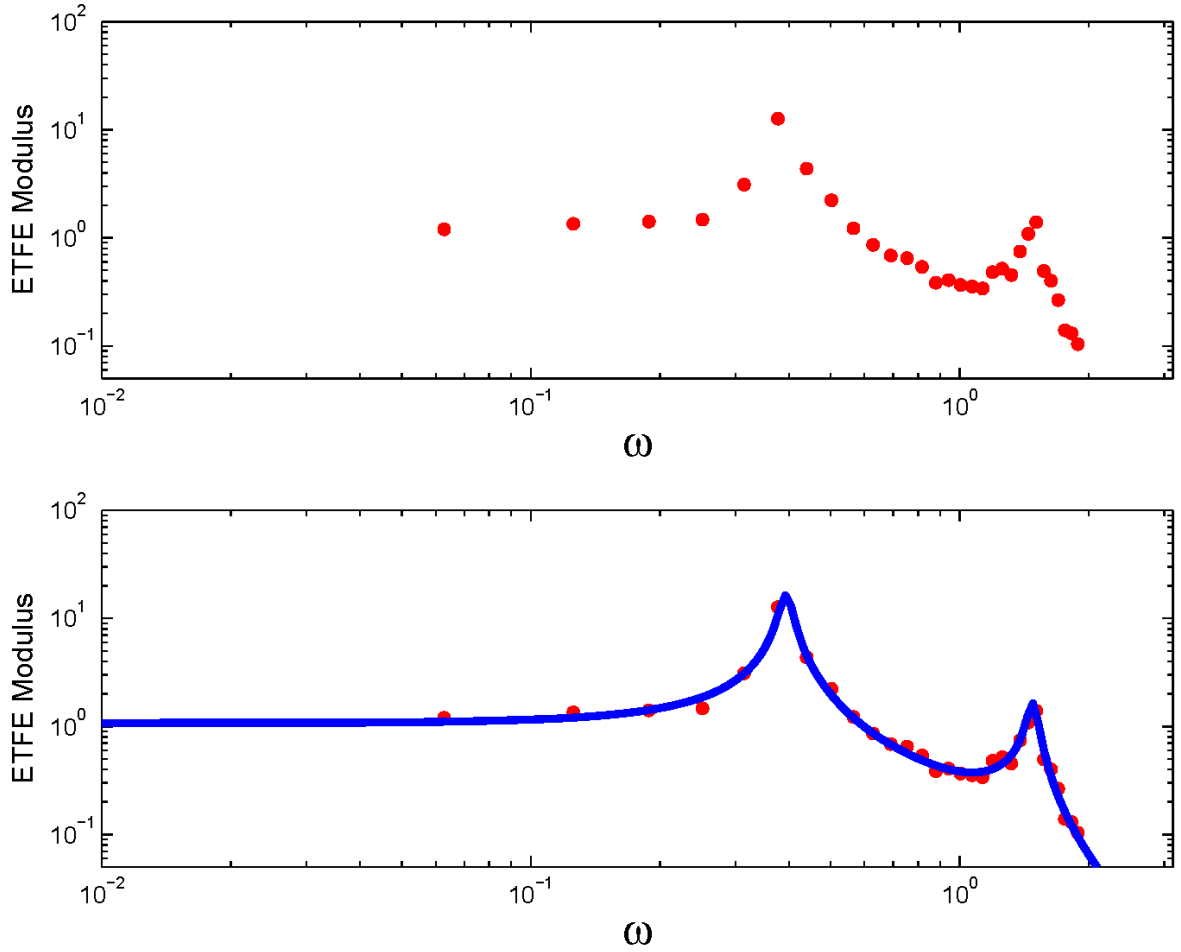
with $e(t)$ a white noise process with variance $\sigma_e^2 = 0.1$.
We collect $N = 10000$ data of this system with two different input signals having the same input power $\mathcal{P}_u = 0.5 = 5\sigma_e^2$

Input signal 1: a multisine of fundamental frequency $\omega_0 = \frac{2\pi}{100} \approx 0.06$ rad/sec (power=0.5)

$$u(t) = \frac{1}{\sqrt{30}} \sum_{k=1}^{30} sin(k\omega_0 t)$$

The ETFE is computed at the 30 harmonics of ω_0 present in $u(t)$

$N_0 = 100$. Excited frequencies: $k\frac{2\pi}{N_0}$, with $k = 1, \cdots 30$.



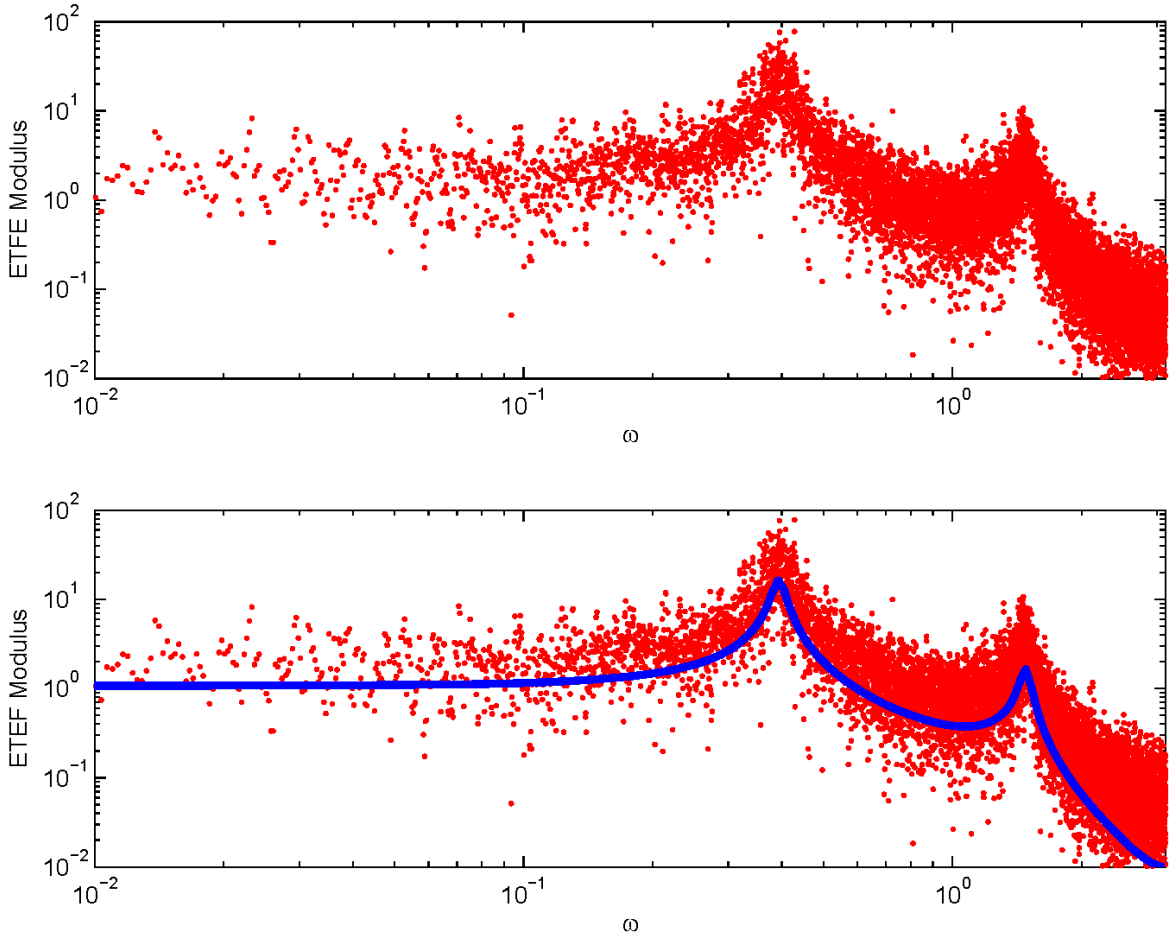
Upper plot: the ETFE at the 30 harmonics of ω_0; Lower plot: the same with the frequency response of $G_0(e^{i\omega})$ (blue).

$\Rightarrow$ ETFE is a good estimate of $G_0(e^{i\omega})$ at the harmonics of ω_0

Input signal 2: a white noise of variance 0.5

The ETFE is computed at all the $\frac{N}{2} = 5000$ frequencies ω_k

$$\omega_k = k\frac{\pi}{5000}, \quad k = 1, \cdots 5000$$



Upper plot: the ETFE at ω_k; Lower plot: the same with the frequency response of $G_0(e^{i\omega})$ (blue).

$\Rightarrow$ ETFE is an erratic and poor estimate of $G_0(e^{i\omega})$

How can we explain this?

For this purpose, we need to understand the statistical properties of the ETFE

- Bias has been analyzed and seems to be OK
- How about the variance?

We know from the expression for $\mathbb{E}\breve{G}_N(e^{i\omega})$ on slide 12 that

$$\breve{G}_N(e^{i\omega}) - \mathbb{E}\breve{G}_N(e^{i\omega}) = \frac{V_N(\omega)}{U_N(\omega)}$$

$$\begin{aligned} var(\breve{G}_N(e^{i\omega})) &:= \mathbb{E}\{|\breve{G}_N(e^{i\omega}) - \mathbb{E}\breve{G}_N(e^{i\omega})|^2\} \\ &= \frac{\mathbb{E}\frac{1}{N}|V_N(e^{i\omega})|^2}{\frac{1}{N}|U_N(e^{i\omega})|^2} \end{aligned}$$

Numerator:

Basic result to relate DFT's (3)

Let v be a stationary stochastic process with $\mathbb{E}v = 0$ and spectral density $\Phi_v(\omega)$, then

$$E\frac{1}{N}|V_N(\omega)|^2 \quad \rightarrow \quad \Phi_v(\omega) \quad \text{for } N \rightarrow \infty.$$

Denominator:

$$\frac{1}{N}|U_N(e^{i\omega})|^2$$

is the periodogram (power spectrum) of the (deterministic) signal u.

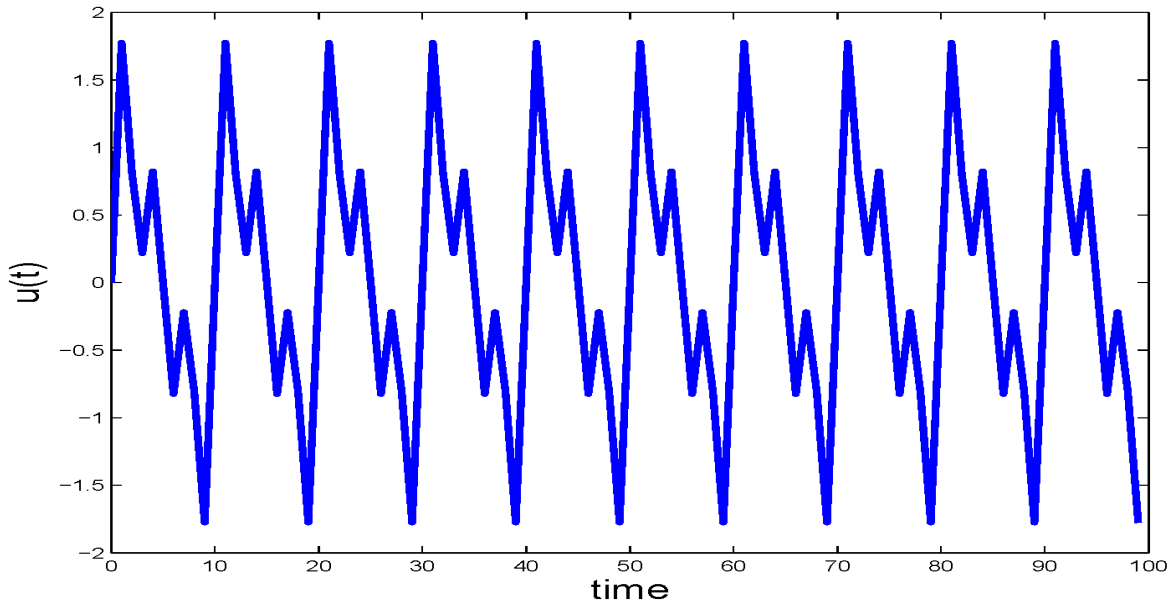
How does it behave when u is periodic?

Illustration

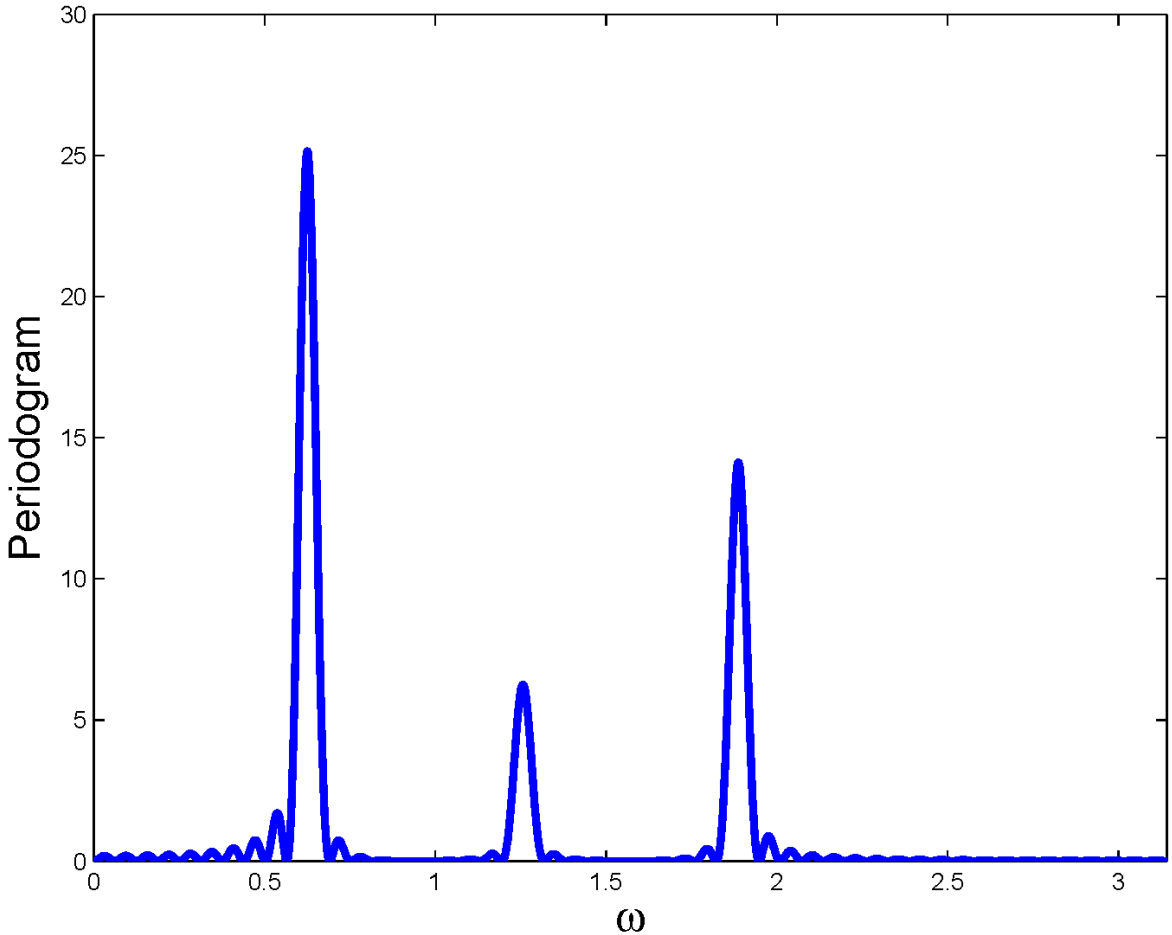
Consider periodic signal:

$$u(t) = sin(0.63t) + \frac{1}{2}sin(1.26t) + \frac{3}{4}sin(1.89t)$$

with fundamental period = 10 time-samples.

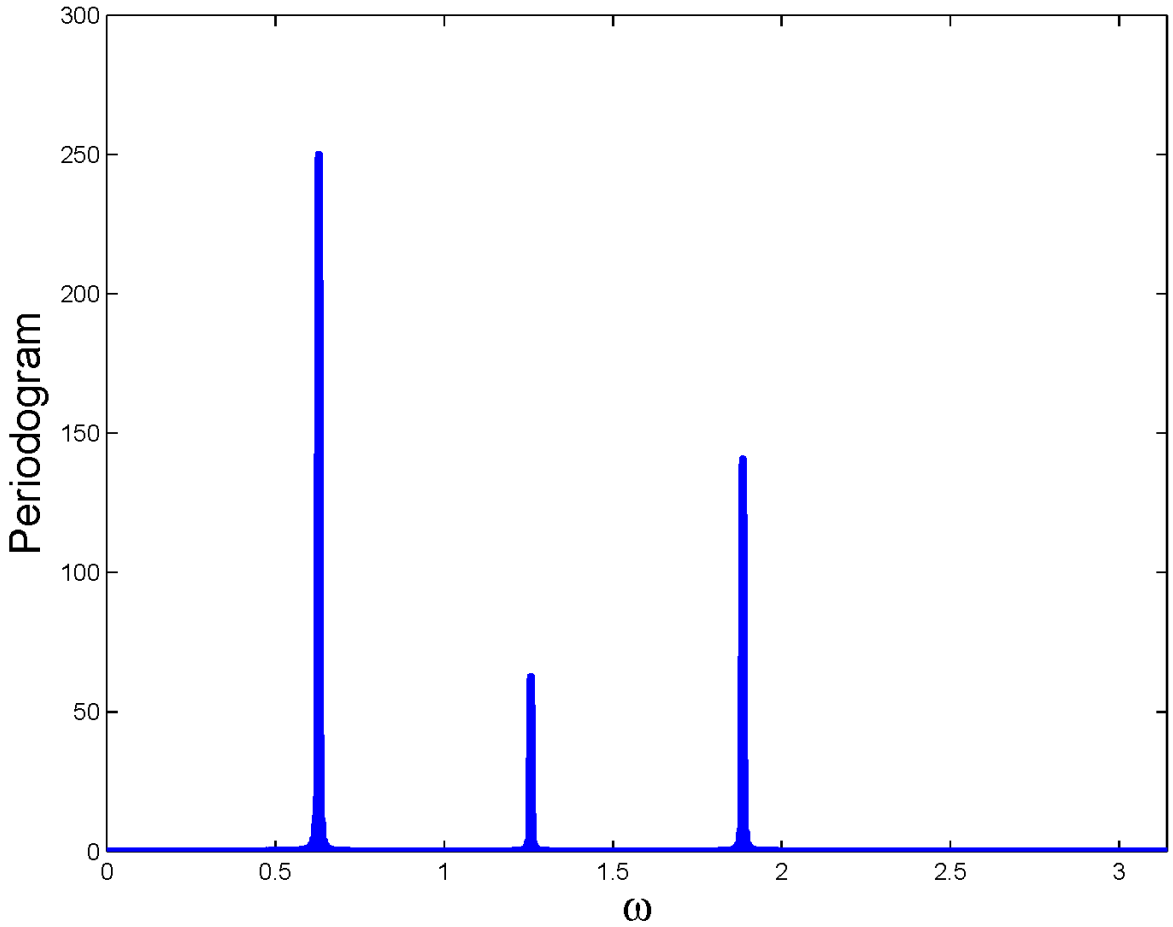


Obtained Periodogram $\frac{1}{N}|U_N(\omega)|^2$ for $\omega \in [0\ \pi]$ for $N = 100$:



The amplitudes of each peak are equal to $\frac{NA_i^2}{4}$ where A_i $(i = 1, 2, 3)$ is the amplitude of the sinusoid of frequency ω_i.

For $N \to \infty$, $\frac{1}{N}|U_N(\omega)|^2$ tends to $\Phi_u(\omega)$.
The periodogram for $N = 1000$:



Delta (Dirac) peak behaviour is understandable, since the finite power of the signal should satisfy

$$\mathcal{P}_u = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) d\omega$$

while $\Phi_u(\omega)$ is only nonzero at a finite number of frequencies.

ETFE result for periodic input signal

Let u be a periodic signal with period N_0 and length $N = rN_0$, r integer. For $\omega = k\frac{2\pi}{N_0}$, $k = 1, \cdots N_0$:

(a) ETFE is asymptotically unbiased, $\mathbb{E}\breve{G}(e^{i\omega}) = G_0(e^{i\omega})$, at those frequencies where $\Phi_u(\omega) \neq 0$;

(b) The asymptotic variance $\dfrac{\Phi_v(\omega)}{\Phi_u(\omega)}$ tends to 0 for increasing values of r.

Consequence: variance $\to 0$ for frequencies in the grid (because Φ_u converges to dirac functions)

Consistent estimate in limited number of f-points

ETFE result for non-periodic input signal

For all ω:

(a) ETFE is asymptotic unbiased, $\mathbb{E}\check{G}(e^{i\omega}) = G_0(e^{i\omega})$, at those frequencies where $\Phi_u(\omega) \neq 0$;

(b) For N large, the variance of $\check{G}$ is $\dfrac{\Phi_v(\omega)}{\Phi_u(\omega)}$ and does not tend to 0.

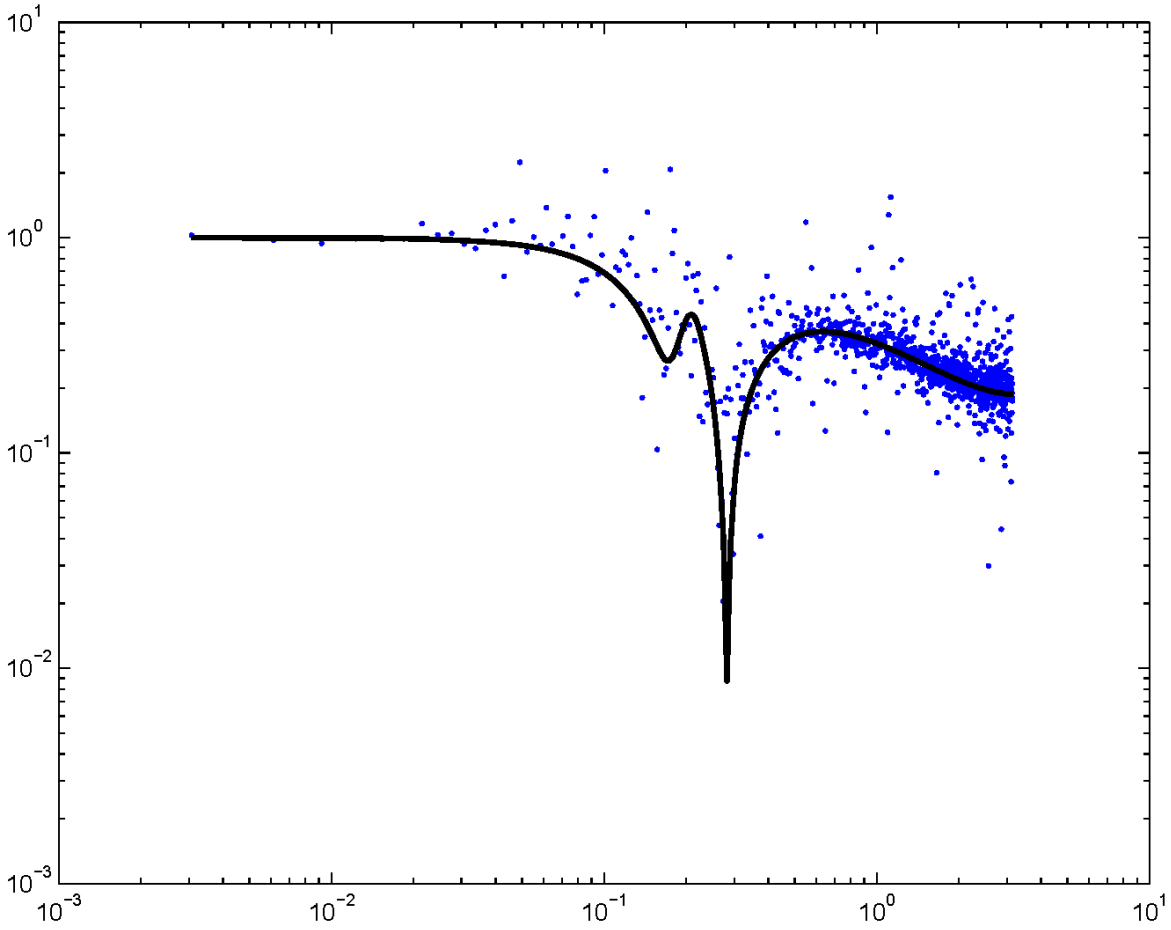
Consequence: variance does not decay for $N \to \infty$ as $\Phi_u(\omega)$ is bounded (does not contain Dirac pulses).

$\Rightarrow$ No consistency

Example

Bode-amplitude plot of data-generating system, and ETFE on the basis of

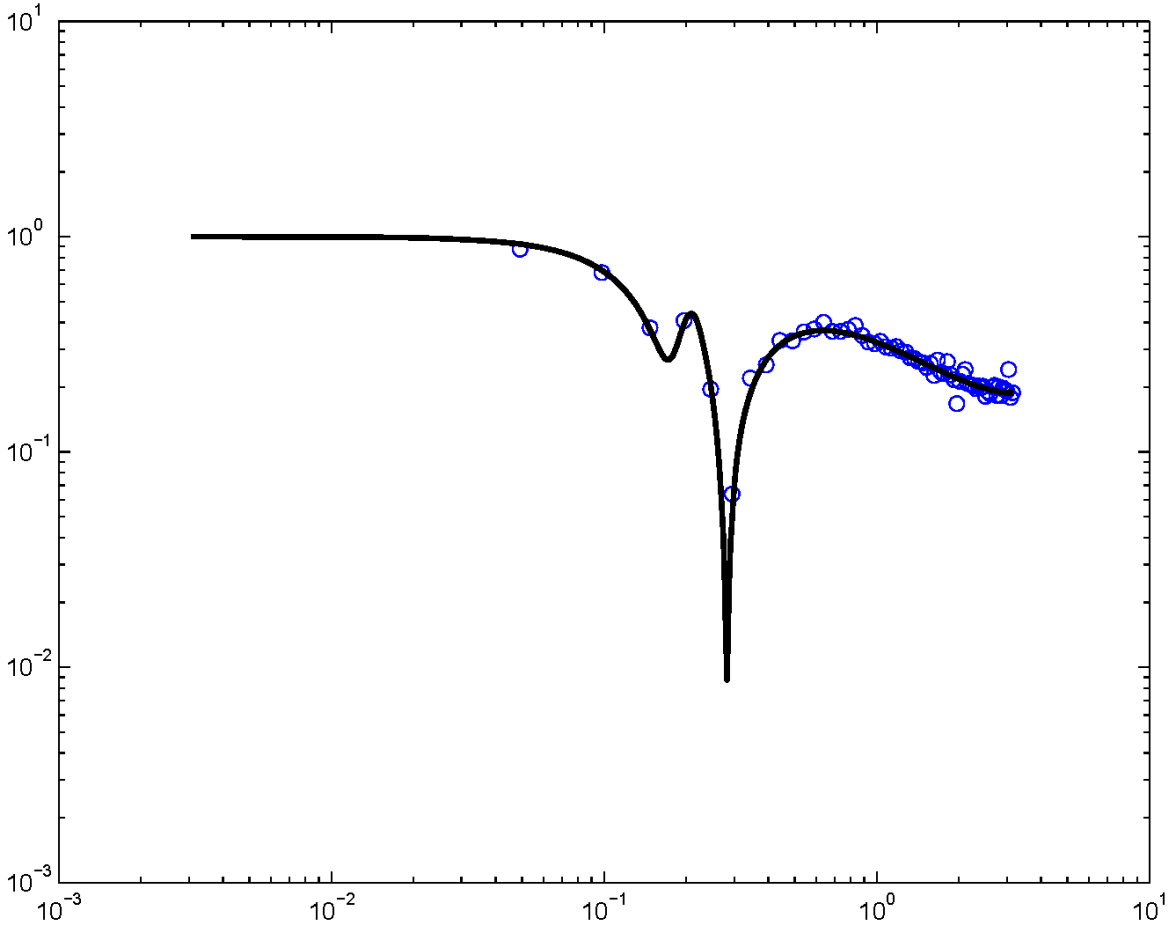
- input signal: one realization of a white noise process
- N=2048



ETFE is obtained in N frequency points in $[-\pi, \pi]$
Smallest frequency is: $2\pi/N$ (resolution)

If u periodic with period N_0, then $\Phi_u(\omega)$ has contributions in N_0 frequency points.
Same example with

- white noise u with $N_0 = 128$
- 16 times repeated: $N = 16 \cdot N_0$



- Information in smaller number of f-points
- variance is reduced

Methods for variance reduction:

- Periodic input signal
  $\Phi_u(\omega)$ unbounded in a limited number of f-points
  instead of $\Phi_u(\omega)$ bounded in infinite number of f-points
- "Smoothing" of ETFE through the use of *windows*
  Averaging over neighbouring f-points
- Periodogram averaging (splitting the data sequence)

ETFE smoothing

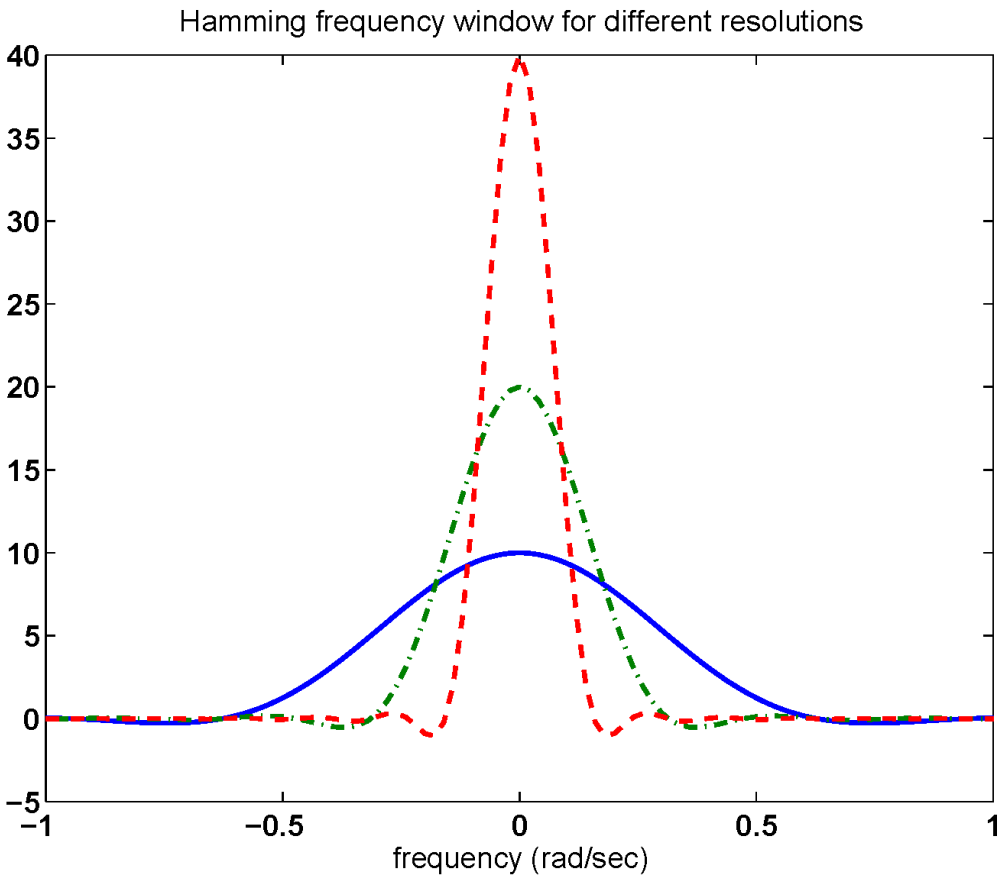
1. Windowing

Smoothing is motivated by:

- ETFE estimates are independent for different ω's
- Averaging over a frequency area where G_0 is constant reduces the variance

$$\hat{G}(e^{i\omega}) = \frac{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega)\breve{G}_N(e^{i\xi})d\xi}{\int_{-\pi}^{\pi} W_\gamma(\xi - \omega)d\xi}$$

with $W_\gamma(\xi)$ a positive real-valued frequency-function. (convolution window)



$W_\gamma(\xi)$ of Hamming window for $\gamma = 10$ (solid), $\gamma = 20$ (dash-dotted) and $\gamma = 40$ (dashed).

γ is measure for the width of the window.

Interpretation:

At each frequency ω_k, the ETFE $\breve{G}(e^{i\omega_k})$ is replaced by a weighted average of values $\breve{G}(e^{i\omega})$ in the neighborhood of ω_k, with a weight determined by the window function that operates on $\omega - \omega_k$.

- Window introduces bias in an attempt to reduce the variance (bias/variance trade-off)
- Choice of window width dependent on expected smoothness of $G_0(e^{i\omega})$
- Window too narrow: variance too large
  Window too wide: possible smoothing of underlying dynamics
  $\rightarrow$ bias

- Convolution in f-domain is approximated because of discrete number of f-points
- Variance of $\breve{G}_N(e^{i\omega})$ has to be taken into account when averaging (see later)

2. Periodogram averaging

Observation:
ETFE of data sequence of length N provides estimates in $\omega = 2\pi k/N$.
When splitting the data in r parts (subintervals) of length N_0, we can

- Estimate an ETFE on each subinterval
- Average the ETFE's over the r intervals

The frequency grid is forced to remain limited to $2\pi k/N_0$, even with increasing data.

This procedure (periodogram averaging, Welch's method) is implemented in Matlab's `tfestimate`.

3. Smoothing through lag windows in time domain

Alternative interpretation of smoothing: spectral analysis

The ETFE can be written as

$$\breve{G}(e^{i\omega}) := \frac{Y_N(e^{i\omega})}{U_N(e^{i\omega})} = \frac{\frac{1}{N} Y_N(e^{i\omega}) U_N^*(e^{i\omega})}{\frac{1}{N} U_N(e^{i\omega}) U_N^*(e^{i\omega})}.$$

It appears that $\frac{1}{N} Y_N(e^{i\omega}) U_N^*(e^{i\omega})$ is an estimate of the cross-spectral density $\Phi_{yu}(\omega)$.

Spectral estimates on the basis of N data

There exists a simple Fourier-transform relation:

$$\hat{\Phi}_{yu}^N(\omega) = \frac{1}{N} Y_N(\omega) U_N^*(\omega).$$

$$\overset{\Leftrightarrow}{\textit{Fourier}}$$

$$\hat{R}_{yu}^N(\tau) := \frac{1}{N} \sum_{t=0}^{N-1} y(t)u(t-\tau), \qquad |\tau| < N \text{ and } 0 \text{ elsewhere}$$

and similarly for $\hat{\Phi}_u^N(\omega) = \frac{1}{N}|U_N(\omega)|^2$ (deterministic)

So an equivalent formulation of the ETFE can be obtained according to:

$$\breve{G}(e^{i\omega}) = \frac{\sum_{\tau=-N}^{N} \hat{R}_{yu}^{N}(\tau) e^{-i\omega\tau}}{\sum_{\tau=-N}^{N} \hat{R}_{u}^{N}(\tau) e^{-i\omega\tau}}$$

Therefore:

$$\text{Convolution window } W_\gamma(\omega) \text{ on } \hat{\Phi}_{yu}^N(\omega)$$

is equivalent with

$$\text{Multiplication window } w_\gamma(\tau) \text{ on } \hat{R}_{yu}^N(\tau)$$

while $w_\gamma(\tau) \overset{\Leftrightarrow}{Fourier} W_\gamma(\omega)$.

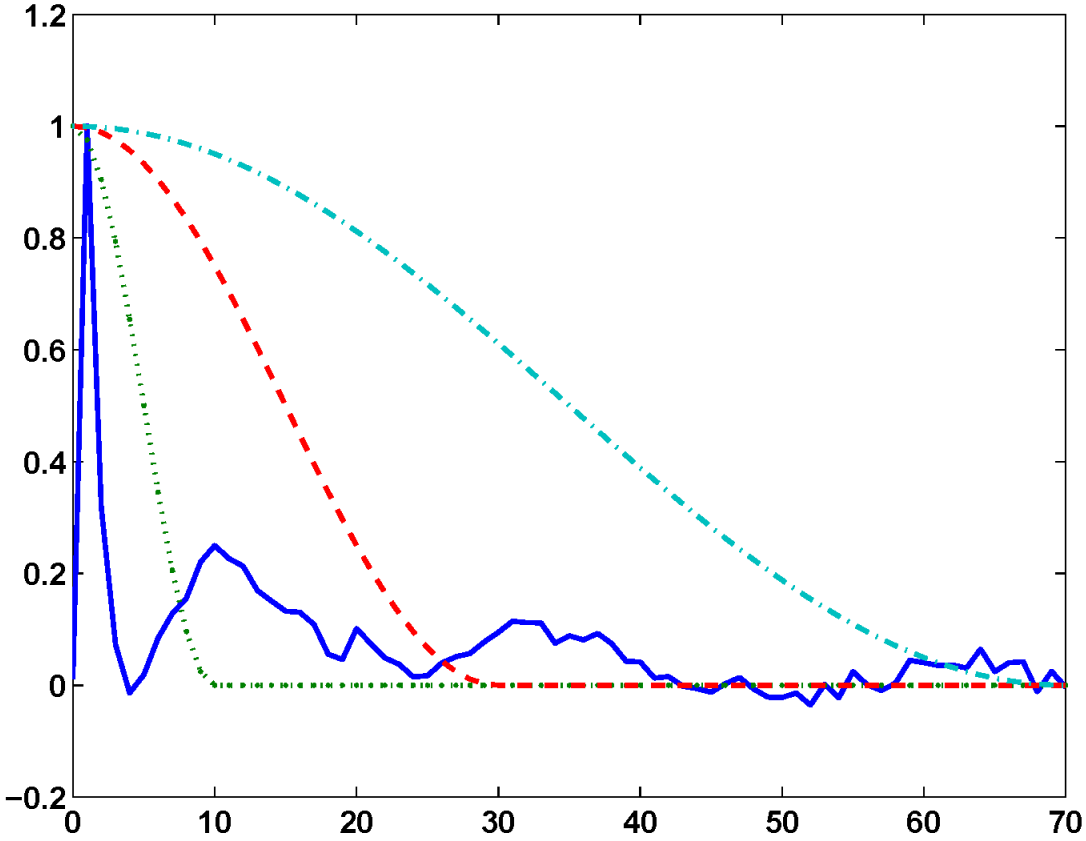
Equivalent smoothing of estimate through a time/lag-window:

$$\hat{\Phi}_{yu}^N(\omega) = \sum_{\tau=-\gamma}^{\gamma} w_\gamma(\tau)\hat{R}_{yu}^N(\tau)e^{-i\omega\tau}$$

with $w_\gamma(\tau)$ a window with width γ: $w_\gamma(\tau) = 0,\ |\tau| > \gamma$.

- γ indicates the maximum time lag τ for which $\hat{R}_{yu}^N(\tau)$ is considered reliable.
- Note that for increasing τ, $\hat{R}_{yu}^N(\tau)$ is based on fewer samples (higher variance).

Hamming lag-window



$\hat{R}^N_{yu}(\tau)$ (solid) together with the Hamming lag-windows $w_{10}(\tau)$ (dotted), $w_{30}(\tau)$ (dashed) and $w_{70}(\tau)$ (dash-dotted).

time-window ("lag window") more narrow $\rightarrow$

$\rightarrow$ γ smaller

$\rightarrow$ stronger "smoothing"

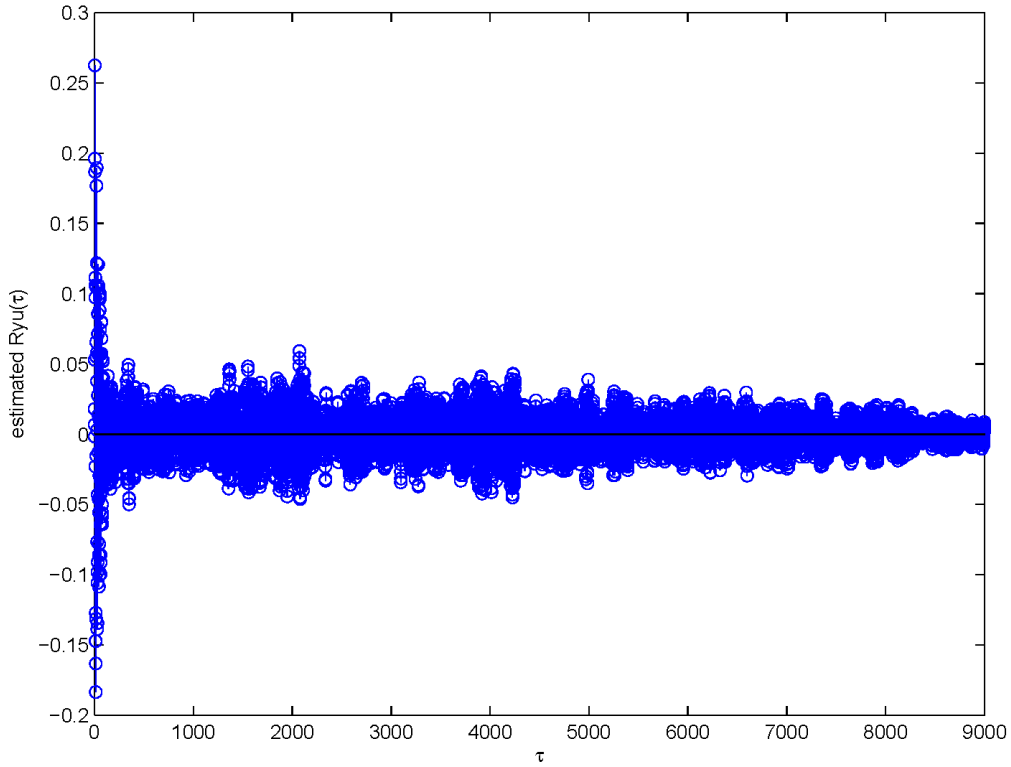
$\rightarrow$ frequency window more wide

- γ sufficiently small w.r.t. N
- $|\hat{R}^N_{yu}(\tau)| << \hat{R}^N_{yu}(0)$ for $\tau > \gamma$.

γ is equal to the smoothing variable M in the Matlab tool ETFE.

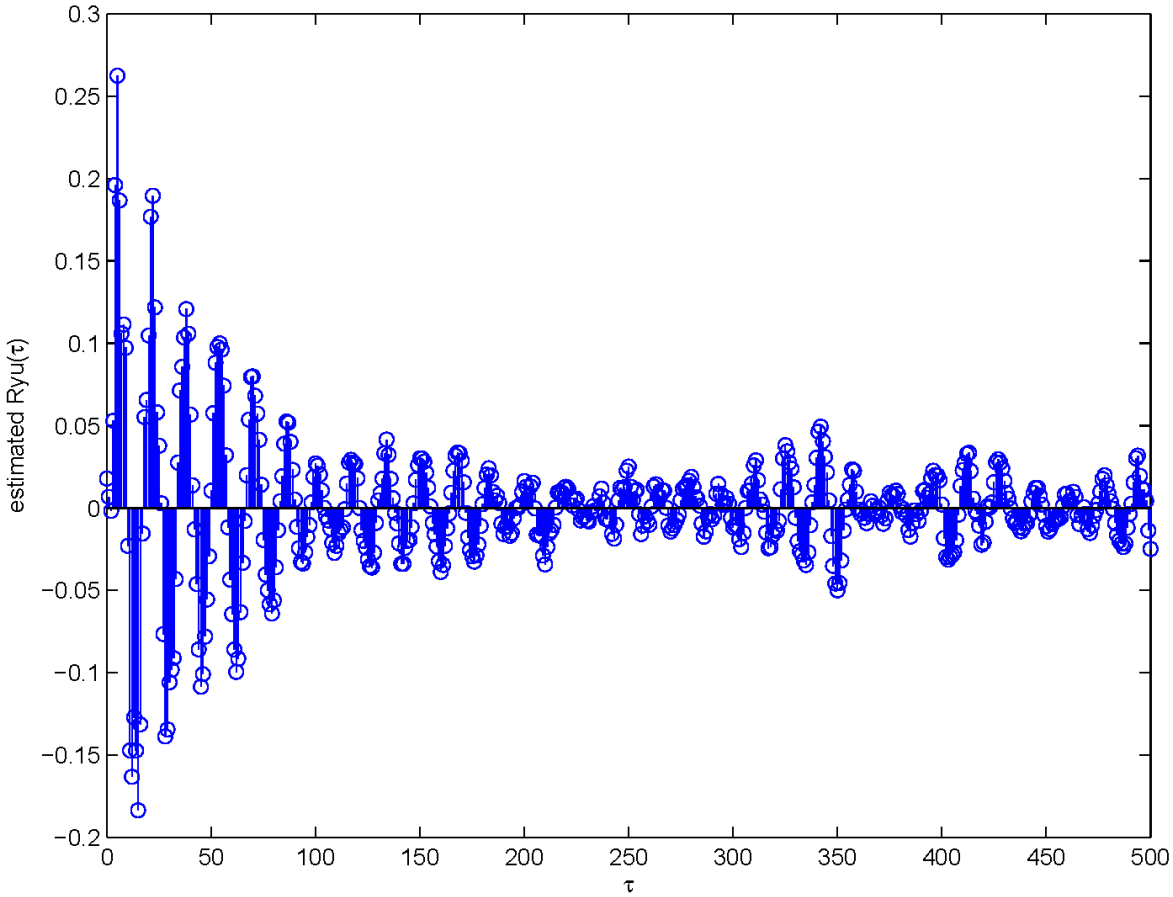
Illustration (cont'd from example, slide 3-13):

We compute $\hat{R}^N_{yu}(\tau)$ with the data generated by the white noise input of variance 0.5



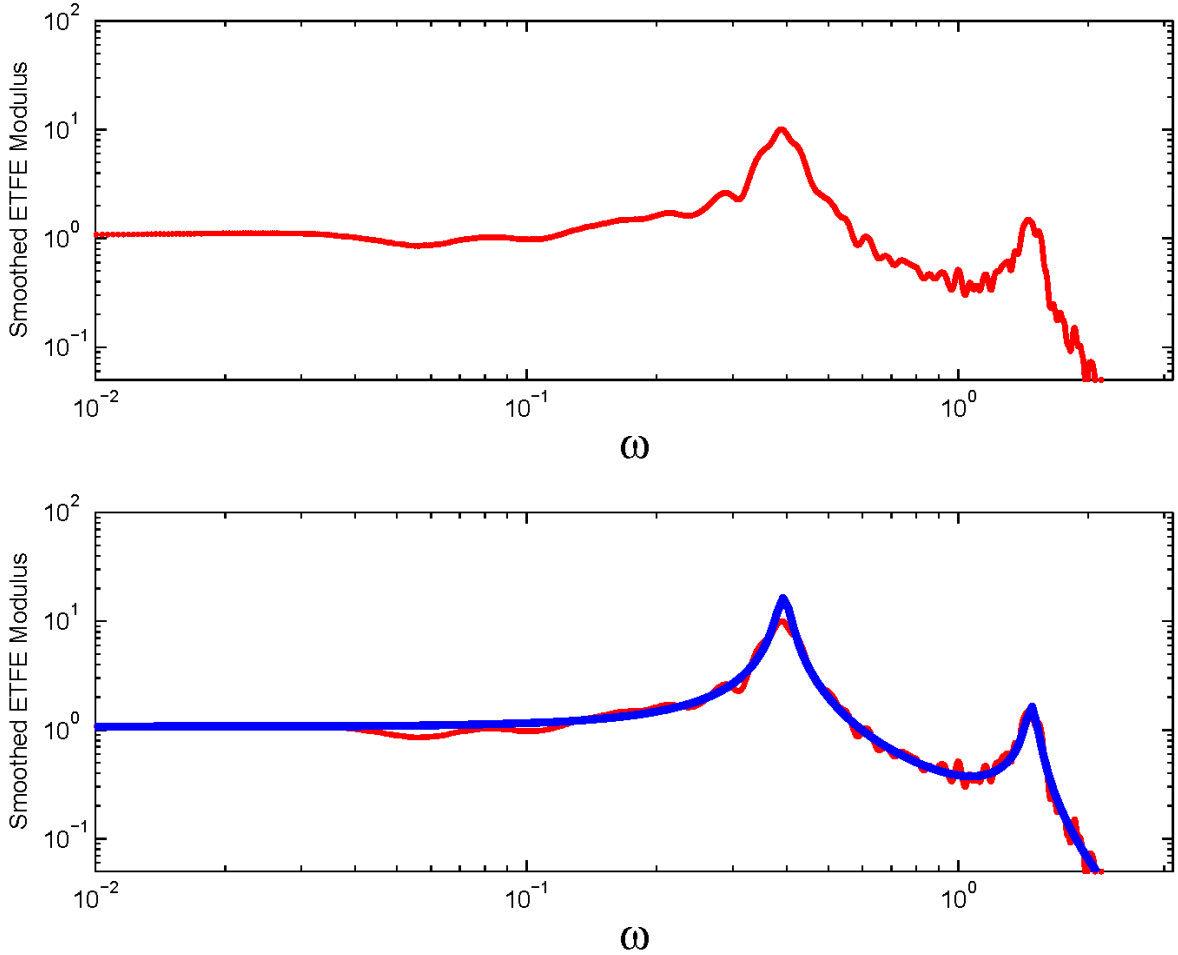
we see the inaccuracy of the estimate: $\hat{R}^N_{yu}(\tau)$ does not tend to 0 for $\tau \to \infty$

Let us focus on the first 500 τ's



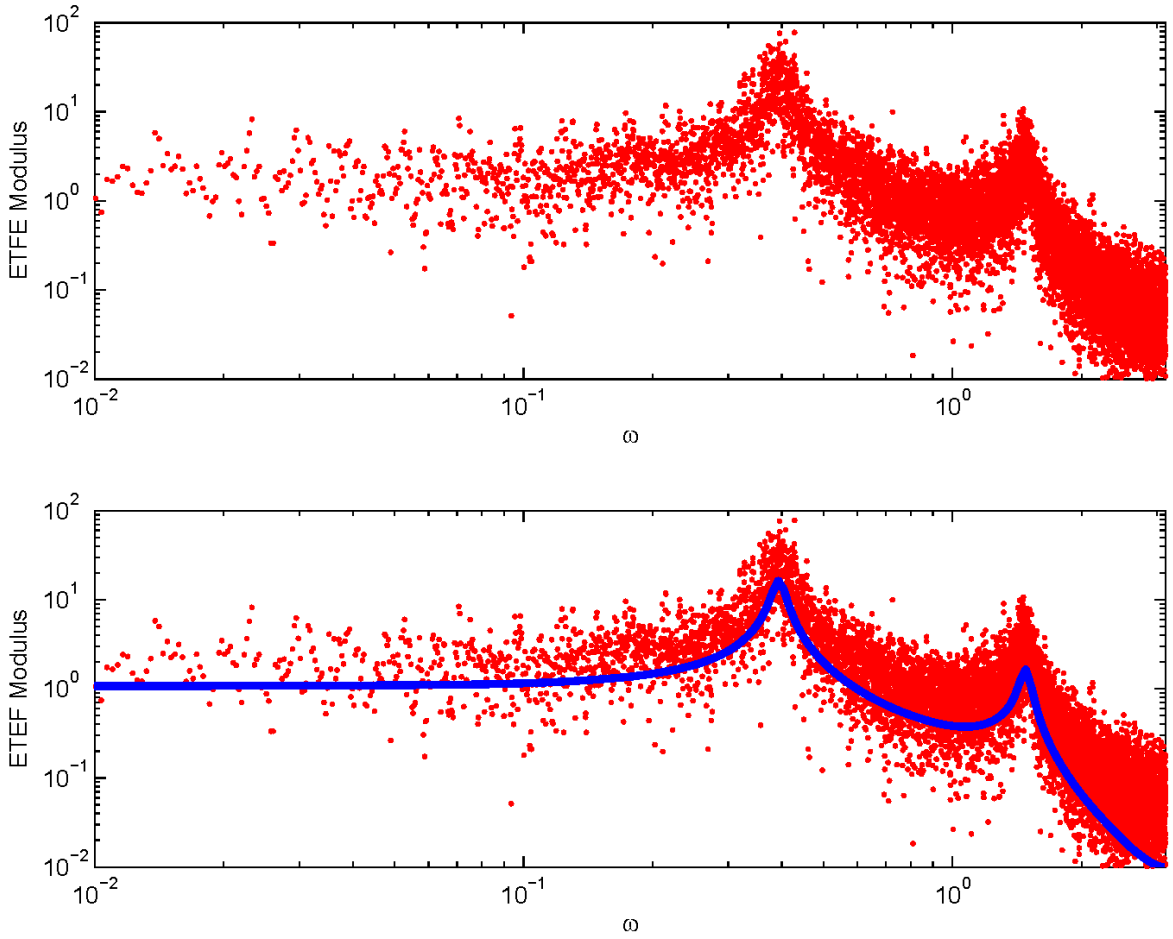
we see that, after $\tau = 100$, $\hat{R}^N_{yu}(\tau)$ increases again which could be unlikely for $R_{yu}(\tau) \Longrightarrow$ we select $\gamma = 100$

Obtained smoothed ETFE with $\gamma = 100$



Upper plot: the amplitude of the smoothed ETFE; Lower plot: the same with the frequency response of $G_0(e^{i\omega})$ (blue)

Compare with the original ETFE:



Difference between ETFE and spectral estimate:

- Without smoothing: identical
- Frequency window (ETFE) versus time-window (Spec.anal).

The windows are related through DFT.
This shows the close relation.

General statement

- Periodic input signals are strongly advised:
  - No bias problems even for finite N if an integer number of periods is applied
  - Variance reduces with data length, without introducing bias
- Care has to be taken when input is non-periodic: smoothing is an "engineering" tool that is non-optimal.

Final remarks: nonparametric estimation

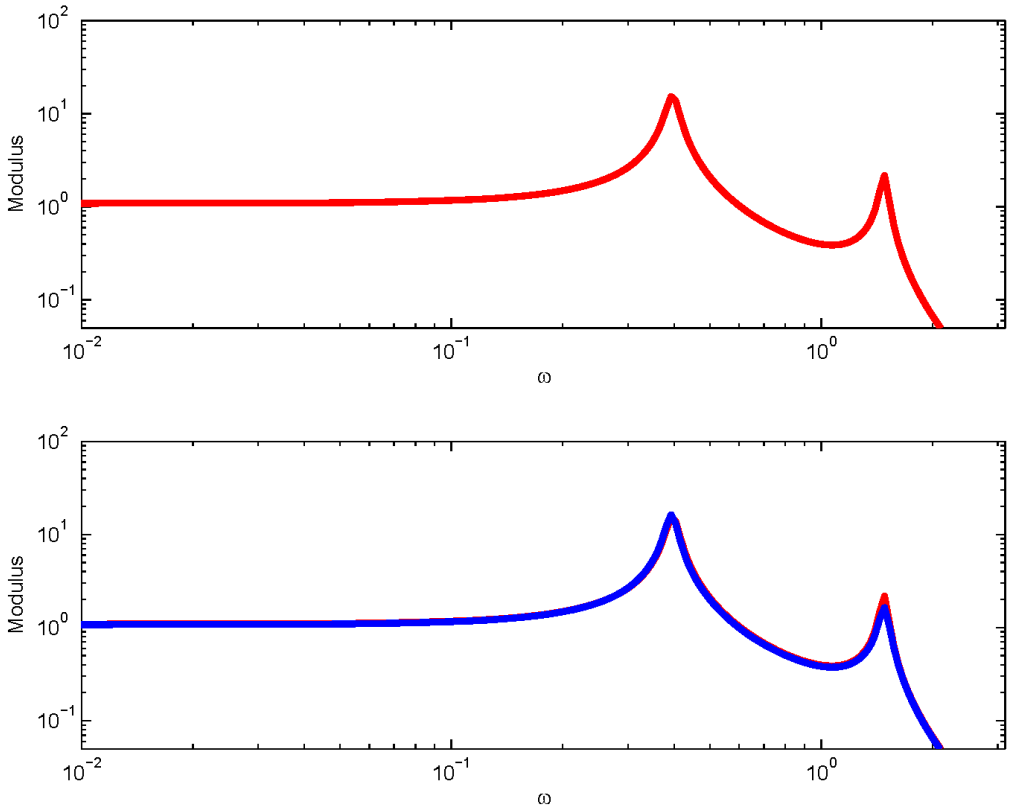
The methods provide an estimate of the frequency response of $G_0(e^{i\omega})$ and not of the rational transfer function $G_0(z)$.
For simulation purposes a parametric model (transfer function) is required.
A controller can be designed on the basis of an f-response (loop shaping). More advanced control design schemes require a low-order (parametric) model.

$\Longrightarrow$ parametric identification (prediction error identification)

- delivers a model of the plant G_0 and information on $\Phi_v(\omega)$
- higher accuracy:

  $var(G(e^{i\omega}, \hat{\theta}_N) \approx \frac{n}{N} \frac{\Phi_v(\omega)}{\Phi_u(\omega)}$ with PEI, where n is model order

Illustration (cont'd): Parametric Identification of G_0 with the 1000 first samples of the white noise input case with variance 0.5



Upper plot: frequency response of the identified model; Lower plot: the same with the frequency response of $G_0(e^{i\omega})$ (blue)

The previous figure has to be compared with the non-smoothed ETFE and the smoothed ETFE
This comparison shows that PEI delivers much better results even with ten times less data points, provided that the right structure of the parametric model is chosen.

Summary

- Non-parametric identification (FRF estimation) provides a relative simple tool for estimating the frequency response of a linear system
- The estimated model can not be used for simulation, but can e.g. be used for f-domain feedback controller tuning
- Periodic excitation has clear advantages (consistency)
- For non-periodic excitation a smoothing operation is required for variance reduction, at the cost of introducing bias; this requires careful (engineering) tuning