Recent Advances on Bootstrap Methods for Spectral Analysis

Abdelhak M Zoubir

Signal Processing Group

Signal Processing Group
Technische Universität Darmstadt
E-Mail: zoubir@spg.tu-darmstadt.de
URL: www.spg.tu-darmstadt.de

Motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

- Spectral analysis is an important tool for science and engineering and has been a subject of study since the 19th Century [Stokes (1879), Schuster (1894)].
- There are three principal reasons for using spectral analysis:
  - to provide useful descriptive statistics,
  - as a diagnostic tool to indicate which further analyses might be relevant,
  - to check postulated theoretical models.
- Frequency analysis is now firmly involved in various areas of electrical engineering, such as communications, radar and sonar.
- Statistical inference for spectra and functions of spectra such coherence and transfer functions is fundamental in these areas.
- Most techniques for statistical inference rely on assumptions [Rosenblatt (1985), Priestley (1981)], which may not hold in practice.

Motivation (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- Let X_t, $t \in \mathbb{Z}$, be a real-valued strictly stationary process with covariance function $c_{XX}(\tau) = \mathrm{E}[X_0 X_{|\tau|}]$, $\tau \in \mathbb{Z}$, and spectral density

$$C_{XX}(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} c_{XX}(\tau) e^{-j\omega\tau}, \qquad -\pi < \omega \leq \pi.$$

- Given the model $X_1, X_2, \ldots, X_T$ for the observations $x_1, x_2, \ldots, x_T$,

we wish to find:

- an estimator $\hat{C}_{XX}(\omega)$ for $C_{XX}(\omega)$, and
- an $100(1-\alpha)\%$ confidence interval for $C_{XX}(\omega)$, for a given α.

Motivation (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- Define the periodogram of X_t,

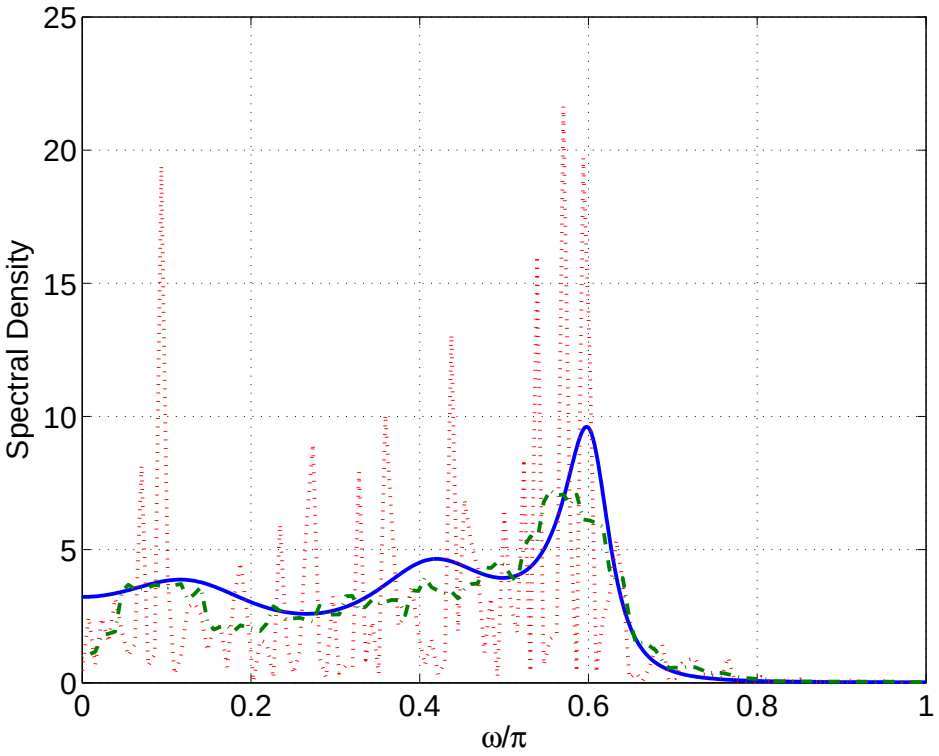
$$I_{XX}(\omega) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} X_t e^{-j\omega t}\right|^2, \qquad -\pi < \omega \leq \pi$$

- and construct the estimator of $C_{XX}(\omega)$ by

$$\hat{C}_{XX}(\omega) = \frac{1}{T \cdot h}\sum_{k=-n}^{n} K\left(\frac{\omega - \omega_k}{h}\right) I_{XX}(\omega_k),$$

where the kernel $K(\cdot)$ is a known symmetric, non-negative, real-valued function, h is its bandwidth and $n = [T/2]$.

Motivation (Cont'd)



Spectral density of an AR(6) process (—), the periodogram (...) and the spectral density estimate (-.-) for $T = 256$, $h = 0.1$ and the Bartlett-Priestley kernel $K(\cdot)$.

Motivation (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- Under some regularity conditions, we have asymptotically for $T \to \infty$,

$$\hat{C}_{XX}(\omega) \sim \frac{C_{XX}(\omega)}{4m+2}\chi^2_{4m+2},$$

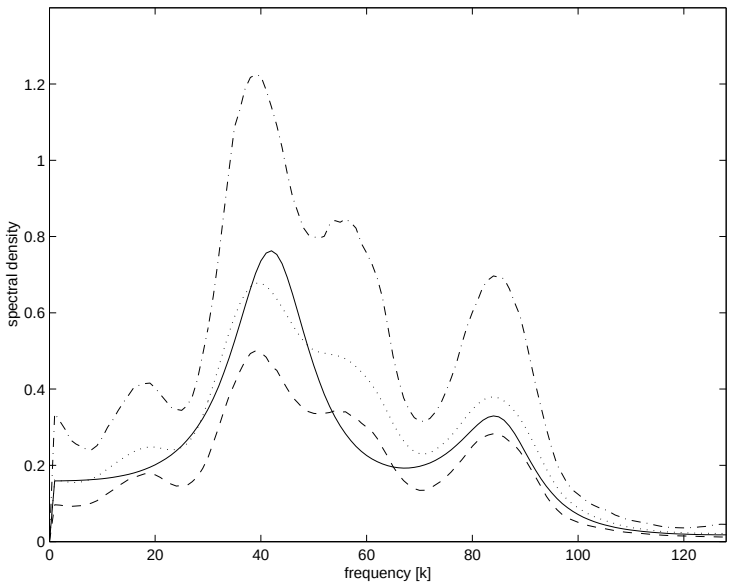
$m = \lfloor (hT-1)/2 \rfloor$.

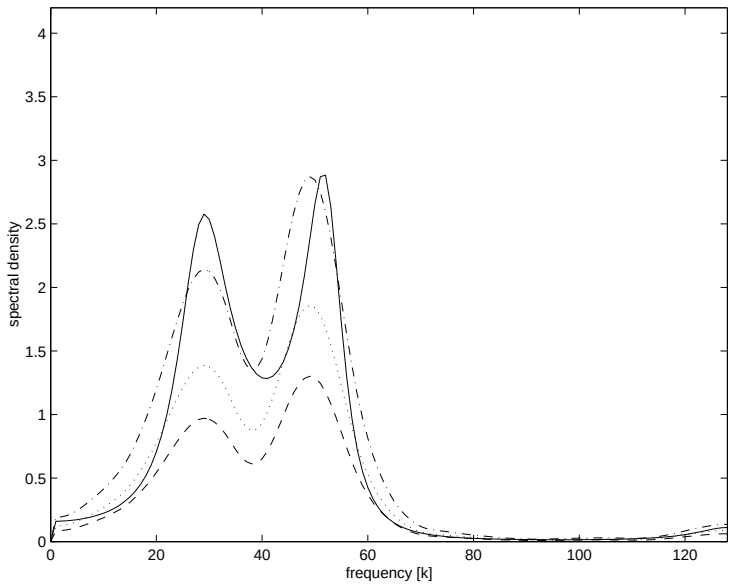
- An $100(1-\alpha)\%$ confidence interval for $C_{XX}(\omega)$ is found as,

$$\frac{(4m+2)\hat{C}_{XX}(\omega)}{\chi^2_{4m+2}(1-\frac{\alpha}{2})} < C_{XX}(\omega) < \frac{(4m+2)\hat{C}_{XX}(\omega)}{\chi^2_{4m+2}(\frac{\alpha}{2})},$$

where $\chi^2_\nu(\alpha)$ is such that $\text{Prob}[\chi^2_\nu < \chi^2_\nu(\alpha)] = \alpha$ [Brillinger (1981)].

Motivation (Cont'd)





95% confidence interval for $C_{XX}(\omega)$ of the Gaussian process $X_t = 0.5X_{t-1} - 0.6X_{t-2} + 0.3X_{t-3} - 0.4X_{t-4} + 0.2X_{t-5} + \varepsilon_t$ (left) and $Y_t = Y_{t-1} - 0.7Y_{t-2} - 0.4Y_{t-3} + 0.6Y_{t-4} - 0.5Y_{t-5} + \zeta_t$ (right), where ζ_t are identically and independently uniformly distributed on $\mathcal{U}[-\pi, \pi)$. The true $C_{XX}(\omega)$ (—) and an estimate based on $T = 256$ (...) are given.

Limitations

TECHNISCHE UNIVERSITÄT DARMSTADT

- We assumed a large T so that $\hat{C}_{XX}(\omega_1), \ldots, \hat{C}_{XX}(\omega_M)$ are independently and identically distributed as χ^2 ramdom variates.
- Asymptotic approaches are often inapplicable:

- real-time applications do not always permit collection of long data segments [Bendat & Piersol (2000)], or
- we restrict the sample size to be small for the (quasi) stationarity assumption.

- analytical solutions for the finite sample case are often tedious or impossible.
- Bootstrap techniques are the alternative to asymptotic theory for short data segments and perform usually better under weaker assumptions [Efron (1979), Hall (1992), Efron & Tibshirani (1993), Davison & Hinkley (2006)].

Contents

TECHNISCHE UNIVERSITÄT DARMSTADT

Contents

TECHNISCHE UNIVERSITÄT DARMSTADT

Introduction (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- The term *bootstrap* is often associated with *The Baron von Münchhausen*.

Courtesy of Prof. Dr. Karl Heinrich Hofmann, TU Darmstadt and Head of Group *Algebra, Geometry and Functional Analysis*

Introduction (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- This analogy may suggest that the bootstrap is able to perform the impossible.
- The bootstrap is not a magic technique that provides a panacea for all statistical inference problems!
- The bootstrap is a powerful tool that can substitute tedious or often impossible theoretical derivations with computational calculations.
- There are situations where the bootstrap fails and care is required [Young (1994)].

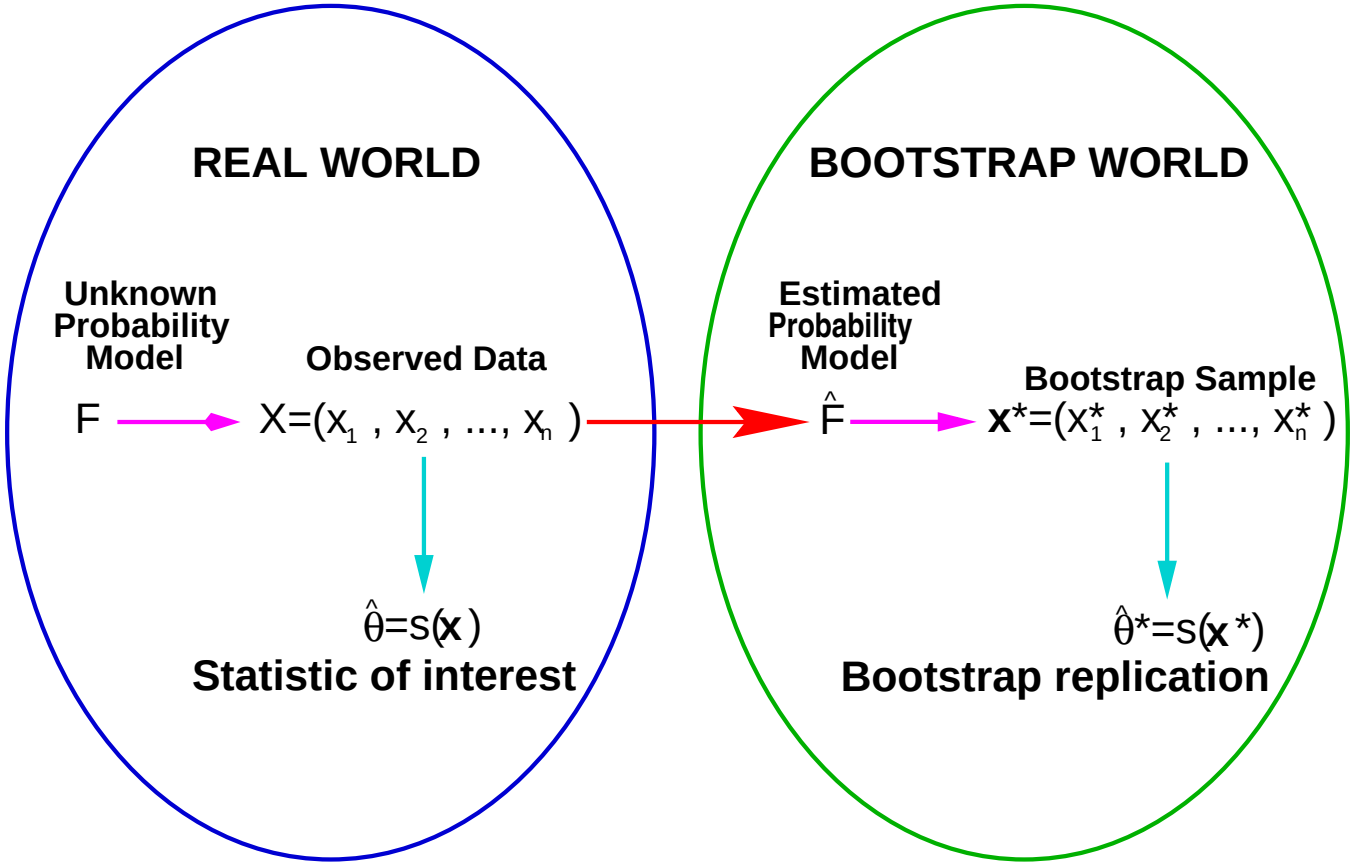
What can I use the bootstrap for?

TECHNISCHE UNIVERSITÄT DARMSTADT

- *The bootstrap is a computational tool for statistical inference*.
- It can be used for:
  - Estimation of statistical characteristics such as *bias*, *variance*, *distribution function of estimators* and thus *confidence intervals*.
  - Hypothesis tests, for example for *signal detection*, and
  - *model selection*.
- When can I use the bootstrap?

When I know little about the statistics of the data and/or I have only a few data so that asymptotic theory does not hold.

The Bootstrap Idea



Basic idea: simulate the probability mechanism of the real world by substituting the unknowns with estimates derived from the data.

Why Does the Bootstrap Work?

TECHNISCHE UNIVERSITÄT DARMSTADT

- Let the distribution of T admit an Edgeworth expansion,

$$G(x) = \mathrm{P}(T \leq x) = \Phi(x) + n^{-1/2} q(x)\phi(x) + O(n^{-1}),$$

  where q is an even quadratic polynomial and Φ, ϕ are standard normal distribution and density function, respectively.
- The bootstrap estimate of G admits an analogous expansion

$$\hat{G}(x) = \mathrm{P}(T^* \leq x|\mathcal{X}) = \Phi(x) + n^{-1/2} \hat{q}(x)\phi(x) + O_p(n^{-1}),$$

  where $\hat{q}$ is obtained from q on replacing unknowns by bootstrap estimates.
- Because $\hat{q} - q = O_p(n^{-1/2})$,

$$\hat{G}(x) - G(x) = \mathrm{P}(T^* \leq x|\mathcal{X}) - \mathrm{P}(T \leq x) = O_p(n^{-1/2})$$

  Thus, the bootstrap approximation is in error by $\boxed{n^{-1}}$ while the normal approximation Φ is in error by $\boxed{n^{-1/2}}$.

A Brief History Recount

TECHNISCHE UNIVERSITÄT DARMSTADT

- It was by marrying the power of *Monte Carlo approximation* with an exceptionally broad view of the sort of problem that bootstrap methods might solve, that [Efron (1979)] famously vaulted earlier resampling ideas out of the arena of sampling methods and into the realm of a universal statistical methodology.
- Arguably, the prehistory of the bootstrap encompasses pre-1979 developments of Monte Carlo methods for sampling.
- Important aspects of bootstrap roots lie in methods for spatial sampling in India in the 1920's [Hubback (1946), Mahalanobis (1946)], even before Quenouille's and Tukey's work on the jackknife [Quenouille (1949, 1956), Tukey (1958)].

The independent data bootstrap

The Bootstrap for independent Data

TECHNISCHE UNIVERSITÄT DARMSTADT

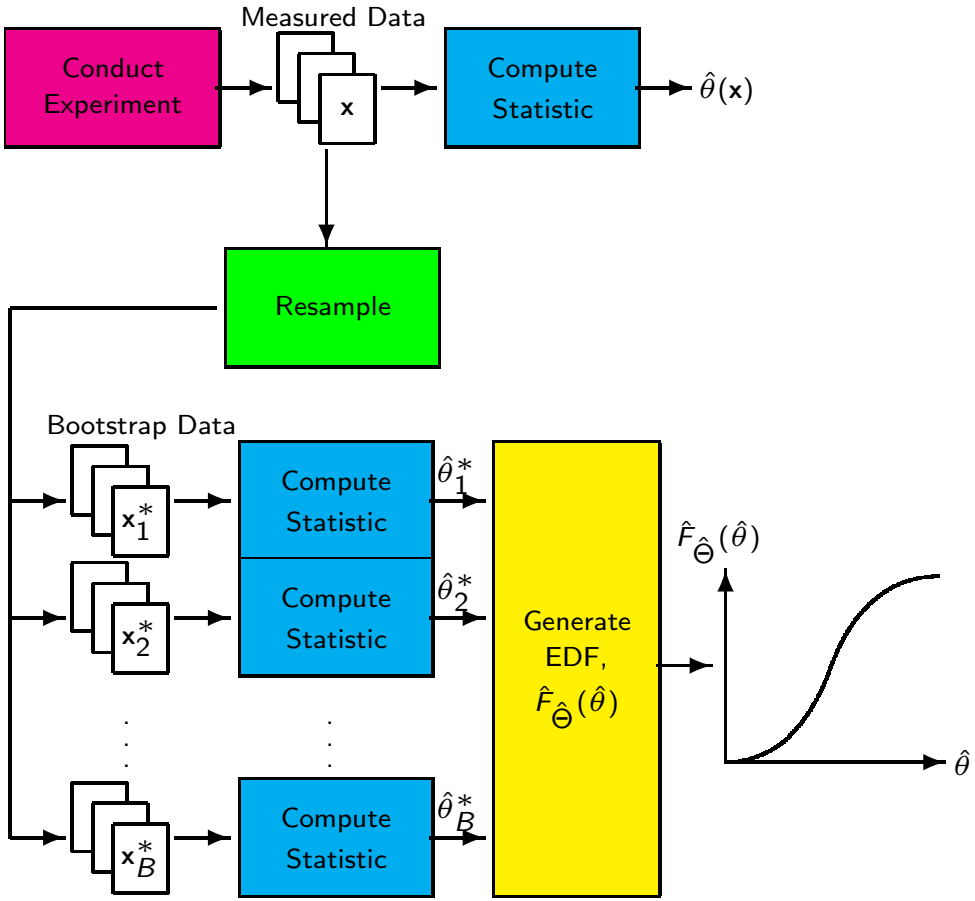
Step 1. Conduct the experiment to obtain the random sample $\mathcal{X} = \{X_1, X_2, \dots, X_n\}$ and find the estimator $\hat{\theta}$ from $\mathcal{X}$.

Step 2. Construct the empirical distribution $\hat{F}_\theta$, which puts equal mass $1/n$ at each observation $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$.

Step 3. From the selected $\hat{F}_\theta$, draw a sample $\mathcal{X}^* = \{X_1^*, X_2^*, \dots, X_n^*\}$, called the bootstrap (re)sample.

Step 4. Approximate the distribution of $\hat{\theta}$ by the distribution of $\hat{\theta}^*$ derived from $\mathcal{X}^*$.

Implementation of the Bootstrap Procedure



Algorithm to compute the approximative distribution function of the estimator $\hat{\theta}$.

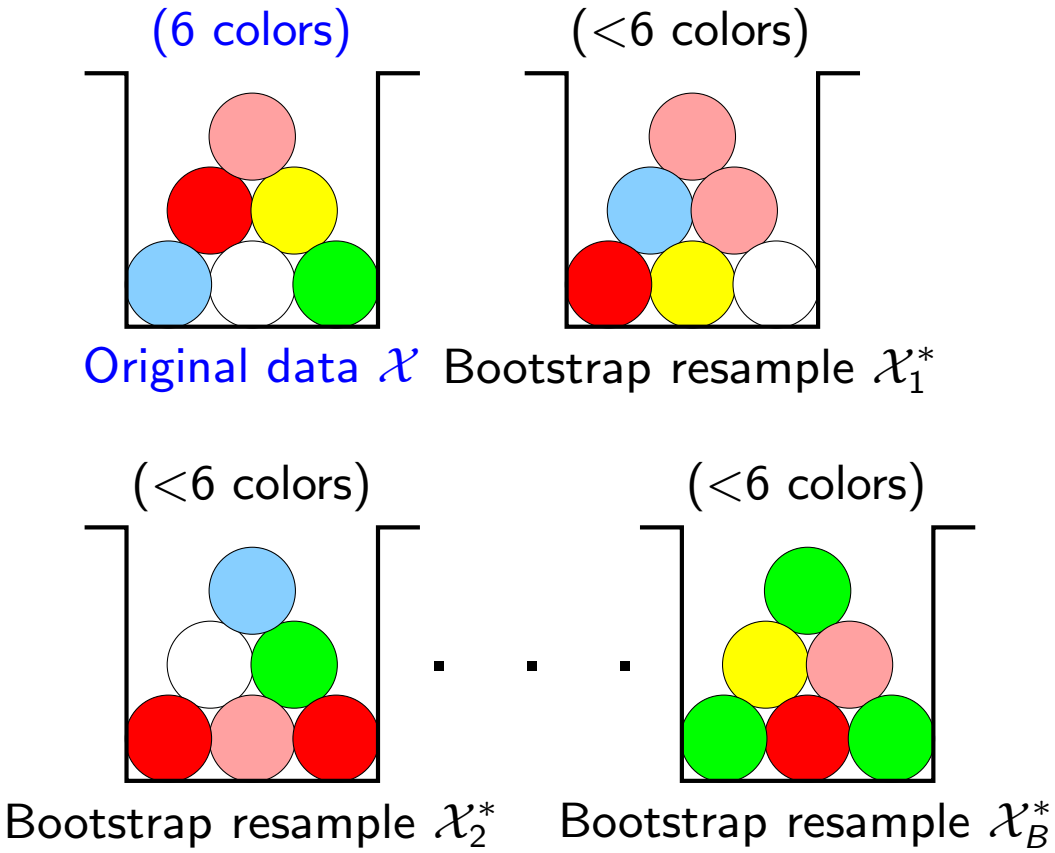
Resampling i.i.d. Data

TECHNISCHE UNIVERSITÄT DARMSTADT

- A resample $\mathcal{X}^* = \{X_1^*, X_2^*, \dots, X_n^*\}$ is an unordered collection of n sample points drawn randomly from $\mathcal{X}$ with replacement, so that each X_i^* has probability n^{-1} of being equal to any one of the X_j's. In other terms, $\mathrm{P}\left(X_i^* = X_j \,|\, \mathcal{X}\right) = n^{-1}, 1 \leq i, j \leq n$.
- That is, the X_i^*'s are independent and identically distributed, conditional on the random sample $\mathcal{X}$, with this distribution [Hall (1992)].

This means that $\mathcal{X}^*$ is likely to contain repeats. As an example, consider $n = 4$ and the collection $\mathcal{X}^* = \{0.5, -3.7, -3.7, 2.8\}$ which should not be mistaken for the set $\{0.5, -3.7, 2.8\}$ because the bootstrap sample has one repeat. Also, $\mathcal{X}^*$ is the same as $\{0.5, -3.7, 2.8, -3.7\}$, $\{-3.7, 0.5, 2.8, -3.7\}$, etc. because the order of elements in the resample plays no role [Hall (1992), Efron & Tibshirani (1993)].

Resampling i.i.d. Data (Cont'd)



Example: Variance Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

Consider the problem of finding the variance $\sigma^2_{\hat{\theta}}$ of an estimator $\hat{\theta}$ of θ, based on the random sample $\mathcal{X} = \{X_1, \dots, X_n\}$ from the unknown distribution F_θ.

- If tractable, one may derive an analytic expression for $\sigma^2_{\hat{\theta}}$.
- For example, for $X_1, \dots, X_n$ identically and independently Gaussian distributed as $\mathcal{N}(\mu, \sigma^2)$ and $\theta = \mu$,

$$\hat{\theta} = \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \sigma^2_{\hat{\mu}} = \frac{\sigma^2}{n}.$$

- Alternatively, one may use asymptotic arguments [Serfling (1980)] to compute an estimate $\hat{\sigma}^2_{\hat{\theta}}$ for $\sigma^2_{\hat{\theta}}$, in which situation the validity conditions for the above may not be fulfilled.

Example: Variance Estimation of μ

TECHNISCHE UNIVERSITÄT DARMSTADT

Observations $\mathcal{X} = \{X_1, \dots, X_n\}$

Resamples $\mathcal{X}_1^*$ $\mathcal{X}_2^*$ ..... $\mathcal{X}_B^*$

Bootstrap statistics $\hat{\mu}_1^*$ $\hat{\mu}_2^*$ ..... $\hat{\mu}_B^*$

Variance estimation $\hat{\sigma}_{\hat{\mu}}^{*2} = \frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\mu}_b^* - \frac{1}{B} \sum_{b=1}^{B} \hat{\mu}_b^* \right)^2$

The independent data bootstrap for dependent data

The Independent Data Bootstrap for Dependent Data

TECHNISCHE UNIVERSITÄT DARMSTADT

- The assumption of i.i.d. data can break down in practice either because the data is not independent or because it is not identically distributed, or both.
- We can still invoke the bootstrap principle if we *knew the model* that generated the data [Efron & Tibshirani (1993), Bose (1988), Kreiß & Franke (1992), Paparoditis (1996), Zoubir (1993)].
- For example, a way to relax the i.i.d. assumption is to assume that the data is identically distributed but not independent such as in autoregressive (AR) models.

An Example: Variance Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

Assume T observations x_t, $t = 0, \dots, T-1$, from

$$X_t + a \cdot X_{t-1} = Z_t \,,$$

where Z_t is i.i.d. noise with $\mathrm{E}[Z_t] = 0$, $c_{ZZ}(u) = \sigma_Z^2 \delta(u)$ and a such that $|a| < 1$.

Goal: Estimate the variance of $\hat{a}$.

An estimate for the variance of $\hat{a}$ exists asymptotically and for Z_t Gaussian:

- Detrend the data and fit the AR(1) model to x_t, $t = 0, \dots, T-1$.
- With $\hat{c}_{xx}(u) = 1/T \sum_{t=0}^{T-|u|-1} x_t x_{t+|u|}$ for $0 \leq |u| \leq T-1$, calculate the Maximum Likelihood Estimate (MLE) of a, $\hat{a} = -\hat{c}_{xx}(1)/\hat{c}_{xx}(0)$, which has approximate variance $\hat{\sigma}_{\hat{a}}^2 = (1-a^2)/T$.

Example: The Bootstrap Variant

TECHNISCHE UNIVERSITÄT DARMSTADT

Step 1. *Calculate residuals*. With an estimate $\hat{a}$ of a, define the residuals $\hat{z}_t = x_t + \hat{a} \cdot x_{t-1}$ for $t = 1, 2, \ldots, T-1$.

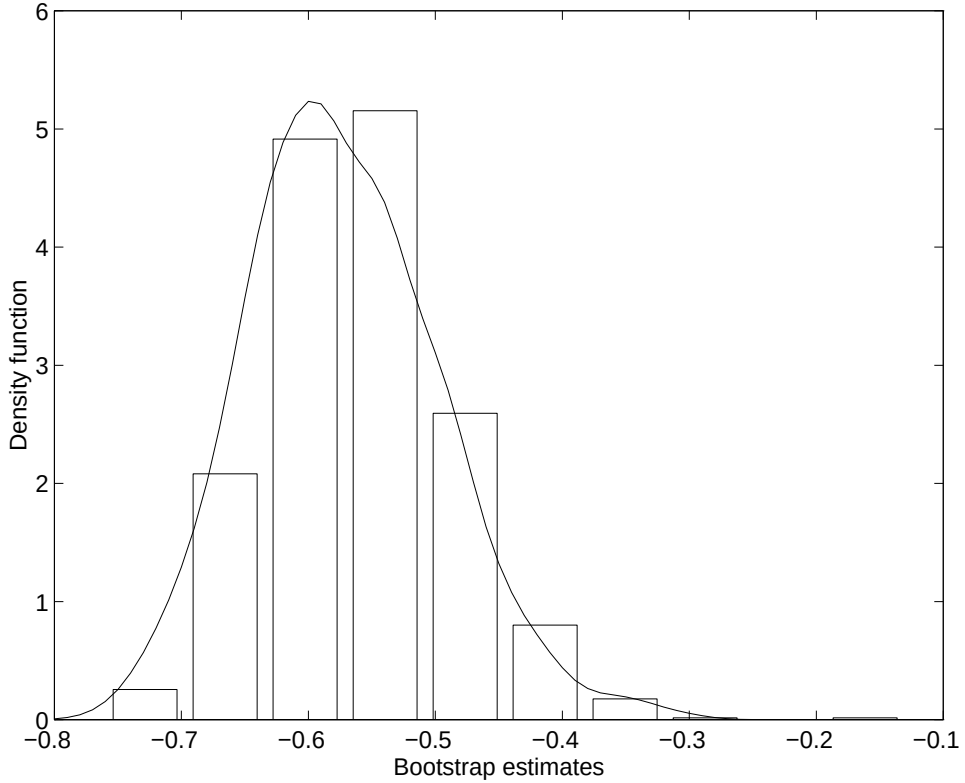
Step 2. *Resampling*. Create a bootstrap sample $x_0^*, x_1^*, \ldots, x_{T-1}^*$ by sampling $\hat{z}_1^*, \hat{z}_2^*, \ldots, \hat{z}_{T-1}^*$, with replacement, from the residuals $\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_{T-1}$, then letting $x_0^* = x_0$, and $x_t^* = -\hat{a}x_{t-1}^* + \hat{z}_t^*$, $t = 1, 2, \ldots, T-1$.

Step 3. *Calculate bootstrap estimates*. Centre the time series $x_0^*, x_1^*, \ldots, x_{T-1}^*$, and compute $\hat{a}^*$ in the same way $\hat{a}$ was obtained but based on $x_0^*, x_1^*, \ldots, x_{T-1}^*$.

Step 4. *Repetition*. Repeat Steps 2–3 B times, to obtain $\hat{a}_1^*, \hat{a}_2^*, \ldots, \hat{a}_B^*$.

Step 5. *Variance estimation*. From $\hat{a}_1^*, \hat{a}_2^*, \ldots, \hat{a}_B^*$, approximate the variance of $\hat{a}$ by $\hat{\sigma}_{\hat{a}}^{*2}$.

Example (Cont'd)



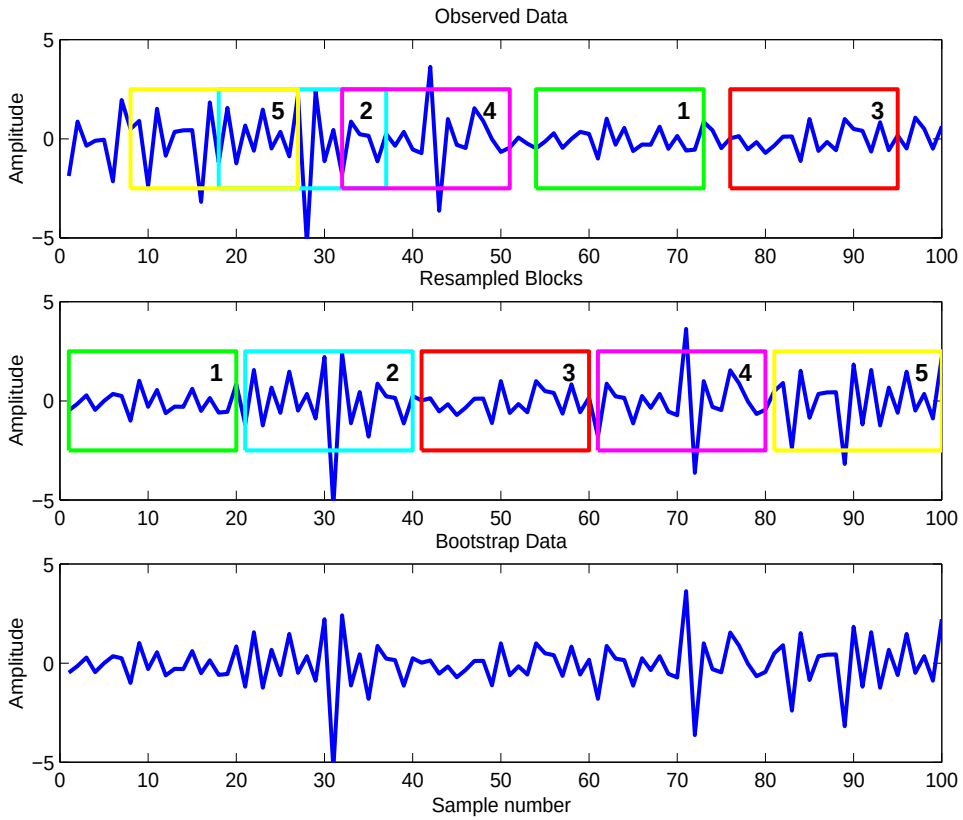
Histogram of $\hat{a}_1^*, \hat{a}_2^*, \ldots, \hat{a}_{1000}^*$ for $a = -0.6$, $T = 128$ and Z_t Gaussian. The MLE for a is $\hat{a} = -0.6351$ and $\hat{\sigma}_{\hat{a}} = 0.0707$. The bootstrap estimate is $\hat{\sigma}_{\hat{a}}^* = 0.0712$ as compared to $\hat{\sigma}_{\hat{a}} = 0.0694$ based on 1000 Monte Carlo simulations.

The Dependent Data Bootstrap (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- If no plausible model such as AR is available for the probability mechanism generating stationary observations, we could make the assumption of weak dependence.
- Strong mixing processes, i.e., *loosely speaking, processes for which observations far apart (in time) are almost independent* [Rosenblatt (1985)], for example, satisfy the weak dependence condition.
- The *moving blocks* bootstrap [Künsch (1989), Liu & Singh (1992), Politis & Romano (1992,1994)] has been proposed for bootstrapping weakly dependent data.

The Moving Blocks Bootstrap



Randomly select blocks of the original data (top) and concatenate them together (centre) to form a resample (bottom). Here, the block size is $l = 20$ and $n = 100$.

Other Block Bootstrap Methods

TECHNISCHE UNIVERSITÄT DARMSTADT

- The *circular block bootstrap* [Politis & Romano (1992), Shao & Yu (1993)] allows blocks which start at the end of the data and wrap around to the start.
- The *blocks of blocks bootstrap* [Politis *et al.* (1992)] uses two levels of blocking to estimate confidence bands for spectra and cross-spectra.
- The *stationary bootstrap* [Politis & Romano (1994)] allows blocks to be of random lengths instead of a fixed length.
- The *tapered block bootstrap* [Paparoditis & Politis (2001)] assigns different weights to the data points according to their position within their block. This results in a bias correction of Künsch's method while preserving asymptotic validity.

Frequency domain bootstrap methods for dependent data

Bootstrapping Spectral Densities

TECHNISCHE UNIVERSITÄT DARMSTADT

- Confidence interval estimation (or hypothesis testing) for spectral densities is encountered in numerous real-life applications, e.g., car engine monitoring.
- Two approaches to bootstrapping spectral densities:
  - Use a time domain approach (block bootstrap) to bootstrap time series and estimate bootstrap spectral density estimates from the time series replicae, or
  - use a technique that applies directly in the frequency domain.
- Bootstrap time series not always mimic the (complicated) dependence structure of the underlying process.
- Frequency domain bootstrap methods do not require to first generate time series replicae to get periodogram resamples.

Frequency Domain Bootstrap (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- Assume X_t, $t \in \mathbb{Z}$ to be a real-valued strictly stationary process, with representation $X_t = \sum_{\tau=-\infty}^{\infty} h_\tau \mathcal{E}_{t-\tau}$, $t \in \mathbb{Z}$.
- Given $X_1, \ldots, X_T$, define the periodogram by

$$I_{XX}(\omega) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} X_t e^{-j\omega t} \right|^2, \qquad -\pi \leq \omega \leq \pi$$

- and construct the estimator of $C_{XX}(\omega)$ by

$$\hat{C}_{XX}(\omega) = \frac{1}{T \cdot h} \sum_{k=-n}^{n} K\left(\frac{\omega - \omega_k}{h}\right) I_{XX}(\omega_k),$$

where the kernel $K(\cdot)$ is a known symmetric, non-negative, real-valued function, h is its bandwidth and $n = [T/2]$.

Frequency Domain Bootstrap (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- The periodogram evaluated at frequencies, $I_{XX}(\omega_k)$, $\omega_k = 2\pi k/T$, $-n \leq k \leq n$, is known to have

$$\mathrm{E}[I_{XX}(\omega_k)] = C_{XX}(\omega_k) + O(T^{-1})$$

$$\mathrm{Cov}\,[I_{XX}(\omega_j), I_{XX}(\omega_k)] = \begin{cases} C_{XX}(\omega_j)^2 + O(T^{-1}), 0 < \omega_j = \omega_k < \pi \\ \frac{1}{T}\eta_4 C_{XX}(\omega_j) C_{XX}(\omega_k) + o(T^{-1}), \omega_j \neq \omega_k, \end{cases}$$

- where $\eta_4 = \kappa_4/\sigma_{\mathcal{E}}^4$ and $\kappa_4 = \mathrm{E}[\mathcal{E}_t^4] - 3\sigma_{\mathcal{E}}^4$ is the fourth order cumulant of $\mathcal{E}_t$ [Anderson (1971)].

The Residual Based Bootstrap

TECHNISCHE UNIVERSITÄT DARMSTADT

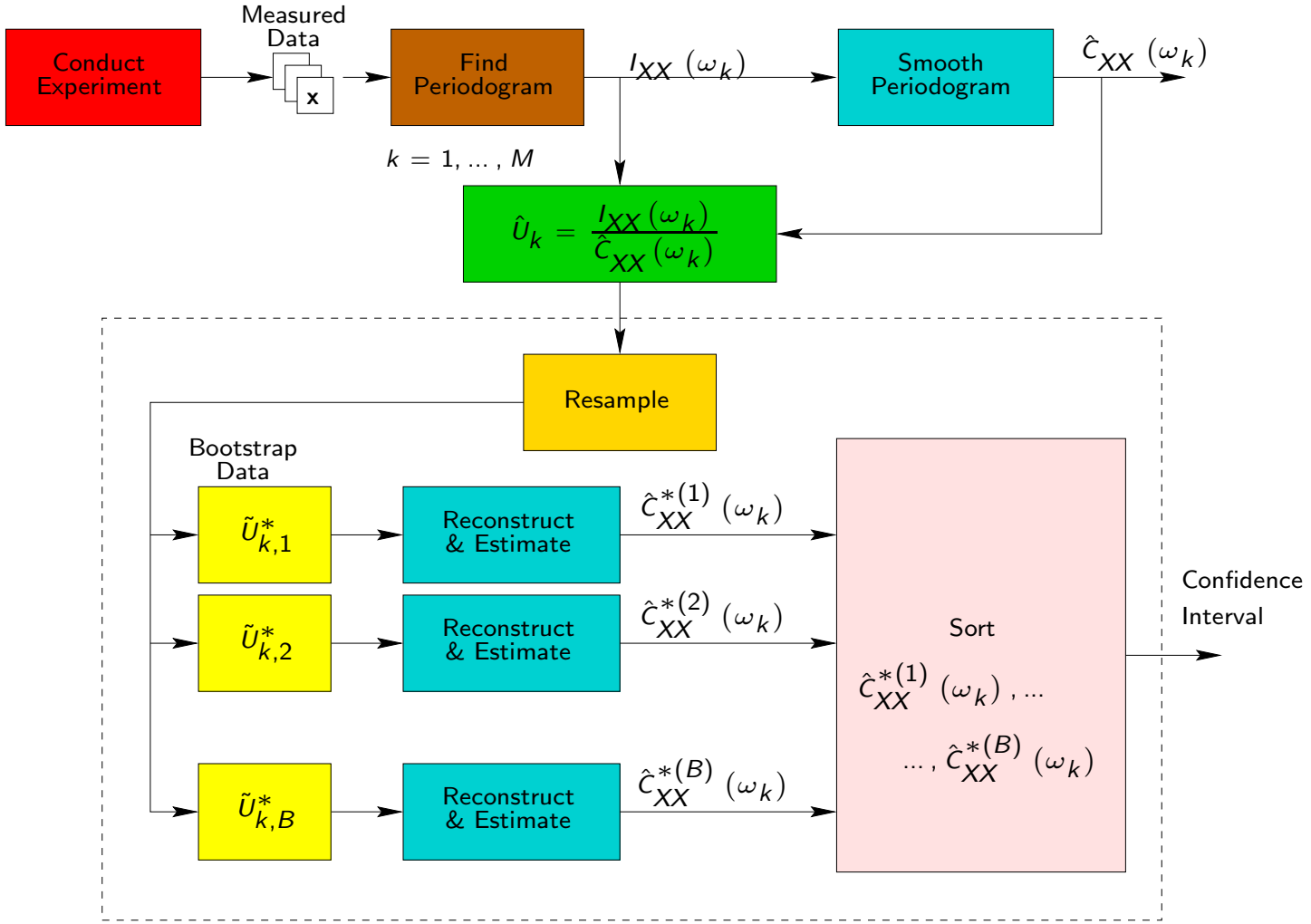
- The residual based method [Franke & Härdle (1992)] is based on the multiplicative regression

$$I_{XX}(\omega_k) = C_{XX}(\omega_k)\frac{I_{\mathcal{E}\mathcal{E}}(\omega_k)}{2\pi\sigma_{\mathcal{E}}^2} + R_T(\omega_k),$$

$$I_{\mathcal{E}\mathcal{E}}(\omega_k) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T}\mathcal{E}_t e^{-j\omega_k t}\right|^2, \quad \omega_k = \frac{2\pi k}{T}, k = 0, \dots, n,$$

- with $R_T(\omega_k)$ satisfying $\max_{\omega_k\in[0,\pi]}\mathrm{E}[R_T(\omega_k)^2] = O(T^{-1})$ [Brockwell & Davis (1991)].
- Ignoring $R_T(\omega_k)$ and taking into account the asymptotic independence of periodogram ordinates, suggests resampling periodograms via mimicking the behavior of $I_{\mathcal{E}\mathcal{E}}(\omega_k)/\sigma_{\mathcal{E}}^2$ and replacing C_{XX} by $\hat{C}_{XX}$.

The Residual Based Bootstrap (Cont'd)



The Residual Based Bootstrap (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- The residual based method approximates the distribution of $L_T(\omega) = \sqrt{Th}(\hat{C}_{XX}(\omega) - C_{XX}(\omega))/C_{XX}(\omega)$ through the distribution of $L_T^*(\omega) = \sqrt{Th}(\hat{C}_{XX}^*(\omega) - \hat{C}_{XX}(\omega))/\hat{C}_{XX}(\omega)$ where

$$\hat{C}_{XX}^*(\omega) = \frac{1}{T \cdot h} \sum_{k=-n}^{n} K\left(\frac{\omega - \omega_k}{h}\right) I_{XX}^*(\omega_k)$$

- Under some regularity conditions (see above and some more on the estimation of $C_{XX}(\omega)$) [Franke & Härdle (1992), Paparoditis (2002)], we have

$$d_2\left(\mathcal{L}\left\{L_T(\omega)\right\}, \mathcal{L}\left\{L_T^*(\omega) \,|X_1, X_2, \ldots, X_T\right\}\right) \to 0 \quad \text{in probability}$$

The Local Bootstrap

TECHNISCHE UNIVERSITÄT DARMSTADT

- An alternative method which does not require prior estimation of the spectral density [Paparoditis & Politis (1999)].
- Recall that for large T, $I_{XX}(\omega)$ is approximately exponentially distributed with parameter $C_{XX}(\omega)$.

Idea: For $C_{XX}(\omega)$ a smooth function of ω, it is expected that the sampling behaviour of the periodogram at any particular frequency ω_s will be similar for periodogram ordinates corresponding to frequencies in a small neighboorhood of ω_s. Thus, periodogram replicae at ω_s can be obtained by locally resampling the periodogram, i.e., by choosing with replacement between periodogram ordinates corresponding to frequencies 'close' to ω_s.

Validity: Under some regularity conditions [Paparoditis (2002)],

$$d_2\left(\mathcal{L}\left\{L_T(\omega)\right\}, \mathcal{L}\left\{L_T^{+}(\omega)\,|X_1, X_2, \ldots, X_T\right\}\right) \to 0 \quad \text{in probability}$$

Limitations

TECHNISCHE UNIVERSITÄT DARMSTADT

- The previous two methods assume independence of the periodogram ordinates $I^*_{XX}(\omega)$ and $I^+_{XX}(\omega)$, justified by the asymptotic independence of $I_{XX}(\omega)$.
- One should take into account the weak dependence structure of the periodogram ordinates.
- One approach is to modify *appropriately* the above procedures.
- Using a non-parametric estimator of η_4, Janas & Dahlhaus (1994) modified $I^*_{XX}(\omega)$ in a way that emulates also the covariance of $I_{XX}(\omega)$.
- It is problematic to nonparametrically estimate the complicated functional η_4 [Dahlhaus & Janas (1996)].

Autoregressive-Aided Periodogram Bootstrap

TECHNISCHE UNIVERSITÄT DARMSTADT

- An alternative is the so-called *autoregressive-aided periodogram bootstrap* [Kreiss & Paparoditis (2003)], a combination of time domain and frequency domain resampling [Zoubir (2010)].

Idea: Use a parametric fit in the time domain to generate periodogram ordinates that imitate the essential features of the data and weak dependence structure of the periodogram, while a non-parametric kernel based correction in the frequency domain is applied to catch features not represented by the parametric fit.

The AR-Aided Bootstrap (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- To see the differences with the two previous methods, using $C_{XX}(\omega) = q(\omega)C_{\mathrm{AR}}(\omega)$, we have under some regularity conditions

$$I_{XX}(\omega_k) = q(\omega_k)C_{\mathrm{AR}}(\omega_k)\frac{I_{\mathcal{E}\mathcal{E}}(\omega_k)}{2\pi\sigma^2_{\mathcal{E}\mathcal{E}}} + R_T(\omega_k)$$

- A similar expression is also valid for $I^*_{XX}(\omega_k)$

$$I^*_{XX}(\omega_k) = \hat{q}(\omega_k)\hat{C}_{\mathrm{AR}}(\omega_k)\frac{I_{\tilde{\mathcal{E}}\tilde{\mathcal{E}}}(\omega_k)}{2\pi\hat{\sigma}^2} + R^*_T(\omega_k)$$

where $I_{\tilde{\mathcal{E}}\tilde{\mathcal{E}}}(\omega_k)$ is the periodogram of the bootstrapped i.i.d. residuals from the AR model and $\max_k \mathrm{E}_*[R^*_T(\omega_k)] = O_p(T^{-1})$.

A Simulation Example

TECHNISCHE UNIVERSITÄT DARMSTADT

- Consider X_t, where U_t are independent uniformly distributed variables on the interval $[-2.5, 2.5]$ and $T = 256$

$$X_t = X_{t-1} - 0.7X_{t-2} - 0.4X_{t-3} + 0.6X_{t-4} - 0.5X_{t-5} + U_t$$

Estimates of coverage probabilities in [%] of a nominal 95% confidence interval (1000 MC).

k	RSD	TBB	AR	AR-AP
26	86.68	77.98	87.66	88.27
27	85.94	77.55	87.54	87.85
28	85.64	77.77	87.62	87.96
52	84.08	78.48	85.76	87.60
53	83.61	77.80	85.88	87.43
54	84.30	77.88	86.44	87.56
95	86.73	76.74	87.72	90.31
96	87.52	77.41	87.78	90.80
97	87.80	77.98	87.85	91.32

A Simulation Example

TECHNISCHE UNIVERSITÄT DARMSTADT

- Consider X_t, where U_t are independent uniformly distributed variables on the interval $[-2.5, 2.5]$ and $T = 256$

$$X_t = X_{t-1} - 0.7X_{t-2} - 0.4X_{t-3} + 0.6X_{t-4} - 0.5X_{t-5} + U_t$$

Estimates of coverage probabilities in [%] of a nominal 95% confidence interval (1000 MC).

k	**RSD**	**TBB**	**AR**	**AR-AP**
26	86.68	77.98	87.66	88.27
27	85.94	77.55	87.54	87.85
28	85.64	77.77	87.62	87.96
52	84.08	78.48	85.76	87.60
53	83.61	77.80	85.88	87.43
54	84.30	77.88	86.44	87.56
95	86.73	76.74	87.72	90.31
96	87.52	77.41	87.78	90.80
97	87.80	77.98	87.85	91.32

Two real-life example of the bootstrap

Confidence Intervals for knock Spectra

TECHNISCHE UNIVERSITÄT DARMSTADT

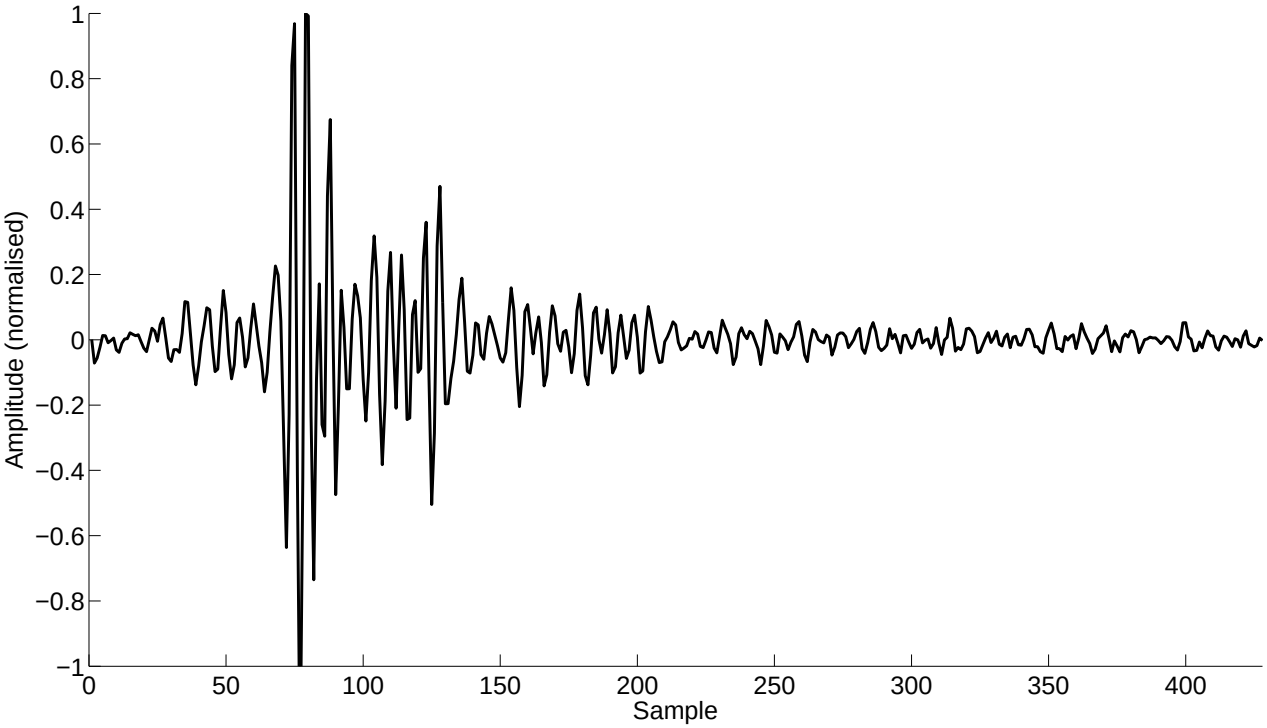
- Special interest continues to a be reduction of fuel consumption in combustion engines due to meeting the demands of new legislations on greenhouse emission.
- A means for lowering fuel consumption in spark ignition engines is to increase the compression ratio. Further increase in compression is limited by the occurrence of knock.
- To attain safe operation at maximum efficiency, engine control systems which detect knock and adapt spark timing of each cylinder separately have found special interest.
- The data was collected at ROBERT BOSCH GMBH on an engine test bed. Measurements from a combustion engine running at 3500 rpm under knocking conditions were recorded and sampled at 100 kHz.

Confidence Intervals for knock Spectra

TECHNISCHE UNIVERSITÄT DARMSTADT

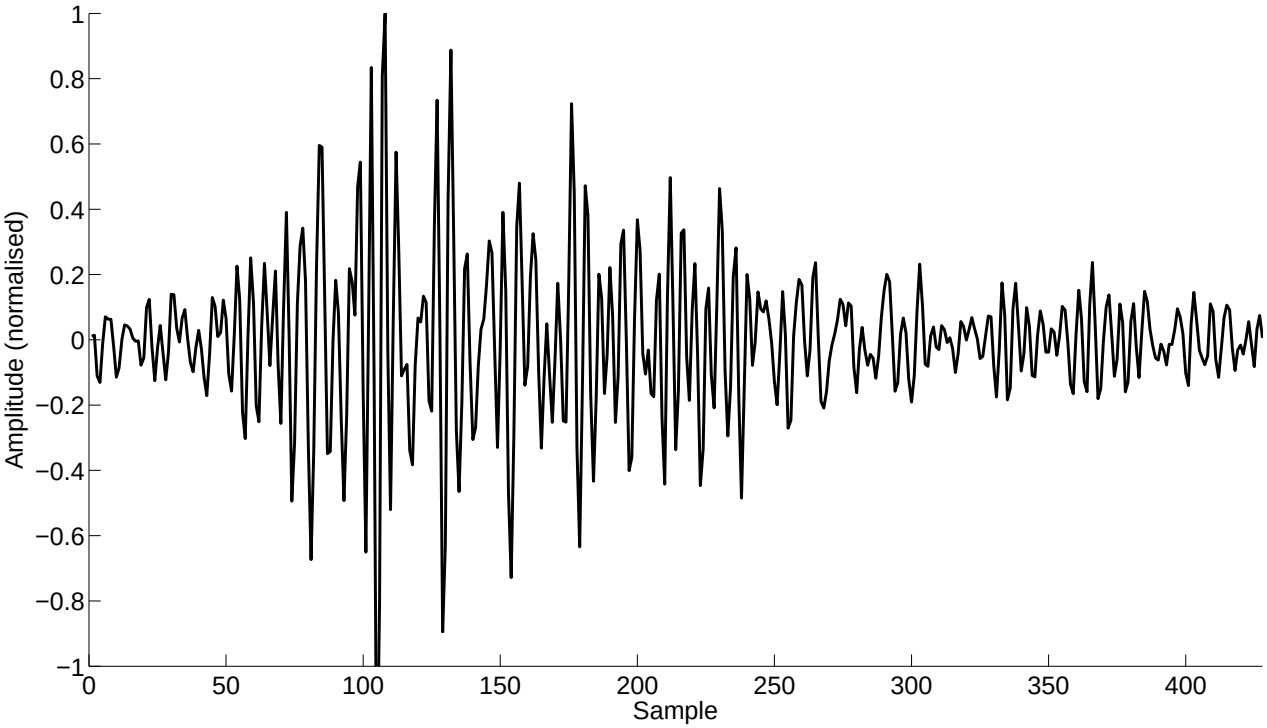
- The estimation of confidence intervals is very helpful for the detection of knock.
- There is a well known and established duality between confidence intervals and hypothesis testing [Lehman (1986)].
- if $\mathcal{I}$ is a confidence interval for an unknown parameter θ, with coverage probability α, then a $(1-\alpha)$-level test of the null hypothesis $\mathsf{H} : \theta = \theta_0$ against $\mathsf{K} : \theta \neq \theta_0$ is to reject H if $\theta_0 \notin \mathcal{I}$.

Confidence Intervals for knock Spectra (Cont'd)



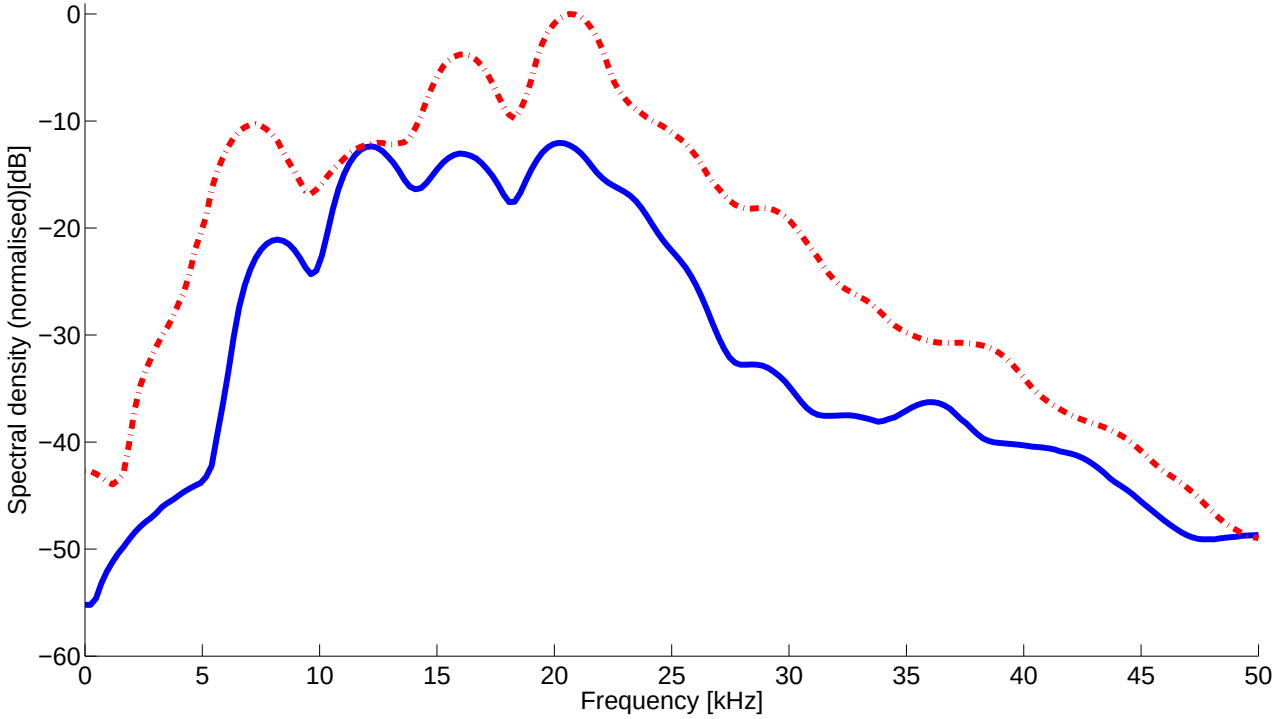
In-cylinder pressure signal (high-pass filtered) of a knocking combustion cycle.

Confidence Intervals for knock Spectra (Cont'd)



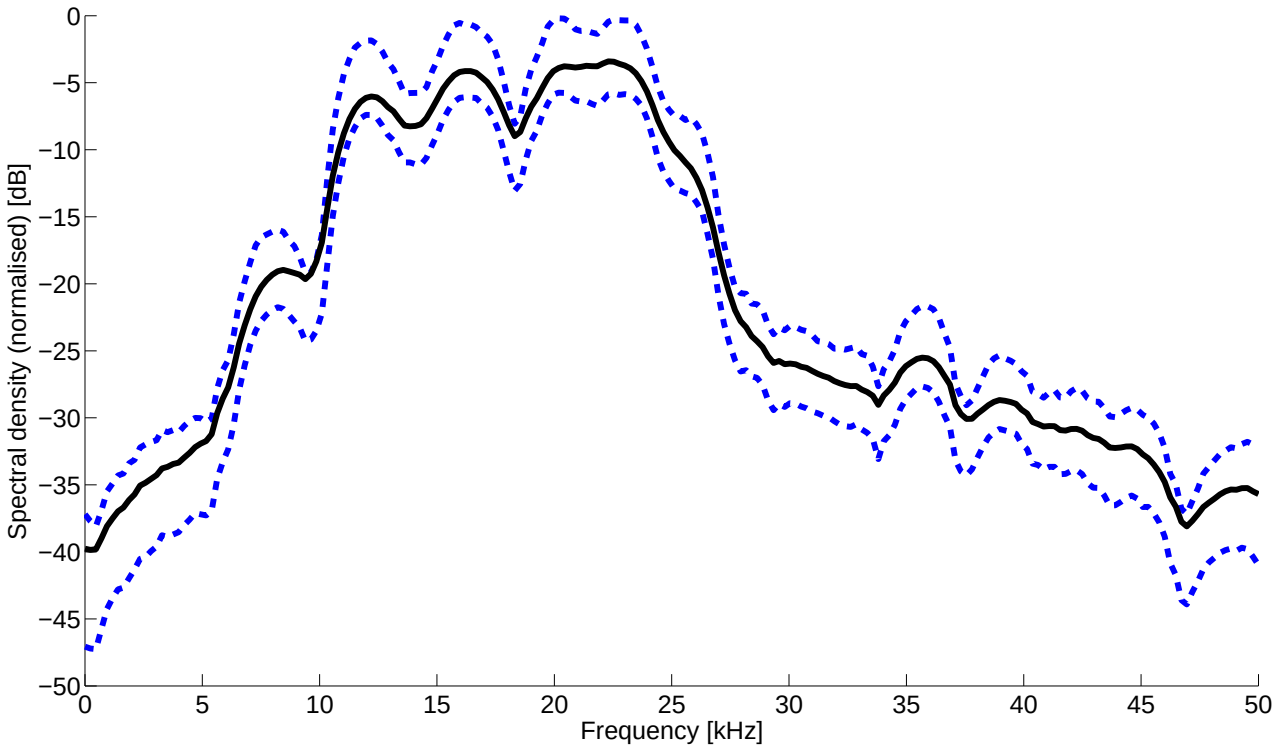
Structure-borne sound signal of a knocking combustion cycle.

Confidence Intervals for knock Spectra (Cont'd)



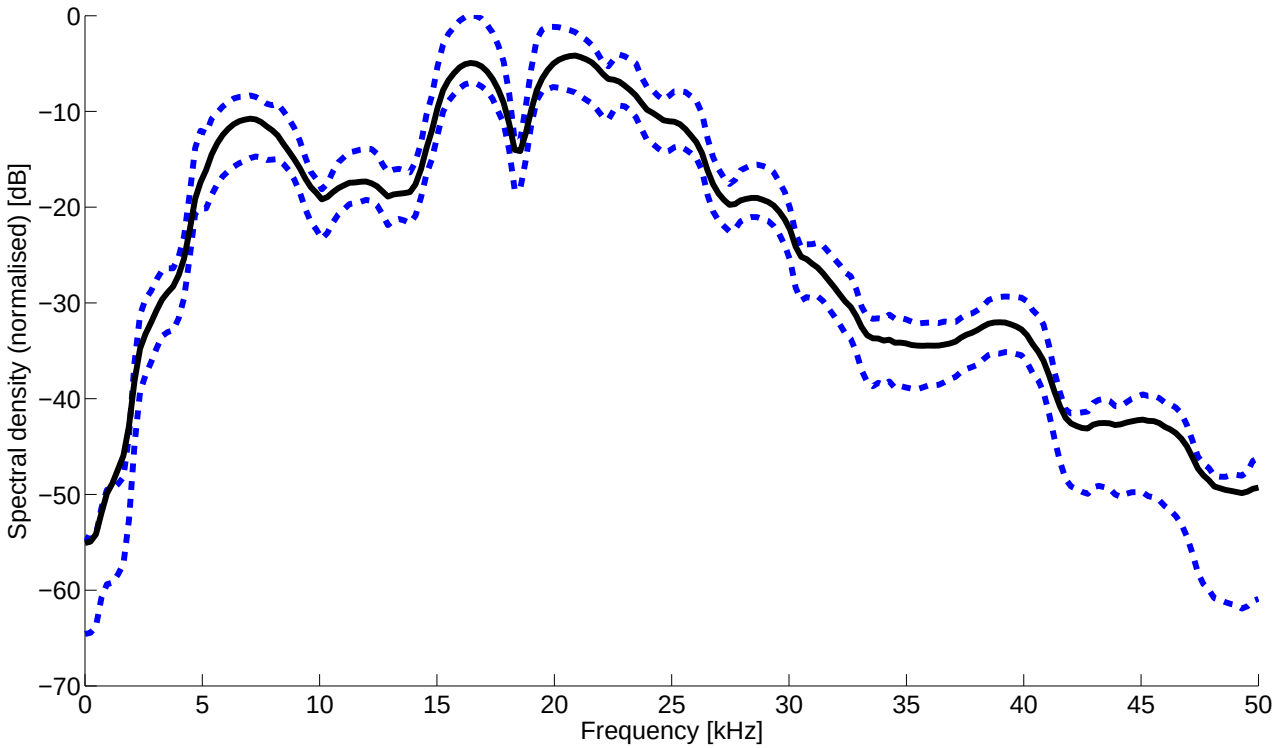
Spectral density estimates for in-cylinder pressure (solid —) and structure-borne sound (dashed-dotted -.-) obtained by averaging 594 periodograms.

Confidence Intervals for knock Spectra (Cont'd)



Estimated 95% confidence interval (dashed line - - -) for the spectral density of in-cylinder pressure, along with the spectral density estimate for the combustion cycle above.

Confidence Intervals for Knock Spectra (Cont'd)



Estimated 95% confidence interval for the spectral density of structure-borne sound (dashed line - –), along with the spectral density estimate for the combustion cycle above

Micro-Doppler Analysis

TECHNISCHE UNIVERSITÄT DARMSTADT

- Doppler often arises in engineering applications due to the relative motion of an object with respect to the measurement system.
- If the motion is harmonic, for example due to vibration or rotation, the resulting signal can be well modeled by an FM process [Huang *et al.* (1990)].
- Estimation of the FM parameters may allow one to determine physical properties such as the angular velocity and displacement of the vibrational/rotational motion which can in turn be used for classification.

Objective: Estimate the FM parameters along with a measure of accuracy, such as confidence intervals.

Micro-Doppler Analysis (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- Assume the following AM-FM signal model:

$$s(t) = a(t)\exp\{j\varphi(t)\}$$

where $a(t;\boldsymbol{\alpha}) = \sum_{k=0}^{q} \alpha_k t^k$ and $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_q)$.

- The phase modulation for a micro-Doppler signal is described by: $\varphi(t) = -D/\omega_m \cos(\omega_m t + \phi)$ and the instantaneous frequency (IF) by:

$$\omega_i(t;\boldsymbol{\beta}) \triangleq \frac{d\varphi(t)}{dt} = D\sin(\omega_m t + \phi),$$

where $\boldsymbol{\beta} = (D, \omega_m, \phi)$ are the FM or micro-Doppler parameters.

- Given $\{x(k)\}_{k=1}^{n}$ of $X(t) = s(t) + V(t)$, the goal is to estimate $\boldsymbol{\beta} = (D, \omega_m, \phi)$ as well as their confidence intervals.

Micro-Doppler Analysis (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- The estimation of the phase parameters is performed using a time-frequency Hough transform (TFHT) [Barbarossa & Lemoine (1996), Cirillo, Zoubir & Amin (2006)].
- Once the phase parameters have been estimated, the phase term is demodulated and the amplitude parameters are estimated via least-squares.
- Given estimates for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the residuals are obtained.
- We whiten the residuals using a fitted AR model. The innovations are re-sampled, filtered and added to the estimated signal term.

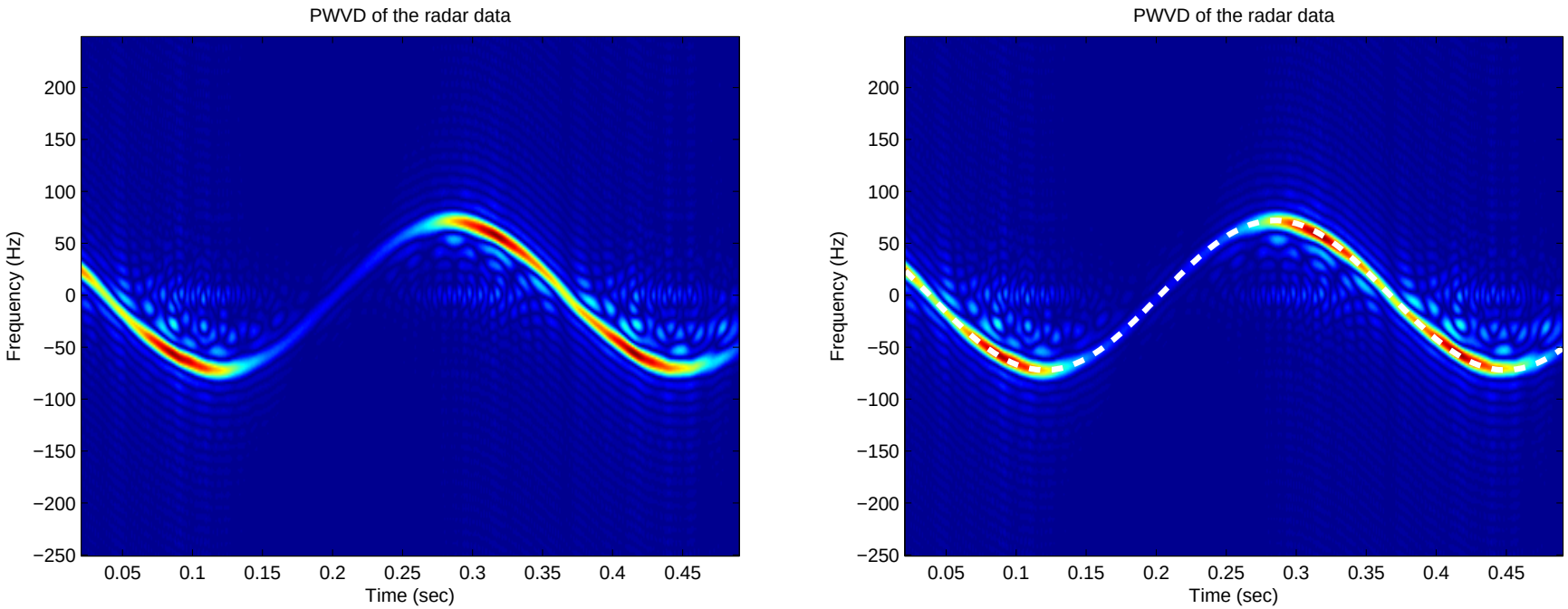
Micro-Doppler Analysis (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- The results shown here are based on an experimental radar system*, operating at carrier frequency $f_c = 919.82$ MHz.
- After demodulation, the in-phase and quadrature baseband channels are sampled at $f_s = 1$ kHz.
- The radar system is directed toward a spherical object, swinging with a pendulum motion, which results in a typical micro-Doppler signature.
- The PWVD of the observations is computed and shown below:

* Courtesy of Prof. Moeness Amin, Director of the Center of Advanced Communications, Villanova University, Philadelphia, USA

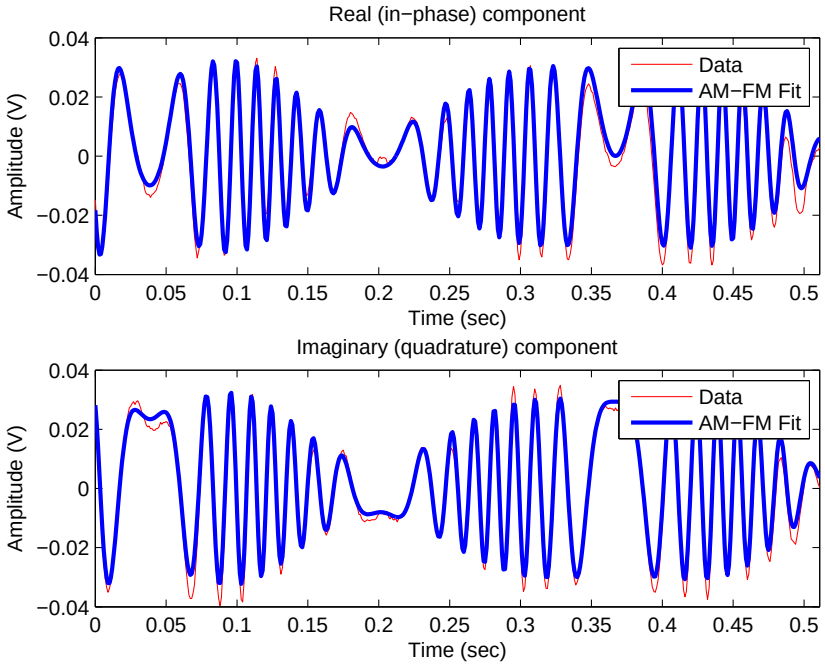
Micro-Doppler Analysis (Cont'd)



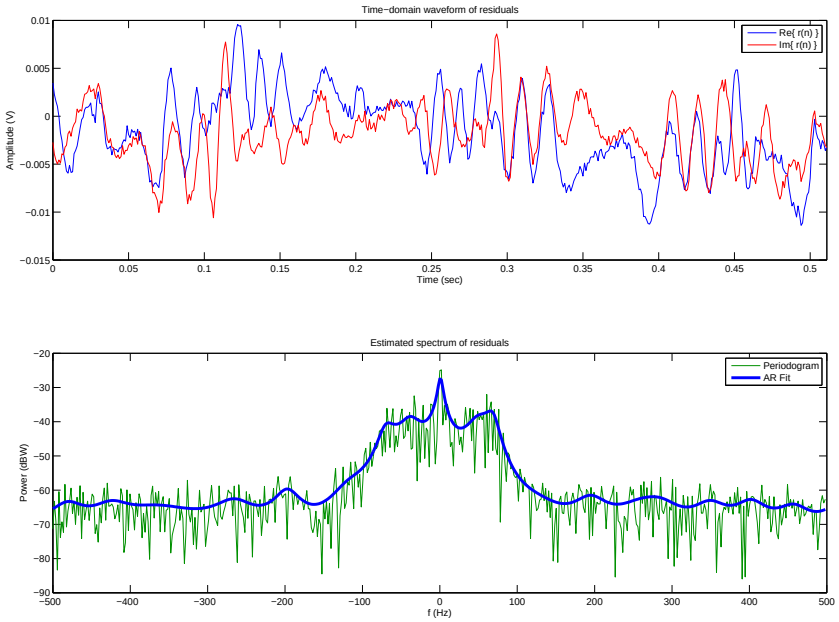
(a) (b)

The PWVD of the radar data (a). The PWVD of the radar data and the micro-Doppler signature estimated using the TFHT (b).

Micro-Doppler Analysis (Cont'd)



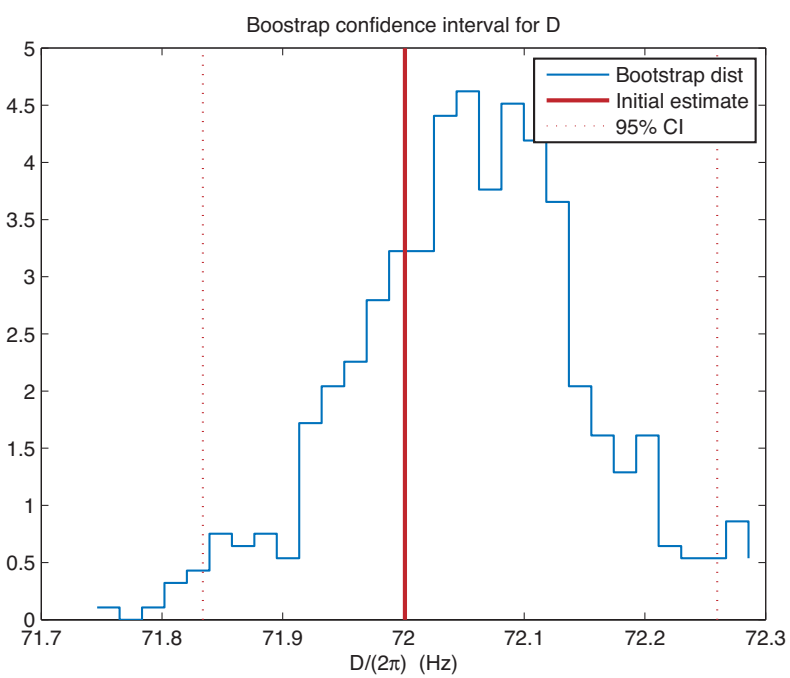
(c)



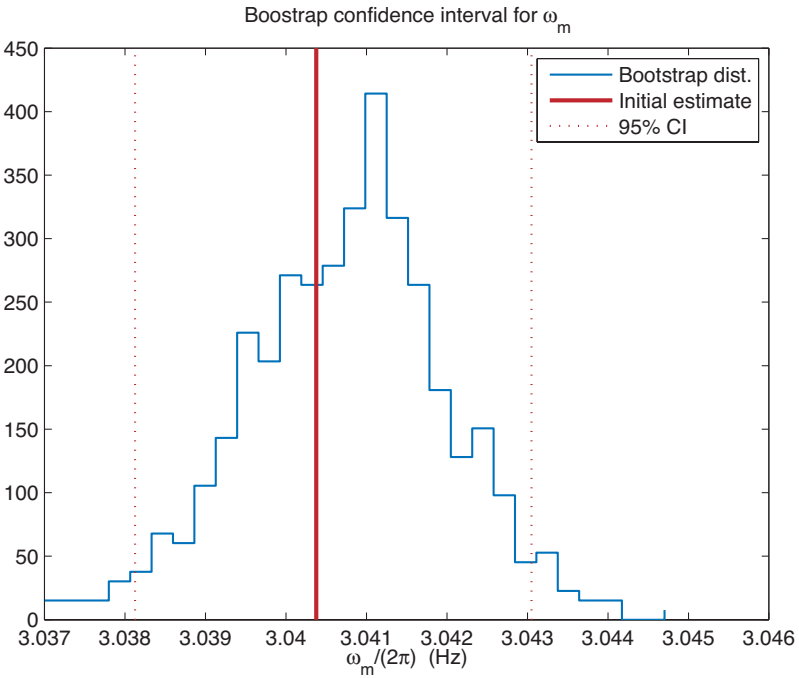
(d)

The real and imaginary components of the radar signal are with their estimated counterparts (c). The real and imaginary parts of the residuals and their spectral estimates (d).

Micro-Doppler Analysis (Cont'd)



(a)



(b)

The bootstrap distributions and 95% confidence intervals for the FM parameters D (a) and ω_m (b) using $B = 500$.

Summary

TECHNISCHE UNIVERSITÄT DARMSTADT

- Many signal processing problems require the computation of quality measures for estimators.
- Most techniques available assume that the size of the available set of sample values is sufficiently large, so that "asymptotic" results can be applied.
- In many signal processing problems this assumption cannot be made because, for example, the process is non-stationary and only small portions of stationary data are considered.
- Bootstrap techniques are an alternative to asymptotic methods and provide good results for many practical problems.

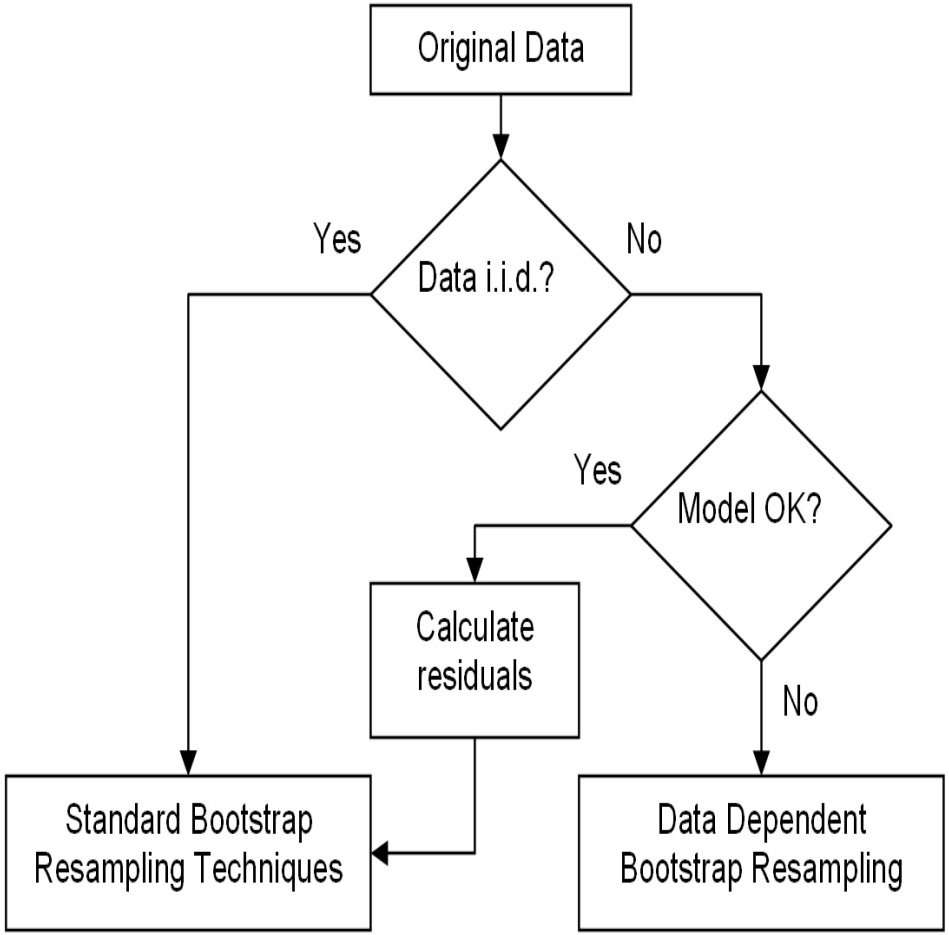
Summary (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

What was not discussed in this presentation:

- Other variants of the bootstrap such as bootstrap bagging and bootstrap bumping [Tibshirani & Knight (1997)].
- Model selection with the bootstrap [Shao & Tu (1995), Zoubir (1999)].
- Hypothesis testing [Hall & Titterington (1989), Hall & Wilson (1991), Hall (1992)] and signal detection [Zoubir (2000)].
- Bayesian approaches to the bootstrap.

Which Method Should I Use?



The Baron von Münchhausen

TECHNISCHE UNIVERSITÄT DARMSTADT

Varney, Allen (2007), "The Extraordinary Adventures of Baron Münchhausen", in Lowder, James, Hobby Games: The Best 100, Green Ronin Publishing, pp. 107109, ISBN 978-1-932442-96-0