Robust Spectral Estimation with Applications

B. Spangl and R. Dutter

University of Natural Resources and Applied Life Sciences, Vienna

and

Vienna University of Technology

Workshop RobHD 2004, Vorau

Contents

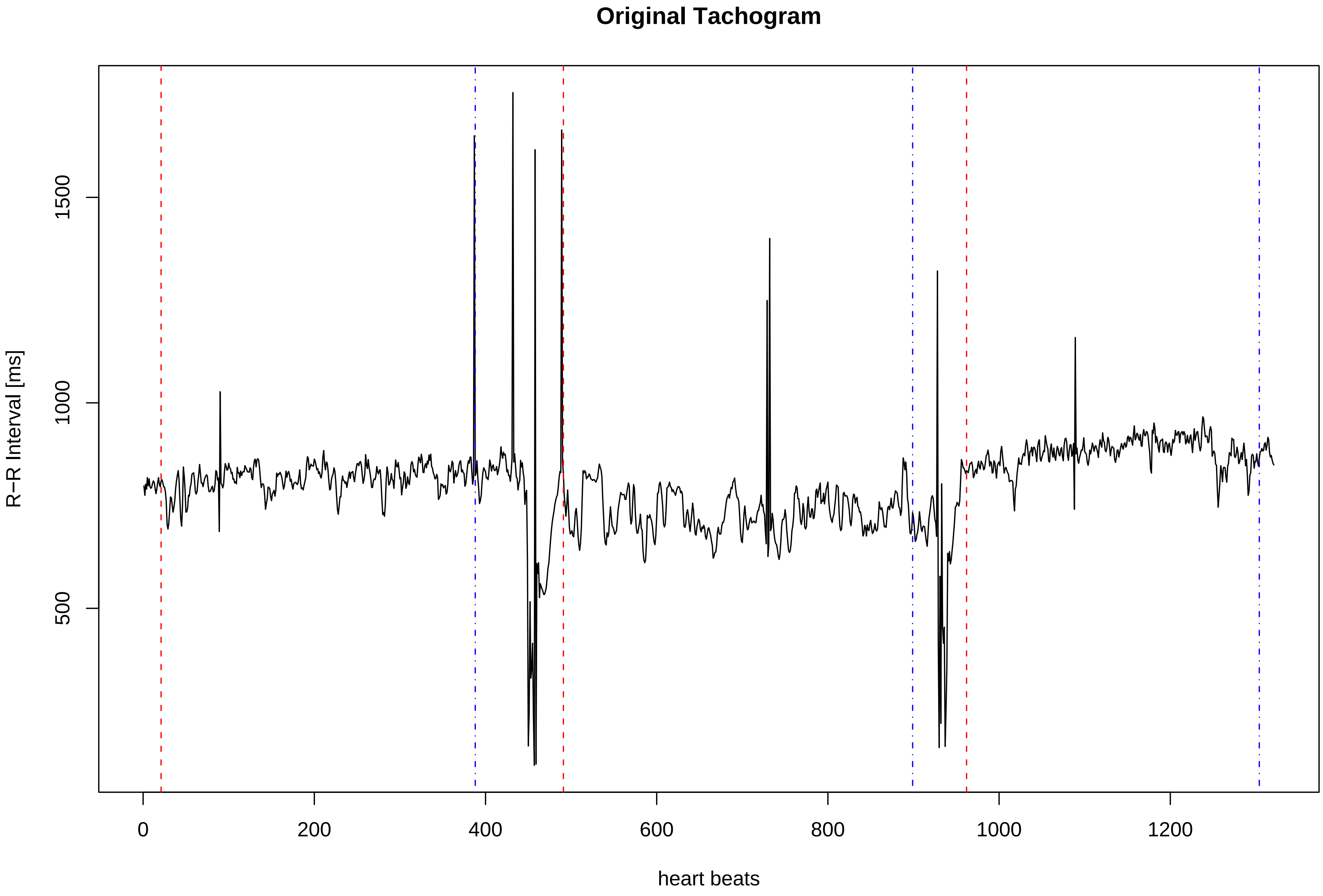
Applications of Spectral Estimation

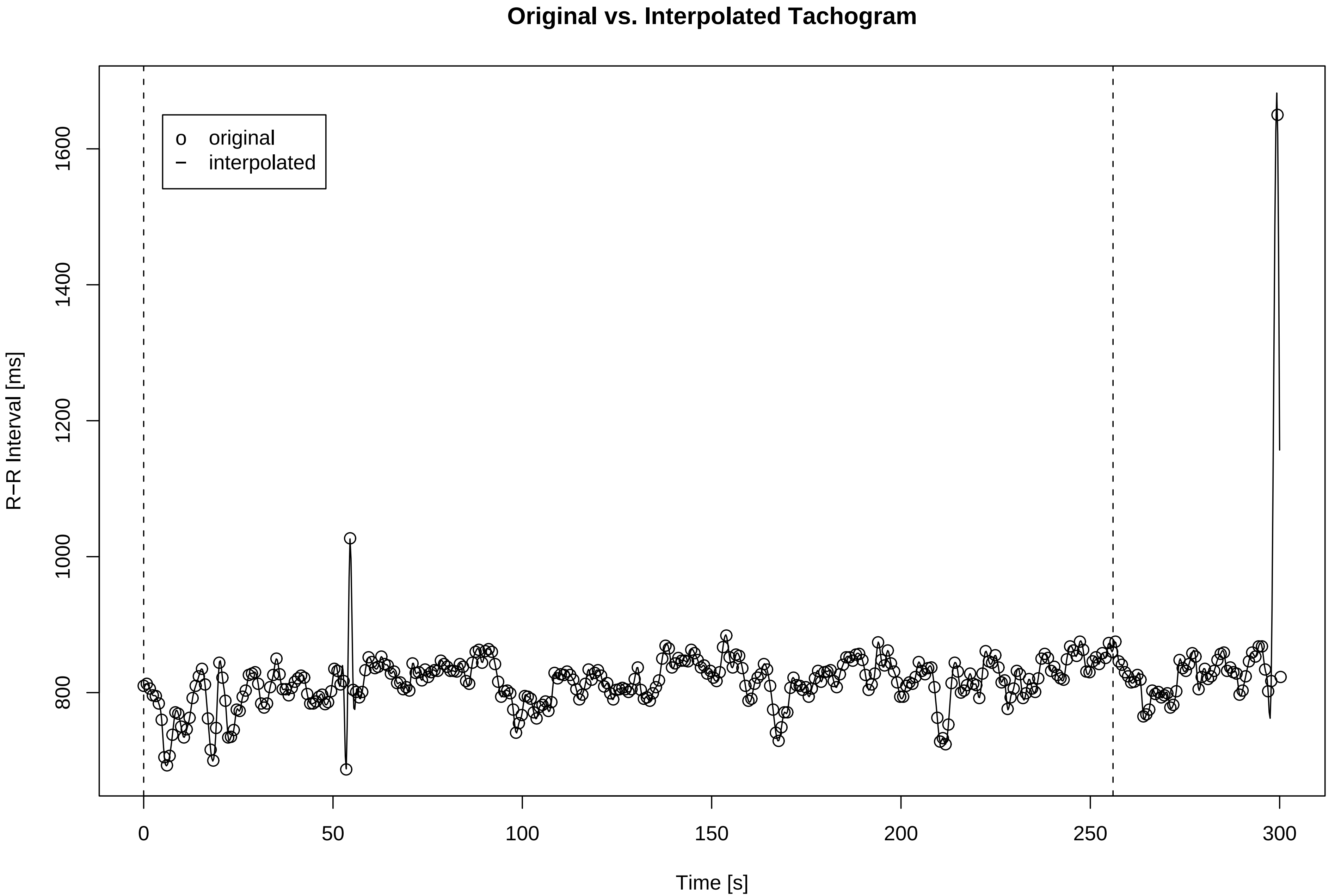
- geophysics
- signal processing
- medicine $\longrightarrow$ analysis of heart rate variability (HRV)

Heart Rate Variability Data

- data: heart rate variability (HRV) recordings
  (tachogram of 1321 successive heart beats)

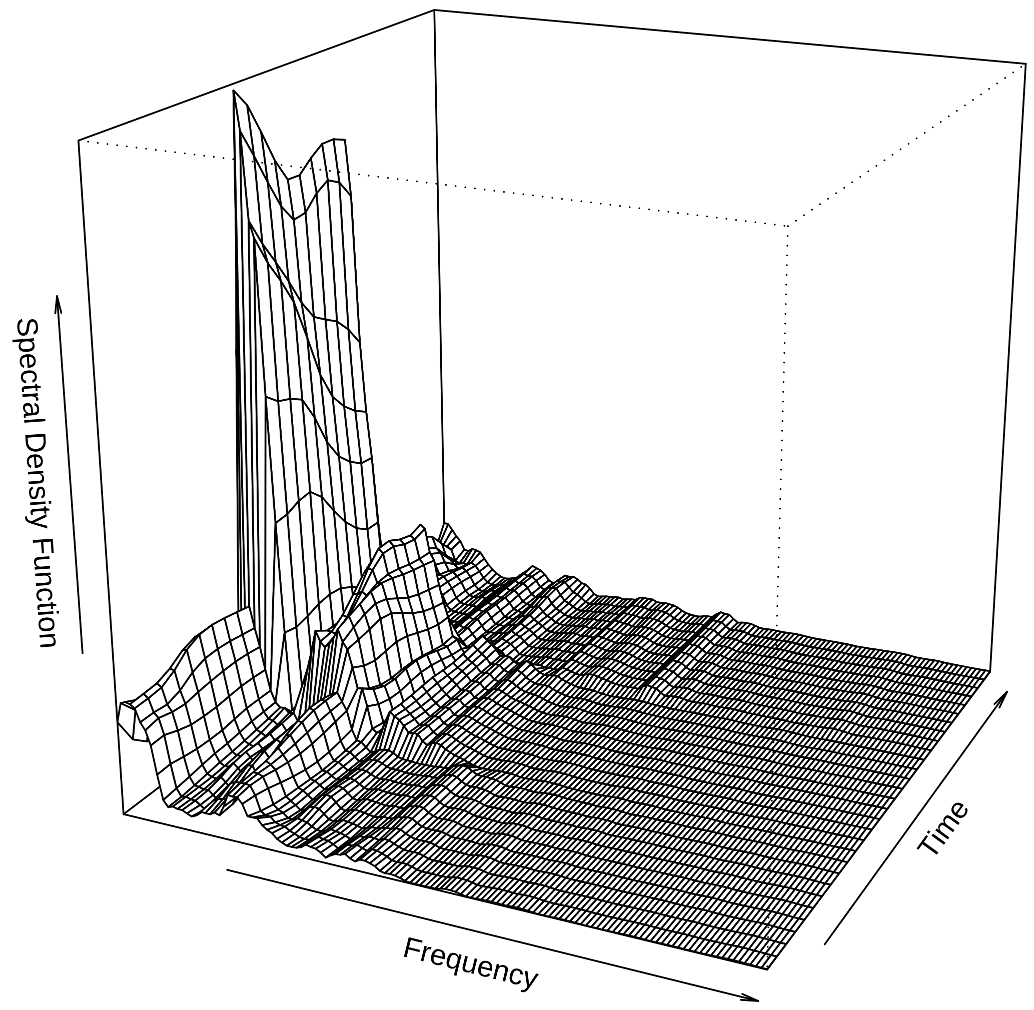
- provided by J. Pumprla and K. Howorka,
  Department of Biomedical Engineering and Physics,
  General Hospital of Vienna


Original Tachogram
R−R Interval [ms]
1500
1000
500
0
200
400
600
800
1000
1200
heart beats



Original vs. Interpolated Tachogram
o original
− interpolated
R−R Interval [ms]
800
1000
1200
1400
1600
0
50
100
150
200
250
300
Time [s]

Short−term Analysis of HRV Data



Maximum: 37.979 *10^3 [ms^2/Hz]

Artificial Example

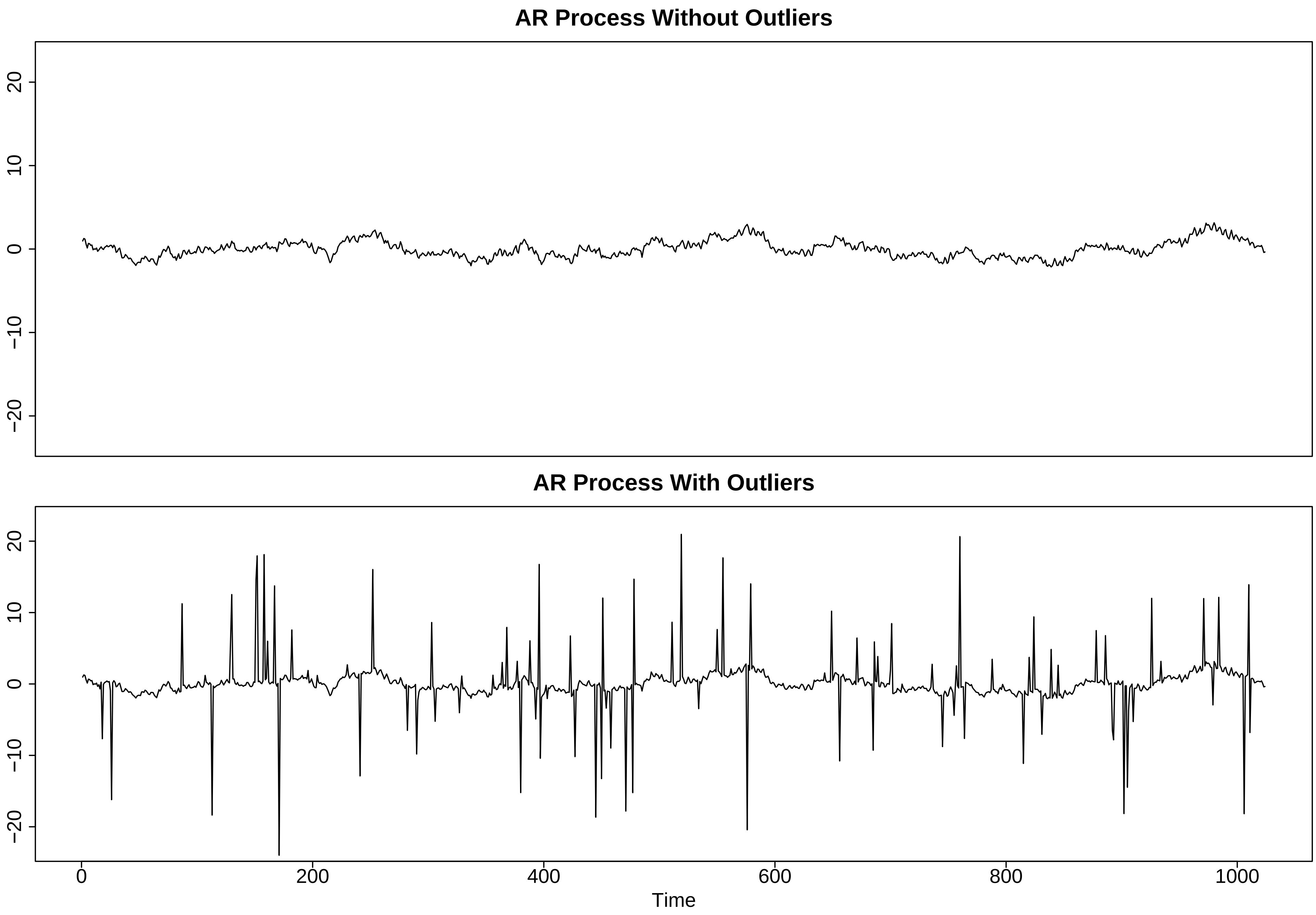
This example is a artificial process with AO composed of the following three autoregressive processes:

$$
\begin{aligned}
u_k &= 0.975 u_{k-1} + \varepsilon_k \ , \\
w_k &= 0.95 w_{k-1} - 0.9 w_{k-2} + \eta_k \ , \\
z_k &= 0.33 z_{k-1} - 0.9 z_{k-2} + \zeta_k \ ,
\end{aligned}
$$

with $\varepsilon_k, \eta_k, \zeta_k \sim N(0,1)$. We standardize u_k, w_k and z_k, and compute

$$
y_k = \sqrt{75} u_k + w_k + z_k \ , \quad k = 1, \ldots, n \ ,
$$

and then standardize y_k again. Additionally we add some noise from $0.9\delta_0 + 0.1 N(0, 10)$.


AR Process Without Outliers
20
10
0
−10
−20
AR Process With Outliers
20
10
0
−10
−20
0
200
400
600
800
1000
Time

Basics

- Second-order Stationarity
- Additive Outliers
- Spectral Representation Theorem
- Non-parametric Estimation
- Parametric Estimation
- Estimates Based on Prewhitening

Additive Outliers (AO)

We say that the process $\{y_t : t \in \mathbb{Z}\}$ has additive outliers (AO) if it is defined by

$$y_t = x_t + v_t \tag{1}$$

where x_t is an ARMA(p, q) process and the v_t are independent identically distributed (iid) with common distribution $F_v = (1 - \varepsilon)\delta_0 + \varepsilon H$ where δ_0 is the degenerate distribution having all its mass at the origin and H is a heavy-tailed symmetric distribution with mean 0 and variance σ^2. Hence, the ARMA(p, q) process x_t is observed with probability $1 - \varepsilon$ whereas the ARMA(p, q) process plus an error v_t is observed with probability ε. We shall also assume that x_t and v_t are independent.

Estimates Based on Prewhitening

Let $\{x_k, k = 1, \ldots, N\}$ denote the observed values of a second-order stationary process with mean zero. Kleiner et al. (1979) prefer the following autoregression prewhitened spectral density estimate which was originally suggested by Blackman & Tukey (1958):

$$\widehat{S}(f) = \frac{\widehat{S}_r^{(lw)}(f)}{|\widehat{H}_p(f)|^2} \,, \tag{2}$$

where $\widehat{S}_r^{(lw)}(f)$ is a lag window spectral estimate of the prediction residuals $r_k = x_k - \sum_{j=1}^{p} \widehat{\phi}_{j,p} x_{k-j}$, $k = p+1, \ldots, N$ and

$$\widehat{H}_p(f) = 1 - \sum_{j=1}^{p} \widehat{\phi}_{j,p} e^{-i2\pi f j \Delta t} \,. \tag{3}$$

One proposal is to robustify both the numerator $\widehat{S}_r^{(lw)}(f)$ and the denominator $\widehat{H}_p(f)$ in (2) which leads to the robust Filter-cleaner approach suggested by Martin & Thomson (1982).

Robust Filter-cleaner (Overview)

Let $y_t = x_t + v_t$ denote the observed process.

- Order selection of the underlying AR(p) process
- Calculating the robust estimates $\widehat{\phi}_p = (\widehat{\phi}_{1,p}, \ldots, \widehat{\phi}_{p,p})^\top$ and $\widehat{\sigma}^2_{\varepsilon,p}$
- Robust Filter-Cleaning of the process y_t
- Calculating the prewhitened spectral estimate $\widehat{S}(f)$

Robust Filter-cleaner (Part I)

Let $\{y_k, k = 1, \ldots, N\}$ denote the observed values of a second-order stationary process with mean zero. The filter-cleaner algorithm rely on the AR(p) approximation of the underlying process x_t, represented in the following state-space form with $t = p+1, \ldots, N$:

$$\boldsymbol{X}_t = \boldsymbol{\Phi} \boldsymbol{X}_{t-1} + \boldsymbol{U}_t \ , \tag{4}$$

where

$$\boldsymbol{X}_t = (x_t, x_{t-1}, \ldots, x_{t-p+1})^\top \ , \tag{5}$$

$$\boldsymbol{U}_t = (\varepsilon_t, 0, \ldots, 0)^\top \ , \tag{6}$$

$$\text{with } \boldsymbol{\Phi} = \begin{pmatrix} \phi_{1,p} & \phi_{2,p} & \cdots & \phi_{p,p} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} , \quad \boldsymbol{Q} = \begin{pmatrix} \sigma^2_{\varepsilon,p} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \text{ and} \tag{7}$$

$$y_t = x_t + v_t = (1, 0, \ldots, 0)\boldsymbol{X}_t + v_t \quad \text{with} \quad \boldsymbol{R} = \text{var}(v_t) = \sigma_0^2 \ . \tag{8}$$

Robust Filter-cleaner (Part II)

The algorithm computes robust estimates $\widehat{\boldsymbol{X}}_t$ of the vector $\boldsymbol{X}_t$ according to the following recursion:

$$\widehat{\boldsymbol{X}}_t = \boldsymbol{\Phi}\widehat{\boldsymbol{X}}_{t-1} + \frac{\boldsymbol{m}_t}{s_t^2}\ s_t\ \psi\left(\frac{y_t - \widehat{y}_t^{t-1}}{s_t}\right) \tag{9}$$

with $\boldsymbol{m}_t$ being the first column of $\boldsymbol{M}_t$, which is computed recursively as

$$\boldsymbol{M}_{t+1} \; = \; \boldsymbol{\Phi}\boldsymbol{P}_t\boldsymbol{\Phi}^\top + \boldsymbol{Q} \tag{10}$$

$$\boldsymbol{P}_t \; = \; \boldsymbol{M}_t - w\left(\frac{y_t - \widehat{y}_t^{t-1}}{s_t}\right)\frac{\boldsymbol{m}_t\boldsymbol{m}_t^\top}{s_t^2}\ . \tag{11}$$

The scale s_t is defined by $s_t^2 = m_{11,t}$ and $\widehat{y}_t^{t-1}$ denotes a robust one-step prediction of y_t based on $\boldsymbol{Y}^{t-1} = (y_1, \ldots, y_{t-1})^\top$, and is given by

$$\widehat{y}_t^{t-1} = (\boldsymbol{\Phi}\widehat{\boldsymbol{X}}_{t-1})_1\ . \tag{12}$$

Finally, the cleaned process at time t is

$$\widehat{x}_t = (\widehat{\boldsymbol{X}}_t)_1\ . \tag{13}$$

Robust Filter-cleaner (Part III)

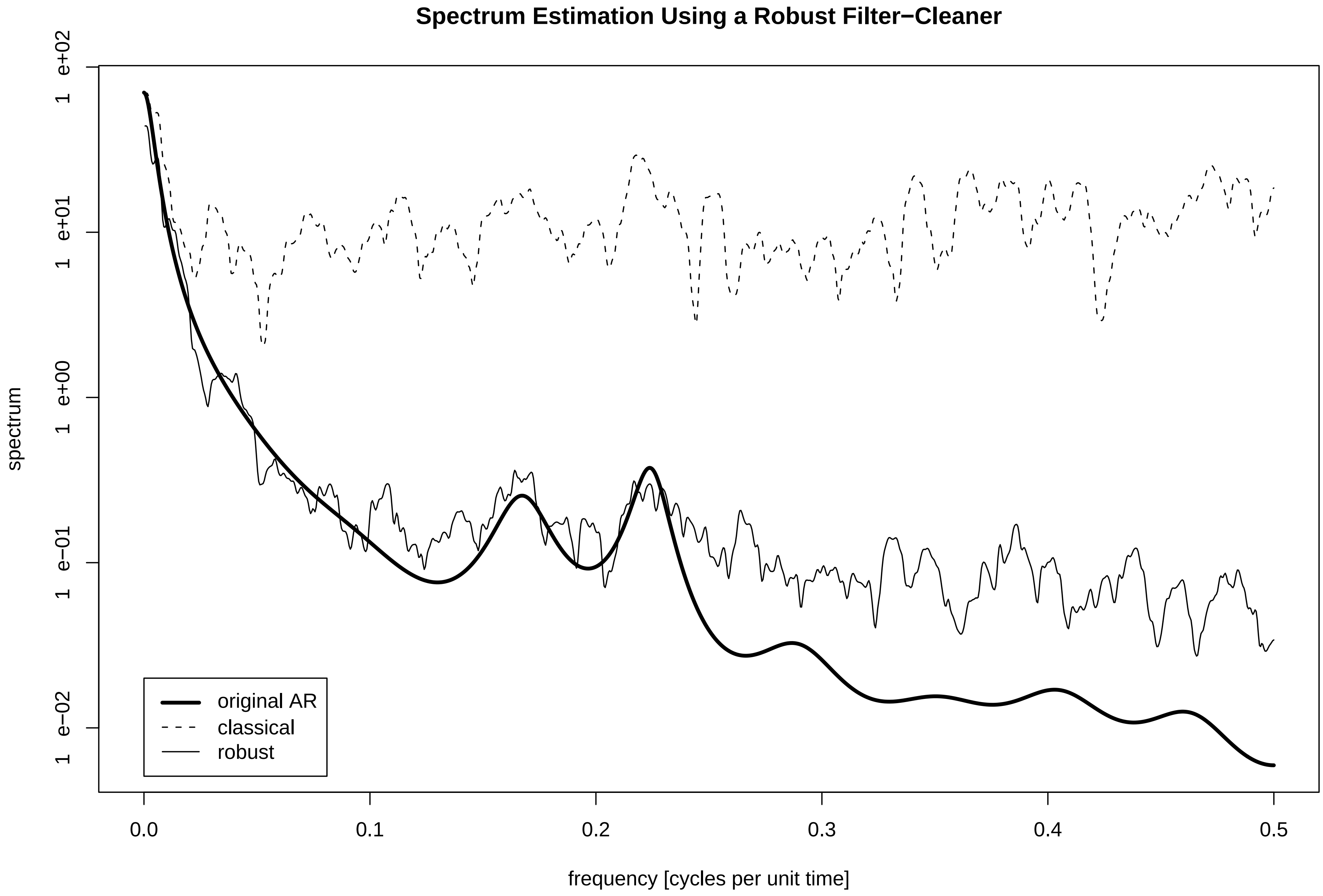
To use the filter-cleaner algorithm we need robust estimates $\widehat{\phi}_p$ and $\hat{\sigma}^2_{\varepsilon,p} = s^2_{\varepsilon,p}$ of $\phi_p = (\phi_{1,p}, \ldots, \phi_{p,p})^\top$ and $\sigma^2_{\varepsilon,p}$.

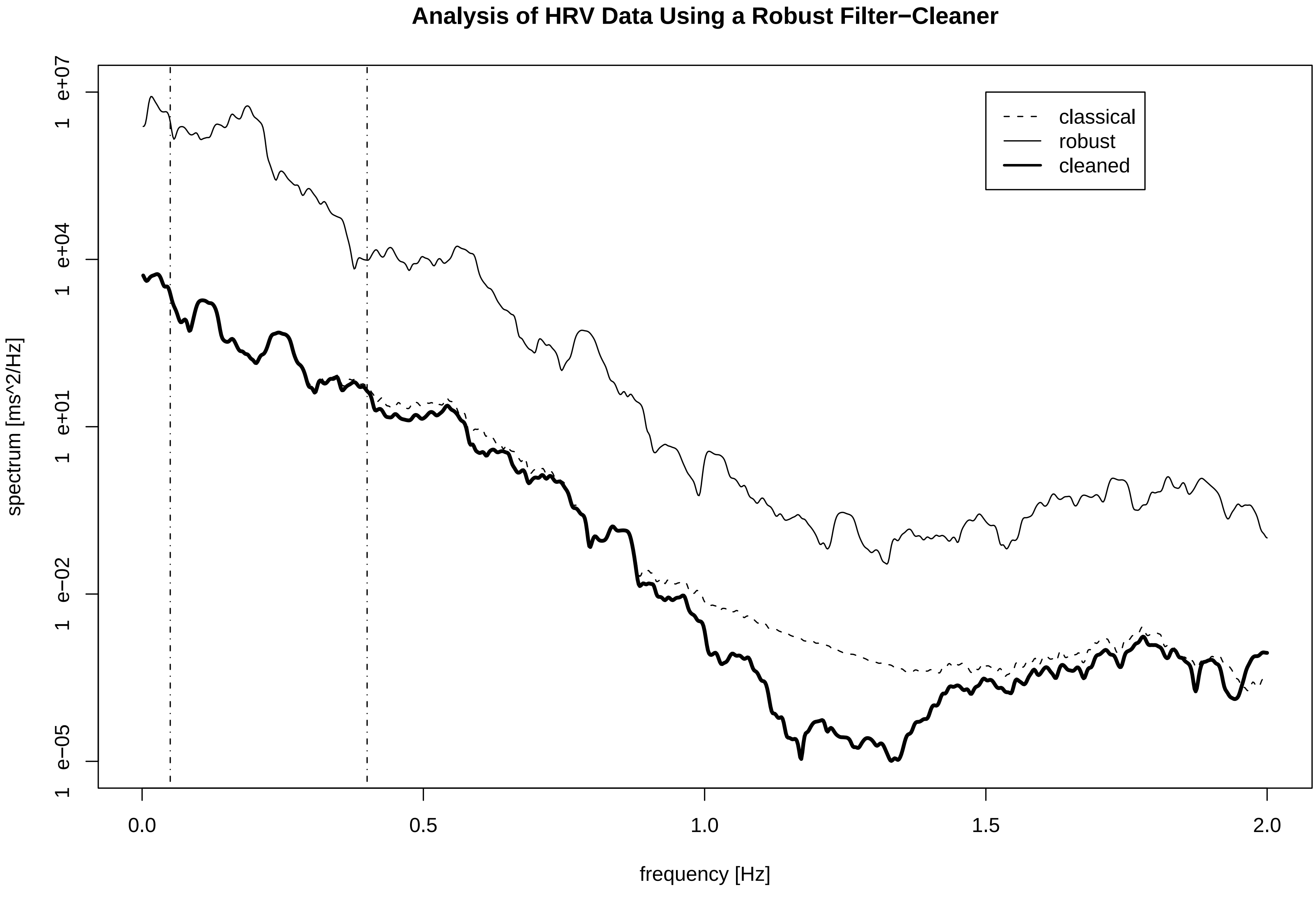
Until now, we have tried different approaches to obtain initial estimates leading to similar results

- using bounded-influence autoregression (BIAR) via iterated re-weighted least squares (IWLS) (cf. Martin & Thomson, 1982) or

- a highly robust autocovariance function estimator (cf. Ma & Genton, 2000) and the Yule-Walker equations or

- robust autoregression using LTS- and LMS-regression.


Spectrum Estimation Using a Robust Filter−Cleaner
spectrum
1 e+02
1 e+01
1 e+00
1 e−01
1 e−02
0.0
0.1
0.2
0.3
0.4
0.5
frequency [cycles per unit time]
original AR
classical
robust



Analysis of HRV Data Using a Robust Filter−Cleaner
spectrum [ms^2/Hz]
1 e+07
1 e+04
1 e+01
1 e−02
1 e−05
0.0
0.5
1.0
1.5
2.0
frequency [Hz]
classical
robust
cleaned

Robust Spectral Estimation via WOSA (Overview)

- WOSA – Welch's Overlapped Segment Averaging
- Definition of Spectral Estimation via WOSA
- Robustification of the WOSA spectral estimator

Robust Spectral Estimation via WOSA (Part 1)

Let x_t, $t = 1, \ldots, N$, be the observed process.

- Splitting the process into N_B overlapping blocks of length N_S
- Calculating direct spectral estimates for different blocks of N_S contiguous data values

$$\widehat{S}_l^{(d)}(f) := \Delta t \left| \sum_{t=1}^{N_S} h_t x_{t+l-1} e^{-i2\pi f t \Delta t} \right|^2 , \quad 1 \le l \le N + 1 - N_S . \quad (14)$$

where $h_1, \ldots, h_{N_S}$ is a data taper.

- Then the WOSA spectral estimator is defined by

$$\widehat{S}^{(WOSA)}(f) := \frac{1}{N_B} \sum_{j=0}^{N_B - 1} \widehat{S}_{jn+1}^{(d)}(f) . \quad (15)$$

Robust Spectral Estimation via WOSA (Part 2)

A robust spectral estimator can be obtained by replacing the sample mean by an M-estimator (cf. Chave et al., 1987), i.e.,

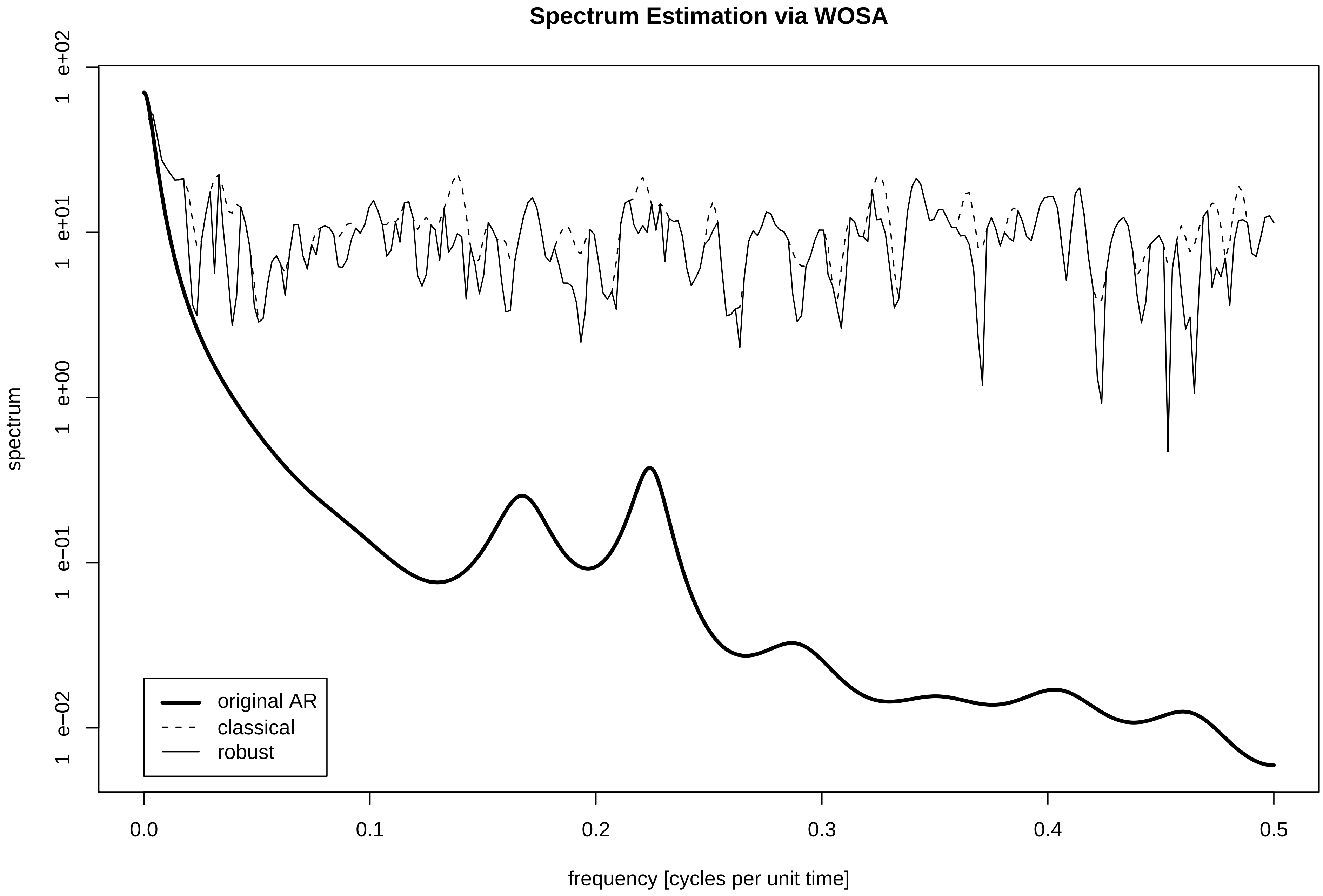
$$\min_{\theta} \sum_{i=1}^{N} \rho\left(\frac{x_i - \theta}{s}\right) \quad \Leftrightarrow \quad \sum_{i=1}^{N} \psi\left(\frac{x_i - \theta}{s}\right) = 0 \;, \tag{16}$$

where the x_i, $i = 1, \ldots, N$, are independent and $\psi(x, \theta) = \frac{\partial}{\partial\theta}\rho(x, \theta)$ is called an influence function. The solution $\widehat{\theta}$ is called an M-estimate.

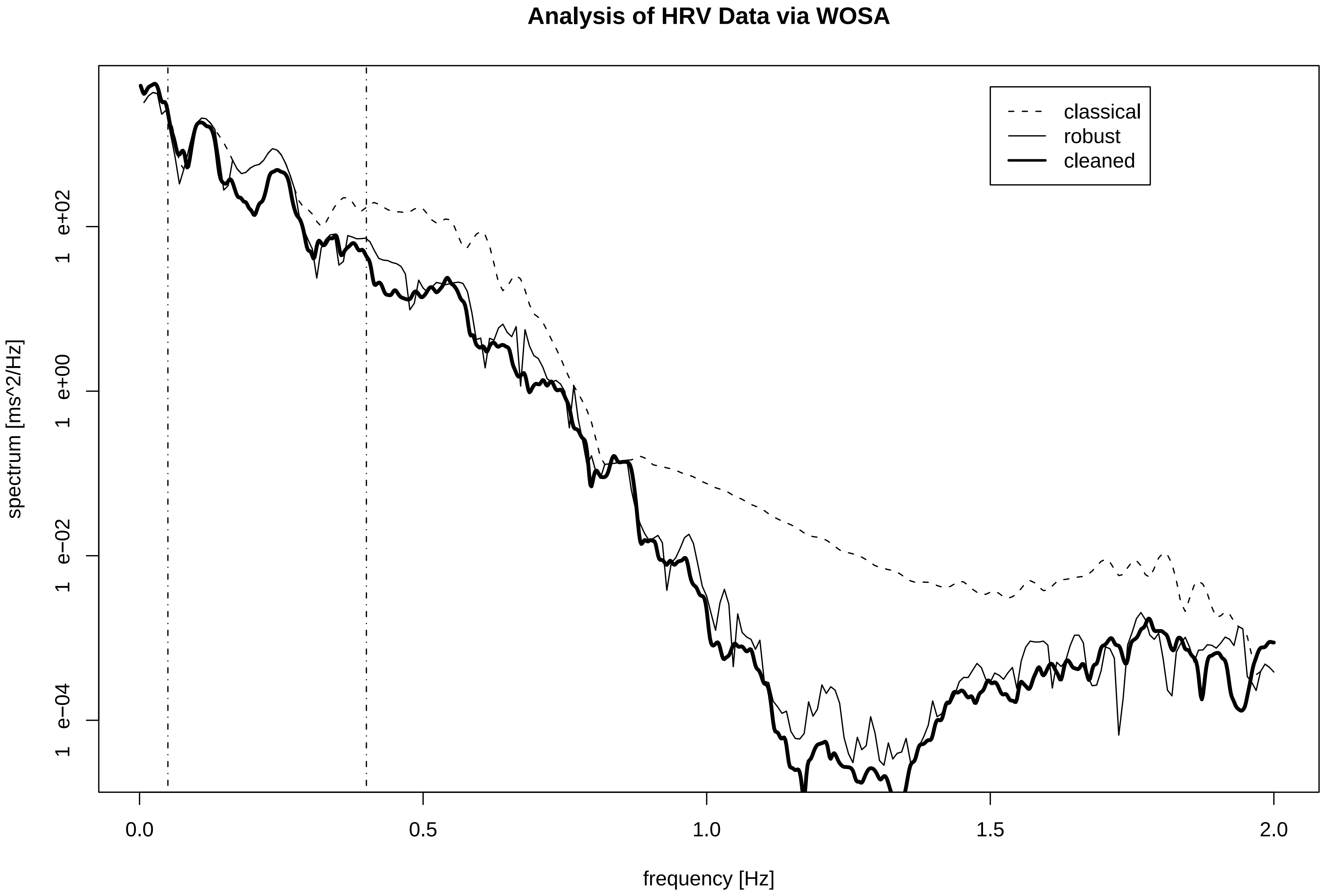
Because outlier contamination can only result in a spectrum that is biased upwards, a special asymmetric influence function is used:

$$\psi(x) = x \exp(-\exp(\beta(x - \beta))) \;. \tag{17}$$

The solution is calculated using iterated re-weighted least squares (IWLS) with proper initial values, e.g., the sample median and a corrected version of the median absolute deviation (MAD).



Spectrum Estimation via WOSA
spectrum
1 e+02
1 e+01
1 e+00
1 e−01
1 e−02
0.0
0.1
0.2
0.3
0.4
0.5
frequency [cycles per unit time]
original AR
classical
robust



Analysis of HRV Data via WOSA
classical
robust
cleaned
spectrum [ms^2/Hz]
1 e+02
1 e+00
1 e−02
1 e−04
0.0
0.5
1.0
1.5
2.0
frequency [Hz]

Summary & Discussion

- Simulated process
  - spectral estimate using the robust filter-cleaner approach: good
  - robust WOSA spectral estimate: poor

- Analysis of HRV data
  - spectral estimate using the robust filter-cleaner approach: poor
  - robust WOSA spectral estimate: good

Further Research

- Further detailed research of the algorithms
- Using a band-pass filter in the case of the analysis of HRV data
- Extension of the spectral estimators by Thomson's multitaper approach