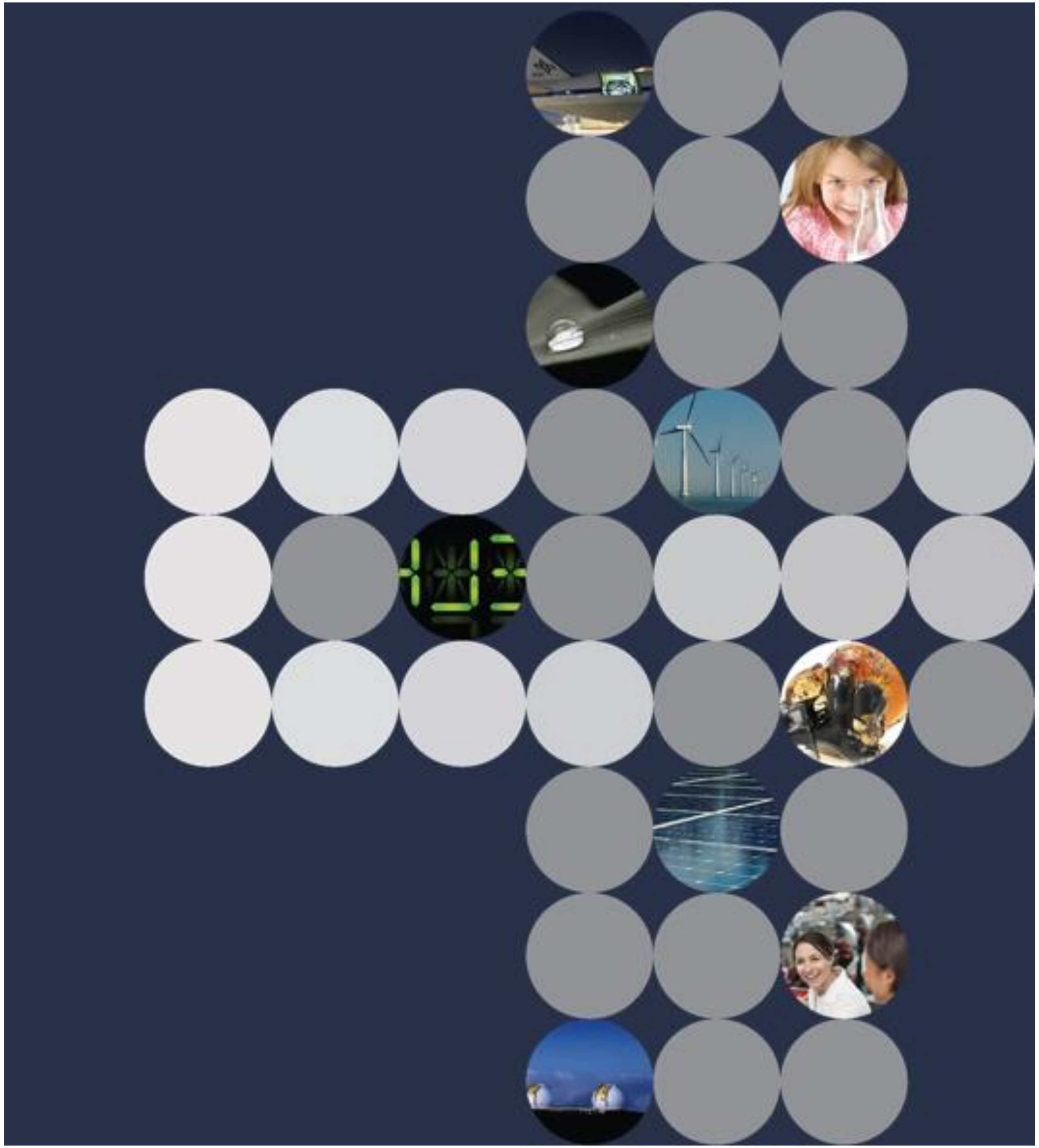
EE512 – Applied Biomedical Signal Processing

Basics II

Philippe RENEVEY
CSEM Signal Processing Group

EPFL csem



Content

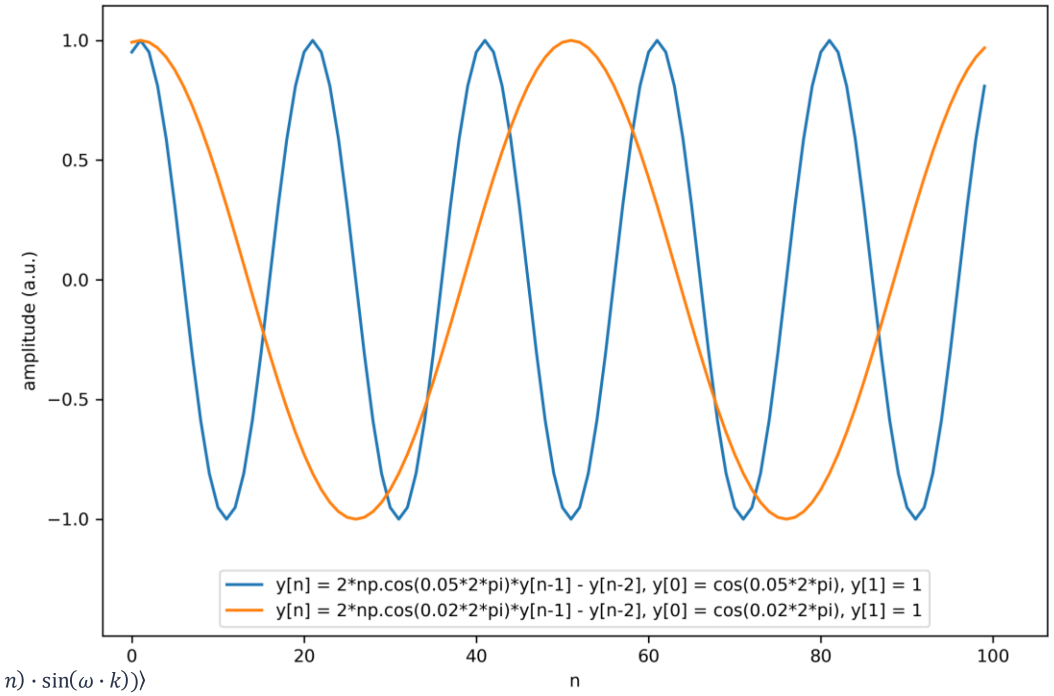
- Deterministic vs Random
- Stochastic process
- Bias, variance and consistence
- Auto correlation and auto covariance
- Stationarity
- Inter / cross correlation
- Power spectral density

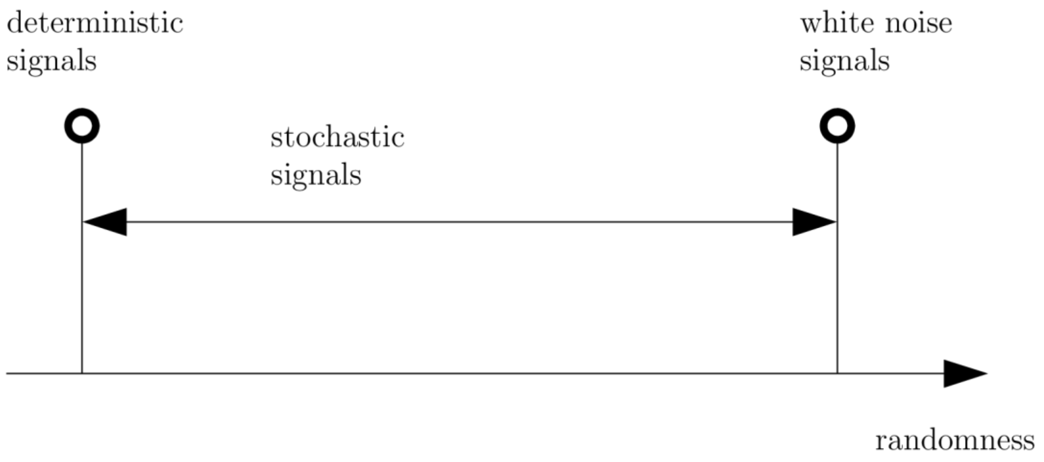
Deterministic vs random

- Deterministic → past and future can be predicted from a small set of measurements
  - $$\begin{cases} y(n) = 2 \cdot \cos(\omega) \cdot y(n-1) - y(n-2) \\ y(-1) = \cos(\omega) \\ y(-2) = \cos(2\omega) \end{cases}$$
  - $y(n) = [\cos(\omega \cdot n)]$
  - $\langle y(n) \cdot y(n-k) \rangle = \frac{1}{2}\cos(\omega \cdot k)$

$$\langle \cos(\omega \cdot n) \cdot \cos(\omega \cdot (n-k)) \rangle$$
$$= \langle \cos(\omega \cdot n) \cdot (\cos(\omega \cdot n) \cdot \cos(\omega \cdot k) + \sin(\omega \cdot n) \cdot \sin(\omega \cdot k)) \rangle$$
$$= \langle cos^2(\omega \cdot n) \rangle \cdot \cos(\omega \cdot k) = \frac{1}{2}\cos(\omega \cdot k)$$

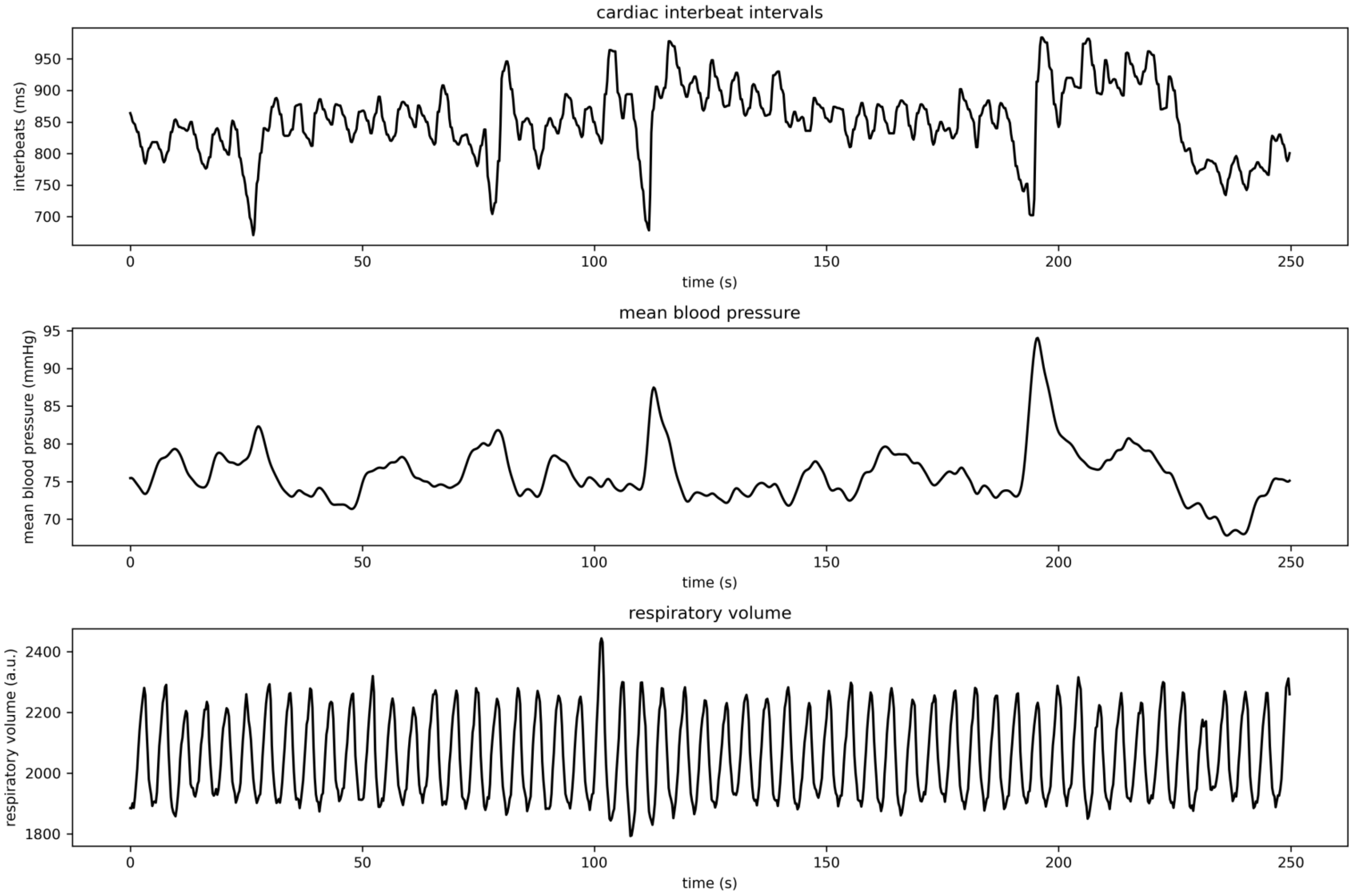
- White Gaussian noise
  - $y(n) \sim \mathcal{N}(\mu, \sigma^2)$
  - μ mean of the Gaussian
  - σ^2 variance of the Gaussian
  - $\langle y(n) \cdot y(n-k) \rangle = \mu^2 \; \forall \, n \neq k$
  - $\langle y(n) \cdot y(n) \rangle = \mu^2 + \sigma^2$





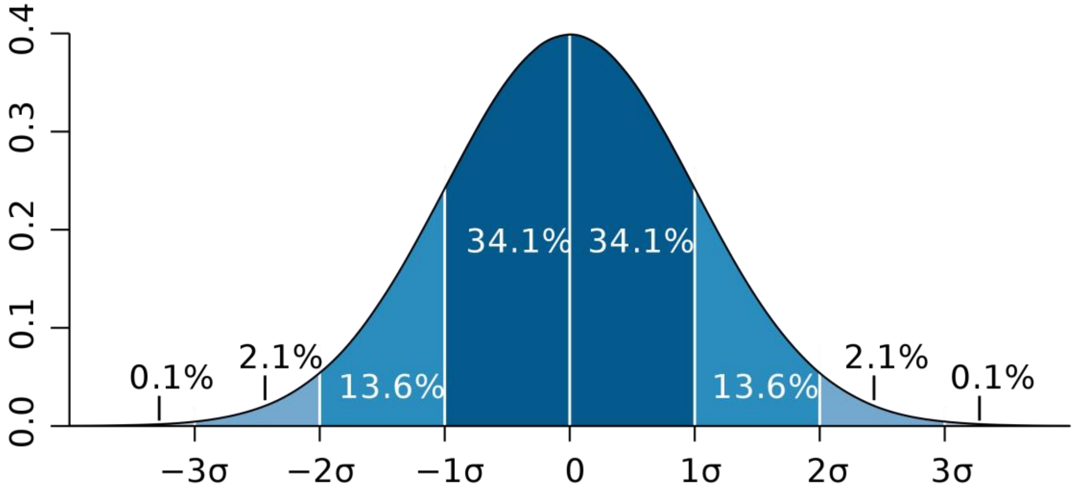
Example of stochastic signals

- signals present a mix between random and deterministic behaviors
- respiratory volume is the closest to deterministic
- mean blood pressure is the closest to a noise

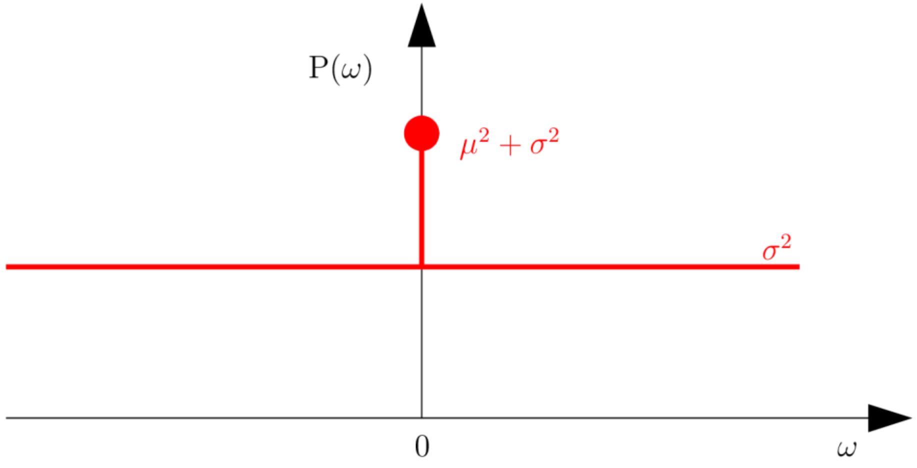


White Gaussian noise

- $y(n) \sim \mathcal{N}(\mu, \sigma^2)$
  - $\text{pdf}\big(y(n)\big) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(y(n)-\mu)^2}{2\sigma^2}}$
  - $\langle y(n) \rangle = \mu$
  - $\langle (y(n))^2 \rangle = \sigma^2 + \mu^2$
  - $\langle y(n) \cdot y(n-k) \rangle = \mu^2 \forall\ k \neq 0$
    - all the samples are independents
  - $\langle Y(\omega) \cdot Y^*(\omega) \rangle = \sigma^2 + \delta(\omega) \cdot \mu^2$
- Mean of Gaussian processes
  - $\frac{1}{K}\sum_{k=1}^{K} \mathcal{N}(\mu, \sigma^2) \sim \mathcal{N}\left(\mu, \frac{1}{K}\sigma^2\right)$

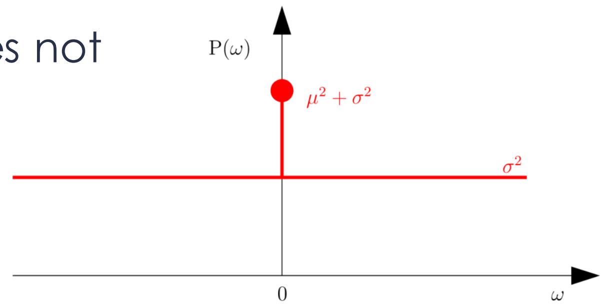


https://upload.wikimedia.org/wikipedia/commons/8/8c/Standard_deviation_diagram.svg

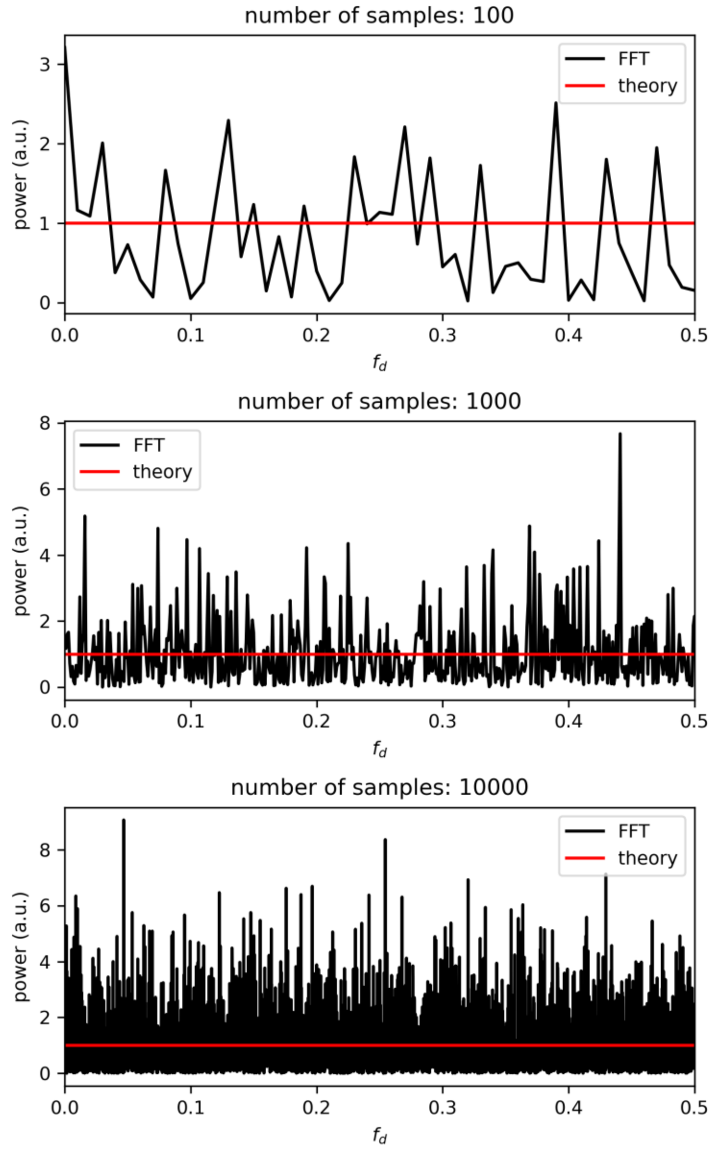


Stochastic processes

- Relationship between consecutive measurements exists but it **can only be analyzed statistically**
- Tools developed for the analysis of deterministic signals are poorly suitable for the analysis of stochastic signals
- In order to get relevant information from the time series **averaging is mandatory**
  - FFT transform **n samples to n samples**
  - Increasing the number of samples does not improve the estimation
  - **FFT is a non-consistent estimator**



FFT of white Gaussian noise $\mathcal{N}(0,1)$

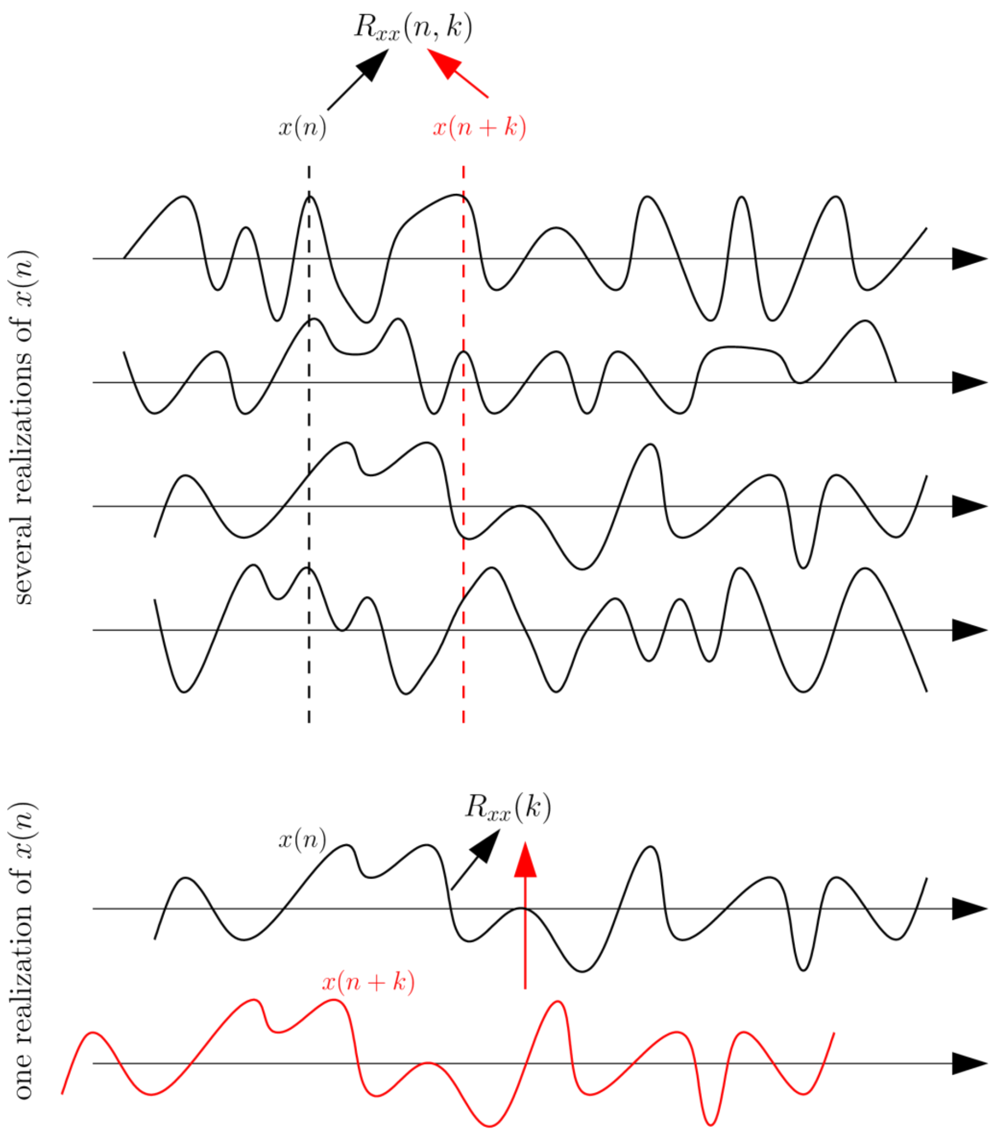


Mean and variance

- $y(n)$ is a stochastic variable
- *mean operator*
  - $\widehat{\mu_y}(N) = \frac{1}{N}\sum_{n=1}^{N} y(n)$
  - $\widehat{\mu_y}(N) \sim \mathcal{N}\left(\mu_y, \frac{\sigma_y^2}{N}\right)$
  - the estimator is **non biased**:
    - $\langle\widehat{\mu_y}(N)\rangle = \mu_y$
  - the estimator is **consistent:**
    - $\lim_{N\to\infty} \mathcal{N}\left(\mu_y, \frac{\sigma_y^2}{N}\right) = \mathcal{N}(\mu_y, 0)$
- *variance operator*
  - $\widehat{\sigma_y^2}(N) = \frac{1}{N}\sum_{n=1}^{N}(y(n) - \widehat{\mu_y}(N))^2$
  - $\widehat{\sigma_y^2}(N) \sim \mathcal{N}\left(\frac{N-1}{N}\sigma_y^2, \frac{2\cdot\sigma_y^4}{N}\right)$
  - the estimator is **biased:**
    - $\left\langle\widehat{\sigma_y^2}(N)\right\rangle = \frac{N-1}{N}\sigma_y^2$
  - the estimator is **consistent:**
    - $\lim_{N\to\infty} \mathcal{N}\left(\frac{N-1}{N}\sigma_y^2, \frac{2\cdot\sigma_y^4}{N}\right) = \mathcal{N}(\sigma_y^2, 0)$
    - asymptotically non biased

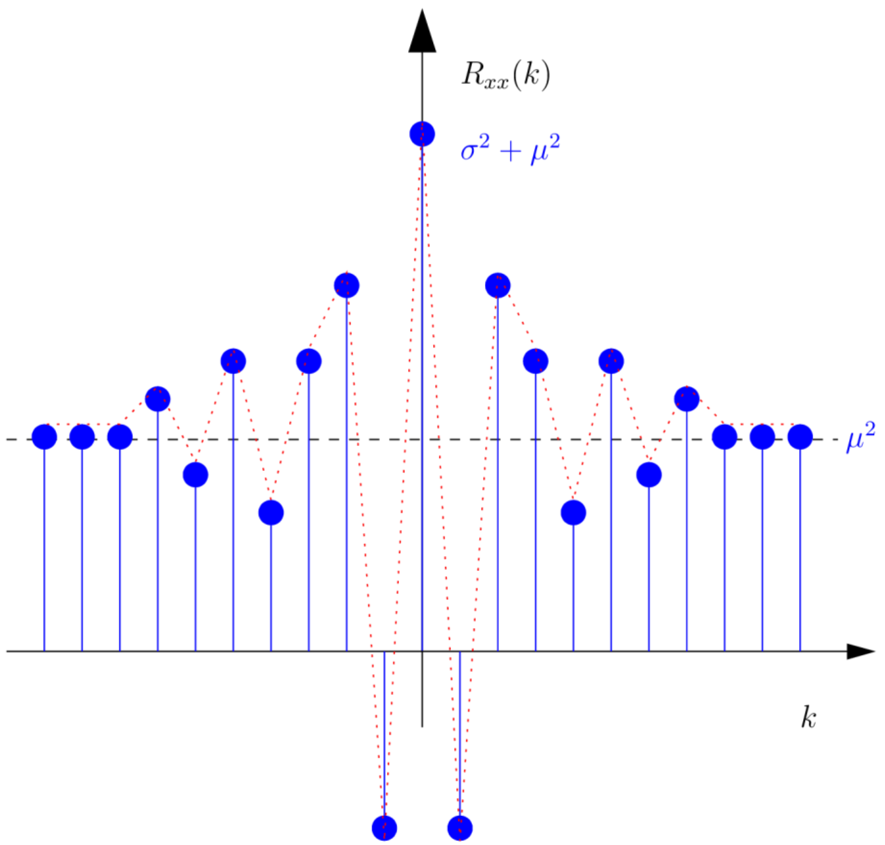
Auto correlation

- *$x(n)$ is a realization of a stochastic process*
- *Its auto correlation is given by*
  - $R_{xx}(n,k) = \langle x(n) \cdot x(n+k) \rangle$
- x is a **stationary process**
  - $R_{xx}(n,k) = R_{xx}(k)$
  - The estimation of the auto correlation for several realization of the process at time *n* is equivalent to the estimation of the auto correlation on one realization independently of the time



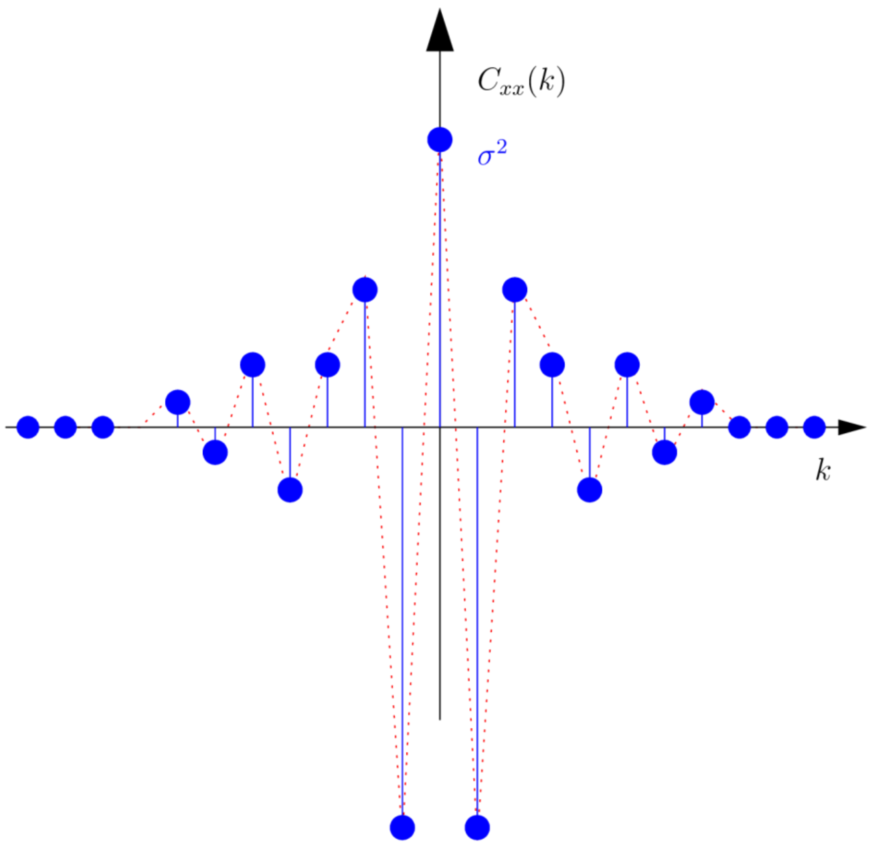
Auto correlation of a stationary process

- $R_{xx}(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n) \cdot x(n+k)$
- $R_{xx}(0) = \mu_x^2 + \sigma_x^2$
  - *power of the signal*
- $R_{xx}(-k) = R_{xx}(k)$
  - *symmetry of the autocorrelation*
- $\lim_{k\to\pm\infty} R_{xx}(k) = \mu_x^2$
  - *with exception of sustained oscillations*
- $|R_{xx}(k)| \leq R_{xx}(0)$



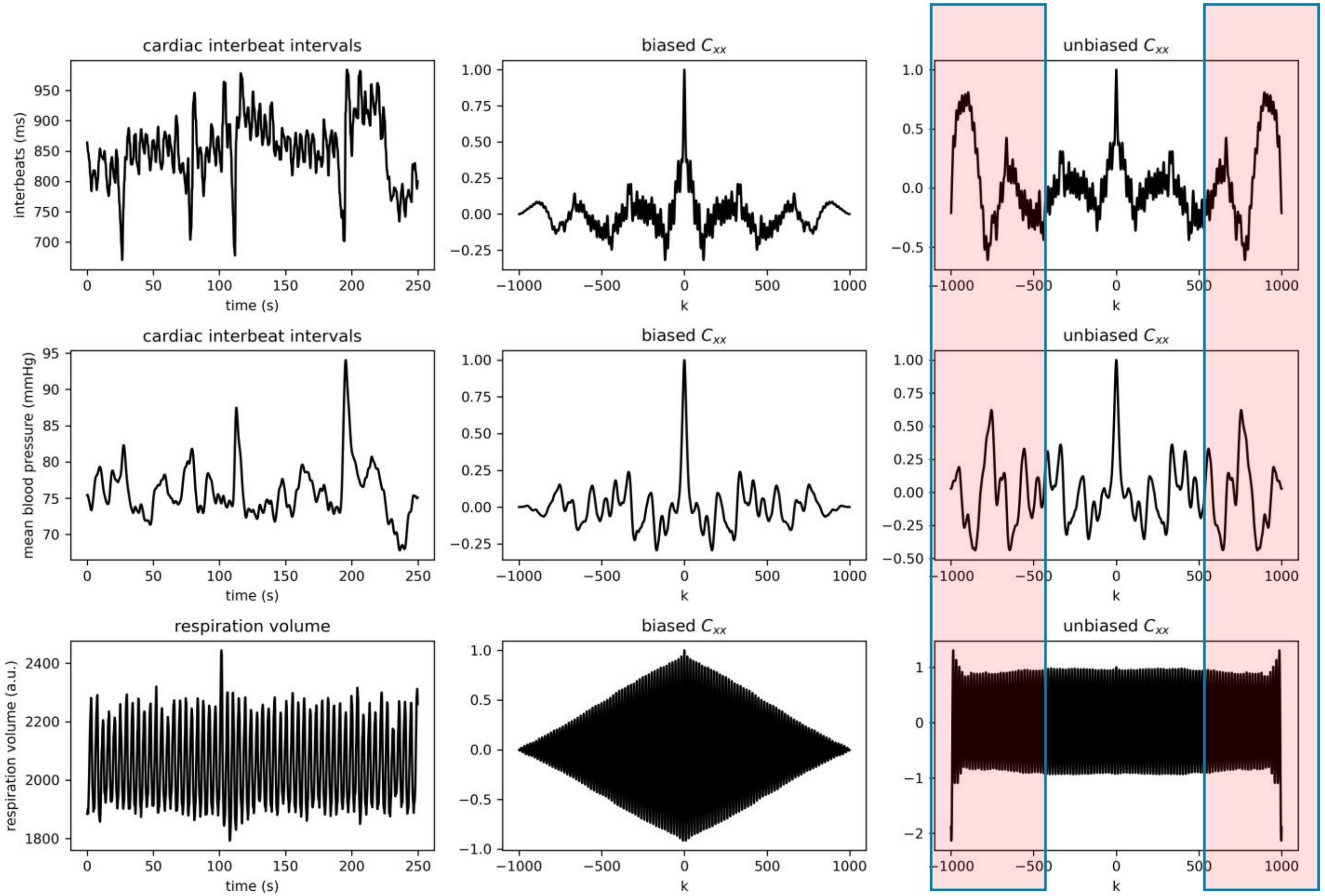
Auto covariance of a stationary process

- $C_{xx}(k) = \frac{1}{N}\sum_{n=0}^{N-1}(x(n) - \mu_x) \cdot (x(n+k) - \mu_x)$
- $C_{xx}(0) = \sigma_x^2$
  - *power of the signal*
- $C_{xx}(-k) = C_{xx}(k)$
  - *symmetry of the autocovariance*
- $\lim_{k \to \pm\infty} C_{xx}(k) = 0$
- $|C_{xx}(k)| \leq C_{xx}(0)$
- $C_{xx}(k) = R_{xx}(k)$ if $\mu_x = 0$



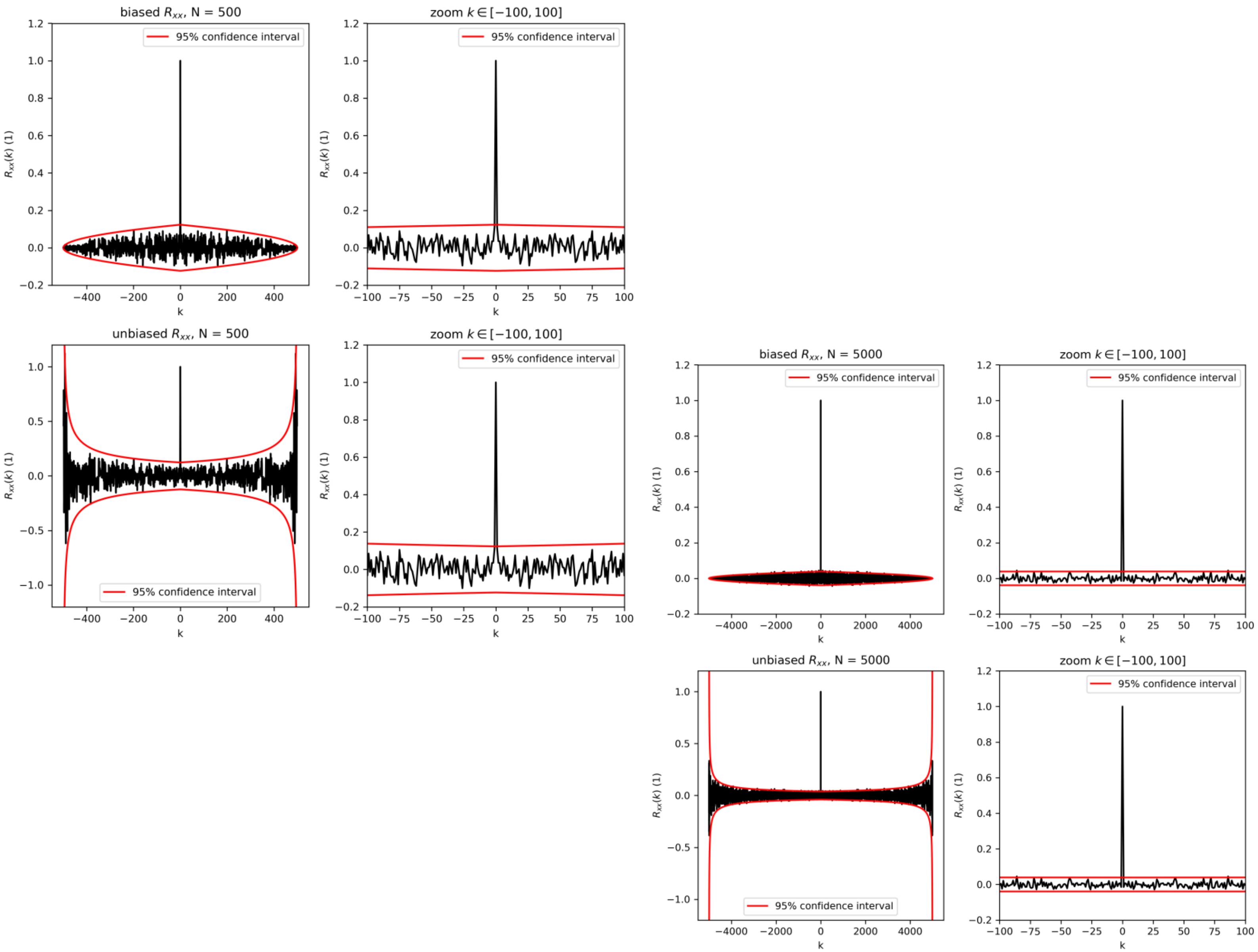
Auto correlation | covariance with a fixed number of samples

- Biased estimator
  - $R_{xx}(k) = \frac{1}{N}\sum_{n=0}^{N-1-k} x(n) \cdot x(n+k)$
- Non biased estimator
  - $R_{xx}(k) = \frac{1}{N-k}\sum_{n=0}^{N-1-k} x(n) \cdot x(n+k)$
- Unbiased estimator have no bias but introduce a large amount of noise at the border values of R_{xx} → generally the biased estimator is preferred for practical applications



Auto correlation of a white noise

- For a white Gaussian noise
  - $R_{xx}(k) = \sigma_x^2 \cdot \delta(k)$
  - $\widehat{\sigma^2}(N) \sim \mathcal{N}\left(\frac{N-1}{N}\sigma^2, \frac{2 \cdot \sigma^4}{N}\right)$
- For $k \neq 0$
  - $|R_{xx}(k)| < 1.96 \cdot \sqrt{\frac{2}{N}} \cdot R_{xx}(0)$
  - This is the **95% confidence interval**
  - (95% of the value must fulfil this criteria for a WGN)



Autocorrelation of a filtered WGN

- Difference equation of the filter

$$y(n) + \sum_{i=1}^{N_a} a_i \cdot y(n-i) = \sum_{i=0}^{N_b} b_i \cdot x(n-i)$$

- Z transform

$$Y(Z) = \frac{B(z)}{A(z)} \cdot X(Z) = H(z) \cdot X(Z)$$

- X is a zero mean white Gaussian noise

$$x(n) \sim N(0, \sigma^2)$$

- Power spectral density σ^2 y
  - $\mathrm{PSD}(y) = Y(z) \cdot Y^*(z) = H(z) \cdot X(Z) \cdot H^*(z) \cdot X^*(Z)$
  - $= X(Z) \cdot X^*(Z) \cdot H(z) \cdot H^*(z)$
  - $= \sigma^2 \cdot H(z) \cdot H^*(z)$

- Autocorrelation
  - $R_{yy}(n) = \mathrm{Z}^{-1}\left(\sigma^2 \cdot H(z) \cdot H^*(z)\right)$
    $= \sigma^2 (h(n) * h(-n)) = \sigma^2 \cdot R_{hh}(n)$

- The autocorrelation of a filtered WGN is the product of the variance of the WGN and the auto correlation of the impulse response

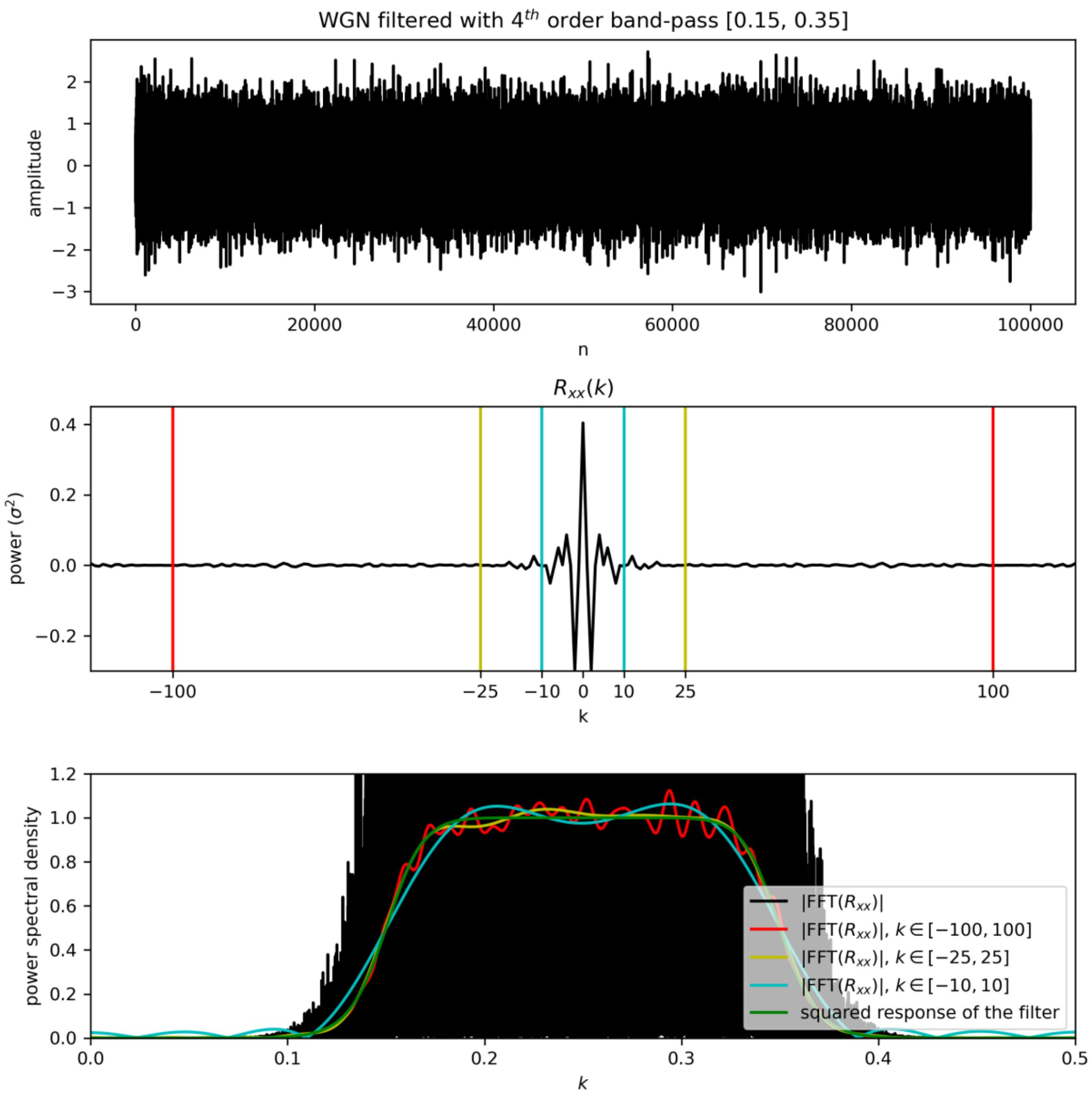
- Example:
  - $H(z) = 1 - a \cdot z^{-1}$
  - $Y(z) = H(z) \cdot X(z)$
  - $x(n) \sim N(0, \sigma^2)$
  - $R_{yy}(k) = \begin{cases} \sigma^2(1+a^2), & k = 0 \\ -\sigma^2 a, & k = \pm 1 \\ 0, & |k| > 1 \end{cases}$

Typical exam question

- A white Gaussian noise $x(n)$ of zero mean and unit variance is filtered by the filter $y(n) = x(n) + \frac{1}{2} \cdot x(n-1) + \frac{1}{4} \cdot x(n-2).$ Compute the non-zero values of the autocorrelation of y.

Power spectral density

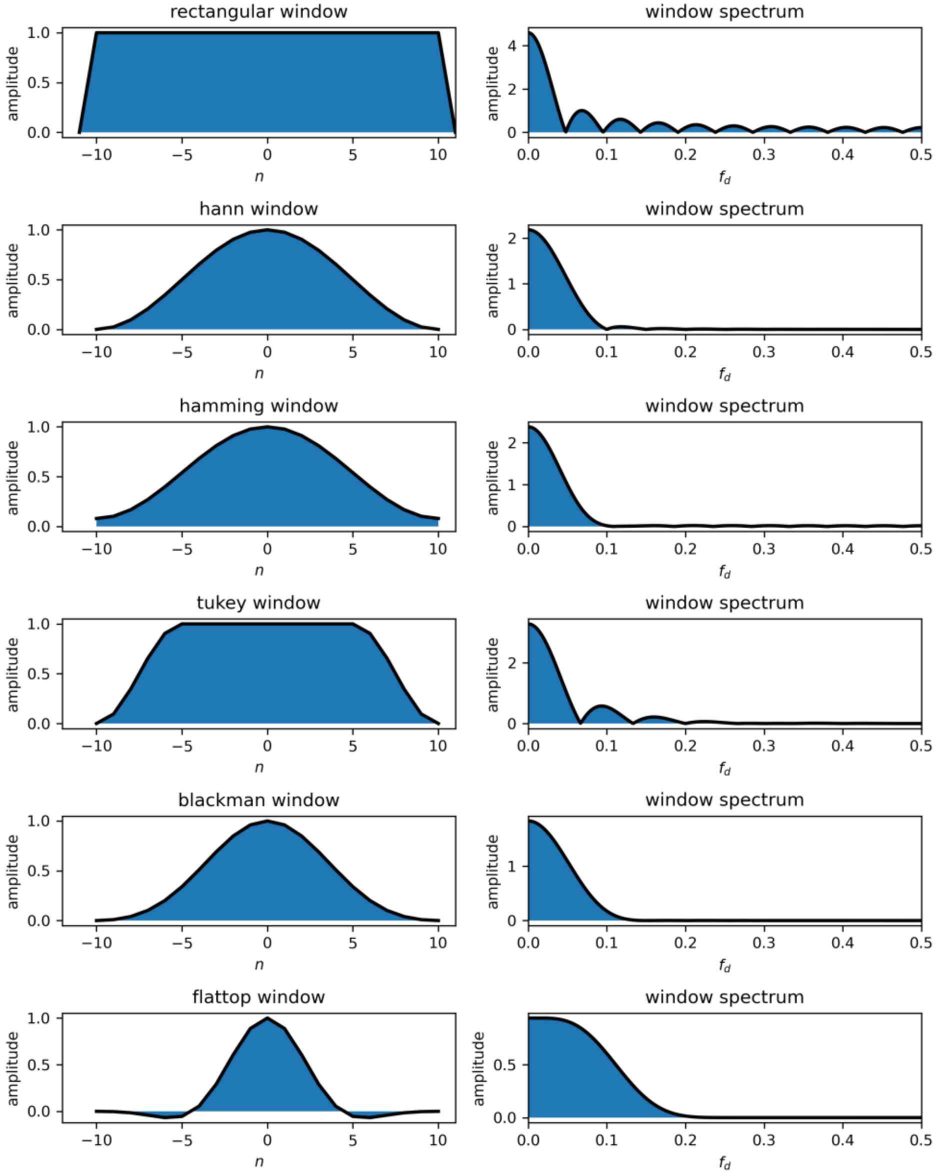
- $\text{FFT}^{-1}\left(X^*(k)\right) = \text{FFT}^{-1}\left(X(-k)\right) = x(-n)$
- $\text{FFT}^{-1}\left(X(k) \cdot X^*(k)\right) = x(n) * x(-n) = R_{xx}(n)$
- Remember:
  - FFT is a non consistent spectral estimator (increase of the point does not decrease noise variance)
  - $R_{xx}(k)$ is poorly defined in its borders due to the reduced number of values used for its estimation
- → **Use only the central part of $R_{xx}(k)$**



EPFL csem

Windowing

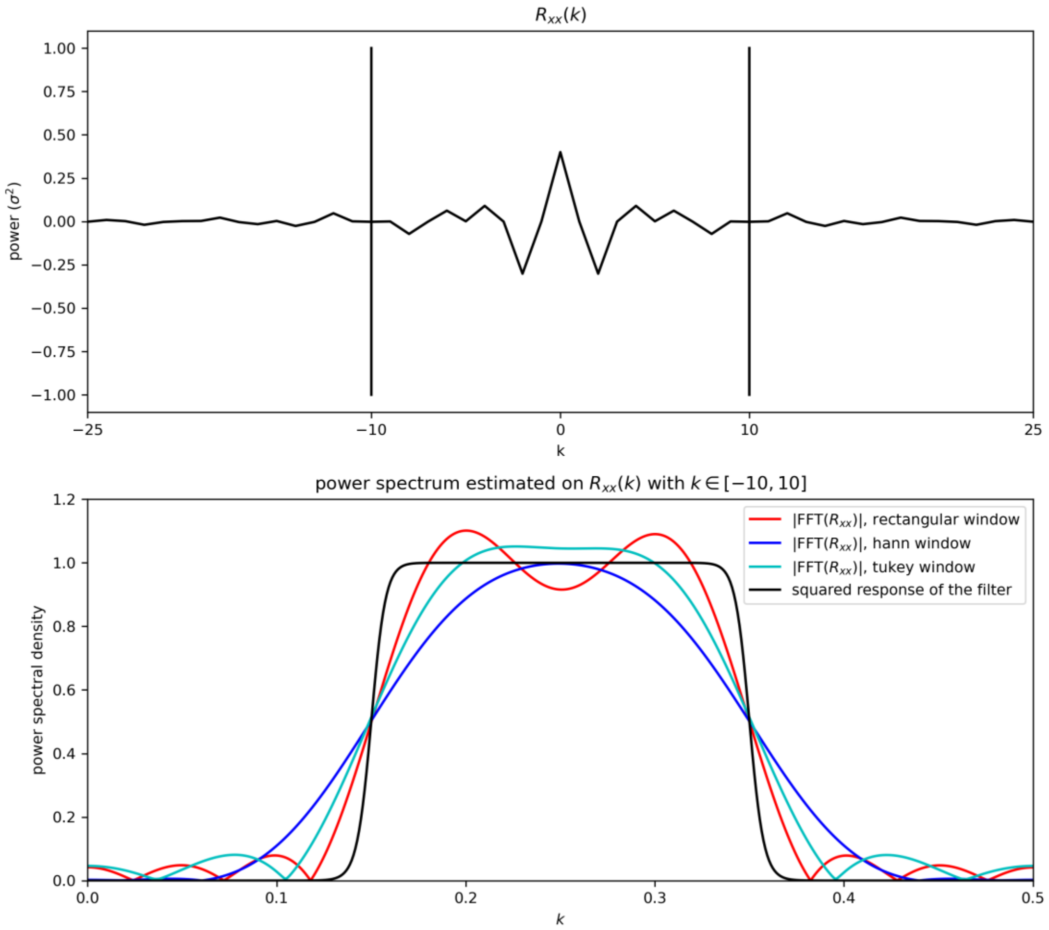
- Truncation of R_{xx} create **oscillations** if $R_{xx} \neq 0$ outside to truncation interval (rectangular window)
- In order to estimate power spectral distribution more accurately
    - $\mathrm{FFT}(w(n) \cdot R_{xx}(n))$
- The selection of the window is a **compromise** between **spectral resolution** and **oscillations**



EPFL csem

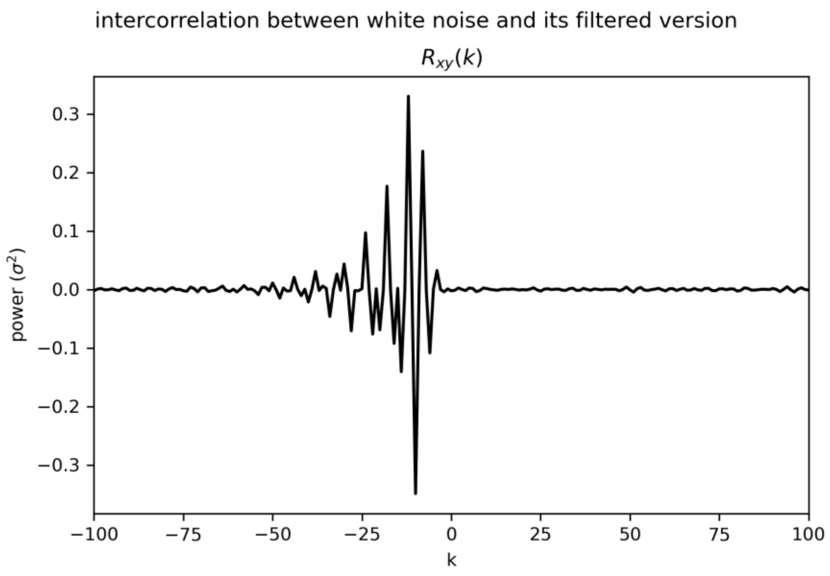
Windows examples

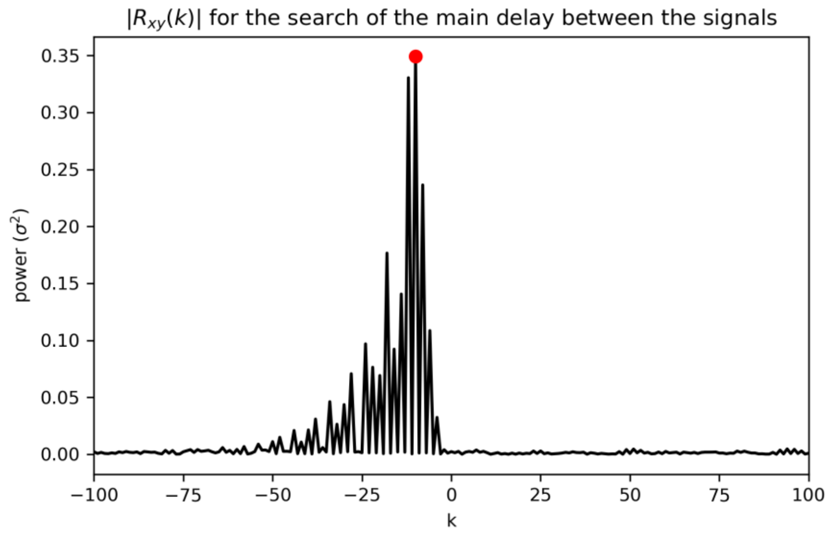
- The selection of the windows permits a compromise between oscillations and spectral resolution
- Generally
  - When small amplitude components are mixed with large amplitude components → minimize oscillations
  - When components of same amplitude are mixed → optimize spectral resolution



Intercorrelation

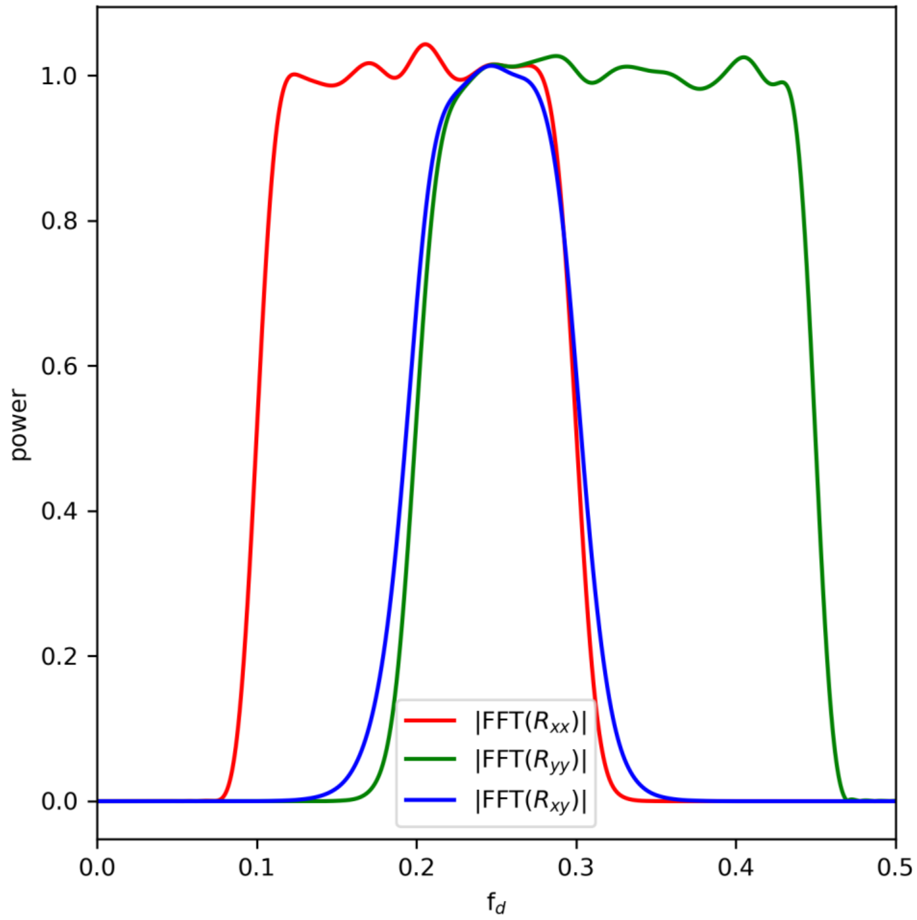
- $R_{xy}(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n) \cdot y(n+k)$
- The intercorrelation measures the similarities between two signal
- The measure is not symmetrical
- $\max(|R_{xy}(k)|)$ gives the main delay between the two signals





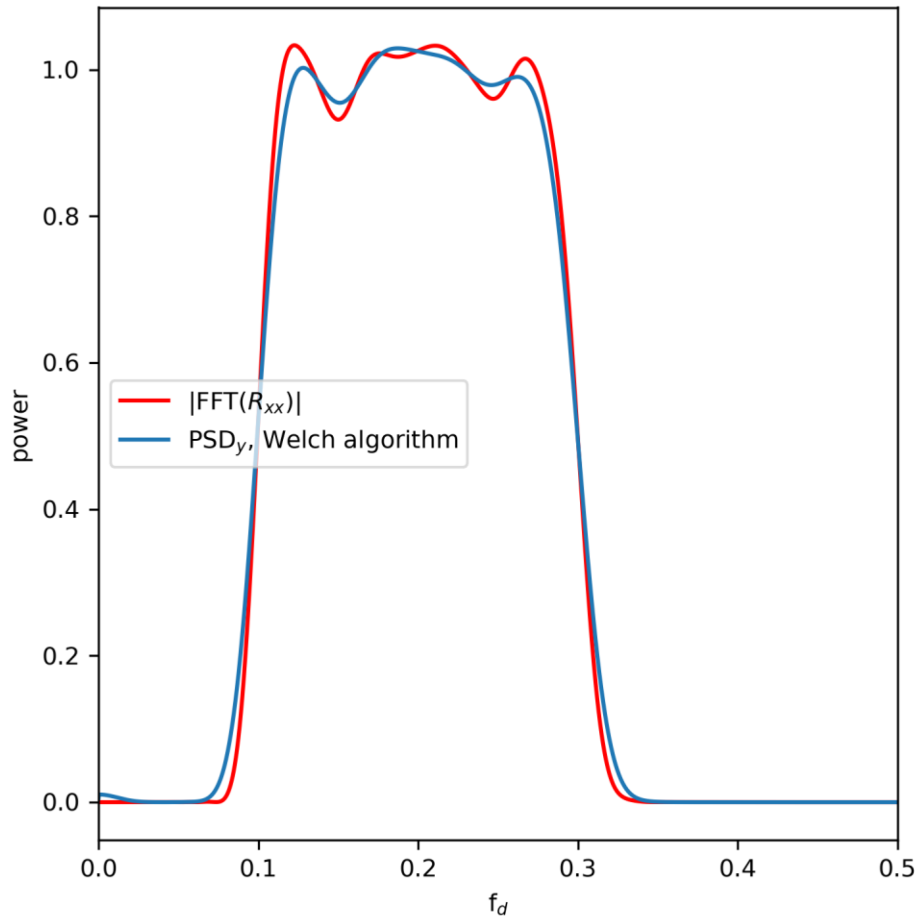
Power inter-spectral density

- $\text{FFT}(R_{xy})$ is the power inter-spectral density
- Its value is large in frequencies where the two signal have a relationship and lower elsewhere
- The processing is similar to power spectral density
  - windowing, central interval of the intercorrelation, ...



Welch power spectral density

- $\text{FFT}(R_{xx}) \leftrightarrow X \cdot X^*$
  - Relation exists between power of the FFT and FFT of the autocorrelation
  - Robust estimation of the power spectral density implied to use the central part of R_{xx}
  - Averaging the FFT using shorter blocks (half of the one used for the central part of R_{xx} gives same results) with 50% overlap
- $\text{PSD}_{Welch}(\text{k}) = \frac{1}{N_{block}} \sum_{i=1}^{N_{block}} (\text{FFT}(win \cdot x_{block}(i)))^2$
- The only difference is that in Welch algorithm shows an effect that corresponds to the square of the window

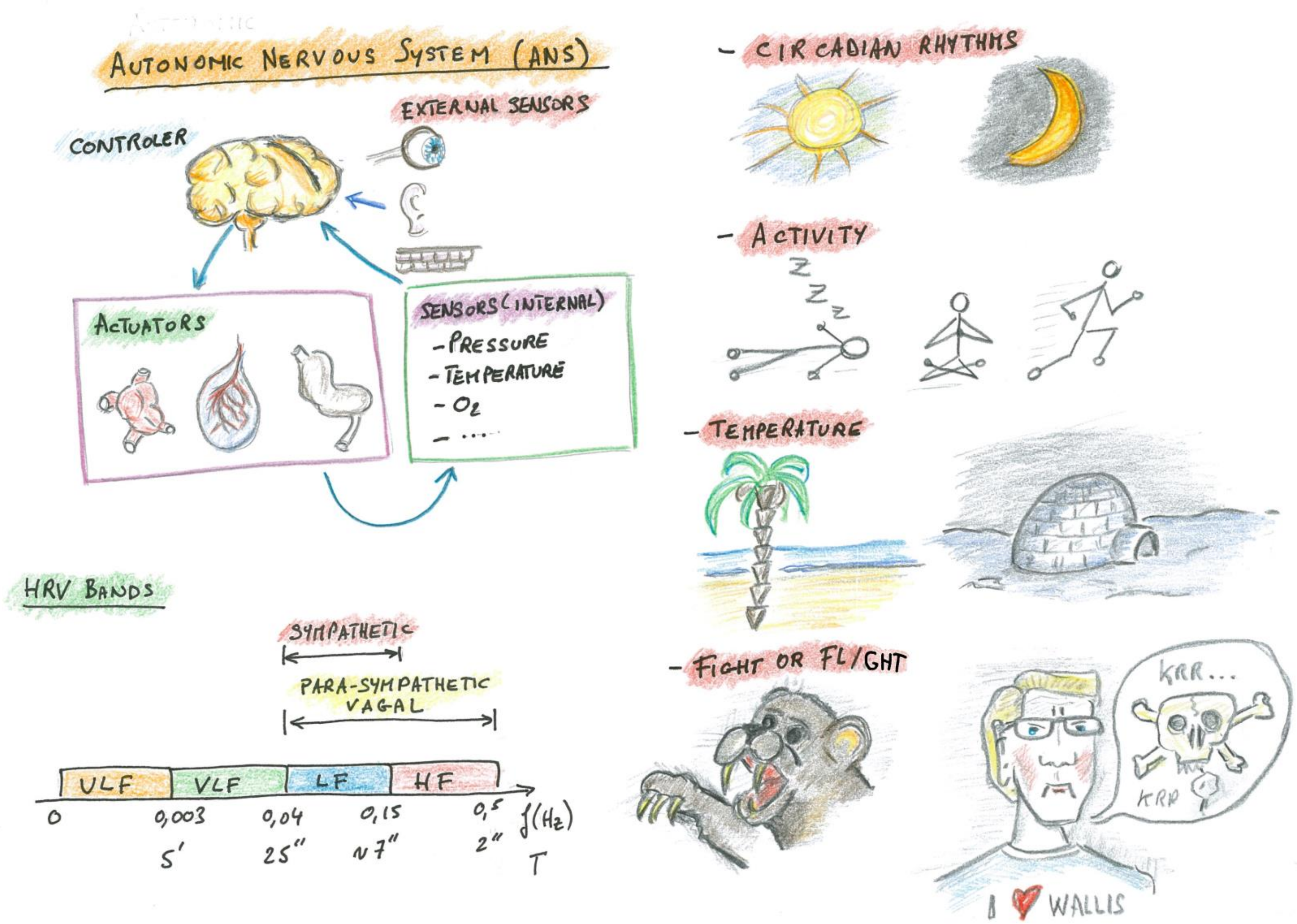


EPFL csem

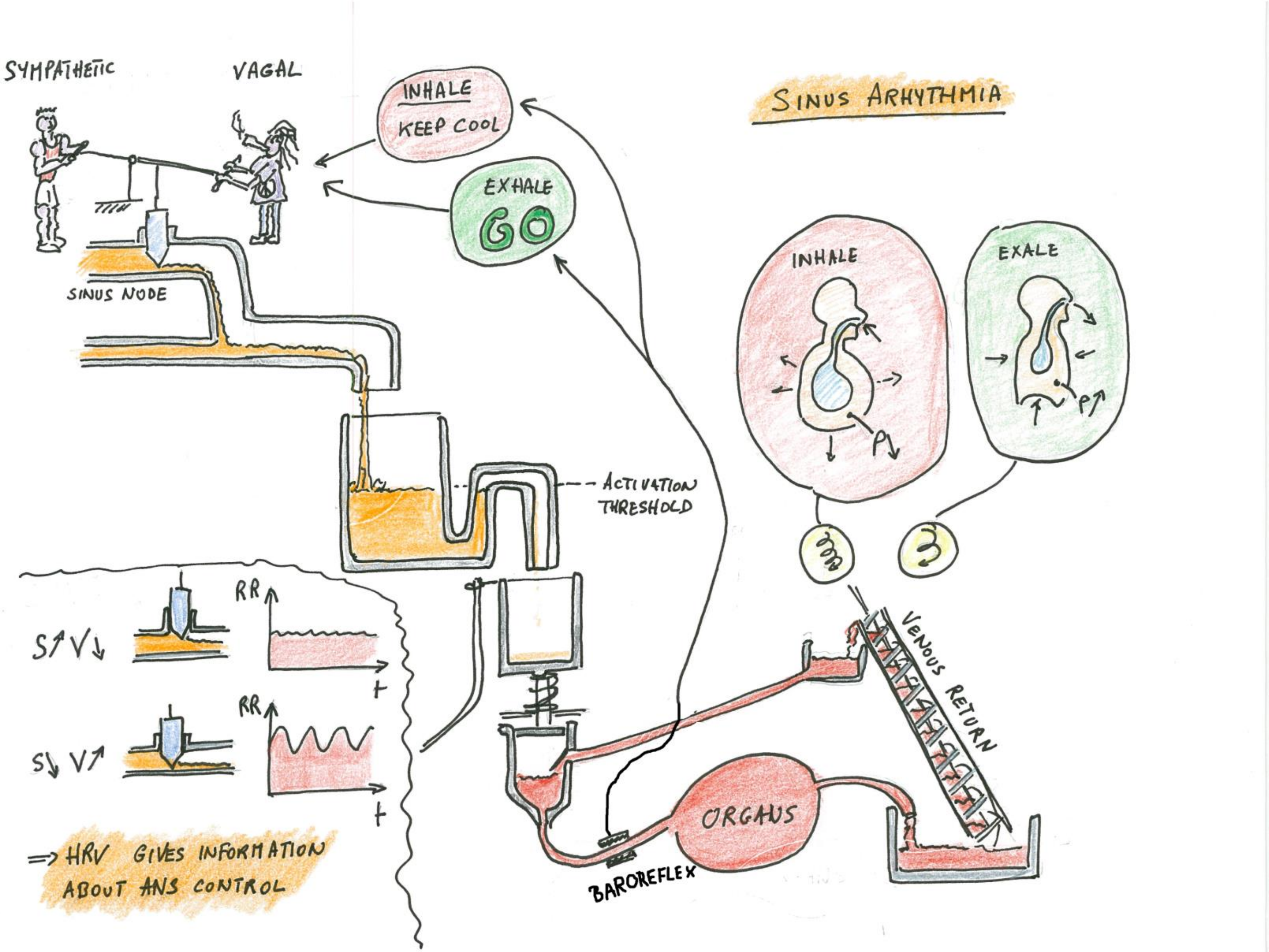
Summary

- Stochastic signals: characteristics between deterministic signals and noise
- The analysis of such signal requires **averaging**
- The autocorrelation permits to obtain a better estimation of the power spectral density than direct FFT
- The selection of the window and the length of the block permits a **trade-off between spectral resolution and oscillations**
- Inter-correlation permits to analyses relationships between different signals

Autonomous nervous system

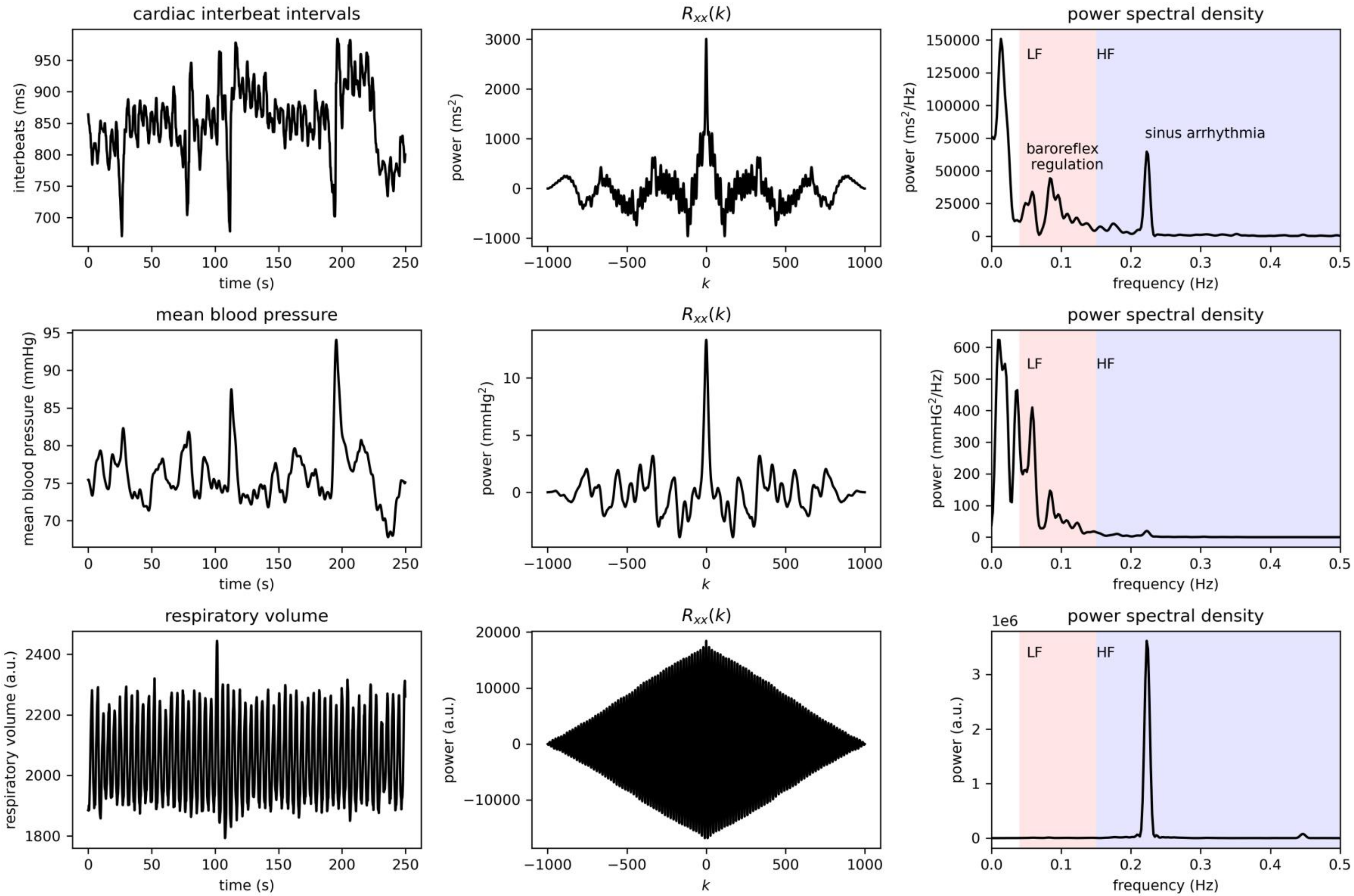


Interaction between breathing and cardiac variability



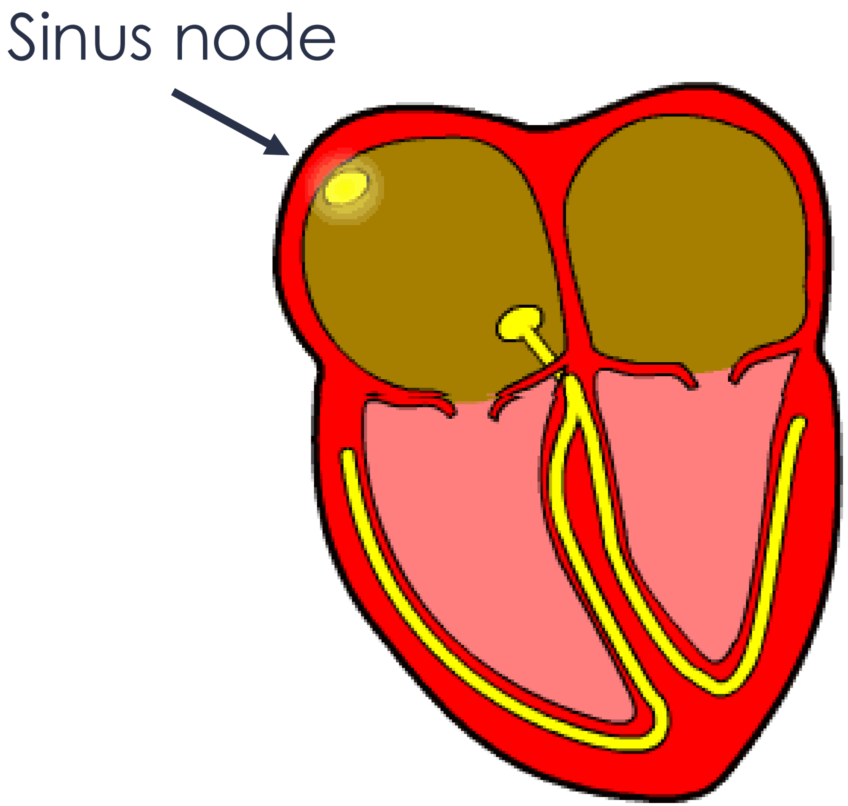
EPFL csem

Power spectral density of ANS control

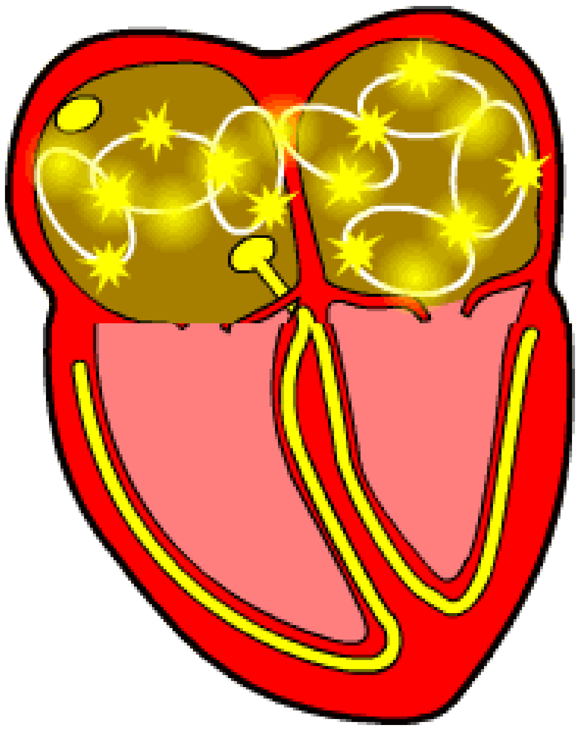


Cardiac arrythmias - Mechanisms

Sinus (normal) rhythm



Atrial fibrillation

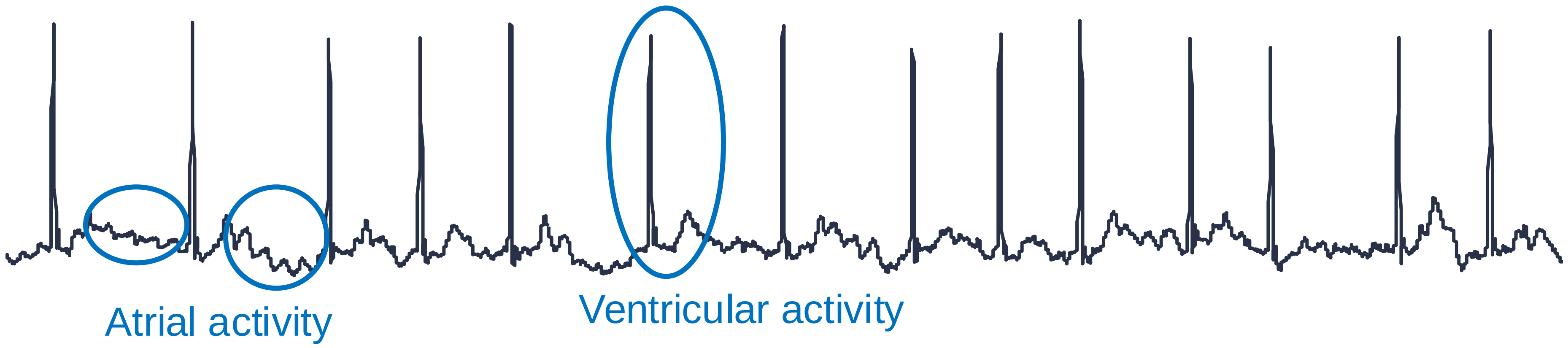


EPFL csem

Electrocardiogram during cardiac arrythmias

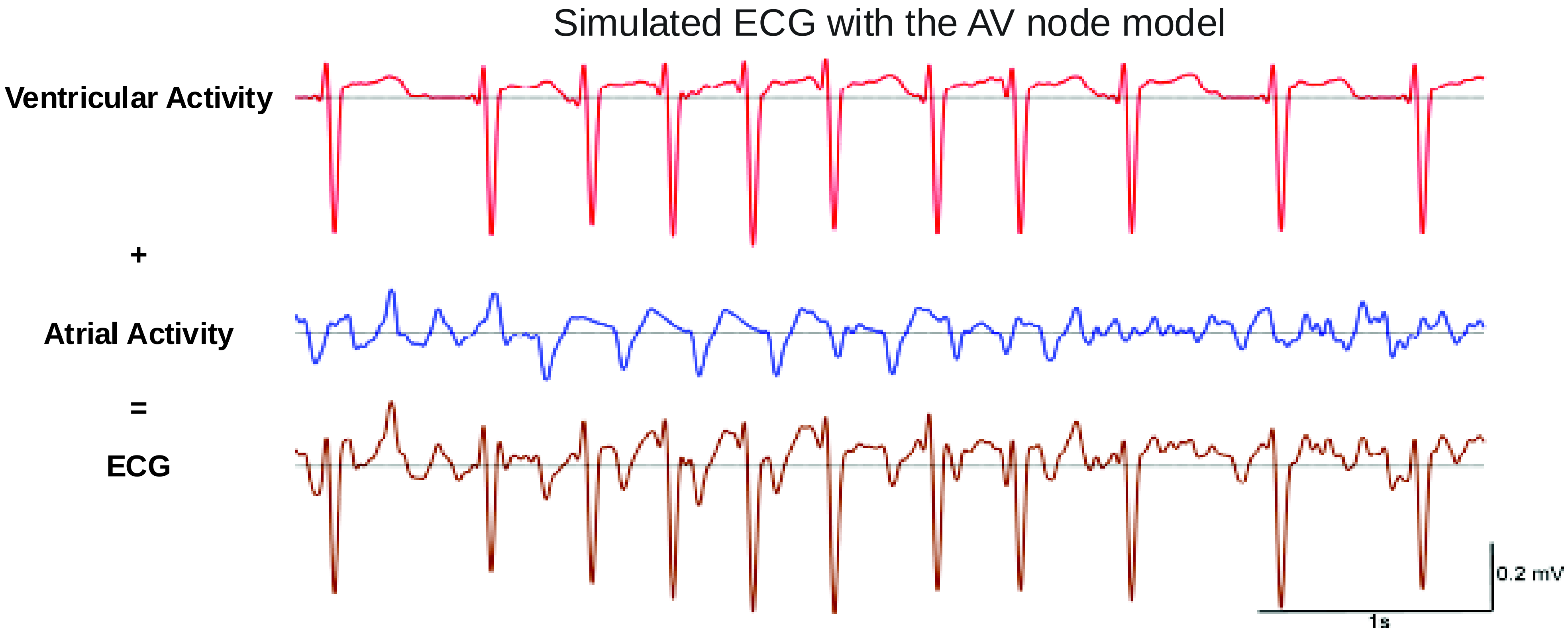
Surface ECG during atrial fibrillation (AF or Afib)

The most common tool used for the **clinical evaluation** of arrythmias



Electrocardiogram - signal processing applications

Simulated 12-lead ECG:



Labo exercises

- 3 exercices
  - m03_ex1_ecg_50_hz.py
  - m03_ex2_ans_control.py
  - m03_atrial_fibrilation.py
- Groups of 3 pax (2 or 4 if mod(num. people,3)≠ 0)
  - One report for the group
    - Names and surnames of group's participants
    - one section per exercise
      - discuss results
      - answer questions
    - naming: **name1_name2_name3_labo_m03.pdf**
    - *optional: at the end of the document free comment about curse and exercises*
- upload the **same report for each person individually** (delay:1 week)

Questions →

Figure →

```python
# import numerical processing library
"""
    The objective of this exercise is that you analyse the code provided and
    make the link with the curse. You have to provide a short report that
    comments and analyse the results. You can use directly the results or adapt
    them to you needs.

"""

# import the numerical library
import numpy as np
# import signal processing library
import scipy.signal as sp
# import ploting library
import pylab as py
py.ion()
py.close('all')

# load the ecg signal
x = np.genfromtxt('respiration.dat')
# sampling frequency of the signal is 500 Hz
fs = 2
# generate correponding time vector
t = np.arange(len(x))/fs

"""
    The signal is a measurement of the breathing obtained by inductance
    plethysmography.

    The objective is to estimate the breathing frequency.
"""

"""
    The Hilbert transforms permits to estimate the instaneous amplitude and
    phase of a narrow band signal.

    Q: Comment the figures.
    Q: Why the envelope does no follow the maxima of the signal

"""

# compute the analytical signal of x (Hilbert transform)
xa = sp.hilbert(x)

# plot the signal
py.figure(1, figsize=[5,5])
py.clf()
py.plot(t, x, label='breathing signal')
py.plot(t, np.abs(xa), label='envelop')
py.xlabel('time (s)')
py.ylabel('amplitude (a.u.)')
py.legend(loc='upper right')
py.title('Breathing signal')

"""
    The raw breathing signal does not fullfil the requirement of narrow band.
    The normal range of frequency for the breathing is within 0.1 to 0.25 Hz.
    The signal is first filtered for this interval.

    Q: Comment the figures
    Q: How is the estimation of the amplitude envelope.

"""


# Analogic limit of the passband frequency
f_pass = np.array([0.1, 0.25])
# Analogic limit of the stopband frequency
f_stop = np.array([0, 0.6])
# Convertion into Nyquist frequency
f_pass_N = f_pass/fs*2
f_stop_N = f_stop/fs*2
# Max attenutation in passband (dB)
```


EPFL csem

Typical exam question

- A white Gaussian noise $x(n)$ of zero mean and unit variance is filtered by the filter $y(n) = x(n) + \frac{1}{2} \cdot x(n-1) + \frac{1}{4} \cdot x(n-2)$. Compute the non-zero values of the autocorrelation of y.

- $R_{yy}(0) = 1^2 + (\frac{1}{2})^2 + (\frac{1}{4})^2 = \frac{21}{16}$
- $R_{yy}(\pm 1) = 1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} = \frac{5}{8}$
- $R_{yy}(\pm 2) = 1 \cdot \frac{1}{4} = \frac{1}{4}$