The Heart Rate Variability under the point of view of statistical physics

Angelo Facchini

Center for the Study of complex Systems,

Dept. of information Engineering,

University of Siena.

a.facchini@unisi.it

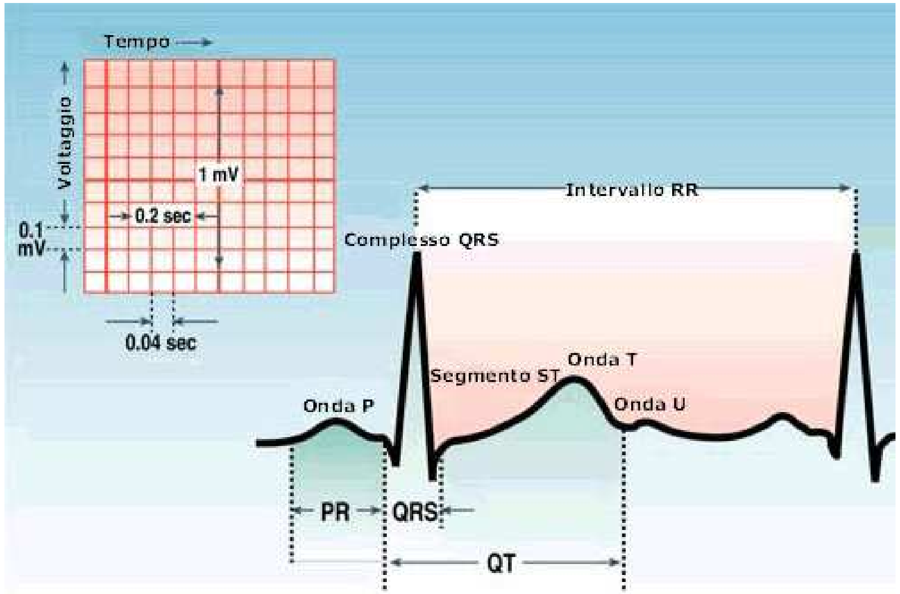
Mny Tnx to Antonio Politi, Roberto Livi, Holger Kantz, Markus Niemann, Friedrich Lenz, Nikolay Vitanov

Outline

- The cardiac dynamics and the RR signal.
- The statistical physics approach.
- Detrended Fluctuation Analysis.
- Multifractal DFA.

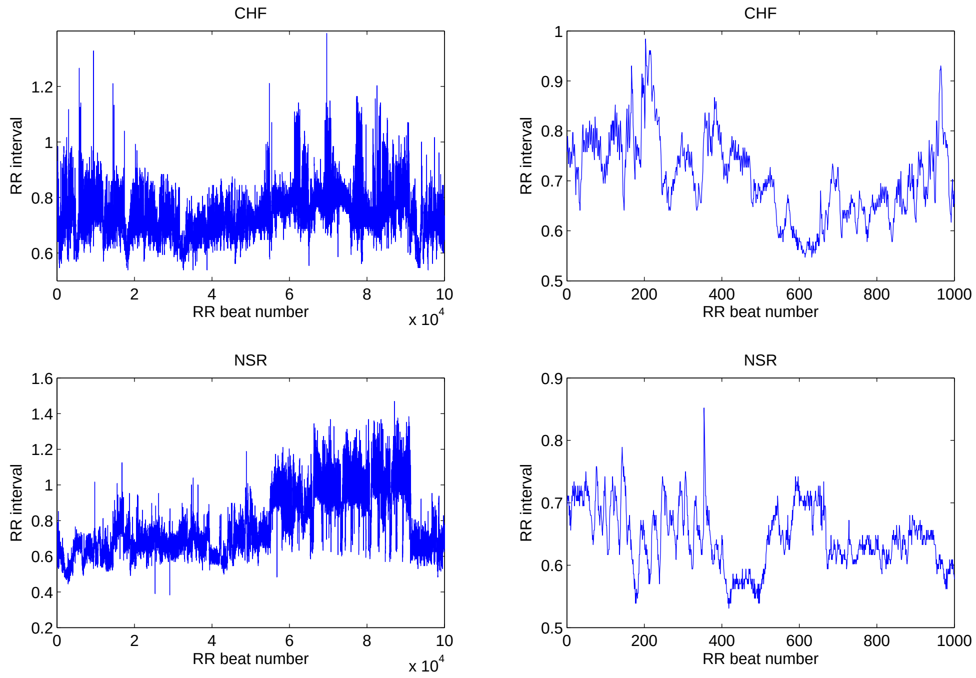
The cardiac dynamics and the RR signal

- P wave: Atrial depolarization.
- QRS Complex: Depolarization of the ventricules.
- ST segment: Connets QRS and T wave.
- T wave: repolarization of the ventricules.
- U wave: Repolarization of papillary muscles.



The RR signal

- Offers a well measurable parameter for the cardiac activity.
- It is defined as the time between two consecutive QRS complexes.



The RR signal: Characteristics

- It is the plot between two consecutive “maxima” of a time series.
- May be considered as a sort of Poincaré section of the system.
- Random oscillations around a mean value.
- Highly non stationary.

The RR as a diffusion process

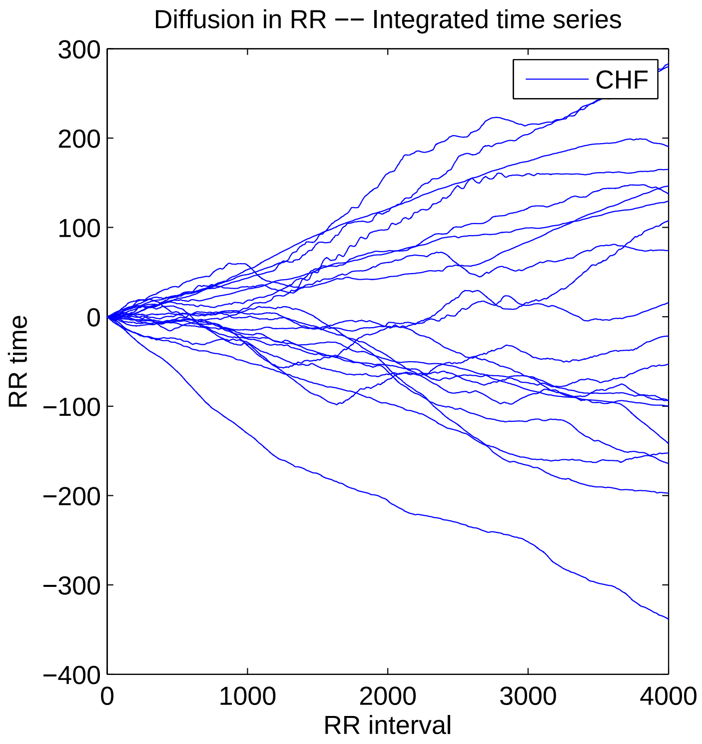
► We use the method of time series analysis derived from statistical physics.

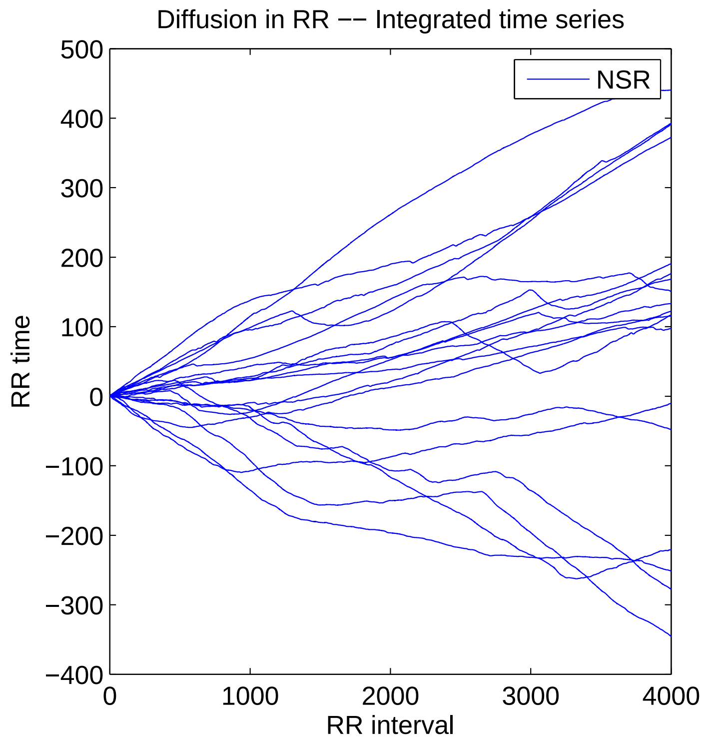
► Study of critical phenomena, where fluctuations at all time scale length appear.

► The RR as a diffusion process.

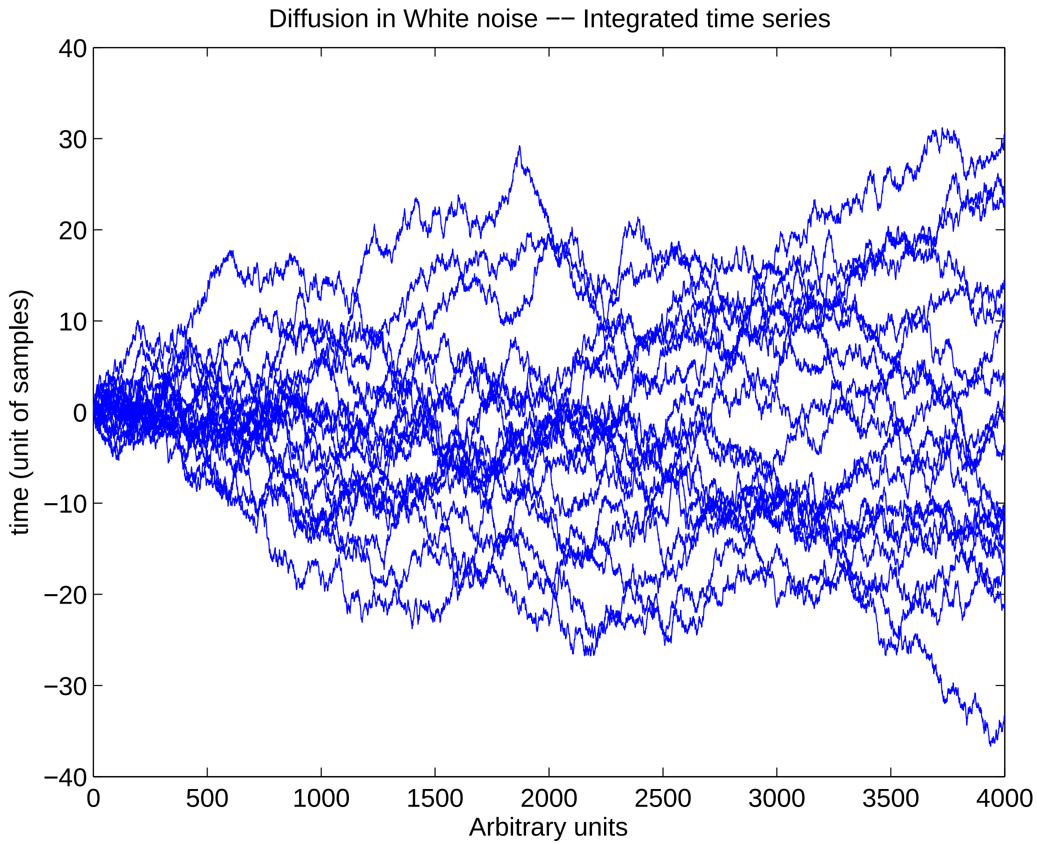
A simple example

▶ We plot rescaled windows of the integrated time series we get a sort of (smooth) random walk.





Let's do it with white noise...



Long-range and short-range correlations

▶ Consider a record x_i, $i = 1, \ldots, N$ of equally spaced measurements.

▶ We are interested in the correlation of x_i and x_{i+s}, for different lags s.

▶ By subtracting the average value $< x >$ we get $\bar{x}_i = x_i - < x >$ and we define the autocorrelation function as:

$$C(s) = \langle \bar{x}_i \bar{x}_{i+s} \rangle = \frac{1}{N-s} \sum_{i=1}^{N-s} \bar{x}_i \bar{x}_{i+s} \tag{1}$$

▶ We speak about *short range correlations* if there exists a time scale s^* for which $C(s)$ decays exponentially, i.e. $C(s) \sim \exp(-s/s^*)$.

▶ For *long-range correlations* $C(s)$ decays as power law: $C(s) \sim s^{-\gamma}$, $\gamma \in (0, 1)$.

Why not?

▶ The computation of the ACF from experimental data is problematic.

▶ Effects of trends of unknown origin, that should be distinguished form the intrinsic fluctuations of the signal.

▶ By example, a moving average with a certain window of width w may introduce a new time scale w and may destroy scaling behaviors happening over times w.

▶ Very often, we do not know the origin of trends in data, and we do not know their scales.

▶ A method for the evaluation of detrended fluctuations over different time scales is needed.

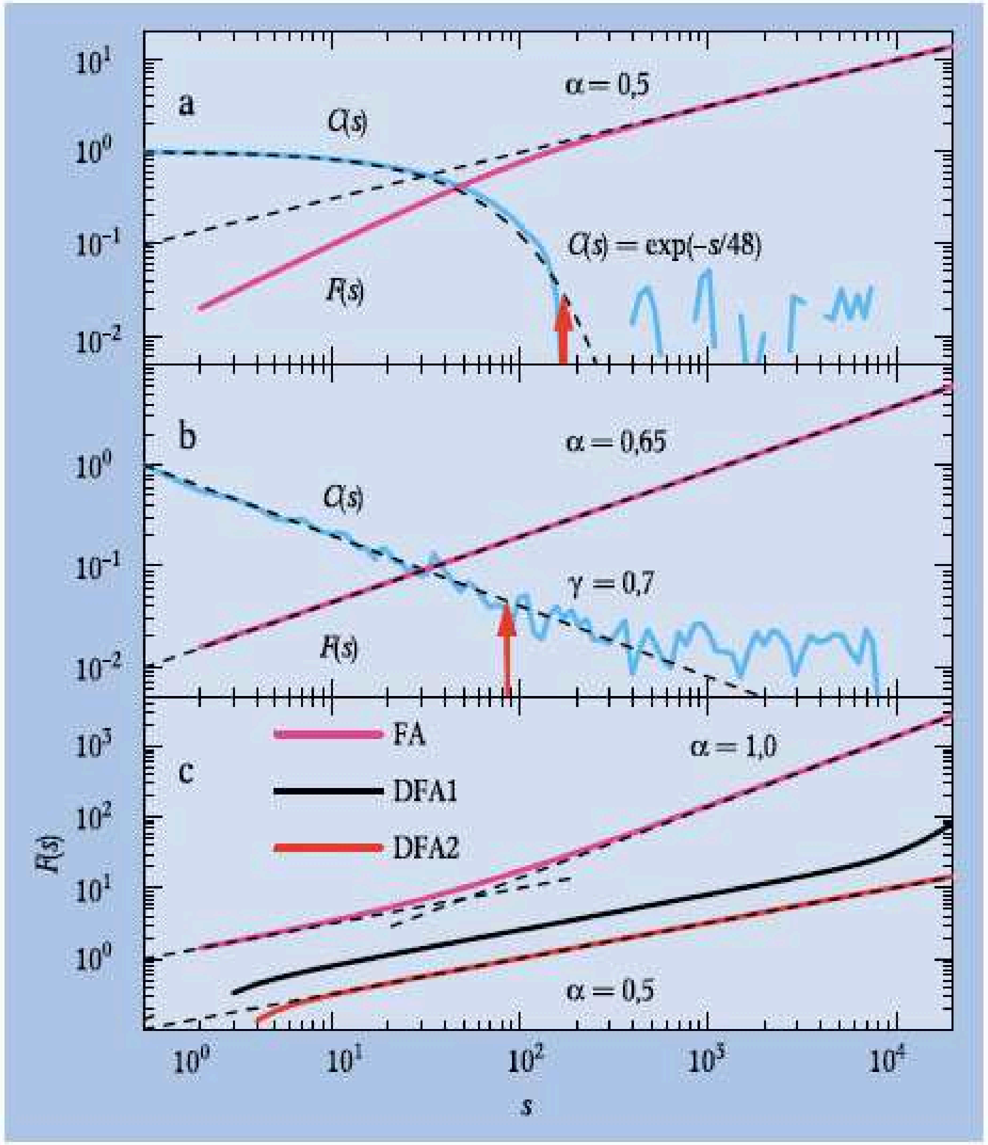


Abb. 2:
Beispiele von Autokorrelationsfunktionen $C(s)$ (blau) und den zugehörigen Fluktuationsfunktionen $F(s)$.
▶ a) Analyse von kurzreichweitig korrelierten Zahlen mit exponentiell abfallendem $C(s)$, die man erhält, indem man die τ_i rekursiv mit einer Wahrscheinlichkeit von 0,98 gleich dem vorhergehenden Wert τ_{i-1} setzt und mit einer Wahrscheinlichkeit von 0,02 als neue unabhängige Zufallszahlen wählt.
▶ b) Analyse von langreichweitig korrelierten Zahlen mit potenzgesetzartig ($\gamma = 0{,}7$) abfallendem $C(s)$, die mit dem in [7] beschriebenen Algorithmus erzeugt wurden.
▶ c) Analyse von unkorrelierten Zahlen, zu denen der Trend $5i/10^6$ addiert ist. $F(s)$ ist über jeweils 500 Reihen aus 10^5 Werten gemittelt, während $C(s)$ jeweils aus nur einer Reihe mit 5×10^4 Werten bestimmt ist. Die roten Pfeile in a) und b) zeigen die kleinsten s an, bei denen das numerische $C(s)$ erstmals negativ wird.

Detrended Fluctuation Analysis: Introduction

► Characterization of long range correlated processes. $C(s) \sim s^{-\gamma}$, $0 < \gamma < 1$.

► Estimation of the power law scaling for non-stationary time series.

► Elimination of trends of different order avoids spurious detections.

► Applications: Cardiac dynamics, temperature recordings, wind time series, econometrics...

DFA: The algorithm

Given a time series $s(i)$ $(i = 1, \ldots, N)$, three steps follow:

1. Integration of $u(i)$ produces the profile function:

$$y(j) = \sum_{i=1}^{j} [u(i) - < u >] \qquad < u > = \frac{1}{N} \sum_{i=1}^{N} u(i) \tag{2}$$

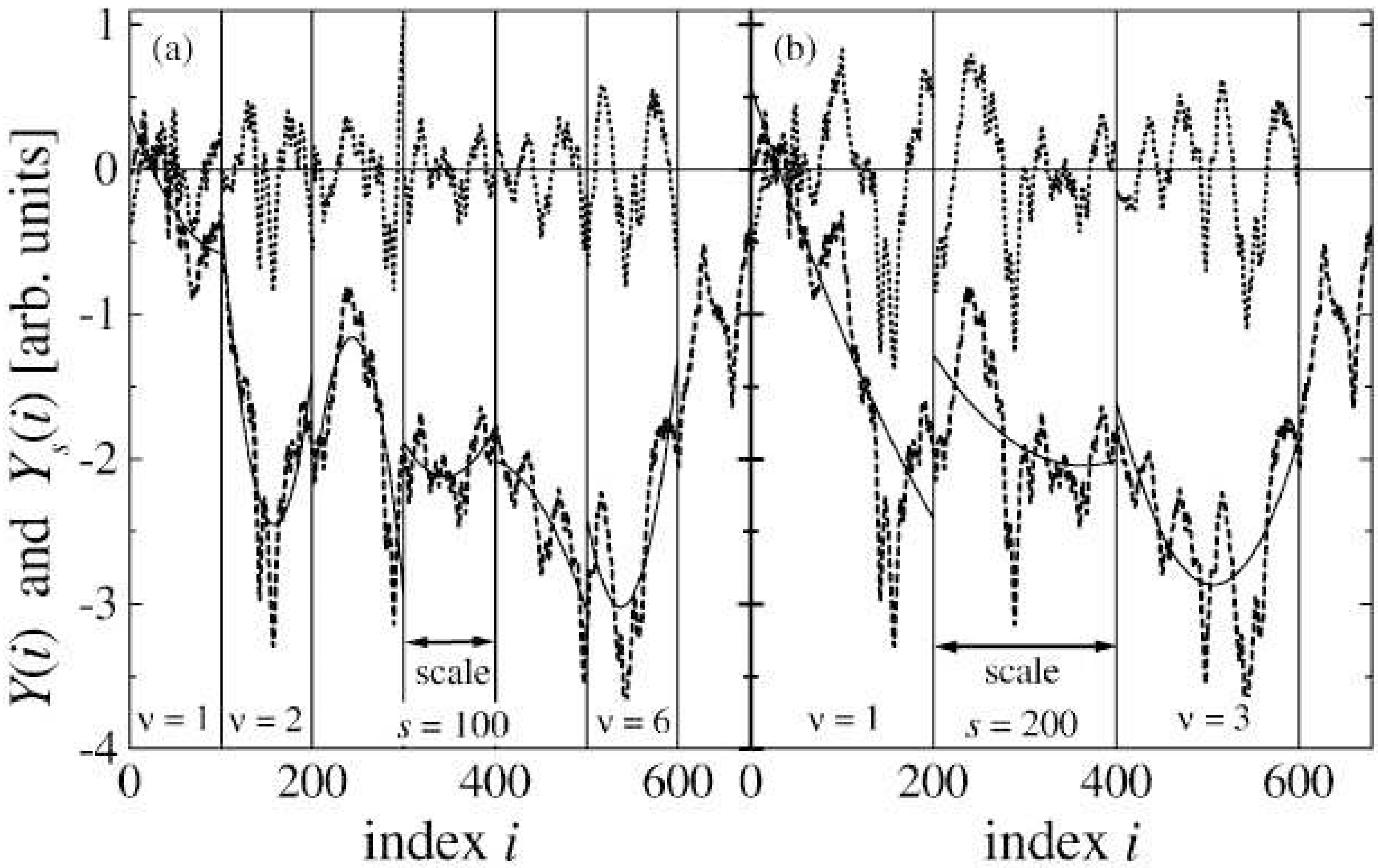
   y is divided in segments of length n

2. In any box, y is fitted by an l order polynomial y^l_{fit}:

$$Y(i) = y(i) - y^l_{fit}(i) \tag{3}$$

3. The fluctuation function is computed:

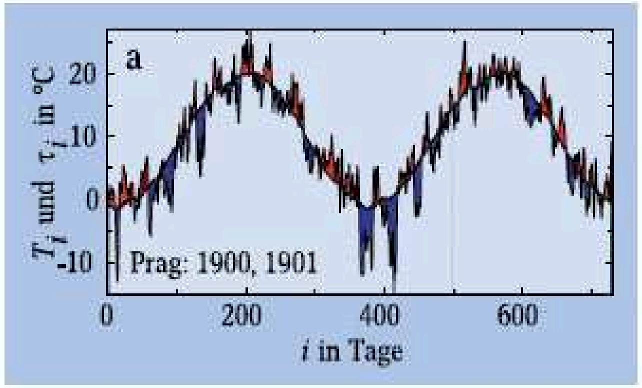
$$F(n) = \left(\frac{1}{N} \sum_{i=1}^{N} Y(i)^2 \right)^{1/2} \tag{4}$$


(a)
(b)
Y(i) and Y_s(i) [arb. units]
scale
scale
ν = 1
ν = 2
s = 100
ν = 6
ν = 1
s = 200
ν = 3
index i
index i
1
0
-1
-2
-3
-4
0
200
400
600
0
200
400
600

- The procedure is repeated for increasing values of n.
- If $F(n) \sim n^H$, H is an estimation of the Hurst exponent.
- The connection to the scaling law of the AC function is given by:

$$H = 1 - \gamma/2, \quad H \in (0.5, 1) \tag{5}$$

$H = 0.5$ is the value of uncorrelated noise, while $H \in (0.5, 1)$ may indicate long range correlations.



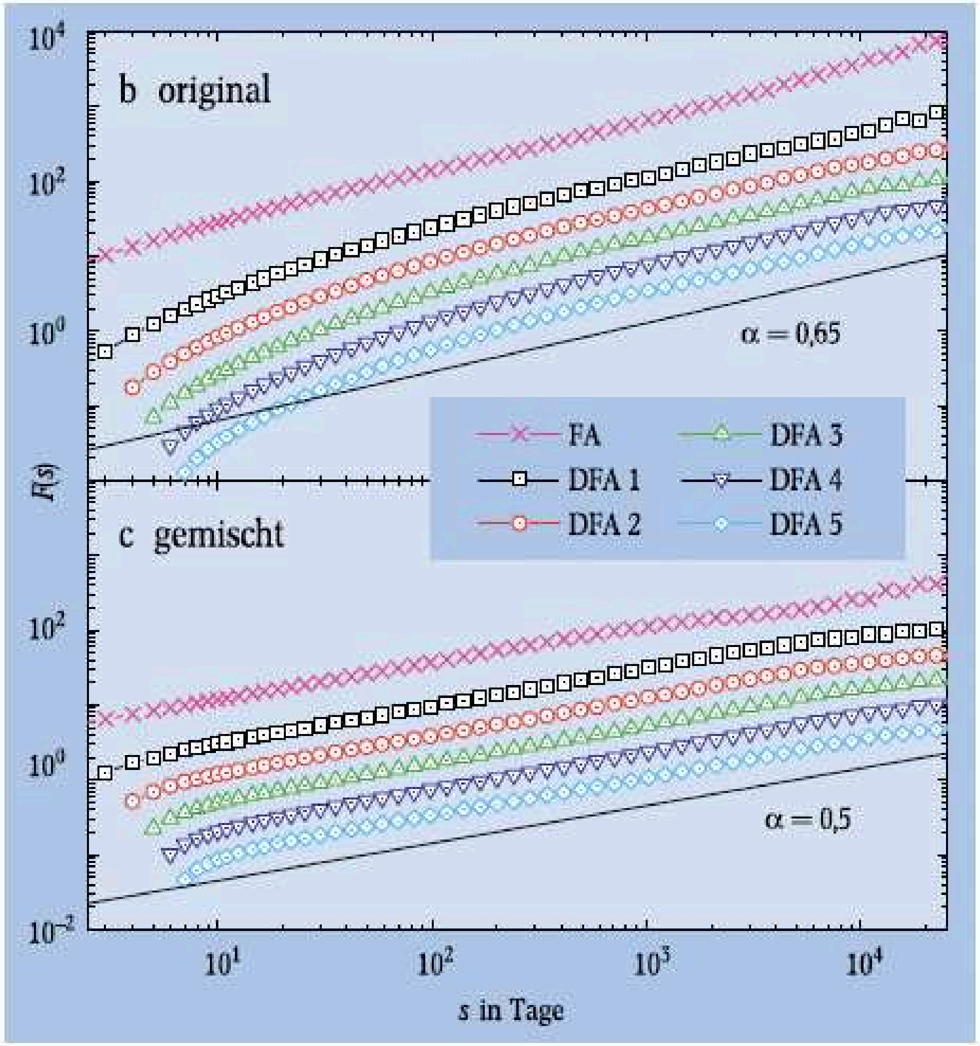


Abb. 1:
Analyse der Prager Temperaturdatenreihe, die aus den gemessenen Tagesmittelwerten von 1775 bis 1992 besteht. Teil (a) zeigt für die Jahre 1900 und 1901 die mittlere Tagestemperatur T_i, wobei die Abweichungen vom Jahresgang rot oder blau eingefärbt sind. Die Teile (b) und (c) zeigen die Fluktuationsfunktionen $F(s)$ für FA (×), DFA1 (□), DFA2 (○), DFA3 (△), DFA4 (▽) und DFA5 (◇) für die Originaldaten bzw. für Daten, die nach der Subtraktion des Jahresganges zufällig gemischt wurden, um die langreichweitigen Korrelationen zu zerstören (in Anlehnung an [3]).

The Hurst Exponent

- Orginally proposed by Hurst, who studied the floods of river Nile.

- Given the time series $x = [x_1, \ldots, x_N]$, one has to study the scaling behavior of the function:

$$S_2 = \langle |x(i + \tau) - x(i)|^2 \rangle_T \sim \tau^{2H} \tag{6}$$

where τ is the time lag and T is an average time such that $T \gg \tau$, and usually is the largest time scale of the system.

- $H \in (0, 1]$.

- H links directly to the fractal dimension via $D = 2 - H$.

- The more H is high, the more x is smooth.

Some history...

- Hurst computed the average yearly flood height and its cumulative deviations.
- Then he computed the maximum value reached by the cumulative deviations and compared it with the minimum, calling R this difference.
- The original formula reads:

$$\log\left(\frac{R}{\sigma}\right) = K \log\left(\frac{N}{2}\right) \quad R = \sigma\left(\frac{N}{2}\right)^{K} \tag{7}$$

- Where N the number of observations (years), σ the standard deviation between time $i-1$ and i. R was the height of the flood.
- He found $K = 0.73$.

DFA: link with the autocorrelation function

► For data without trends and zero offset $\bar{x}_i = x_i$, and $Y_s(i) = Y_i$ for $i \leq s$.

► In each segment ν the mean square displacement can be computed as:

$$\begin{aligned}\left\langle Y^2(i)\right\rangle = \left\langle \sum_{k=1}^{i} x_k^2 \right\rangle + \left\langle \sum_{k\neq j}^{j,k\leq i} x_j x_k \right\rangle = \\ i\left\langle x^2\right\rangle + \sum_{k\neq j}^{j,k\leq i} C(|k-j|) = i\left\langle x^2\right\rangle + 2\sum_{k=1}^{i-1}(i-k)C(k)\end{aligned} \tag{8}$$

▶ For large i, the second term can be approximated:

$$\sum_{k=1}^{i-1} C(k) \sim \sum_{k=1}^{i} k^{-\gamma} \sim \int_1^i k^{-\gamma}\, dk \sim i^{1-\gamma} \quad \text{and} \quad \sum_{k=1}^{i-1} kC(k) \sim i^{2-\gamma} \tag{9}$$

▶ If the data are PL correlated with $\gamma \in (0,1)$, this term dominates, giving

$$\langle Y^2(i)\rangle \sim i^{2-\gamma} \tag{10}$$

▶ Using the same approximation for $F(s)$ gives:

$$F(n) \sim n^{1-\gamma/2} \quad H = 1 - \gamma/2 \tag{11}$$

▶ If the data are short-range or uncorrelated, one gets $\langle Y^2(i)\rangle \sim i$, then $H = 1/2$

▶ Practically, H is extracted by linear fit of $F(n)$ in a log-log plot.

DFA: tips & tricks

Despite the DFA seems to be a powerful method, one should consider:

▶ Numerical stability: The core of the method is the detrending algorithm, usually is a polynomial fit of a certain order.

▶ The least-square routine must be **stable** especially when dealing with large scales (up to 1e5 points). The best I found is the `e02adf` `FORTRAN` routine, provided by NAG.

▶ Strong systematic trends, like linear or periodic, may influence the computation of H, and shuold be removed before the analysis.

▶ The linear fit of the fluctuation function *must be very good*, and the $F(s)$ shoul be always inspected.

▶ Always use different polynomial orders: the analysis is wrong if different values of H are obtained.

Workarounds and pitfalls

► When pre-detrending is not possible (e.g. when the origin of strong trends is unknown) one may differentiate the time series.

► Then $H = H_{diff} + 1$. Sometimes works, sometimes not...

► $C(s) \sim s^{-\gamma} \rightarrow H_{DFA} \in (0, 1)$

► Values of $H > 1$ may occur, and are normally considered as a bias of the method. When obtaining such results...please...*pay attention!*.

► Always use different DFA routines.

Application of DFA on HRV data: Outline

▶ Cardiac pathology: Congestive hearth failure.

▶ Control sample: Healthy subjects with Normal Sinusal Rythm.

▶ The DFA is applied to study the scaling behavior of RR time series, both on large and small scales.

▶ The data come from the Physionet Bank[a].

[a] http://www.physionet.org

Correlated regions in HRV during sleep

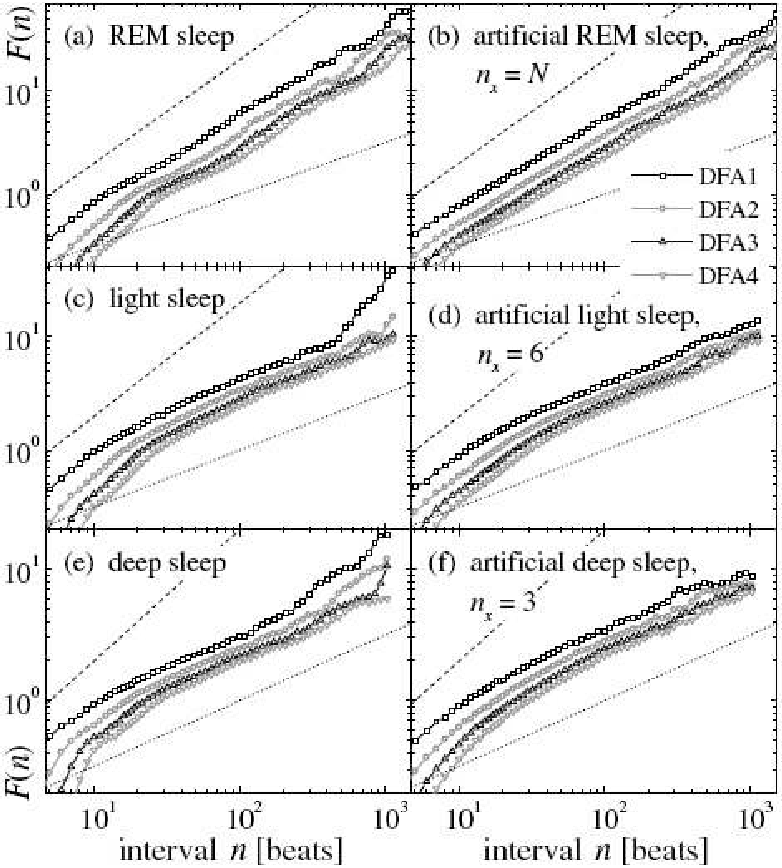


FIG. 3. The fluctuation function $F(n)$ for (a) REM sleep, (c) light sleep, and (e) deep sleep obtained from DFA1 (squares), DFA2 (grey circles), DFA3 (triangles up), DFA4 (grey triangles down) compared with results from artificial data (b),(d),(f). The artificial data consist of correlated random numbers with $\alpha = 0.85$ below $n_\times$. Above $n_\times$ the data are uncorrelated. The lengths and numbers of the control sequences are identical with the lengths and numbers of the sleep stages in the real recordings. The best $n_\times$ values representing the real data are $n_\times = N$ for REM sleep (b), $n_\times = 6$ for light sleep (d), and $n_\times = 3$ for deep sleep (f). To allow for a comparison of the data, the standard deviation of the τ_i values was set equal to 1 in all cases.

Correlated regions in HRV during sleep

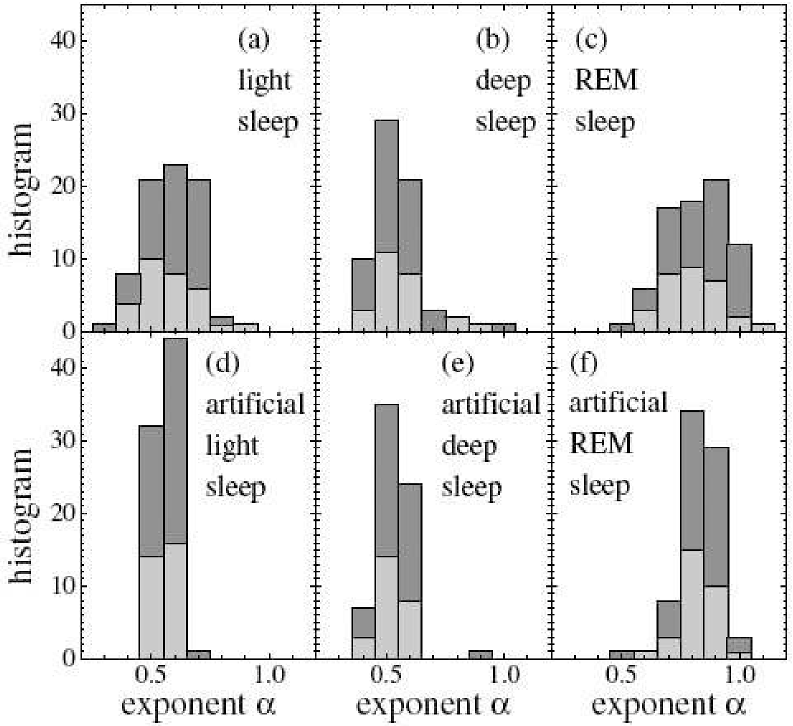
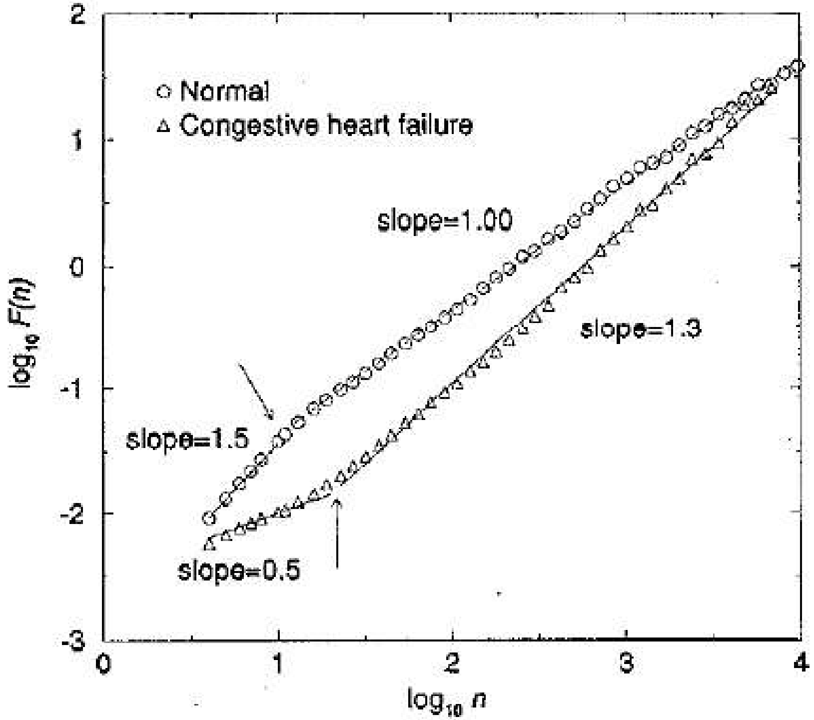


FIG. 4. (a)-(c) Histograms of the fluctuation exponents α obtained from linear fits to log-log plots of $F(n)$ versus n in the regime $70 < n < 300$ for (a) light sleep, (b) deep sleep, and (c) REM sleep. The fitting range has been chosen to be above the regime of short-range correlations related to breathing and below the n values where the statistical errors become too large due to the finite length of the sleep stages. The data in (a) are based on all 30 records from healthy subjects (light grey) and on all 47 records from patients with moderate sleep apnea (dark grey). In (b) ten records have been dropped since they where too short while in (c) only one record was too short and has been dropped. (d)–(f) The corresponding results for artificial control data sets. The artificial data for the sleep stages have been generated in the way described below Fig. 3 and in the text.

Crossover phenomenon

► the fluctuation function show a systematic cross-over. Small scales show an angle α_1, while larger scale show α_2.



► By building an (α_1, α_2) diagram it is possible to distinguish NSR and CHF subjects.

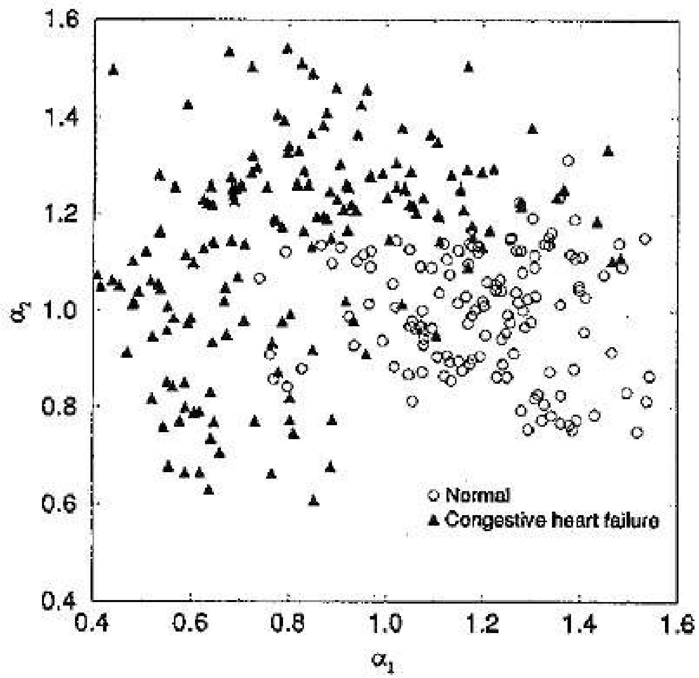


FIG. 5. Scatter plot of scaling exponents α_1 vs α_2 for the healthy subjects (○) and subjects with congestive heart failure (△). The α's were calculated from interbeat interval data sets of length 8192 beats. Longer data set records were divided into multiple data sets (each with 8192 beats). Note good separation between healthy and heart disease subjects, with clustering of points in two distinct "clouds."

Multifractal DFA

▶ Extension of the DFA, for $q \in \mathbb{R}$ one gets several Fluctuation functions:

$$F_q(n) = \left(\left\langle \frac{1}{N} \sum_{i=1}^{N} Y(i)^2 \right\rangle_\nu^{q/2} \right)^{1/q} \tag{12}$$

$<>_\nu$ is the average over the $\nu = N/n$ segments.

▶ (Very) Roughly speaking, it is a generalization of equation 6:

$$S_q = \langle |x(i+\tau) - x(i)|^q \rangle_T \sim \tau^{qH(q)} \tag{13}$$

▶ $F_q(n) \sim n^{h(q)}$. For $q = 2$, $H = h(2)$.

▶ We speak about a spectrum of Hurst exponents $H(q)$.

▶ $f(\alpha)$ has a parabolic form.

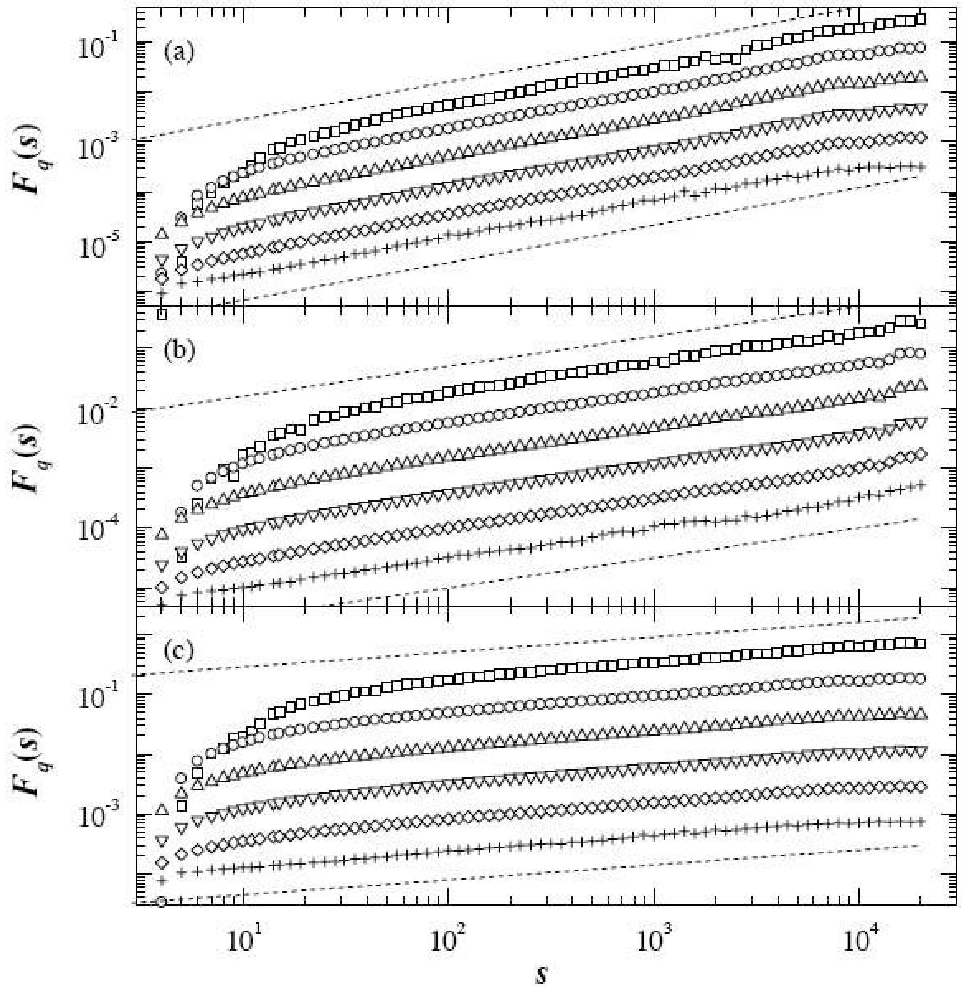
▶ A method for characterizing a multifractal set is the *singularity spectrum* $f(\alpha)$, related to $h(q)$ via a Legendre transform:

$$\alpha = H(q) + qH'(q) \quad f(\alpha) = q[\alpha - H(q)] + 1 \tag{14}$$

▶ The maximum of $f(\alpha)$ shows at which α is positioned the most statistically significant part of the time series, i.e. the subsets with maximum fractal dimension.

MF-DFA: application on FGN

- When analyzing white noise or Fractional Gaussian Noise, the value of $H(q) = H(2)\ \forall q$.
- The fluctuation functions have the same slope.



MF-DFA: application on the binomial MF series

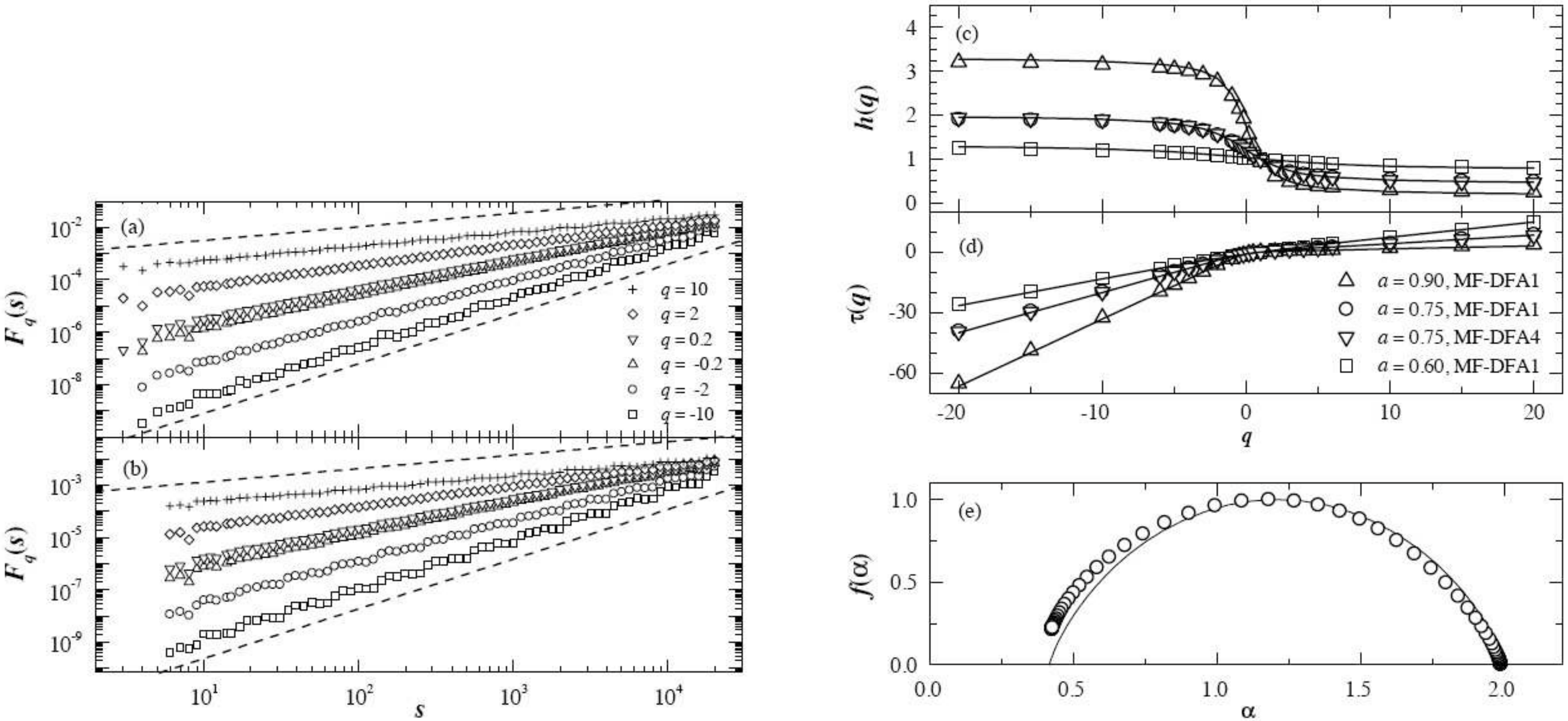
- The series is defined as:

$$x_k = a^{n(k-1)}(1-a)^{n_{max}-n(k-1)} \tag{15}$$

with $a \in (0.5, 1)$ and $n(k)$ is the number of ones in the binary representation of k

▶ The slope of the fluctuation functions increases with q.

▶ The spectrum of the Hurst exponents is monotonous decreasing curve.



MF-DFA: Tips, triks, workarounds and pitfalls

The same as for DFA...furthermore:

▶ For negative q the algorithm is more unstable, then **always** look at the fluctuation functions.

▶ Repeat the analysis by shuffling the time series: it should be $\alpha^* = 0.5$ and $W_\alpha \sim 0.05$.

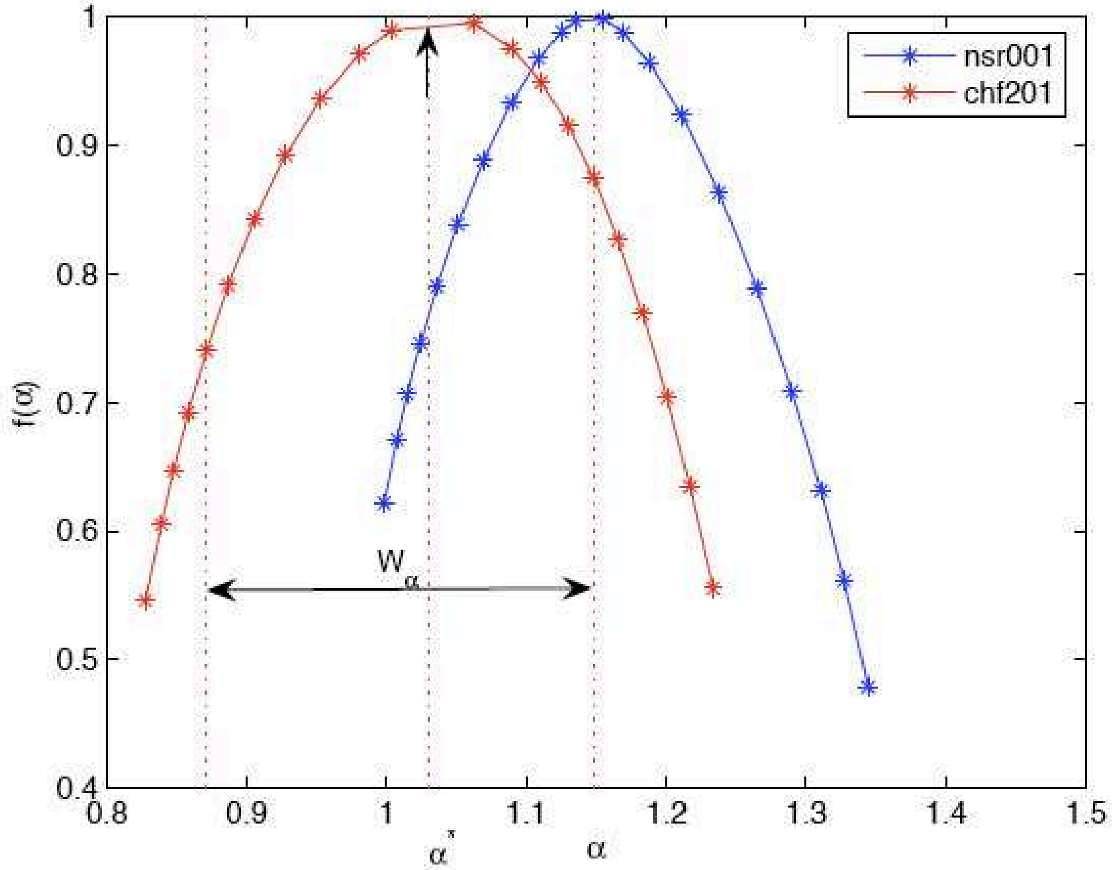
▶ Beware from spurious MF due to finite size. For a good analysis one should use at least 5e5 points.

▶ Less points are possible, but the fit of the $F_q(s)$ **must** be very good.

MF-DFA for the analysis of HRV data

- ▶ The RR time series show different self-similarity at different scales.
- ▶ Characterization of the $f(\alpha)$ spectra using two indicators:
  - α^*: the value of α for which the spectrum has its maximum. It indicates the subset with maximum fractal dimension among all subsets of the time series.
  - W_α: The width of $f(\alpha)$ computed in the range $q \in [-3, 3]$.
  - For monofractal signals, the $f(\alpha)$ spectrum is singular, but, because of the finite size a residual MF is always found.

Characterization of HRV using α^* and W_α



References

[1] J. Kantelhardt et al., Detecting long range correlations with Detrended Fluctuation Analysis, Physica A, 295, pp. 441-454, 2001.

[2] N.K. Vitanov, K. Sakai, E.D. Yankulova, On correlations and fractal characteristics of time series, J. Theor. Appl. Mechanics, 35, pp. 73-90, 2005.

[3] K. Hu, P.C. Ivanov, Z. Chen, P. Carpena H.E. Stanley, Effects of trends on Detrended Fluctuation Analysis, Phys. Rev. E, 64, 011114, 2001.

[4] Z. Chen, P.C. Ivanov, K. Hu, H.E. Stanley, Effect of nonstationarities on Detrended Fluctuation Analysis, Phys. Rev. E, 65, 41107, 2002.

[5] P. Absil, R. Sepulchre, A. Bilge, P. Gérard. Nonlinear analysis of cardiac rhythm fluctuations using DFA method, Physica A, 272, pp. 235-244, 1999.

[6] A. Bunde, J. Kantelhardt, Langzeitkorrelationen in der Natur: von Klima, Erbgut und Herzrhytmus, Physikalische Blätter, 57, 2001.

[7] A. Bunde et al., Correlated and uncorrelated regions in heart rate fluctuations during sleep, Phys. Rev. Lett., 17, 2000.

[8] C.K. Peng, S. Havlin H.E. Stanley, A. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series, Chaos, 5, pp. 82-87, 1995.

[9] C.K. Peng et al., Long-Range Anticorrelations and non-gaussian behavior of the heartbeat, Phys. Rev. Lett., 70, 1343, 1999.

[10] J.W. Kantelhardt et al., Multifractal Detrended Fluctuation

Analysis, Physica A, 316, pp. 87-114, 2002.

[11] H.E. Stanley et al., Statistical physics and physiology: monofractal and multifractal approaches, Physica A, 270, pp. 309-324, 1999.

[12] D. Makoviec et al., Long-range dependencies in heart rate signals – Revisited, Physica A, 369, pp. 632-644,2006.