

# The fluctuation function of the detrended fluctuation analysis - Investigation on the AR(1) process

Marc Höll

Max-Planck-Institut für Physik komplexer Systeme, Dresden

02-04-2014



MAX-PLANCK-GESELLSCHAFT

# Table of contents

- 1 Range of Correlations
- 2 Detrended Fluctuation Analysis
- 3 Fluctuation function of the AR(1) process
- 4 Summary

# Range of correlations

- Stationary time series:  $x_i$  with  $i = 1, \dots, N$
- Autocorrelation:

$$C(s) = \frac{\text{Cov}(x_i, x_{i+s})}{\text{Cov}(x_i, x_i)} = \frac{\langle (x_i - \bar{x})(x_{i+s} - \bar{x}) \rangle}{\langle (x_i - \bar{x})^2 \rangle}$$

- Different classes of autocorrelation structures can be distinguished with respect to the form of their **decay** for large time lags  $s$ :
  - short-range correlations:
    - exponential decay

$$C(s) \sim e^{-\frac{s}{s_c}}$$

- long-range correlations:
    - algebraical decay

$$C(s) \sim s^{-\gamma} \text{ with } 0 < \gamma < 1$$

- Problem: Given data set  $\Rightarrow$  Determine  $C(s)$  directly?
- Consider a stationary time series with a non-stationary trend
- Example: AR(1) process with linear trend

$$x_i = \underbrace{\sum_{k=2}^N a^{i-k} \eta_k}_{\text{AR(1) process}} + \underbrace{bi}_{\text{trend}}$$

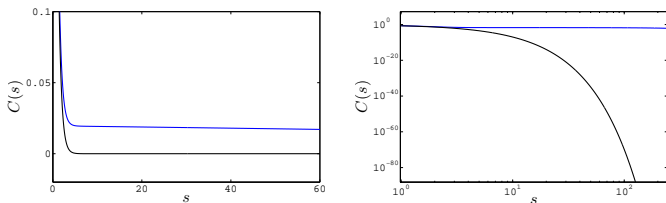


Figure: Stationary process (black) and with trend  $b = 5 \cdot 10^{-4}$  (blue)

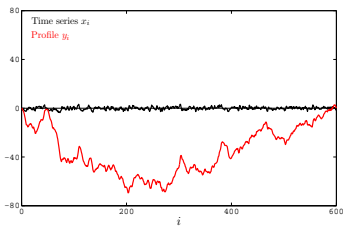
- Trend gives impression of long-range correlations and hides the true nature of the noise

# Detrended Fluctuation Analysis

- Estimating the autocorrelation function  $C(s)$  from empirical data is effected by non-stationarities like trends and observational noise
- In 1994 Peng et al.<sup>1</sup> suggested **Detrended Fluctuation Analysis (DFA)** to indirectly gain information about the correlation structure imposed on a time series
- It has been applied to such diverse fields of interest as
  - DNA
  - human gait
  - heart rate dynamics
  - long-time weather records
  - economical time series
  - ...

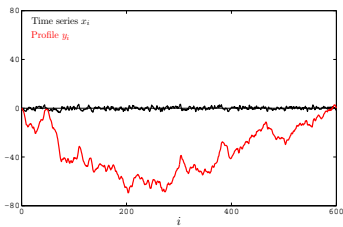
---

<sup>1</sup>C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Phys. Rev. E 49 (1994) 1685



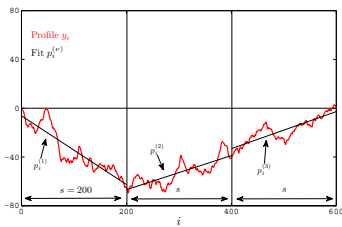
(1) Profile of the time series

$$y_i = \sum_{k=1}^i (x_k - \langle x \rangle)$$



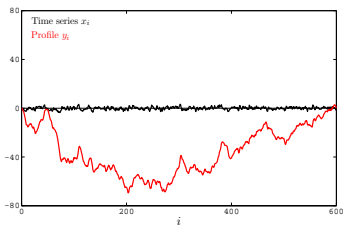
(1) Profile of the time series

$$y_i = \sum_{k=1}^i (x_k - \langle x \rangle)$$



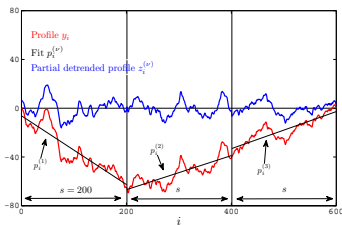
(2) Partial detrended profile

$$z_i^{(\nu)} = y_i - p_i^{(\nu)}$$



(1) Profile of the time series

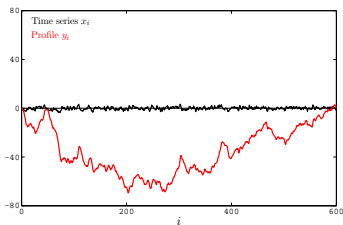
$$y_i = \sum_{k=1}^i (x_k - \langle x \rangle)$$



(2) Partial detrended profile

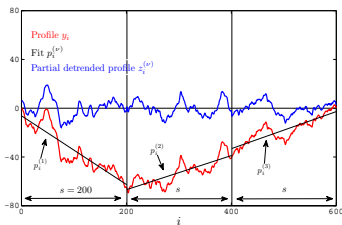
$$z_i^{(\nu)} = y_i - p_i^{(\nu)}$$





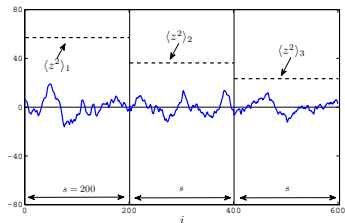
(1) Profile of the time series

$$y_i = \sum_{k=1}^i (x_k - \langle x \rangle)$$



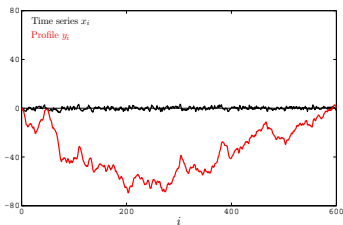
(2) Partial detrended profile

$$z_i^{(\nu)} = y_i - p_i^{(\nu)}$$



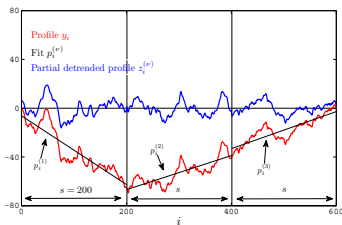
(3) Variance in segment  $\nu$

$$\langle z^2 \rangle_\nu = \frac{1}{s} \sum_\nu \left( z_i^{(\nu)} \right)^2$$



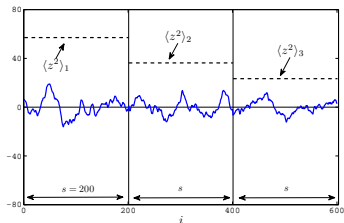
(1) Profile of the time series  

$$y_i = \sum_{k=1}^i (x_k - \langle x \rangle)$$



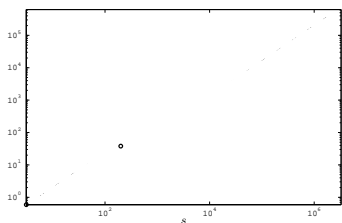
(2) Partial detrended profile  

$$z_i^{(\nu)} = y_i - p_i^{(\nu)}$$



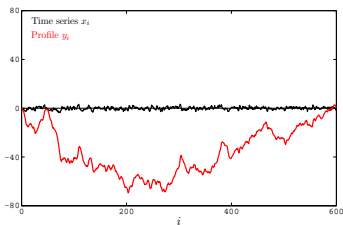
(3) Variance in segment  $\nu$   

$$\langle z^2 \rangle_\nu = \frac{1}{s} \sum_\nu (z_i^{(\nu)})^2$$



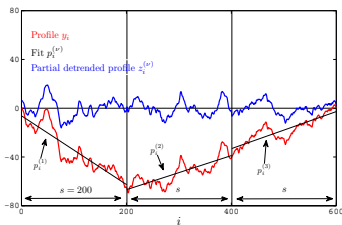
(4) Fluctuation function  

$$F(s) = \frac{1}{K} \sum_{\nu=1}^K \langle z^2 \rangle_\nu$$



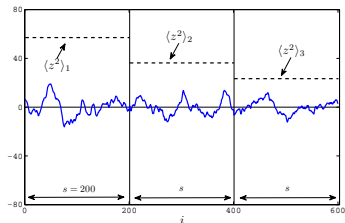
(1) Profile of the time series  

$$y_i = \sum_{k=1}^i (x_k - \langle x \rangle)$$



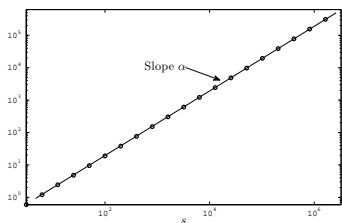
(2) Partial detrended profile  

$$z_i^{(\nu)} = y_i - p_i^{(\nu)}$$



(3) Variance in segment  $\nu$   

$$\langle z^2 \rangle_\nu = \frac{1}{s} \sum_\nu \left( z_i^{(\nu)} \right)^2$$



(4) Fluctuation function  

$$F(s) = \frac{1}{K} \sum_{\nu=1}^K \langle z^2 \rangle_\nu$$

- The fluctuation function increases according to a power law

$$F(s) \sim s^\alpha$$

- Relationship between the fluctuation function  $F(s)$  and the autocorrelation  $C(s)$ :

- (1) Short-range correlations

$$\alpha = \frac{1}{2}$$

- (2) Long-range correlations ( $C(s) \sim s^{-\gamma}$  with  $0 < \gamma < 1$ )

$$\alpha = 1 - \frac{\gamma}{2} \text{ with } \frac{1}{2} < \alpha < 1$$

- (3) Other cases

$$\alpha = 1 \Rightarrow \frac{1}{f} - \text{noise}$$

$$\alpha = \frac{3}{2} \Rightarrow \text{Brownian motion}$$

- Problem: There exists a proof of these relationship only for the case of no trends  $p_i^{(\nu)} = 0$  (Fluctuation Analysis FA)
- With trends (DFA): no proof
  - Testing with artificial generated time series<sup>1</sup>
  - No analytical expressions for the fluctuation function, only one for the asymptotic behaviour:  
Fractional Gaussian Noise <sup>2</sup> ( $H = \frac{1}{2} \rightarrow$  Brownian motion)

$$F(s) \sim \left( \frac{2}{2H+1} + \frac{1}{H+2} - \frac{2}{H+1} \right) s^{2H} \text{ with } 0 < H < 1$$

$$\alpha = H + 1$$

---

<sup>1</sup>J.W. Kantelhardt, E. Koscielny-Bunde, H.H.A. Rego, S. Havlin, A. Bunde, Physica A 295 (2001) 441

<sup>2</sup>M.S. Taqqu, V. Teverovsky, W. Willinger, Fractals 3 (1995) 785

- No proof of the DFA
- Are there any known problems?
- For instance Maraun et al.<sup>1</sup> applied the DFA on AR(1) processes and have shown that the fluctuation function  $F(s)$  suggests long-range correlations
- The reason for this discrepancy is the undersized length of the time series

---

<sup>1</sup>D. Maraun, H.W. Rust, J. Timmer, Nonlinear Processes in Geophysics (2004)  
11:495-503

# Fluctuation function of the AR(1) process

- First order autoregressive process AR(1)

$$x_i = ax_{i-1} + \eta_i \text{ with } -1 < a < 1 \text{ and } \eta_i \sim \mathcal{N}(0, 1)$$

- Autocorrelation:

$$C(s) = a^s = e^{-\frac{s}{s_c}}$$

with characteristic correlation time  $s_c = -\frac{1}{\ln a}$

- DFA assumes a fluctuation parameter  $\alpha = \frac{1}{2}$
- Other cases
  - $a = 0 \rightarrow$  Gaussian White noise (GWN)  $\Rightarrow \alpha = \frac{1}{2}$
  - $a = 1 \rightarrow$  Random Walk (RW)  $\Rightarrow \alpha = \frac{3}{2}$

- Profile

$$\begin{aligned}
 y_i &= \sum_{k=1}^i x_k \\
 &= \sum_{k=1}^i \eta_k \frac{1 - a^{1+i-k}}{1 - a}
 \end{aligned}$$

- Partial detrended profile

$$\begin{aligned}
 z_i^{(\nu)} &= y_i - (A + Bi) \\
 &= \sum_{k=1}^s \eta_k \left( \frac{1 - a^{1+i-k}}{1 - a} \Theta(i - k) + B_{k,i} \right)
 \end{aligned}$$

$$\text{with } B_{k,i} = \frac{4s+2-6i}{-s^2+s} \sum_{j=k}^s \frac{1-a^{1+j-k}}{1-a} + 6 \frac{s+1-2i}{s^3-s} \sum_{j=k}^s j \frac{1-a^{1+j-k}}{1-a}$$



## Fluctuation function of the AR(1) process

$$F^2(s) = \frac{1}{15(a-1)^7(a+1)^2} \frac{a^{2s}G_2(s)+a^2G_1(s)+G_0(s)}{s^4-s^2}$$

with the polynomials

$$G_2(s) = g_{23}s^3 + ga_{22}s^2 + g_{21}s + g_{20}$$

$$G_1(s) = g_{12}s^2 + g_{11}s + g_{10}$$

$$G_0(s) = g_{05}s^5 + g_{04}s^4 + g_{03}s^3 + g_{02}s^2 + g_{01}s + g_{00}$$

and the coefficients

$$g_{23} = 15a^4(a-1)^3$$

$$g_{22} = -60a^4(a-1)^2(a+1)$$

$$g_{21} = 15a^4(5a^3 + 9a^2 - 9a - 5)$$

$$g_{20} = -30a^4(a^3 + 5a^2 + 5a + 1)$$

$$g_{12} = -60a^2(a-1)^3(a+1)^2$$

$$g_{11} = \dots$$

- Behaviour for large  $s$

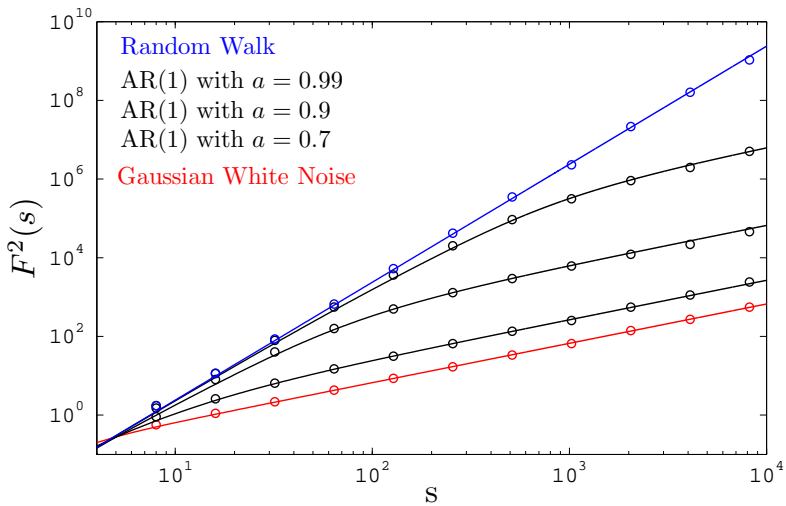
$$F^2(s) \sim \frac{1}{15(a-1)^2} s$$

- Fluctuation function of the Gaussian White Noise

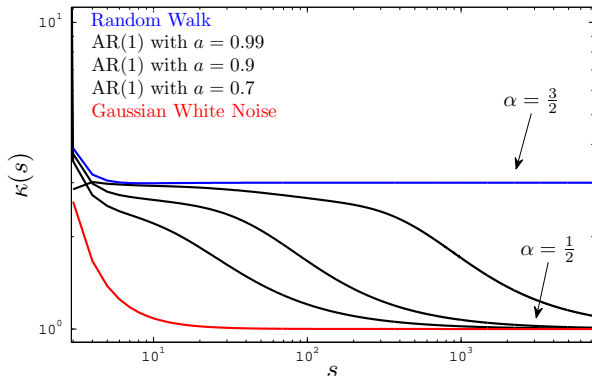
$$F^2(s) = \frac{s^2 - 14}{15s} \sim \frac{1}{15} s$$

- Fluctuation function of the Random Walk

$$F^2(s) = \frac{s^4 + s^2 - 20}{420s} \sim \frac{1}{420} s^3$$



- Slope of  $F^2(s)$  in the log-log plot  $\kappa(s) = \frac{d \log F^2(s)}{d \log s} = \frac{s}{F^2(s)} \frac{dF^2(s)}{ds}$



- We observe for the AR(1) process:
  - (1) Region with  $\alpha = \frac{1}{2}$  after crossover  $s_x$
  - (2) There exists a scaling region with  $\alpha > \frac{1}{2}$  before  $s_x$

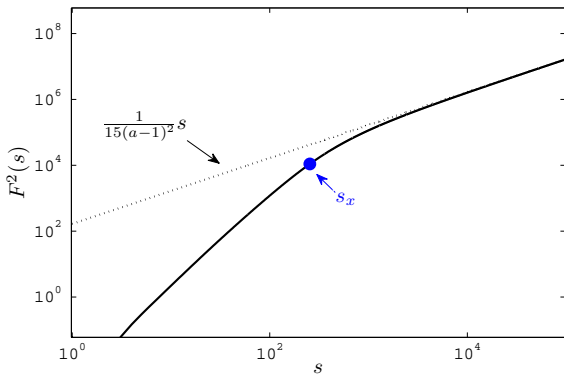
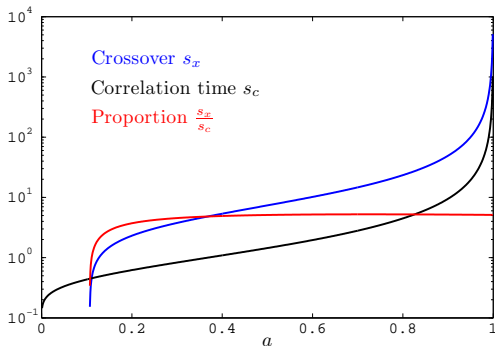


Figure:  $F^2(s)$  and its asymptotic behaviour with  $a = 0.98$

Crossover point:

$$s_x = \frac{\sqrt{-19a^4 + 30a^3 + 68a^2 + 30a - 4}}{-a^2 + 1}$$

Comparison between the crossover  $s_x$  and the characteristic correlation time  $s_c$  (remember  $C(s) = a^s = e^{-\frac{s}{s_c}}$ ):



- Rule of thumb:  $s < N/10$
- Proportion  $s_x/s_c \approx 5$
- DFA is data consuming and needs about 5 times more data

- Analytical expression for the fluctuation function of the AR(1) process
- Region with  $\alpha > 1/2$  (long-range) and  $\alpha = 1/2$  (short-range)
- Crossover point  $s_x$
- DFA is data consuming