Detrended fluctuation analysis revisited

Marc Höll[1], Yu Zhou[2], Holger Kantz[1]

1 Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, 01187 Dresden, Germany

2 The Chinese University of Hong Kong, Hong Kong, China

MAX-PLANCK-GESELLSCHAFT

Long-range correlations in time series

- Long-range correlations (LRC) is an important phenomenon found in many diverse fields (hydrology, climate, finance, DNA sequencing, networks, ...)
- Sum of autocorrelation function diverges $\sum_{s=1}^{\infty} C(s) = \infty$
- Autocorrelation function $C(s)$ decays slowly with time lag s
- Usually described by a power law with correlation exponent $0 < \gamma < 1$

$$C(s) \sim s^{-\gamma}$$

- A prominent data model is Fractional Gaussian Noise (FGN) with Hurst parameter $H = 1 - \gamma/2$ or, in discrete time an ARFIMA(0,d,0) process with $d = (1-\gamma)/2$
- Due to the lack of an internal time scale, LRC is often related to self-similar processes

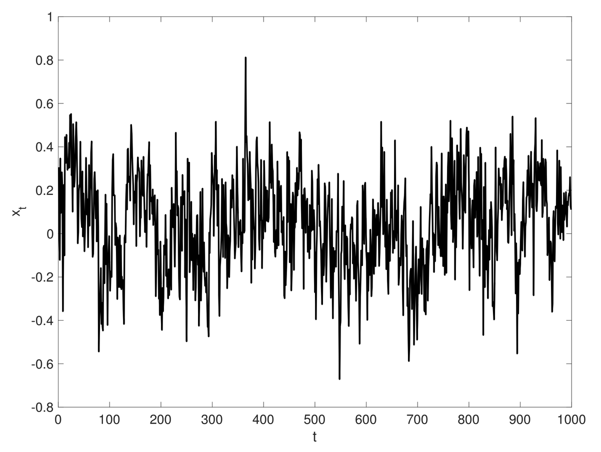

Figure 1: Single realisation of FGN with $H = 0.9$

Practical problems of the estimator of the autocorrelation function

- Given a time series $\{x_t\}_{t=1}^N$, the estimator of the autocorrelation function

$$\hat{C}(s) = \frac{N}{N-s} \frac{\sum_{t=1}^{N-s}(x_t - \bar{x})(x_{t+s} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2}$$

has two practical problems, especially for short time series:

(P1) Statistical uncertainty $\rightarrow$ difficult to observe power law scaling

(P2) Estimator only meaningful for stationary time series

($\mathcal{P}$2) Possible detection of LRC due to external trends

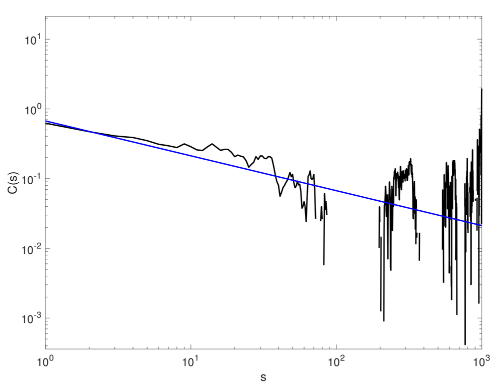

Figure 2: $C(s)$ of FGN with $H = 0.9$: Estimation (black) and theoretical (blue)

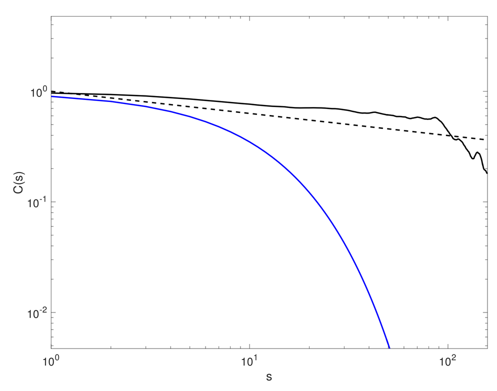

Figure 3: $C(s)$ of AR(1) with $a = 0.9$ and added linear trend with slope 0.02: Estimation (black), theoretical (blue) and reference line $s^{-0.2}$ (dashed)

The fluctuation function of a time series

- The fluctuation function as an integral transform of $C(r)$ with kernel L for stationary noise

$$F^2(s) = \int_0^s C(r) L(r,s) \qquad (1)$$

- We claim for the kernel L:

(S1) F increases according to a power law with fluctuation parameter $\alpha = \alpha(\gamma)$

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

(S2) External trends are deleted by L, only noise is taken into account

- Estimation of F via segmentation of the time axis into K segments ν of length s

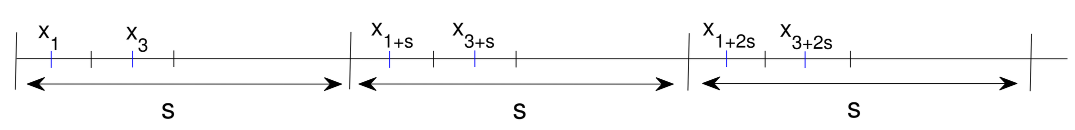

Figure 4: Estimation of Eq.(1)

- In each segment ν the kernel L acts on the data and yields pairs $x_{j+(\nu-1)s} x_{j+(\nu-1)s+r} L(r,s)$
- Averaging over all segments

$$C(r)L(r,s) \approx \frac{1}{K} \sum_{\nu=1}^{K} x_{j+(\nu-1)s} x_{j+(\nu-1)s+r} L(r,s)$$

Detrended fluctuation analysis (DFA)

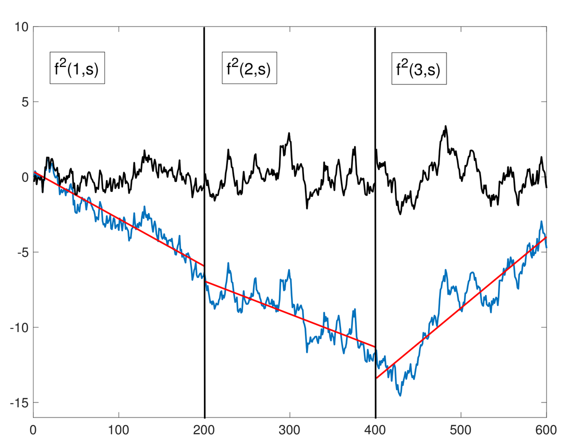

Figure 5: DFA1 with y_t and p_t

- Widely used method for the detection of LRC in nonstationary time series
- The fluctuation function of DFAm with profile $y_t = \sum_{k=1}^{t}(x_k - \bar{x})$ and locally fit p_t of order m is

$$\hat{F}^2(s) = \frac{1}{K}\sum_{\nu=1}^{K} f^2(\nu, s) = \frac{1}{K}\sum_{\nu=1}^{K}\left(\frac{1}{s}\sum_{t=1+(\nu-1)s}^{\nu s}(y_t - p_t)^2\right)$$

Relationship between F and C

(1) DFA as estimator of Eq. (1):

– Using the original series to express the locally detrended profile $y_t - p_t = \sum_{k=1+(\nu-1)s}^{\nu s} x_k \mathcal{P}_{t,k}^{(\nu)}$, the fluctuation function of DFAm is

$$\hat{F}^2(s) = \frac{1}{K}\sum_{\nu=1}^{s}\frac{1}{s}\left(x_{j+(\nu-1)s}^2 L(0,s)/2 + 2\sum_{r=1}^{s-1} x_{j+(\nu-1)s} x_{j+(\nu-1)s+r} L(r,s)\right)$$

with kernel

$$L(r,s) = \frac{2}{s}\sum_{k=1}^{s-r}\sum_{t=1}^{s}\mathcal{P}_{t,k}^{(1)}\mathcal{P}_{t,k+r}^{(1)} \text{ with } \mathcal{P}_{t,k}^{(1)} = \Theta(t-k) - \frac{1}{\det\mathbb{S}}\sum_{i=0}^{m} t^i \sum_{j=0}^{m}(\text{adj}\mathbb{S})_{i+1,j+1}\sum_{h=k}^{s} h^j$$

with adj$\mathbb{S}$ as adjugate matrix of the $(m+1)\times(m+1)$ matrix $\mathbb{S}$ with elements $S_{ij} = \sum_{k=1}^{s} k^{i+j-2}$

– For detrending order $m = 1$ the kernel is

$$L(r,s) = \frac{1}{15(s^4 - s^2)}\left(3r^5 + (-20s^2 + 5)r^3 + (-15s^4 + 35s^2 - 8)r + 2s(s^4 - 5s^2 + 4)\right)$$

(2) Taking the ensemble mean of the original $\hat{F}^2$ of DFA:

– Crucial condition: Variances in all segments are statistically equal

$$\langle f^2(\nu, s)\rangle = \langle f^2(\omega, s)\rangle$$

– Taking the ensemble mean of $\hat{F}^2$ using the crucial condition

$$\langle \hat{F}^2(s)\rangle = \frac{1}{K}\sum_{\nu=1}^{K}\langle f^2(\nu,s)\rangle = \langle f^2(1,s)\rangle = \frac{1}{s}\sum_{t=1}^{s}\langle (y_t - p_t)^2\rangle = \frac{1}{s}\sum_{t=1}^{s}\sum_{k,r=1}^{s}\langle x_k x_r\rangle \mathcal{P}_{t,k}^{(1)}\mathcal{P}_{t,r}^{(1)}$$

$$= \langle x^2\rangle\left(L(0,s)/2 + \sum_{r=1}^{s-1} C(r) L(r,s)\right) \text{ for stationary noise}$$

The fluctuation function of stationary noise

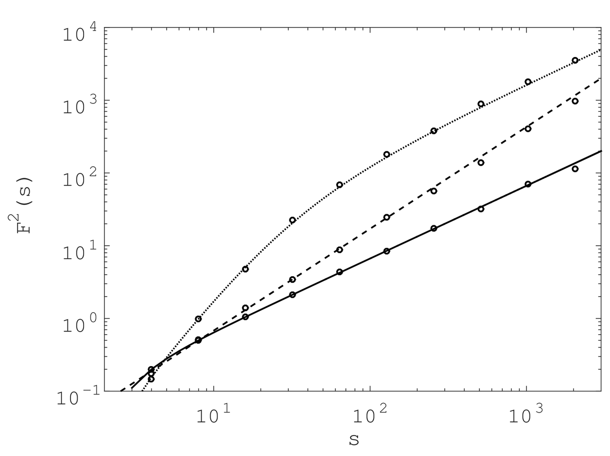

Figure 6: From bottom to top: white noise, ARFIMA(0,d,0) and AR(1) with $a = 0.8$

- The asymptotic behaviour of $F \sim s^{\alpha}$:
  - $\alpha = 1/2$ $\Rightarrow$ SRC with $C(s) \sim e^{-s/s_c}$
  - $1/2 < \alpha < 1$ $\Rightarrow$ LRC with $\alpha = 1 - \gamma/2$
- Crossover point for SRC $s_\times = c_m \frac{e^{1/s_c}}{e^{2/s_c - 1}}$ with $c_0 = 6$, $c_1 = 15$ and $c_2 = 70/3$

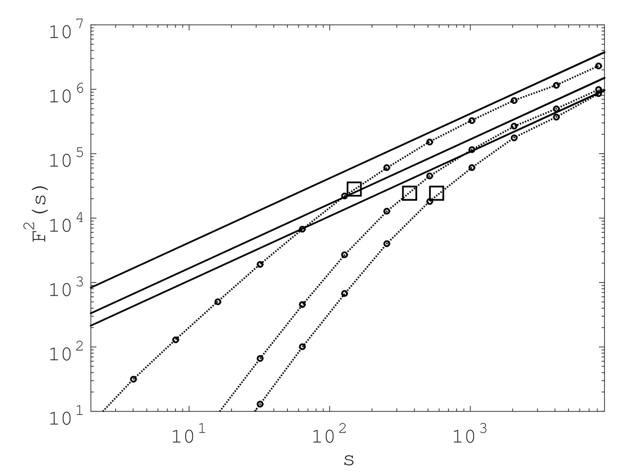

Figure 7: AR(1) with $a = 0.95$ and different order of detrending $m = 0, 1$ and 2 (from bottom to top)

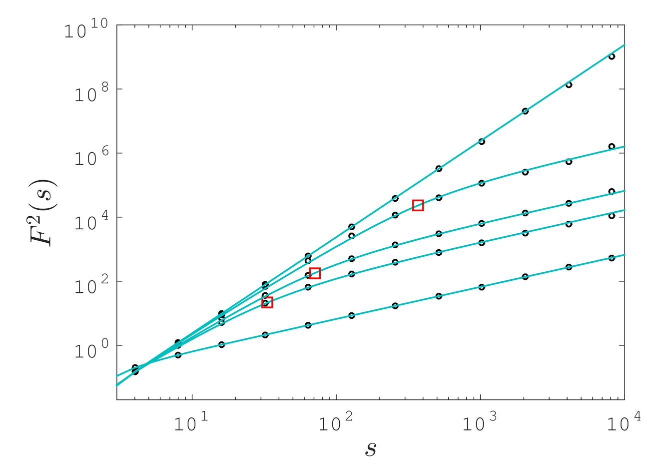

Figure 8: AR(1) with s_c (from bottom to top: 0, 0.45, 9.5, 49.5 and random walk) with $m = 1$

Scaling of intrinsic diffusion-like noise

- The validation of DFA for Fractional Brownian Motion (FBM) is based on fulfillment of the crucial condition:

	FGN	FGN + mt	FBM	$\langle f^2(1,s)\rangle$ for BM with $H = 1/2$
DFA0	$\alpha = H$	$\alpha = 1$	$\alpha = 1$	$\alpha = 3/2$
DFA1	$\alpha = H$	$\alpha = 1$	$\alpha = H + 1$	$\alpha = 3/2$
DFA2	$\alpha = H$	$\alpha = H$	$\alpha = H + 1$	not calculated

- Checking crucial condition directly not always possible $\rightarrow$ check autocovariance difference between them (here we compare segment $\nu = 1$ with 2)

$$D_{ij} = \langle x_{i+s} x_{j+s}\rangle - \langle x_i x_j\rangle, \text{ for BM it is } D_{ij} = s$$

- Then the ensemble mean of the second variance is $\langle f^2(2,s)\rangle = \langle f^2(1,s)\rangle + \langle \mathcal{F}(\mathbb{D})\rangle$
- Crucial condition fulfilled $\langle f^2(1,s)\rangle = \langle f^2(2,s)\rangle \iff \langle \mathcal{F}(\mathbb{D})\rangle = 0$
- For FBM the autocovariance difference is removed $\langle \mathcal{F}(\mathbb{D})\rangle = 0 \Rightarrow$ Crucial condition fulfilled
- DFA does not remove nonstationarity of FBM, it removes the difference between segments
- Crucial condition also fulfilled for stationary noise with added trend

$$\Longrightarrow$$

The crucial condition provides a unified perspective to refresh the understanding of how DFA works on nonstationarity

References

1. C.-K. Peng et al., Phys. Rev. E **49**, 1685 (1994)
2. M. Höll and H. Kantz, Eur. Phys. J. B **88**, 126 (2015)
3. M. Höll and H. Kantz, Eur. Phys. J. B **88**, 327 (2015)
4. M. Höll, Y. Zhou and H. Kantz, Phys. Rev. E, *submitted*

s