## Fluctuation analysis

See Matlab help for `fluctuationAnalysis.m`. We recommend using only this wrapper function!

*Definitions*

| | |
|---|---|
| $x(t)$ | stems from a fractional Gaussian noise (fGn) process |
| $y(t) = \sum_{s=1}^{t} x(s)$ | the cumulative sum of $x(t)$ resembles fractional Brownian motion (fBm) |

## Conventional detrended fluctuation analysis (DFA) – `fluctuationAnalysis(…,'DFA')`

Let $N$ denote the number of samples in the time series, that are split into $M$ non-overlapping segments $i = 1 \dots M$ of length $\Delta t$ each; $M = \lfloor N/\Delta t \rfloor$. Per segment $i$ the fluctuation strength is computed as squared difference between $y_i(t)$ and its trend $y_i^{\text{trend}}(t)$ (in the linear case this is just the regression line of $y_i(t)$ over the interval $t = 1 \dots \Delta t$:

$$F_i(\Delta t) = \sqrt{\frac{1}{n} \sum_{t=1}^{\Delta t} \left[ y_i(t) - y_i^{\text{trend}}(t) \right]^2} \quad (1)$$

This average fluctuation strength for segment size $n$ reads

$$\bar{F}(\Delta t) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} F_i^2(\Delta t)} \quad (2)$$

In the case of scale-free correlation this average value can be expressed as the power law:

$$\log \bar{F}(\Delta t) = \alpha \log \Delta t + C_0 \quad (3)$$

The scaling parameter $\alpha$ is the primary outcome measure of DFA. In the case of the scale-free processes with the aforementioned power law, $\alpha$ resembles the Hurst exponent leading to the interpretation:

| | |
|---|---|
| $0 < \alpha < 0.5$ | $x(t)$ – or $y(t)$ – contains anti-persistent fluctuations |
| $\alpha = 0.5$ | $x(t)$ is uncorrelated white Gaussian noise – or $y(t)$ resembles Brownian motion |
| $0.5 < \alpha < 1$ | $x(t)$ – or $y(t)$ – contains persistent fluctuations |
| $1 \leq \alpha$ | $x(t)$ is non-stationary (strictly speaking, DFA is only defined for $0 < \alpha < 1$! |

i. Compute $F_i(\Delta t)$ as in (1) after (linear) detrending $y(t)$ in every segment.
ii. Compute the mean $\bar{F}(\Delta t)$ as in (2) – see `detrendedFluctuationAnalysis.m` – this function also provides the confidence interval around $\bar{F}(\Delta t)$.
iii. Compute the regression line for (3): $\log \bar{F}(\Delta t) = \alpha \log \Delta t + C_0$ – we use `polyfit.m` – and also provide $R^2$ as goodness-of-fit estimator.

## DFA with non-parametric model selection – `fluctuationAnalysis(…,'DFA+')`

In order to test the power law (3) against alternatives one can use not only the mean value $\bar{F}(\Delta t)$ but evaluate the density of fluctuations over the consecutive segments, i.e. the density of $F_i(\Delta t)$, with which one can estimate the log-likelihood for a certain model to generate fluctuations (on a log-scale) as function of $\log \Delta t$.

i. Compute $F_i(\Delta t)$ as in (1)
ii. Estimate the density $p[F(\Delta t)]$ – we use a kernel source density estimator – see `detrendedDensities.m`
iii. Fit any arbitrary model into the density by maximizing the log-likelihood – here we use a nonlinear least squares solver. Outcome is the maximum log-likelihood $\mathcal{L}_{\max}$ and the corresponding model parameters that generate $\mathcal{L}_{\max}$.
Compute the AICc and BIC for the selection

$$\text{AIC}_c = -2 \log \mathcal{L}_{\max} + 2K + \frac{2K(K+1)}{M-K-1}$$
$$\text{BIC} = -2 \log \mathcal{L}_{\max} + K \log M$$

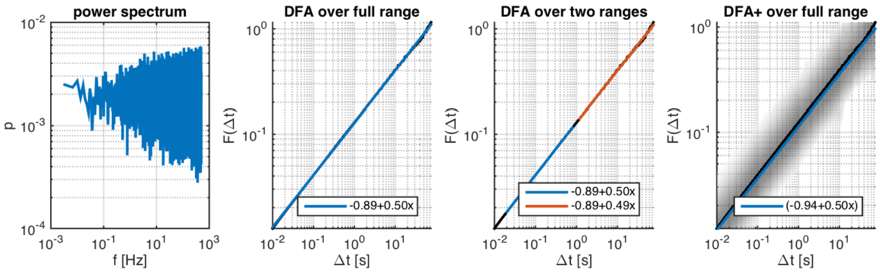
with $K$ being the number of parameter in the model and, recall, $M$ is the number of segments for estimating the density function – see `modelSelection.m`.
iv. Report the best fitting model (if the linear model is not the best one, then $\alpha = \text{NaN}$).

## **Examples** – `DFAexamples`

*Procedures*

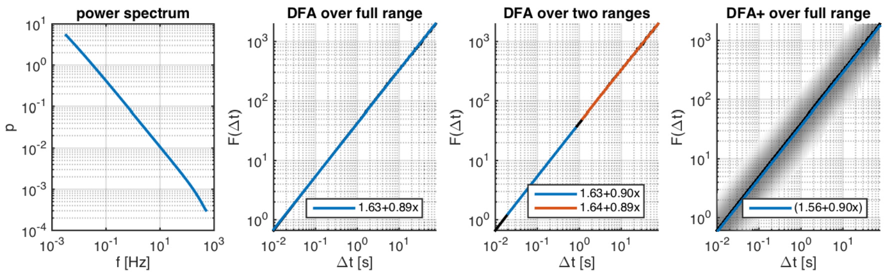
- Generate an fGn signal (design the power spectrum and exploit the Wiener-Khinchin theorem), subsequently randomize the Fourier phase (that preserves the power spectrum and, hence, the auto-correlation function as well as the mean square displacement).
- Plot the power spectrum (left = first panel; mean over trials)
- Plot $\bar{F}(\Delta t)$ (black) and the regression line (blue, conventional DFA; second panel)
- Plot $\bar{F}(\Delta t)$ (black) plus regression lines (colored) for two separate ranges (from 10ms to 800ms and 1200ms to ¼ of the recording time interactive correction of DFA; third panel)
- Plot $p[F(\Delta t)]$ (gray patches), the expectation value $\mathbb{E}[F(\Delta t)]$ (black) plus model fit (blue, DFA plus model comparison; right = fourth panel).

*A. Gaussian white noise*



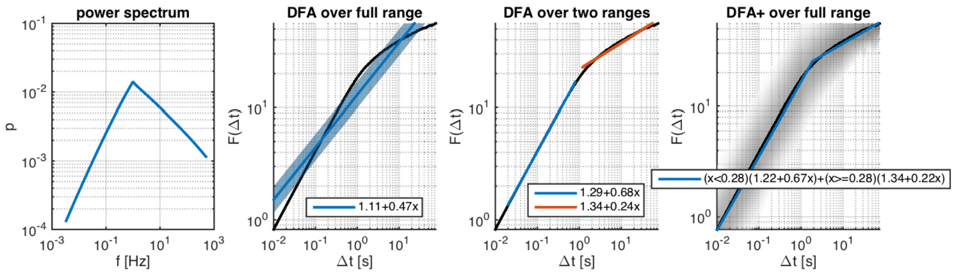

--> all algorithms seem to work well.

*B. Fractional Gaussian noise with Hurst exponent 0.9.*



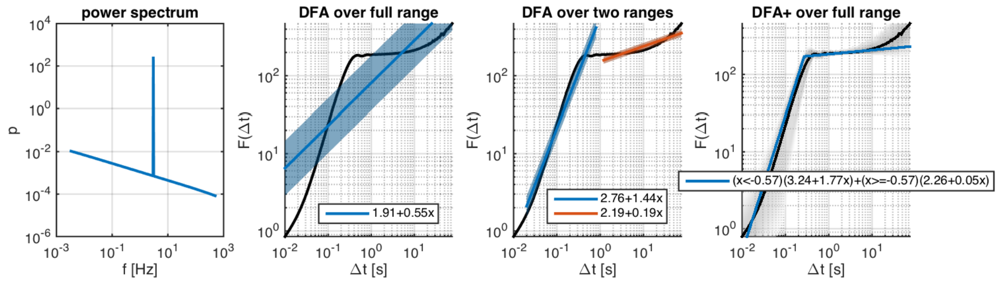

--> again all algorithms seem work well.

*C. Two-partite signal with Hurst exponent 0.7 until 1s, and Hurst exponent 0.05 afterwards.*




--> DFA over the full range give incorrect results, model selection finds a piecewise linear model.

*D. fGn with Hurst exponent 0.9 plus superimposed sinusoidal signal with frequency 3 1/3 Hz.*




--> DFA over the full range give bogus results, model selection finds a two-partite piecewise linear model; if recording duration is increase this will change to a three-partite model.

**Typical output for case C** - note that results depend on the rng-seed and other parameter settings which, in the code of DFAexample are chosen to suit the selected references – those are certainly not optimal from a data analysis perspective

*Sampling rate is fs =1 kHz, total duration is 300s.*

**DFA (whole range)**

```
fluctuationAnalysis: Elapsed time is 15.452894 seconds.
total range of assessment: epoch sizes [0.01 ... 75]
------------------------------------------------------------
fit results over epoch sizes [0.01 ... 75]
-> model: 1.11+0.47x (R2=0.944922)
------------------------------------------------------------
```

*We evaluate the range from 10ms to 75s to guarantee that (i) segments contain enough samples (low bound) and (ii) that we have enough independent segments (upper bound).*
*Result is clearly incorrect as the input signal does not display a single power law but rather two regimes.*

**DFA (2 ranges)**

```
fluctuationAnalysis: Elapsed time is 14.867622 seconds.
total range of assessment: epoch sizes [0.01 ... 75]
------------------------------------------------------------
fit results over epoch sizes [0.02 ... 0.8]
-> model: 1.29+0.68x (R2=0.999830)
------------------------------------------------------------
fit results over epoch sizes [1.2 ... 75]
-> model: 1.34+0.24x (R2=0.978504)
------------------------------------------------------------
```

*Same data but now we split 'by hand' the range of assessment in two, from 20ms to 800ms and from 1200ms to 75s.*
*The result fine: the initial part contains an $\alpha$-exponent of about 0.7 (persistent) followed by an $\alpha$-exponent of about 0.25*

**DFA + model selection**

```
fluctuationAnalysis: Elapsed time is 112.823717 seconds.
total range of assessment: epoch sizes [0.01 ... 75]
------------------------------------------------------------
fit results over epoch sizes [0.01 ... 75]
   | log-lh |   AICc |     BIC | model
------------------------------------------------------------
 1 |    248 |   -493 |    -485 | (c(1)+c(2)x)
 2 |   -238 |    480 |     488 | (c(1)+c(2)x^2)
 3 |    443 |   -881 |    -869 | (c(1)+c(2)x+c(3)x^2)
 4 |    -61 |    126 |     134 | (c(1)+c(2)x^3)
 5 |    324 |   -643 |    -631 | (c(1)+c(2)x+c(3)x^3)
 6 |    -62 |    131 |     142 | (c(1)+c(2)x^2+c(3)x^3)
 7 |    444 |   -879 |    -864 | (c(1)+c(2)x+c(3)x^2+c(4)x^3)
 8 |    248 |   -489 |    -477 | (c(1)+c(2)exp(c(3)x))
 9 |    374 |   -743 |    -731 | real((1/log(10))log(c(1)(1-exp(-c(2)exp(log(10)x)))+c(3)))
10 |    448 |   -888 |    -872 | (x<c(4))(c(1)+c(2)x)+(x>=c(4))(c(1)+(c(2)-c(3))c(4)+c(3)x)
11 |    448 |   -882 |    -855 | (x<c(5))(c(1)+c(2)x)+(x>=c(5)&x<c(6))(c(1)+(c(2)-c(3))c(5)+c(3)x)+(x>=c(6))(c(1)+(c(2)-
c(3))c(4)+(c(3)-c(4))c(6)+c(4)x)
------------------------------------------------------------
max. log-likelihood -> model 11: (x<-0.22)(1.23+0.68x)+(x>=-0.22&x<0.19)(1.21+0.56x)+(x>=0.19)(1.26+(0.56-0.24)0.19+0.24x)
minimum AICc                  -> model 10: (x<0.26)(1.22+0.67x)+(x>=0.26)(1.33+0.23x)
minimum BIC                   -> model 10: (x<0.26)(1.22+0.67x)+(x>=0.26)(1.33+0.23x)
------------------------------------------------------------
```

*DFA + model selection over the range 10ms to 75s; we include 11 models (1-8 are just polynomials starting from linear to cubic, 9 resembles an Ornstein-Uhlenbeck process, and 10-11 are piecewise linear models.*

*In this example both AICc and BIC favor the two-partite piecewise linear model.*

## Selected references

*General*

*General*


- Ton & Daffertshofer, Model selection for identifying power-law scaling, *Neuroimage* 136:215-26, 2016, doi:10.1016/j.neuroimage.2016.01.008.


*Posture related*


- Collins & De Luca, Random walking during quiet standing, *Phys. Rev. Lett.* 73(5):764, 1994.
- Newell, K et al., Stochastic processes in postural center-of-pressure profiles, *Exp. Brain. Res.* 113:158, 1997.


*Artifact related*


- Hu et al., Effect of trends on detrended fluctuation analysis, *Phys. Rev. E* 64, 011114, 2001.


## Notes

*The corresponding Matlab toolbox also provides a parametric model selection procedure, which assumes that the residual errors after model fit are normally distributed – this allows for estimating the log-likelihood via the simple sum of squares (cf. $R^2$ when fitting a polynomial). This procedure is faster but certainly not recommended because on cannot guarantee the normality of the residuals' density. In fact, in most if not all tested examples, this is not the case.*

## Supported functions

| | |
|---|---|
| `DFAexamples.m` | *Examples shown in this documents* |
| `DFAmoreExamples.m` | *More examples including the (failure of) parametric model selection* |
| | |
| `fluctuationAnalysis.m` | *Main function* |
| `reportDetails.m` | *Text report + graphics of results* |
| | |
| `diffusionAnalysis.m` | *Conventional diffusion analysis (SDA)* |
| `detrendedFluctuationAnalysis.m` | *Conventional detrended fluctuation analysis (DFA)* |
| | |
| `defineModelCatalogue.m` | *Set the model to compare* |
| `detrendedDensities.m` | *Estimate the fluctuation densities* |
| `modelSelection.m` | *Non-parametric model comparison (DFA+) (recommended)* |
| `modelSelectionGauss.m` | *Parametric model comparison (DFA-)* |
| | |
| `fftfgn.m` | *Generate an fGn sequence with given Hurst exponent* |
| `psd2signal.m` | *Generate a signal with given power spectrum* |
| `psdfgn.m` | *Generate an fGn with given (set of) Hurst exponent* |
| `randomizeFourierPhase.m` | *Randomize the Fourier phase* |