Statistical investigation of anomalous diffusion processes

Agnieszka Wyłomańska

Faculty of Pure and Applied Mathematics
Hugo Steinhaus Center
Wroclaw University of Science and Technology

Agenda

Agenda

- Introduction and motivation

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes
    a) FBM

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes
  a) FBM
  b) FLSM

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes
  a) FBM
  b) FLSM
  c) CTRW

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes
  a) FBM
  b) FLSM
  c) CTRW
- Selected tools for anomalous diffusion recognition

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes
  a) FBM
  b) FLSM
  c) CTRW
- Selected tools for anomalous diffusion recognition
  a) Distribution testing

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes
  a) FBM
  b) FLSM
  c) CTRW
- Selected tools for anomalous diffusion recognition
  a) Distribution testing
  b) TAMSD-based method

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes
  a) FBM
  b) FLSM
  c) CTRW
- Selected tools for anomalous diffusion recognition
  a) Distribution testing
  b) TAMSD-based method
  c) Codifference - based method

Agenda

- Introduction and motivation
- Selected anomalous diffusion processes
  a) FBM
  b) FLSM
  c) CTRW
- Selected tools for anomalous diffusion recognition
  a) Distribution testing
  b) TAMSD-based method
  c) Codifference - based method
- Conclusions

Introduction and motivation

- In recent years, anomalous diffusion has been identified as a ubiquitous phenomenon in complex and biophysical systems.
- The issue of distinguishing between normal and anomalous diffusion concerns many fields of physics.
- Most of the methods are based on the mean square displacement (MSD) analysis.
- Normal (classical) diffusion: $< X^2(t) > \sim t$
- Anomalous diffusion: $< X^2(t) > \sim t^\beta$ ($\beta < 1$ - subdiffusion, $\beta > 1$ - superdiffusion).
- It is not clear which model applies to a particular system. Information which is essential when diffusion-controlled processes are considered.

Introduction and motivation

- The MSD can be obtained either by performing an average over an ensemble of particles, or by taking the temporal average over a single trajectory (time average MSD, TAMSD).

$$M_N(\tau) = \frac{1}{N-\tau} \sum_{j=1}^{N-\tau} (X(j+\tau) - X(j))^2 \tag{1}$$

- Recent advances in single molecule spectroscopy enabled single particle tracking experiments following individual particle trajectories. These require temporal moving averages.

Anomalous diffusion processes - FBM

- FBM is a generalization of classical Brownian motion.
- Most of its statistical properties are characterized by Hurst exponent $0 < H < 1$.
- FBM has found may applications.
- FBM is zero-mean Gaussian process

$$B_H(t) = \int_{-\infty}^{\infty} \left\{ (t-u)_+^{H-1/2} - (-u)_+^{H-1/2} \right\} dB(u).$$

$(x)_+ = max(x, 0)$.

- FBM is $H-$self similar $B_H(ct) \stackrel{d}{=} c^H B_H(t)$, $c > 0$.

Anomalous diffusion processes - FBM

- FBM has stationary increments.
- For the second moment $\langle B_H^2(t) \rangle = \sigma^2 t^{2H}$, $\sigma > 0$.
- For $H < 1/2$ FBM gives the subdiffusive dynamics, for $H > 1/2$ FBM gives the superdiffusive one.
- For $H > 1/2$ the increments are positively correlated and exhibit long-range dependence.
- For $H < 1/2$ the increments are negatively correlated and exhibit short-range dependence.
- Fractional Gaussian noise (FGN) $b_H(t) = dB_H(t)/dt$.

Anomalous diffusion processes - FLSM

- FLSM is a generalisation of FBM
- $$L_H^\alpha(t) = \int_{-\infty}^{\infty} \left\{ (t-u)_+^{H-1/\alpha} - (-u)_+^{H-1/\alpha} \right\} dL_\alpha(u),$$
  where $L_\alpha(t)$ is a Lévy $\alpha-$stable motion $0 < \alpha \leq 2$ and $0 < H < 1$
- FLSM is $\alpha-$stable.
- For $\alpha = 2$ it becomes FBM.
- FLSM is $H-$self similar and has stationary increments.
- For $H > 1/\alpha$ FLSM exhibits positive, long-range dependence.
- For $H < 1/\alpha$ FLSM exhibits negative dependence.
- Fractional Lévy stable noise (FLSN) $l_H^\alpha(t) = dL_H^\alpha(t)/dt$.

Anomalous diffusion processes - CTRW

The CTRW model is defined as:

$$Y(t) = \sum_{i=1}^{N(t)} X_i,$$

where $N(t)$ is a a counting process

$$N(t) = \max\left\{k \geq 0 : \sum_{i=1}^{k} T_i \leq t\right\}.$$

Here T_i $i = 1, 2, ...$ form a sequence of positive i.i.d. random variables which can be seen as waiting times between consecutive jumps X_i.

Anomalous diffusion processes - CTRW

If we assume the waiting time sequence constitutes sample from $\alpha-$stable distribution with index of stability $\alpha \in (0,1)$ while jumps are i.i.d. random variables with finite moments then

$$\frac{Y(nt)}{n^{\alpha/2}} \xrightarrow{d} B(S_{\alpha}^{-1}(t)),$$

where $B(\cdot)$ is a standard Brownian motion and $S_{\alpha}^{-1}(\cdot)$ is inverse stable subordinator defined as

$$S_{\alpha}^{-1}(t) = \inf\left\{\tau : \ U(\tau) > t\right\}.$$

$\{U(\tau)\}$ denotes strictly increasing stable Lévy motion, i.e., a Lévy stable process with stationary independent increments of the following Laplace transform

$$E(e^{-kU(\tau)}) = e^{-\tau k^{\alpha}}, \ 0 < \alpha < 1.$$

Anomalous diffusion processes - CTRW

The CTRW with heavy-tailed waiting times is described also by FFPE

$$\frac{\partial w(x,t)}{\partial t} = {}_0 D_t^{1-\alpha} \left[\frac{1}{2} \frac{\partial^2}{\partial x^2} \right] w(x,t),$$

$w(x,0) = \delta(x)$. The operator ${}_0D_t^{1-\alpha}$, $0 < \alpha < 1$ is the fractional derivative of the Riemann-Liouville type.

The MSD of CTRW equals $\frac{t^\alpha}{\Gamma(\alpha+1)}$, which is characteristic for subdiffusive dynamics.

Tools for anomalous diffusion recognition - distribution testing

- We assume data are stationary (like increments of FBM or FLSM).
- Problem: testing class of distribution - heavy or light tailed.
- Recognition of stable distribution with Lévy index close to 2.
- Discriminating between light- and heavy-tailed distributions with limit theorem.

Recognition of stable distribution with Lévy index close to 2

- We address the problem of recognizing alpha-stable Lévy distribution with Lévy index close to 2 from experimental data.
- We are interested in the case when the sample size of available data is not large, thus the power law asymptotics of the distribution is not clearly detectable, and the shape of empirical probability density function is close to a Gaussian.
- We propose a testing procedure combining a simple visual test based on empirical fourth moment with the Anderson-Darling and Jarque-Bera statistical tests.
- We apply our method to the analysis of turbulent plasma data.

Recognition of stable distribution with Lévy index close to 2

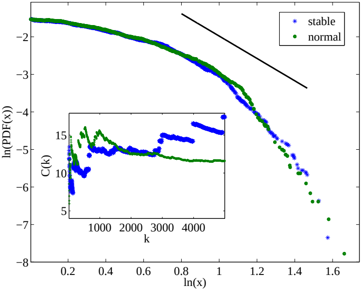

Figure: **The two empirical PDFs calculated for simulated samples from Lévy stable distribution with parameters $\alpha = 1.98$, $\beta = 0$, $\sigma = 1$ and $\mu = 0$, and from Gaussian distribution with $\mu = 0$ and $\sigma^2 = 2$. Inset: Empirical cumulative fourth moment for the samples considered.**

Recognition of stable distribution with Lévy index close to 2

- The method is based on the empirical fourth moment and Anderson-Darling and Jarque-Bera statistical tests.
- The empirical fourth moment

$$C(k) = \frac{1}{k}\sum_{i=1}^{k}(x_k - \overline{x})^4, \quad k = 1, 2, ..., n$$

- AD statistic, similar as KS, measure the distance between empirical and theoretical (tested) CDFs.
- For testing the Gaussianity we propose to use the standard Jarque-Bera (JB) test. The JB statistic is defined as

$$J = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right). \tag{2}$$

Recognition of stable distribution with Lévy index close to 2

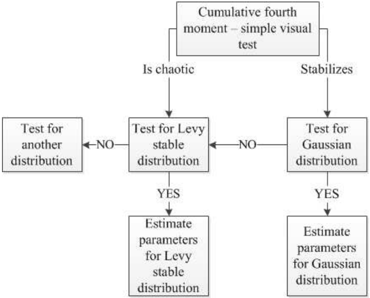

Figure: **Schematic algorithm for recognition of Lévy stable distribution with Lévy index close to 2.**

Recognition of stable distribution with Lévy index close to 2

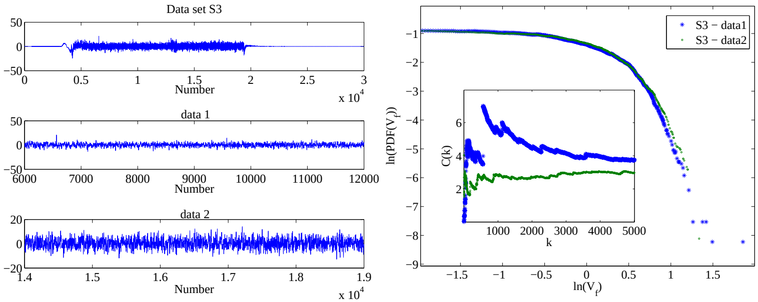

Figure: **The empirical time series from plasma physics (left panel). The right tail of the empirical PDFs and ECFMs for the examined dataset (right panel). Data S2 are Gaussian, data S1 - stable with $\alpha = 1.98$.**

Discrimination algorithm - introduction

We consider the normalized sum of n continuous i.i.d. random variables X_i, $i = 1, \dots, n$ with CDF $F(x)$ and PDF $f(y) = F'(y)$:

$$Y_n = \frac{1}{B_n} \sum_{i=1}^{n} (X_i - A_n).$$

- When F has first and second moments m_1 and m_2, we set $B_n = n^{1/2}$ and $A_n = m_1$ and by CLT, Y_n converges to $N(0, m_2)$.
- If F has a third moment m_3, then $f_n(y) - f(y) = O(n^{-1/2})$, where $f_n(y)$ is the density function of Y_n.
- If $m_3 = 0$, then the rate of convergence is $o(n^{-1/2})$.
- if F does not have a third moment, but if $F(x) = O(|x|^{\alpha})$ as $x \to -\infty$ and $F(x) = 1 - O(x^{\alpha})$ as $x \to \infty$ with $2 < \alpha \leq 3$, then $f_n(y) - f(y) = O(n^{-(\alpha-2)/2})$.

Discrimination algorithm - introduction

When F does not have both first and second moments, the distribution of the Y_n may still converge. A necessary and sufficient condition for this is

$$F(x) = \begin{cases} (c_1 + r_1(x))|x|^{-\alpha} & \text{if } x < 0, \\ (c_2 + r_2(x))|x|^{-\alpha} & \text{if } x > 0, \end{cases}$$

with $0 < \alpha \leq 2$, c_1 and c_2 positive constants, $r_1(x) \to 0$ as $x \to -\infty$ and $r_2(x) \to 0$ as $x \to \infty$. When this condition holds and $0 < \alpha < 2$ we can set $B_n = n^{1/\alpha}$ and by the Generalized Central Limit Theorem the limit is a stable distribution.

Discrimination algorithm

We consider two samples of observations of length N:
$\{x_1, x_2, \ldots, x_N\}$ and $\{y_1, y_2, \ldots, y_N\}$.

1. We divide the dataset into non-overlapping consecutive blocks of length $K = 1, 2, \ldots, 10$. Next, we sum the values within each block and obtain aggregated data of length $[N/K]$. Finally, we estimate the index of stability α for the constructed data via the regression method.
2. We plot the estimated index of stability with respect to $K = 1, 2, \ldots, 10$.
    1. If the estimated values converge to 2, then the data are light-tailed and belong to the domain of attraction of the Gaussian law.
    2. If the estimated values converge to $\alpha < 2$, then the data are heavy-tailed and belong to the domain of attraction of the non-Gaussian stable law.

Discrimination algorithm

- One can enhance this procedure by calculating box plots for the estimated values α. This is intended to help to access if the differences in convergence are statistically justified. The box plot provides a statistical information about the distribution of the values.
- But, how to create box plots from a single dataset? The idea is to generate more samples from one sample (from its empirical distribution function). This procedure is called bootstrapping in statistics. The bootstrapping is done for the whole dataset.

Example. Gaussian vs. non-Gaussian stable with α close to 2

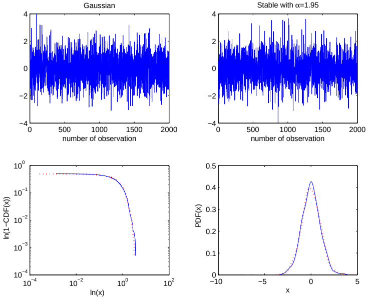

Figure: **Simulated samples from the Gaussian distribution with $\mu = 0$ and $\sigma = \sqrt{2}$ (top left panel) and the symmetric stable distribution with $\alpha = 1.95$ and $\sigma = 1$ (top right panel), and their empirical tails (1-CDFs) in log-log scale (bottom left panel) and PDFs (bottom right panel).**

Example. Gaussian vs. non-Gaussian stable with α close to 2

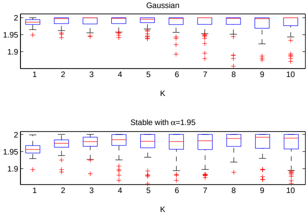

Figure: **Estimated α values for the Gaussian sample from the top left panel of Fig. 4 (top panel) and for the symmetric non-Gaussian stable sample from the top right panel of Fig. 4 (bottom panel).**

Classic two-sample Kolmogorov-Smirnov test does not reject the null hypothesis of common distributions, with p-value equal to 0.44.

Example. Student's t vs. stable

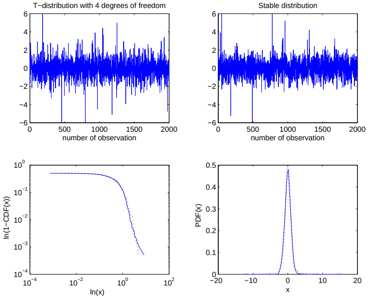

Figure: **Simulated samples from the Student's t distribution with 4 degrees of freedom (top left panel) and stable distribution with $\alpha = 1.85$, $\sigma = 0.77$, $\beta = 0.15$, and $\mu = 0.01$ (top right panel), and their empirical tails in log-log scale (bottom left panel) and PDFs (bottom right panel).**

Example. Student's t vs. stable

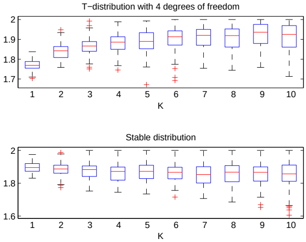

Figure: **Estimated α values for the Student's t sample from the top left panel of Fig. 6 (top panel) and for the stable sample from the top right panel of Fig. 6 (bottom panel).**

The two-sample Kolmogorov-Smirnov test does not reject the null hypothesis of common distributions, with p-value equal to 0.2.

Discrimination algorithm

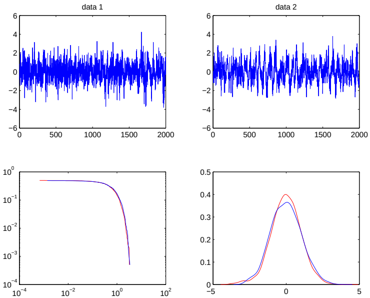

Figure: **Plasma data for the torus radial position $r = 9.5$: data1 (top left panel), data2 (top right panel), and their empirical tails (bottom left panel) and PDFs (bottom right panel).**

Discrimination algorithm

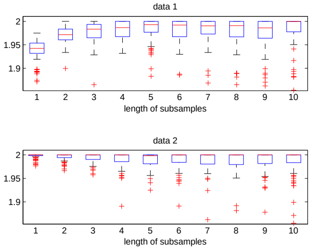

Figure: **Estimated α values for the data1 (top panel) and for the data2 (bottom panel). The box plots were constructed from 100 bootstrap samples of length 2000. Data1 -stable with $\alpha = 1.94$, data 2 - Gaussian.**

Tools for anomalous diffusion recognition - TAMSD - based method

- Test for FBM based on the TAMSD
- From the theory of quadratic forms for Gaussian vectors it is known that $(N-\tau)M_N(\tau)$ for FBM has the generalized χ^2 distribution

$$(N-\tau)M_N(\tau) \stackrel{d}{=} \sum_{j=1}^{N-\tau} \sigma\lambda_j(H,\tau)U_j,$$

where U_j's are i.i.d. chi-squared with 1 degree of freedom random variables and weights $\lambda_j(H,\tau)$ are the eigenvalues of the matrix $\tilde{\Sigma}(D,H,\tau)$ responsible for the covariance of the vector $\{B_H(1+\tau)-B_H(1), B_H(2+\tau)-B_H(2), \ldots, B_H(N)-B_H(N-\tau)\}$.

Tools for anomalous diffusion recognition - TAMSD - based method

- If we know the exact distribution (and hence quantiles) of the sum $\sum_{j=1}^{N-\tau} \sigma\lambda_j(H,\tau)U_j$ we are able to construct the confidence interval
$$M_N(\tau) \in \left[\frac{\sigma Q_{c/2}}{N-\tau}, \frac{\sigma Q_{1-c/2}}{N-\tau}\right] \quad (3)$$
with confidence level $1-c$, where Q_p is the quantile of order p of generalized chi-squared distribution.
- Such results are the basis for construction of a new rigorous statistical test for FBM.
- The null hypothesis (H_0) is that the analyzed vector of observations $\{X(1), X(2), ..., X(N)\}$ forms a trajectory of FBM with given σ and H parameters.

Tools for anomalous diffusion recognition - TAMSD - based method

- We suggest to apply a two-sided test. In the two-sided case we reject H_0 hypothesis if the test statistic is extreme, either larger than an upper critical value or smaller than a lower critical value with a given significance level c (with probability c it's inside the critical region or equivalently outside the acceptance region).
- The acceptance region of the test is given by equation (3).
- If $M_N(\tau)$ falls into the acceptance region with given significance level α, we say that the test does not reject the FBM hypothesis with self-similarity index H, otherwise we reject the null hypothesis of FBM with H. We assume σ is known.

Tools for anomalous diffusion recognition - TAMSD - based method

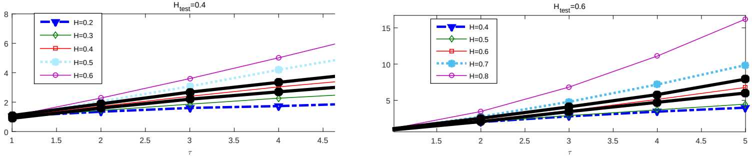

Figure: **Quantiles of $M_N(\tau)$ calculated on the basis of 500 simulations (thick dotted line) and mean TAMSD values (calculated on the basis of 500 realizations) for FBM for different H's. On the top panel the reference value $H = 0.4$ and on the bottom panel $H = 0.6$.**

Tools for anomalous diffusion recognition - TAMSD - based method

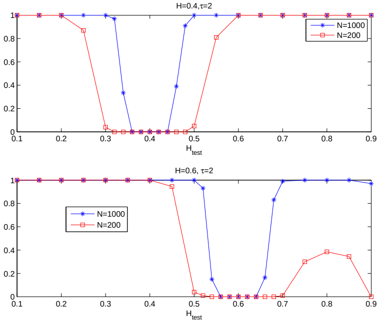

Figure: **Power of the test calculated from the test statistic $M_N(2)$ for FBM with $H = 0.4$ (top panel) and $H = 0.6$ (bottom panel) for $N = 200$ and $N = 1000$.**

Tools for anomalous diffusion recognition - TAMSD - based method

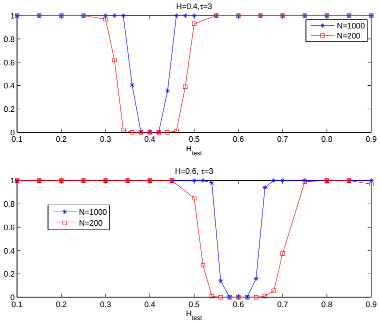

Figure: **Power of the test calculated from the test statistic $M_N(3)$ for FBM with $H = 0.4$ (top panel) and $H = 0.6$ (bottom panel) for $N = 200$ and $N = 1000$. For CTRW power is 1 for all values of H_{test}.**

Tools for anomalous diffusion recognition - TAMSD - based method

- The knowledge about exact distribution of TAMSD was a starting point to calculate the exact distribution of estimator anomalous diffusion exponent for FBM (β)

$$\hat{\beta} \stackrel{d}{=} \frac{\sum_{i=1}^{n} \log(\tau_i) \log\left(\frac{(N-1)\sum_{j=1}^{N-\tau_i} \lambda_j(\tau_i) U_j}{(N-\tau_i)\sum_{k=1}^{N-1} \lambda_k(1) U_k}\right)}{\sum_{i=1}^{n} \log^2(\tau_i)},$$

- Main properties of $\hat{\beta}$ are calculated: it is asymptotically unbiased, its variance tends to zero, it is consistent.
- This can help to test if real data can by modeled by using FBM.

Tools for anomalous diffusion recognition - codifference

- Correlation and spectral analysis represent the standard tools to study interdependence in statistical data.
- For the stochastic processes with heavy-tailed distributions such that the variance diverges, these tools are inadequate.
- Stochastic processes with diverging variance, like alpha-stable Lévy motion, FLSM or Lévy flights, are ubiquitous in nature and finance.
- What is the measure of interdependence for the processes with infinite variance?
- The alternative measures of dependence are rarely discussed in application-oriented literature.

Tools for anomalous diffusion recognition - codifference

- There are know few alternative measures of dependence.
- We analyze codifference (CD).
- It is based on the characteristic function of a given process, therefore it can be used not only for alpha-stable processes.
- The codifference in the Gaussian case reduces to the classical covariance, so it can be treated as the natural extension of the well-known measure.
- It is easy to evaluate the empirical codifference which is based on the empirical characteristic function of the analyzed data.
- The codifference is closely related to the so-called dynamical functional used to study ergodic properties of stochastic processes.

Tools for anomalous diffusion recognition - codifference

We introduce the codifference in terms of the characteristic function. We recall, that the characteristic function always exists for any real-valued random variable and determines the probability distribution in a unique way. The characteristic function of random variable X is defined as follows:

$$\Phi_X(k) =< \exp(ikX) >\equiv \int_{-\infty}^{\infty} \exp(iky) f_X(y) dy, \tag{4}$$

The codifference of two jointly SαS, $0 < \alpha \leq 2$, random variables X and Y is defined as follows:

$$CD(X, Y) = \sigma^{\alpha}_{X-Y} - \sigma^{\alpha}_{X} - \sigma^{\alpha}_{Y}, \tag{5}$$

where σ_X, σ_Y and σ_{X-Y} denote, respectively the scale parameters of X, Y and $X - Y$.

Tools for anomalous diffusion recognition - codifference

The codifference can be also defined in the language of the characteristic function:

$$CD(X, Y) = \ln(< exp\{iX\} >) + \ln(< exp\{-iY\} >) - \ln(< exp\{i(X-Y)\} \tag{6}$$

Thus, the definition given in Eq.(5) can be extended to a more general class of random variables, and in the further analysis we use the representation given in (6).

Tools for anomalous diffusion recognition - codifference

The codifference possesses several useful properties:

- it is always well-defined, since the definition of $CD(X,Y)$ is based on the characteristic functions of appropriate random variables;
- if the random variables are symmetric, then $CD(X,Y) = CD(Y,X)$;
- if X and Y are independent and jointly SαS, then $CD(X,Y) = 0$ for $0 < \alpha \leq 2$. On the other hand, $CD(X,Y) = 0$ implies that X and Y are independent for $0 < \alpha < 1$ and in the Gaussian case $\alpha = 2$. When $1 < \alpha < 2$, $CD(X,Y) = 0$ does not imply that X and Y are independent.

Tools for anomalous diffusion recognition - codifference

The above properties confirm that the codifference is an appropriate mathematical tool for measuring the dependence between alpha-stable random variables as well as random variables from more general class of distributions (e.g., infinitely divisible). In the literature one can also find the generalized codifference which is defined as:

$$GCD(X, Y; \theta_1, \theta_2) =$$

$$\ln(< exp\{i\theta_1 X\} >) + \ln(< exp\{i\theta_2 Y\} >) - \ln(< exp\{i(\theta_1 X + \theta_2 Y)\} >),$$

where $\theta_1, \theta_2 \in R$.

Tools for anomalous diffusion recognition - codifference

For a stochastic process $\{X(t)\}$, the measure $CD(X(t), X(s))$ called autocodifference is defined as:

$$CD(X(t), X(s)) = \ln(< exp\{iX(t)\} >) +$$

$$\ln(< exp\{-iX(s)\} >) - \ln(< exp\{i(X(t) - X(s))\} >).$$

Tools for anomalous diffusion recognition - codifference

- For Gaussian based processes
  $CD(X(t), X(s)) = -cov(X(t), X(s))$
- CD is especially important for infinite variance processes (the covariance does not exists).
- FBM

$$CD(B_H(t), B_H(s)) = \frac{k(H)}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}).$$

- FGN

$$CD(b_H(t), b_H(s)) = \frac{k(H)}{2}(|t-s+1|^{2H} - 2|t-s|^{2H} + |t-s-1|^{2H}).$$

Tools for anomalous diffusion recognition - codifference

- FLSM

$$CD(L_H^\alpha(t), L_H^\alpha(s)) = k(H,\alpha)\left(|t-s|^{\alpha H} - |t|^{\alpha H} - |s|^{\alpha H}\right).$$

- FLSN

For large t the autocodifference has a power law form:

If either $0 < \alpha \le 1$, $0 < H < 1$ or $1 < \alpha < 2$, $1 - \frac{1}{\alpha(\alpha-1)} < H < 1$, $H \neq 1/\alpha$, then for $t \to \infty$

$$CD(l_H^\alpha(t), l_H^\alpha(0)) \sim Ct^{\alpha H - \alpha}.$$

If $1 < \alpha < 2$ and $0 < H < 1 - \frac{1}{\alpha(\alpha-1)}$, then for $t \to \infty$

$$CD(l_H^\alpha(t), l_H^\alpha(0)) \sim Dt^{H-1/\alpha-1}.$$

How to estimate codifference from real data

We define an estimator of autocodifference in the form

$$\widehat{CD}(X(t), X(s)) = \ln(\hat{\phi}(1, 0, X(t), X(s))) + \ln(\hat{\phi}(0, -1, X(t), X(s))) \\ - \ln(\hat{\phi}(1, -1, X(t), X(s))),$$

where $\hat{\phi}(u, v, X(t), X(s))$ is an estimator of the characteristic function:

$$\hat{\phi}(u, v, X(t), X(s)) =< \exp\{i(uX(t) + vX(s))\} > . \qquad (7)$$

If $\{x_k, k = 1, ..., N\}$ is realization of a stationary process $\{X(t)\}$, then the estimator of the characteristic function takes the form:

$$\hat{\phi}(u, v, X(t), X(s)) = \frac{1}{N} \sum_{k=1}^{N-|t-s|} \exp(i(ux_{k+|t-s|} + vx_k)). \qquad (8)$$

How to estimate codifference from real data

For a nonstationary process we are not able to estimate empirical autocodifference by using only a single trajectory, therefore the above estimator requires modification. Suppose, we have M trajectories of a nonstationary process $\{X(t)\}$. Let us take a sample $\{x_k^t, k = 1, ..., M\}$ being a realization of a random variable $X(t)$, that is the values of the process $\{X(t)\}$ taken at a fixed time t, and a sample $\{x_k^s : k = 1, ..., M\}$ composed from the values of the process $\{X(t)\}$ taken at a fixed time s. By construction, both samples consist of independent identically distributed random variables. Thus, in the nonstationary case the estimator of characteristic function $\phi(u, v, X(t), X(s))$ is defined as:

$$\hat{\phi}(u, v, X(t), X(s)) = \frac{1}{M} \sum_{k=1}^{M} \exp(i(ux_k^t + vx_k^s)). \tag{9}$$

How to estimate codifference from real data

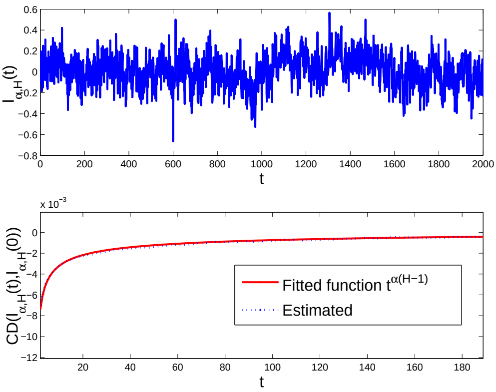

Figure: **The trajectory of fractional Lévy noise $I_H^{\alpha}(t)$ with parameters $\alpha = 1.95, H = 0.8$ (top panel), together with the estimator of autocodifference $CD(I_{\alpha,H}(t), I_{\alpha,H}(0))$ and fitted power function $t^{\alpha(H-1)}$ (bottom panel).**

Real data analysis - plasma data

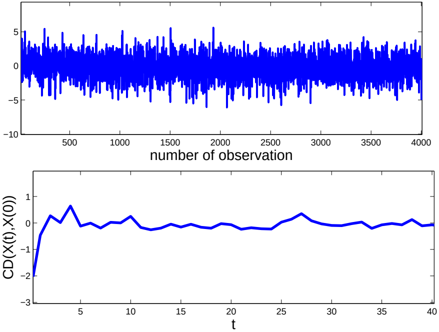

Figure: **The examined time series of the ion saturation current fluctuations (in mA) registered in the U-3M torsatron at the small torus radial positions $r = 9.9cm$ (top panel), and the estimator of autocodifference (bottom panel).**

Real data analysis - plasma data

- We propose to model the process by using white Lévy noise. We confirm this assumption by applying the Anderson-Darling goodness-of-fit test which indicates that the data come from the Lévy stable distribution (p-value is equal to 0.88).
- Estimated $\hat{\alpha} = 1.95$.
- On the basis of autocodifference we can also estimate the scale parameter σ of the analyzed series. As the result we obtain $\hat{\sigma} = 1.05$. We then compare this value with the estimates of σ parameter made by regression and McCulloch methods. The obtained values are $\hat{\sigma} = 1.06$ and $\hat{\sigma} = 1.03$, respectively, which is in good agreement with the value obtained from the autocodifference.

Conclusions

- Anomalous diffusion phenomena is visible in different data.
- The classical anomalous diffusive models are: FBM, FLSM, CTW.
- It is important to recognize the theoretical model appropriate for real data data.
- The presented methods are useful in testing anomalous diffusion behavior for real data.
- In the literature one can find different approaches useful in this problem.

References

- Sikora G., Teuerle M., Wylomanska A., Grebenkov D.: Statistical properties of the anomalous scaling exponent estimator based on time averaged mean square displacement, Phys. Rev. E 96, 022132, 2017
- Sikora G., Burnecki K., Wylomanska A.: Mean-squared displacement statistical test for fractional Brownian motion, Phys. Rev. E 95, 032110, 2017
- Burnecki K., Wylomanska A., Chechkin A.: Discriminating between light- and heavy-tailed distributions with limit theorem, PLoS ONE 10(12): e0145604. doi:10.1371/journal.pone.0145604, 2015

References

- Wylomanska A., Chechkin A., Sokolov I. M., Gajda J.: Codifference as a practical tool to measure interdependence, Physica A 421, 412-429, 2015
- Burnecki K., Wylomanska A., Beletskii A., Gonchar V., Chechkin A.: Recognition of stable distribution with Levy index alpha close to 2, Phys. Rev. E 85, 056711, 2012.

Thank you for your attention!